The younger age profile of COVID-19 deaths in developing countries*

Juan Pablo Chauvin[†1], Annabelle Fowler[2], and Nicolás Herrera L.[3]

[1,3]*Research Department, Inter-American Development Bank*
[2]*Interfaculty Initiative in Health Policy, Harvard University*

October 30, 2020

Abstract

This paper examines why a larger share of COVID-19 deaths occurs among young and middle-aged adults in developing countries than in high-income countries. Using novel data at the country, city, and patient levels, we investigate the drivers of this gap in terms of the key components of the standard Susceptible-Infected-Recovered framework. We obtain three main results. First, we show that the COVID-19 mortality age gap is not explained by younger susceptible populations in developing countries. Second, we provide indirect evidence that higher infection rates play a role, showing that variables linked to faster COVID-19 spread – such as residential crowding and labor informality – are correlated with younger mortality age profiles across cities. Third, we show that lower recovery rates in developing countries account for nearly all of the higher death shares among young adults, and for almost half of the higher death shares among middle-aged adults. Our evidence suggests that lower recovery rates in developing countries are driven by a higher prevalence of preexisting conditions that have been linked to more severe COVID-19 complications, and by more limited access to hospitals and intensive care units in some countries.

*We are grateful to Daniel Da Mata, Julia Dennett, Stephen Marcus, and an anonymous reviewer for helpful comments and suggestions, to Julio Trecenti and Haydée Svab for providing outstanding research support, and to Raveena Sarwal for her assistance with data collection. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

[†]Corresponding author. Email: juancha@iadb.org.

1 Introduction

Researchers have widely documented that the risk of developing severe complications and dying from COVID-19 increases sharply with age (Davies et al., 2020; Levin et al., 2020). While early work found mortality age profiles to be very similar among high-income nations (Goldstein and Lee, 2020), recent research suggests that younger cohorts face a relatively higher risk of dying from the disease in developing countries than in rich countries (Demombynes, 2020; Laxminarayan et al., 2020). This paper confirms this finding, and concludes that the the gap is the result of both higher rates of infection, and lower rates of recovery. These findings expand our understanding of how age-specific vulnerability to the pandemic varies, which is key to informing whether and how the tailoring of lockdown and reopening policies to local characteristics can make such measures more effective.

We start by showing that young and middle-age adults account for a larger share of COVID-19 deaths in developing (i.e. low- and middle-income) countries than in high-income countries. We do this by assembling a cross-country data set on mortality by age group, which includes 18 high-income countries and 13 developing countries (see Data Appendix B for the country list and data sources). We find that deaths from COVID-19 are, on average, 5 percentage points higher for those aged 20 to 39, and 23 percentage points higher for those aged 40 to 59 in developing countries. Similar patterns have been independently documented by Demombynes (2020) in a smaller cross-section of countries, and by Laxminarayan et al. (2020) and Philip et al. (2020) in studies focusing on India.

The rest of the paper assesses alternative explanations for these differences. We structure the analysis with a simple conceptual framework based on the classic Susceptible-Infected-Recovered (SIR) model of spread of disease (Kermack and McKendrick, 1927), evaluating the role of age-specific differences in the size of susceptible populations, rates of infection, and rates of recovery in developing and high-income countries.

The younger age profile of COVID-19 deaths in developing countries is not explained by the fact that these countries have, on average, younger populations. We show this with a simple regression of age-group shares of COVID-19 deaths on age-group and developing-country identifiers, along with their interactions. If we include the share of each age group in each country's population as a control, the measured gap between developing and high-income countries remains largely the same. Moreover, the differences in death rates for non-elderly cohorts are also not explained by COVID-19 testing intensity, or by the timing of school closures or other social-distancing policies.

Given that the age-breakdown of susceptible populations does not account for the observed pattern, the main drivers are likely differences in the rates of recovery, the rates of infection, or both. Our study of recovery rates shows that that the prevalence of preexisting

conditions accounts for part of the gap. Controlling for two related measures – the rate of all-cause mortality for each age group and the share of each age group in the population at high risk of developing severe COVID-19 symptoms – fully closes the COVID-19 mortality gap between the two tiers of countries for individuals aged 20 to 39. The size of this gap also shrinks for those aged 40 to 59. Nonetheless, a statistically significant, 12-percentage-point difference remains unaccounted for, supporting the conjectures of Davies et al. (2020) and Walker et al. (2020), who anticipated that the risk profile of COVID-19 may be different in developing countries due to higher morbidity rates among younger age groups.

We turn next to patient-level data to explore at a more granular level the differences in age-specific COVID-19 mortality and the role of access to healthcare across four countries representing two income tiers. We combine publicly available non-identified individual-level data about COVID-19 patients from Colombia, Canada, Mexico and the US. The data show whether these patients were hospitalized, whether they were placed in an intensive care unit (ICU), and whether they died. We use these data to estimate age-specific probabilities of dying – and of accessing inpatient care conditional on receiving a positive diagnosis. We also find evidence of a younger age profile for COVID-19 deaths at lower income levels in this sample: A COVID-19 patient aged 40 to 49 in a developing country has statistically the same average probability of dying as a COVID-19 patient aged 60 to 69 in a rich country.

Consistent with our previous cross-country analysis, the patient-level results show that middle-aged adults experience lower recovery rates in lower-income countries. The findings also suggest that the lower recovery rates of younger cohorts in Mexico and Colombia are likely linked to more limited access to inpatient care. Conditional on receiving a positive COVID-19 diagnosis, non-elderly adults in these countries are more likely to die regardless of whether they are hospitalized. At the same time, while COVID-19 patients are *more* likely to be hospitalized in these countries, they are *less* likely to receive care in the ICU. These findings are consistent with both a higher prevalence of critical cases, and an insufficient capacity to provide life-saving intensive care for all patients who need it. These age-specific differences between developing and high-income countries hold equally for men and women, even though, in both country tiers, men are more likely to die from the disease.

The inability of the cross-country and patient-level analysis to fully explain the differences in COVID-19 deaths rates among middle-aged adults suggests that differences in infection rates likely play a role. Though the available data do not allow for reliable cross-country comparisons of age-specific infections, we take another approach within these constraints. We assemble a city-level data set for Brazil, the hardest-hit developing country in terms of confirmed COVID-19 deaths at the time of writing. We use this Brazilian

data set to regress the differences between elderly and non-elderly shares of COVID-19 deaths on city characteristics that the literature has linked to higher local prevalence of the disease.

We find that the rates of COVID-19 deaths among non-elderly populations are more likely to be closer to those of the elderly populations in cities that have higher informality rates and larger shares of workers with high school education. The coefficients on these variables are similar regardless of whether we control for the total number of deaths among the elderly population – implying that the differences are driven by higher mortality rates among the younger cohorts. This is consistent with prior research arguing that the virus is likely to spread faster in countries with higher labor informality. Informal workers tend to have less formal schooling, less purchasing power and lower levels of precautionary savings – making them more likely to live hand-to-mouth, and less likely to comply with mobility restriction policies (Alon et al., 2020; Busso et al., 2020; Hausmann and Schetter, 2020). Our work suggests that these risk factors disproportionately affect non-elderly individuals, who are more likely to be economically active. In contrast, after controlling for other covariates we find no statistically significant associations with the elderly-non-elderly difference in death shares from other labor market variables, including the average length of commute, the ability to work from home, and socioeconomic characteristics such as poverty, inequality, and the presence of large populations of disadvantaged ethnic minorities.

The higher share in COVID-19 deaths of non-elderly adults is also related to housing conditions in cities. We find that cities with higher residential overcrowding and lower access to piped water tend to have higher mortality among the non-elderly population. At the same time, holding these and other covariates constant, cities with a large share of their population living in favelas (informal neighborhoods) register higher mortality among the elderly. We do not find a statistically significant connection between the COVID-19 mortality age gaps and city-level measures of co-habitation of different generations in the same household.

This paper contributes to a branch of the COVID-19 pandemic literature that investigates how vulnerability to COVID-19 varies with age. This literature has previously established that the probability of dying from the disease increases sharply at older ages (Brotherhood et al., 2020b; Davies et al., 2020; Dowd et al., 2020; Ioannidis et al., 2020). While comparisons across high-income countries have found very similar mortality age profiles across nations (Goldstein and Lee, 2020), researchers have shown that young and middle-aged adults represent a larger share of COVID-19 deaths in countries with lower income levels (Demombynes, 2020; Laxminarayan et al., 2020; Philip et al., 2020). Our paper independently documents this discrepancy by using more detailed and comprehensive data. In addition,

it provides what is, to our knowledge, the most complete explanation to date of the drivers of this gap.

Second, we contribute to the broader literature on the different impacts of the COVID-19 pandemic between countries of different income levels. Prior research has argued that the health and economic effects of COVID-19 are likely to be more severe in lower-income countries for a wide variety of reasons, including: higher labor informality rates (Alfaro et al., 2020; Busso et al., 2020); higher reliance on daily income for basic subsistence (Alon et al., 2020; Hausmann and Schetter, 2020); and worse housing conditions (Brotherhood et al., 2020a; Brown et al., 2020). At the same time, researchers have suggested that social distancing policies may be less cost-effective in poorer countries, partly because their younger populations imply fewer COVID-19 deaths (Alon et al., 2020; Barnett-Howell and Mobarak, 2020; Walker et al., 2020). Our work shows that having a relatively young population is less of an advantage for these countries than previously has been thought. The non-elderly populations in these countries have a higher incidence of preexisting conditions and less access to potentially life-saving treatment and care. Moreover, the higher informal employment rates and residential overcrowding that are characteristic of lower-income countries puts people at greater risk of exposure to COVID-19 – and these risks appear to disproportionately affect non-elderly adults.

Finally, our paper is relevant for the literature on optimal lockdown and reopening policies. Models used to forecast the total number of COVID-19 infections and deaths oftentimes rely on age differences in mortality rates (Canabarro et al., 2020; Lyra et al., 2020), and researchers have argued that optimal lockdown and reopening policies should differentiate by age groups because of their disparate risk levels (Acemoglu et al., 2020; Alon et al., 2020). By showing that these age-specific risks vary significantly – between countries of different income levels, and between cities with different living conditions within a single country – our work suggests that optimal policies may depend on very local contexts. Further, the measures of age-specific mortality risks and healthcare access provided by this paper can be used to calibrate such structural models in countries with lower income levels.

The paper is organized as follows: Section 2 presents and discusses the country-level evidence on the younger age profile of COVID-19 deaths in developing countries. Section 3 presents patient-level evidence on differences in rates of recovery, and on the role of healthcare access. Section 4 explores other potential drivers of age gaps in mortality rates using city-level evidence from Brazil. Section 5 concludes.

2 Cross-country evidence

Young and middle-aged adults are more likely to die from COVID-19 in developing countries than in higher-income countries. Appendix Figure C1 (a) provides a visual illustration of this pattern; it plots the distributions of age-specific shares of COVID-19 deaths in the hardest-hit developing and high-income countries. This is in line with a recent study that shows, in a sample of 26 countries, that age-specific mortality rates increase at a lower rate in low and middle-income than in high-income nations (Demombynes, 2020).

What might explain these differences in outcomes for non-elderly adults? The intuition of the basic epidemic SIR model —originally proposed by Kermack and McKendrick (1927)— guides our analysis. If an epidemic evolves according to the dynamics of the standard model, at any given point in time the total number of deaths D can be represented as $D = (1 - \gamma)I$, where I is the infected population, and γ the recovery rate of disease (the share of the infected that have not died since the beginning of the epidemic, under the assumption that recovered individuals cannot be reinfected). In turn, the infected population can be characterized by $I = \beta S$, where S is the susceptible population (which we assume to be everyone in the country), and β the infection rate (the share of the susceptible population that became infected since the beginning of the epidemic). It follows that for any age group a from country group c, the total number of COVID-19 deaths can be written as

$$D_{ac} = (1 - \gamma_{ac})\beta_{ac}S_{ac}$$

Thus, differences in the age profile of COVID-19 deaths across countries can be explained by differences in the age profile of the population as a whole (S_{ac}), the rate at which specific age groups are infected (β_{ac}), the rate at which the infected individuals in each age group die from the disease ($1 - \gamma_{ac}$), or some combination of these.

We evaluate the role of these possible explanations using the data set at the age group–country level, with four age groups per country (0-19, 20-39, 40-59, and 60 or older). Countries differ in the age categories they use to report age-specific mortality rates, and we choose these ranges to maximize the number of countries that can be included in the analysis, conditional on the country having reported at least 50 confirmed COVID-19 deaths. Our sample includes 18 high-income countries and 13 developing countries for which the age-specific mortality information was gathered between May 5 and July 12, 2020. We use ordinary least squares (OLS) to estimate the following cross-sectional regression:

$$Y_{a,c} = \alpha_0 + \sum_{a=1}^{3} \alpha_{1a} \left(a_{a,c} \times d_c\right) + \sum_{a=1}^{3} \alpha_{2a} a_{a,c} + \alpha_3 d_c + \alpha_{\mathbf{c}}^{'} \mathbf{C_c} + \varepsilon_{a,c} \tag{1}$$

where $Y_{a,c}$ is the share of age group a in total COVID-19 deaths of country c, $a_{a,c}$ are indicators for the age groups mentioned above (we use the population aged 60 or older as the reference group), d_c is a developing country indicator (low- or middle-income country according to the 2019 World Bank classification), $\mathbf{C_c}$ is a vector of country-level controls, $\varepsilon_{a,c}$ is the error term, and the αs are parameters. We cluster the standard errors at the country level. As part of the analysis we also use this regression to estimate developing–high-income country differences in the age profile of all-cause mortality, defining $Y_{a,c}$ as the share of age group a in all deaths.

2.1 Age structure, testing, and policy responses

Table 1 reports the regression results. Column 1 estimates equation 1 without the interaction terms to show that, as has been widely documented in other studies (Davies et al., 2020; Dowd et al., 2020; Ioannidis et al., 2020), younger cohorts represent a smaller share of COVID-19 deaths than the older reference group does. Column 2 reports our baseline results, estimating the full specification in equation 1 controlling for GDP per capita, population, the total number of COVID-19 cases, and the total number of COVID-19 deaths. The estimates of the interaction coefficients $\hat{\alpha}_{1a}$ are all positive and statistically significant, confirming that, on average, the COVID-19 mortality age gap between the elderly and younger age groups is smaller in developing than in high-income countries. Figure 1 (top-left) reports the age-specific predicted shares that correspond to this regression, showing that the share in COVID-19 deaths of those aged 60 and older is 28 percentage points lower in developing countries than in high-income countries. This corresponds to a 5 percentage point larger share in developing countries for the group aged 20 to 39, and to a 23 percentage point larger share for those aged 40 to 59. As a reference, we report the predicted shares by age group from a regression without controls in Appendix Table C1.

The fact that younger cohorts account for a larger share of COVID-19 deaths in developing countries could be explained by the fact that they tend to have younger populations (Dudel et al., 2020), and therefore more young, susceptible individuals ($S_{yd} > S_{yh}$, where y denotes young, d developing, and h high-income countries). This explanation, however, can account for only a small part of the observed gap. Column 3 in Table 1 introduces the shares of each age group in the overall population as controls. The estimates of all the coefficients of interest change only slightly; all estimates remain statistically significant. This can also be seen in the distribution of age-group shares of COVID-19 deaths adjusted by the age-group shares of the overall population for each country, which we report in Appendix Figure C1 (b). Demombynes (2020), also investigates this issue using a counterfactual exercise that compares the age distribution of deaths in developing countries with a

Table 1: Differences between developing and high-income countries in the share of COVID-19 deaths by age group

	Dependent variable: Share of age group in national COVID-19 deaths						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Developing country indicator	-0.00	-0.29***	-0.25***	-0.27***	-0.25***	-0.16***	-0.12**
	(0.00)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Age groups indicators (*reference group is age 60+*)							
Ages 0-19	-0.82***	-0.94***	-0.94***	-0.94***	-0.93***	-0.46***	-0.33**
	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.08)	(0.12)
Ages 20-39	-0.79***	-0.94***	-0.94***	-0.94***	-0.93***	-0.45***	-0.36***
	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)	(0.09)	(0.11)
Ages 40-59	-0.67***	-0.89***	-0.89***	-0.89***	-0.89***	-0.44***	-0.41***
	(0.05)	(0.01)	(0.01)	(0.01)	(0.01)	(0.08)	(0.08)
Age groups $\times$ *Developing country*							
Ages 0-19		0.29***	0.23***	0.25***	0.23***	0.15***	0.12**
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Ages 20-39		0.34***	0.29***	0.31***	0.29***	0.19***	0.12*
		(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.07)
Ages 40-59		0.52***	0.48***	0.51***	0.48***	0.31***	0.25***
		(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)
Observations	124	124	124	120	116	116	116
R^2	0.91	0.98	0.98	0.98	0.98	0.98	0.99
Country characteristics	No	Yes	Yes	Yes	Yes	Yes	Yes
Shares of age group in total population	No	No	Yes	Yes	Yes	Yes	Yes
Testing controls	No	No	No	Yes	Yes	Yes	Yes
Policy controls	No	No	No	No	Yes	Yes	Yes
Share in all-cause deaths	No	No	No	No	No	Yes	Yes
Share in high-risk of severe complications	No	No	No	No	No	No	Yes

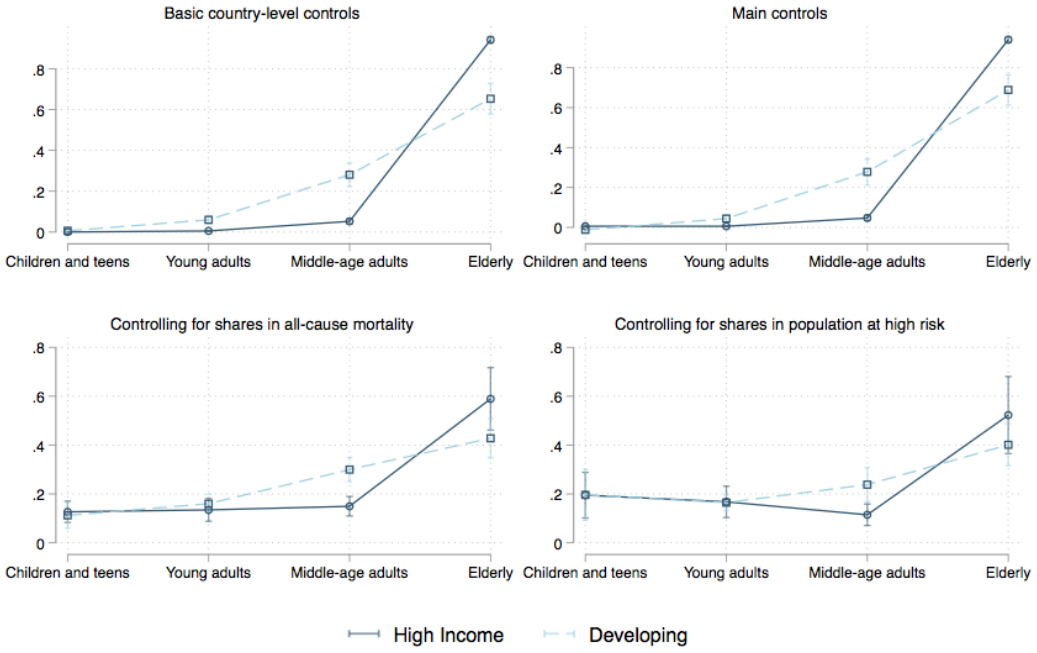
Notes: Regressions at the country-age group level on a sample of 31 countries (13 developing and 18 high-income). Robust standard errors clustered at the country level in parentheses. Country characteristics controls include GDP per capita, population, the total number of COVID-19 cases, and the total number of COVID-19 deaths, all in logarithms. Testing controls include the number of tests per capita conducted, and the positivity rate of these tests at the time when the data on deaths by age were collected. Policy controls include the the number of days between the first confirmed case and school closures, and the number of days and between the first case and the issue of stay at home orders, both in logarithms. All regressions include a constant. *** p<0.01, ** p<0.05, * p<0.1.

hypothetical distribution built with actual age-specific death rates in developing countries and the average share of each age group in the population of high-income countries. This analysis concludes that differences in the age structure of populations only account for part of the differences in the age-group shares of COVID-19 deaths between developing and high income countries.

Another explanation for the younger age profile of COVID-19 mortality in developing countries is the relatively lower prevalence of testing. For a given spread of the disease

in a country, fewer tests imply a larger number of under-recorded infections. As long as mortality rates also increase with age among the undetected infected population, countries with less testing will disproportionately under count deaths in older cohorts. In order to address this concern, column 4 in Table 1 includes the number of tests conducted per capita, and the positivity rate of these tests at the time when the data on deaths by age were collected. The magnitude of the point estimates of the interaction terms increases slightly, and the age gap remains significant.

Figure 1: Predicted age-group shares in COVID-19 deaths in developing and high-income countries



Notes: This figure reports the predicted age-group shares in COVID-19 deaths corresponding to the regression results reported on Table 1 (columns 2, 5, 6 and 7, respectively). The reported 95% confidence intervals are constructed using the Delta method. See Appendixes A and B for further details.

An additional possibility is that higher mortality rates among younger cohorts could be the result of delays in implementing containment measures, particularly those related to mobility. Such delays could have disproportionately affected younger cohorts because they are more likely to commute regularly for work or school. With this in mind, column 5 adds two measures of the timing of non-pharmaceutical interventions in the country: the number of days between the first confirmed case and school closures, and days between the first case and the issue of stay at home orders. None of these controls are able to explain away the main result (Figure 1, top-right).

2.2 Mortality and pre-existing conditions

Since the divergence of the age-mortality profile of COVID-19 between developing and high-income countries is not explained by differences in the number of susceptible individuals or by differences in testing, it is likely driven by differences in age-specific rates of infection β_{ac}, and/or rates of recovery γ_{ac}. Both of these parameters may vary with income.

Official data on confirmed cases are insufficient to produce reliable estimates of cross-country differences in age-specific infection rates. These data overlook infection rates of the population that was not tested. The characteristics of the individuals who are tested may be different than those of the broader population, and the accuracy of testing may vary systematically with income and other correlates of infection risk (Manski and Molinari, 2020). These issues may be overcome if accurate seroprevalence tests – which detect the presence of antibodies likely related to prior infections – become more widely available, and if large-scale, random testing strategies are implemented in multiple countries (Stock, 2020). Such data were unavailable at the time of writing.

Existing data, however, can give us a good sense of the role that differences in recovery rates γ_{ac} play in the observed age-mortality profiles of these two country groups. Specifically, differences in COVID-19 mortality could reflect broader underlying health differences across developing and high-income nations. Developing countries tend to have high morbidity rates across the age distribution, and researchers have argued that this may counteract the lower COVID-19 mortality typically associated with younger populations (Davies et al., 2020; Walker et al., 2020).Two sets of facts support this conjecture. First, the age profile of COVID-19 deaths resembles the age profile of all-cause deaths (Goldstein and Lee, 2020; Demombynes, 2020). This can be seen in Appendix Figure C2 (left), which depicts the predicted age-group shares in all-cause deaths from an OLS estimate of equation 1. Second, non-elderly individuals in developing countries tend to have a higher incidence of some of the preexisting conditions associated with severe COVID-19 complications, including chronic kidney diseases, chronic neurological disorders, HIV/AIDs, and tuberculosis (see Appendix Figure C3).

We gauge the extent to which age differences in COVID-19 deaths across countries reflect broader mortality differences by adding age-group shares in all-cause deaths to our main regression as a control (Table 1, column 6). This leads to substantive changes. The point estimates of all the coefficients of interest shrink substantively – by almost half in most cases – although they maintain their sign and statistical significance. The developing–high-income country gap in the corresponding predicted probabilities shrinks accordingly. As illustrated in Figure 1 (bottom-left), the gap in the cohort aged 20 to 39 shrinks by half. The gap in the group aged 40-59 also shrinks by 3 percentage points but remains large

and statistically-significant. This is consistent with the results in Levin et al. (2020) who, using data from large-scale seroprevalence studies in high-income countries, find that the infection fatality rate of COVID-19 not only is unusually high among the elderly, but also among healthy middle-aged adults. Our cross-country results suggest that the increased vulnerability of this non-elderly age group is more pronounced in developing countries.

The higher COVID-19 mortality rates among younger cohorts in developing countries can be explained, at least in part, by higher overall morbidity rates, including a higher prevalence of conditions associated with severe coronavirus symptoms (Davies et al., 2020). To further explore this mechanism, we use estimates of the fraction of the population with preexisting conditions that put them at high risk of severe COVID-19 complications (Clark et al., 2020) (see Appendix A for additional details). Whereas lower income levels tend to be associated with having *smaller* shares of the population in the high-risk category in cross-country comparisons (Appendix Figure C4, left), the gap between the shares of the elderly (60+) and non-elderly (younger than 60) in the high-risk population is highly correlated with the same gap in the shares of COVID-19 deaths (same figure, right).

While differences in pre-existing conditions appear to fully explain the developing-high income gap in the age profile of all-cause deaths, they account for only part of the age gap in COVID-19 deaths. The predicted age-group shares in all-cause mortality from an OLS estimate of equation 1 no longer shows a statistically significant difference between developing and rich countries once we include age-specific shares in the high-risk group as a control (Appendix Figure C2, right). Even so, when we add the high-risk-shares control to our main COVID-19 deaths regression (Table 1, column 7), the direction and significance of the main coefficients of interest remain largely in place, albeit with a relatively small reduction in the size of the point estimates of interest. As shown in Figure 1 (bottom-right), while the gap in the cohort aged 20 to 39 disappears after introducing this control, a statistically significant 12-percentage-point gap remains between developing and high-income countries in the predicted share of the 40-59 age group in COVID-19 deaths.

We also examine the age profiles of pandemic-related deaths in developing and rich countries by gender. The COVID-19 literature has documented that men are more likely to die from the disease than women (Klein et al., 2020; Laxminarayan et al., 2020; Mohamed et al., 2020). To assess whether these differences are related to the observed gap in the age profile of deaths across countries, we estimate a triple-interaction model, using a variation of equation 1 that includes a male indicator and interactions with the other regressors (see Appendix A for details). The results are reported in Appendix Table C2. Our estimates of the coefficients on the triple interaction of the age group, developing-country and male identifiers are close to zero and not statistically significant, suggesting that the developing -

high-income country gap in COVID-19 deaths is largely uncorrelated with gender.

We perform additional robustness tests on the specification with all the controls, all reported in Appendix Table C3. Column 1 reproduces, as a reference, the results from the main specification including all controls. Column 2 adds as a control the date when the data used for each country was last updated. Column 3 drops countries in which the data may not be fully comparable with the rest, for example because they provide information from only part of the country. Columns 4 and 5 restrict the sample to countries in which more than half of the population live in localities with 1 million or fewer residents, or in locations with fewer 500,000 residents, respectively. Notwithstanding small differences in the size of the point estimates, the findings of the main specification remain unchanged.

3 Patient-level evidence

The cross-country analysis shows that the developing-high-income gap in COVID-19 mortality among the population younger than 40 is mostly explained by a higher prevalence of pre-existing health conditions in low and middle-income countries. Even after accounting for differences in these pre-existing conditions, however, a significant gap remains among middle-aged adults. This could reflect differences in recovery rates (γ_{ac}) that are related to other factors such as access to healthcare, or differences in infection rates β_{ac}, or both.

We investigate the role of differences in access to healthcare using a multi-country, patient-level dataset. Of the eight nations for which we have publicly available patient-level COVID-19 data, two developing (Colombia and Mexico) and two high-income countries (Canada and the US) have information on confirmed cases (regardless of hospitalization status), whether each individual was hospitalized, and whether hospitalized patients were placed in an ICU. We use these data sets to reconstruct patients' clinical histories, and to estimate a set of age-group-specific conditional probabilities and see how these differ across developing and high-income countries. Specifically, we use a Probit model to estimate a variant of equation 1, according to the following regression:

$$Y_{i,a,c} = \delta_0 + \sum^{4} \delta_{1a} \left(a_{i,a,c} \times d_c \right) + \sum^{4} \delta_{2a} a_{i,c,a} + \delta_3 d_c + \delta_{\mathbf{i}}^{'} \mathbf{C_i} + \varepsilon_{i,a,c} \tag{2}$$

where $Y_{i,a,c}$ is an indicator variable measured for individual i from age group a in country c, $a_{i,a,c}$ are age group indicators, d_c is a developing country indicator, $\mathbf{C_i}$ is a vector of individual-level controls, $\varepsilon_{i,a,c}$ is an error term, and the δs are parameters. The individual-level data allow us to consistently define age groups at 10-year intervals. The population aged 80 and older is left as the reference group. In the main specification we exclude

from all samples individuals below age 40 because the number of deaths beneath this age threshold is insufficient to consistently compute predicted probabilities in this set of countries. Standard errors are clustered at the country level.

The indicator outcomes and the sample over which we estimate the model are defined according to the the conditional probability of interest. For example, to estimate the variation between the two tiers of countries for the age-group-specific probability of death conditional on receiving a positive diagnosis and on being hospitalized, we define $Y_{i,a,c} = \mathbb{1}_{\{dead\}}$, and restrict the sample to individuals who received a positive diagnosis and were hospitalized.

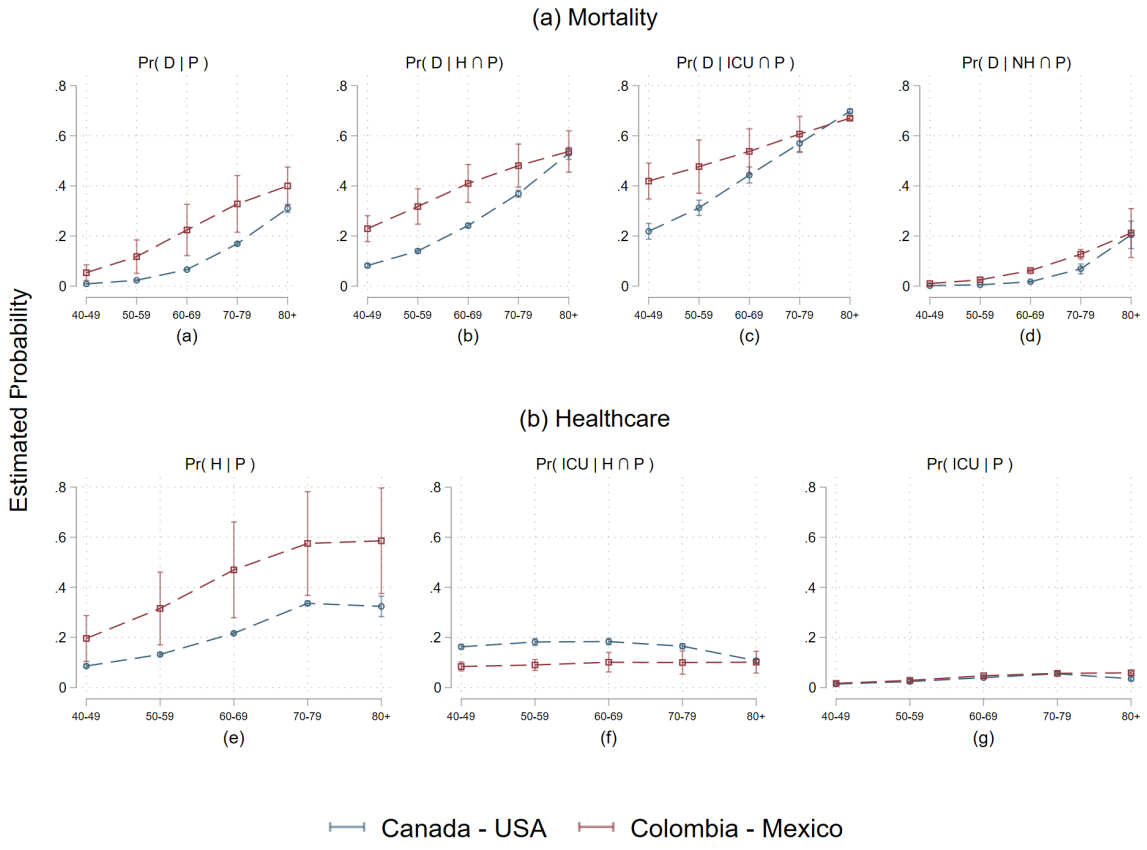
We estimate probabilities of death due to COVID-19 conditional on four different states: the probability of death conditional on receiving a positive diagnosis ($\Pr(D|P)$); the probability of death conditional on being hospitalized after a positive diagnosis ($\Pr(D|H \cap P)$); the probability of death conditional on not being hospitalized after a positive diagnosis $\Pr(\Pr(D|NH \cap P))$; and the probability of death conditional on being placed in an ICU after a positive diagnosis ($\Pr(D|ICU \cap P)$).

We also estimate probabilities of COVID-19 patients receiving different types of healthcare conditional on three different states: the probability of being hospitalized conditional on a positive diagnosis ($\Pr(H|P)$); the probability of being placed in an ICU conditional on a positive diagnosis ($\Pr(ICU|P)$); and the probability of being placed in an ICU conditional on having a positive diagnosis and being hospitalized ($\Pr(ICU|H \cap P)$).

Figure 2 shows the predicted margins of each of these seven different conditional probabilities by age group in developing (Colombia and Mexico) and high-income countries (Canada and the US), obtained with Probit estimates from regression 2. Full regression estimates are reported in Appendix Table C4.

In addition, we can calculate some (but not all) of these probabilities for four additional countries (Brazil, the Czech Republic, the Philippines, and South Korea). To test for robustness, we replicate the conditional probability analysis using information from all countries that have the necessary data for each conditional probability (Appendix Figure C9) and obtain similar results. Appendix Figures C10 and C11 estimate these conditional probabilities separately for each country to explore heterogeneity. Descriptive statistics for all eight countries are reported in Appendix Table C5. We also report the estimated predicted probabilities and standard errors from this sample. This information can be used to calibrate structural models of disease spread and mortality in developing countries (see Appendix tables C6 and C7).

Figure 2: Comparisons of age-specific probabilities of hospitalization, ICU admission, and death due to COVID-19



Notes: The figure shows the estimated conditional probabilities of death, hospitalization, or entering an ICU by country income level and their 95% confidence intervals. Regressions estimated at individual level using a Probit model. Robust standard errors clustered at the country level.

3.1 Age-group differences in recovery

Consistent with our cross-country analysis, the patient-level results suggest that middle-age adults in developing countries experience lower recovery rates γ_{ac} than their peers in rich countries. Conditional on testing positive for COVID-19, the probability of death is substantively larger in developing countries than in high-income countries across all age groups. For younger cohorts in developing countries, the gap is even larger, nearly triple the level for peers in high-income countries. (Figure 2, (a)). In Colombia and Mexico, a person with a positive diagnosis who is in the 40-49 age range has the same average statistical probability of dying from COVID-19 as a person in Canada or the US who is two decades older – that is, in the 60-69 in age range. While both Colombia and Mexico have

higher non-elderly mortality than Canada and the US, the size the average gap is largely driven by the unusually high COVID-19 mortality in Mexico (Appendix Figure C10).

As discussed, these mortality age profile results could potentially be explained by differences in testing coverage. To explore this possibility empirically in the subsample for which we have patient-level information, we start by looking at how country differences in age-specific shares of COVID-19 deaths have evolved since the beginning of the pandemic. Testing levels and positivity rates have changed significantly over time (Appendix Figure C5). We examine the evidence to see whether testing levels drive the differences in the deaths rates for different age groups. Our evidence suggests that the testing is not a driver. Even in the wake of the early weeks of the pandemic, when the numbers of cases and deaths were still small in all these countries, the shares of each age group in the cumulative number of cases and deaths of their respective countries remain stable. Moreover, the country rank in terms of these shares also remains constant across the period (Appendix Figure C6). One exception is the share of the groups aged 40-49 and 50-59 in Mexico, which saw a noticeable increase towards the end of the observed time period studied, without altering the country's relative position in the ranking.

To further probe the role of testing in our results, we replicate the analysis in Figure 2 using the cumulative number of age-specific deaths at similar national testing levels. Specifically, for each country we use the data from the most recent date in which testing was within an overlapping range – within 5.58 and 5.92 tests per 100,000 population. This specific range was chosen because the four countries in our sample have observations within it (see Appendix A for further details). The observed pattern is essentially the same as when we use the most recent data for each country, although the point estimates for Canada and the US tend to be marginally larger than before, as are the standard errors for the Colombia and Mexico estimates (Appendix Figure C7). We perform a similar exercise using data from dates in which positivity rates were similar for these countries (Appendix Figure C8). The age profile of the point estimates is similar to that in the main results, but in this case the standard errors for the Colombia and Mexico estimates increase substantially, making the difference with respect to Canada and the US not statistically significant for some age groups. However, we note that the only time in which the Mexican positivity rates were close to those of other countries over this period was in early April, a time during which the number of deaths was still small, and the estimates of the share of each age group in these deaths were still very noisy as a result (Appendix Figure C6). Overall, the evidence suggests that the observed differences in the mortality age profile are not an artifact of testing gaps between upper-middle-income and high-income countries.

3.2 Disease severity and access to healthcare

To what extent are these differences attributable to lack of access to healthcare in developing countries? The evidence suggests that, though limited healthcare resources do play a role, much of the mortality gap between developing and high-income countries in younger age groups is driven by differences in the severity of COVID-19 complications.

We first note that the average mortality gap among the younger cohorts is larger when considering only patients who were hospitalized, rather than the wider group of people who received a COVID-10 diagnosis (Figure 2, (b)). The inverse relationship between the gap and age becomes even more pronounced when we restrict the sample to individuals who were placed in ICUs (Figure 2, (c)). These differences are primarily driven by Mexican non-elderly deaths. The size of the confidence intervals reflects the fact that post-hospitalization mortality in the US is large, comparable to Colombia in most age groups, and significantly higher than in Canada (Appendix Figure C10). We consider three alternative explanations for this pattern: differences in the composition of the hospitalized group, in the quality of healthcare services, and in the underlying prevalence of severe complications among younger cohorts.

Selection and sorting into healthcare

One possibility is that, among younger cohorts, healthcare admissions tend to be more biased towards severely ill patients in developing countries than in high-income countries, even if the underlying distribution of severe COVID-19 complications is the same for each age group. This could arise from sorting. Economic or cultural reasons may lead people in younger cohorts to make the decision to seek professional healthcare at a higher symptom-severity threshold in poorer countries than in rich countries. Selection issues could play a role. If the demand for beds in hospitals and/or ICUs outweighs existing capacity, which is more likely to happen in countries with resource-constrained health systems, healthcare providers may be forced to restrict hospital admissions only to the more severely ill.

If hospitalized patients come from the same age-specific distribution of the severity of COVID-19 complications in both country tiers, sorting or selection of individuals from the higher end of the distribution into hospitalization in developing countries would imply two things. First, both hospitalized and non-hospitalized individuals would have a higher average severity of COVID-19 complications – and therefore higher mortality rates $\Pr(D|H \cap P)$ and $\Pr(D|NH \cap P)$ – in developing countries, which we do observe. Second, non-elderly cohorts that tested positive for the disease would have a lower probability of hospitalization ($\Pr(H|P)$) in these countries. We find that the probability of hospitalization

across all age groups is *larger* in developing than in high-income countries, which mitigates selection concerns.

Effectiveness of healthcare services

A second possibility is that healthcare services have been less effective at saving lives in developing than in rich countries. This in turn could be due to lower quality in healthcare services provided by hospitals operating within capacity, or due to lower overall capacity – in terms of the number of personnel, the availability and quality of equipment, and other healthcare infrastructure – all of which may lead health services unable to cope with unusually high demand.

Differences in overall mortality rates may also be due to differences in quality (e.g., quality of training of medical personnel, the types medical protocols employed, the types of medicine available for use, etc.). In several developing countries private healthcare facilities offer standards of care comparable to those in rich countries, but only a fraction of the population has access to health insurance and can afford these services (Menezes-Filho and Politi, 2020). However, unless these differences have age-specific biases within country type, they are likely to be captured by the developing country indicator in our specification, and they are unlikely to explain why the mortality gap between these country groups shrinks with age. While the average probability of death conditional on testing positive and being hospitalized (Figure 2, (b)) is around three times larger for the population aged 40-49 in developing than in rich countries, this difference is smaller in older age groups, and it completely disappears in the population aged 80 and older. A very similar, and even more pronounced pattern is observed when looking at the probability of death conditional on testing positive and entering an ICU (Figure 2, (c)).

In contrast, capacity constraints in healthcare systems could in principle explain the different slope of the COVID-19 mortality age profile between developing and rich countries. In particular, if in developing countries there is a relatively smaller availability of critical care facilities or equipment – such as ICU beds or ventilators – it could force providers to ration their use. If the rationing criteria favor patients with more severe complications, those who meet these criteria are more likely to be older, leaving people who are younger disproportionately less access to more intensive care. The evidence suggests that such rationing is indeed taking place. As shown in Figure 2 (f), non-elderly hospitalized COVID-19 patients have on average a significantly *lower* probability of being placed in an ICU in Colombia and Mexico than in the US and Canada.

It is important to note, however, that access to ICUs varies widely from one country to another (Appendix Figure C11). Younger hospitalized COVID-19 patients have a signifi-

cantly higher likelihood of being placed in an ICU in Canada than in the US; they also are more likely to be placed in an ICU in Colombia than in Mexico. In Mexico, access to ICUs is notoriously low for all age groups. By contrast, in Brazil – a country that is not included in the regressions from Figure 2 due to lack of data on non-hospitalized COVID-19 patients – hospitalized COVID-19 patients in younger cohorts are as likely to be placed in an ICU as in Canada, and patients aged 60 or older are significantly more likely to be placed in an ICU than in all other countries in the sample. In fact, when Brazil is included, the average difference between developing and high-income countries in the probability of accessing an ICU if hospitalized is not statistically significant (Appendix Figure C9).

Prevalence of COVID-19 severe complications

A third, non-exclusive explanation for the different age profiles of COVID-19 deaths in developing and high-income countries is that their underlying distributions of the severity of the disease's complications are different. Even if both the probability of being hospitalized at each threshold of symptom severity and the quality of healthcare were the same in both country types, younger populations in developing countries would still have higher mortality rates if more people in their age groups experienced more severe complications than their peers in high-income countries.

Our results suggest that this is the case. People from all age groups under age 80 who receive a positive COVID-19 diagnosis are equally likely to enter an ICU regardless of the income level of their country (Figure 2, (g)). However, people in these age groups are more likely to be hospitalized (Figure 2, (e)), and those who are hospitalized are *less* likely to be transferred to an ICU (Figure 2, (f)). This is consistent with the interpretation that, in developing countries, a larger share of the positive cases develops severe COVID-19 complications and requires ICU care. While healthcare facilities appear to have the capacity to admit most of these patients, the demand for ICU beds by younger cohorts appears to exceed ICU capacity to a greater degree in developing countries than it does in wealthier countries. This can also explain why younger cohorts in developing countries have both a higher average hospitalization rate and a higher average probability of death, both for COVID-19 patients who were hospitalized and for those who were not hospitalized.

Why may people under 80 develop more severe complications from COVID-19 in developing countries than in high-income countries? One possible cause could be that a greater prevalence of preexisting conditions among younger cohorts in developing countries puts people in these age groups at relatively greater risk of experiencing more severe COVID-19 complications. As our cross-country comparisons showed, the greater prevalence of pre-existing conditions fully explains the gap in younger age groups, but only partially

explains the gap among mature adults (aged 40 and older). Here we further probe the role of this mechanism. This additional investigation is possible because two of the countries we analyzed, Mexico and the US, have individual-level data on preexisting conditions. To facilitate comparison, we construct a dummy variable that indicates whether the person has any of the preexisting conditions associated with severe COVID-19 complications, and we re-estimate regression 2 for these two countries with and without this control. The resulting predicted probabilities (with and without preexisting condition controls) are reported in the first and second columns of Appendix Figure C12, respectively. In line with our prior findings, a gap remains between the two countries' mature adult populations after introducing the preexisting condition controls. However, in contrast with the country-level analysis, the gap in the older age groups remains largely unaffected. This could be explained by differences in the definitions of the precondition controls; the patient-level variable considers a much smaller set of preexisting conditions due to data availability.

Though there are alternative explanations for the existence of differences in the distribution of severe COVID-19 complications across developing and high-income countries, we are unable to test them with the existing public data. For example, patients in developing countries could be exposed to larger viral loads from longer or more frequently repeated commutes via public transportation, or from crowded informal dwellings and neighborhoods (Brown et al., 2020). They could also have less timely or less medically appropriate responses to infections, if, for example, they are more likely to continue to work in the presence of symptoms, or if they are more likely to use untested and potentially harmful treatments. Additional data and research are needed to understand the relative contributions of these and other potential mechanisms, and to better inform the policy responses of developing-country governments.

3.3 Differences by gender and ethnic minority status

To conclude our patient-level analysis, we turn to exploring heterogeneity by gender and ethnic minority status in the patient-level data. We do this by estimating triple-interaction models, expanding the specification in equation 2 to incorporate an indicator for males and for ethnic minorities, respectively (see Appendix A for details). In line with the results from the cross-country analysis, we find that gender does not play a role in the age-profile gap of COVID-19 deaths between lower- and high-income countries (Appendix Figure C13, panel (a)). Point estimates in both country types show that men are more likely to die from COVID-19 and to receive healthcare for the disease than women, although the difference is not statistically significant in developing countries. However, the age-profile slopes of both genders track each other within country type.

We also evaluate heterogeneity by ethnicity. Various U.S. studies suggest that poor populations (Weill et al., 2020) and ethnic minorities (Brown and Ravallion, 2020; Benitez et al., 2020; Wiemers et al., 2020) experience higher infection rates. We construct at ethnic minority indicator for the countries for which this information is available (Colombia, Mexico and the US, see details in Appendix B, Table 2), and we use this variable to estimate the triple interaction model. The results point to more limited access to healthcare for ethnic minorities in Colombia and Mexico relative to the US (Appendix Figure C13, panel (b)). The probabilities of death conditional on testing positive and conditional on being hospitalized are very similar for minorities and non-minorities in both country tiers. The severity of COVID-19 complications appears to be higher for ethnic minorities in the US, where they are more likely to be hospitalized if diagnosed with the virus and to die if placed in an ICU relative to non-minorities. By contrast, in Colombia and Mexico the severity of complications from the disease appears to be *lower*, at least among younger cohorts, for whom there is a smaller average probability of death if placed in an ICU, relative to non-minorities. However, conditional on testing positive for COVID-19, minorities in Colombia and Mexico have a lower probability of being hospitalized, and, if they are hospitalized, they have a higher probability of dying than the their non-minority peers within the country. This is in line with a recent study in Brazil (Baqui et al., 2020), which shows that individuals from ethnic minorities hospitalized with COVID-19 also have higher mortality rates, which the researchers link to lower access to healthcare, particularly ICUs.

4 City-level evidence

The preceding analyses shows that young and middle-aged adults are more likely to die from COVID-19 in developing than in high-income countries, and that these differences are partially driven by a lower recovery rate γ_{ac}, associated with a higher incidence of preexisting morbidities, and more limited access to inpatient healthcare. However, a significant difference in mortality remains among middle-age adults after accounting for these factors, suggesting that differences in infection rates βac also play a role.

The growing COVID-19 literature has pointed out several reasons why the disease could spread more rapidly in developing than in rich countries. These include a larger informal economy and a higher reliance on daily income for subsistence, which makes compliance with lockdowns more difficult (Busso et al., 2020; Hausmann and Schetter, 2020), low income and other socioeconomic disadvantages that may lead to worse overall health conditions and reduce access to healthcare (Benitez et al., 2020; Wiemers et al., 2020), lower quality and capacity of their health systems (Canabarro et al., 2020; Walker

et al., 2020), and housing overcrowding which, along with the lack of basic services, limit preventive social distancing and hygiene practices (Brown et al., 2020). If these risk factors affect younger groups disproportionately, they may also be behind observed differences in the COVID-19 age profiles of infections in developing and high-income countries. The results of a large-scale seroprevalence study in 133 Brazilian cities (Hallal et al., 2020) are consistent with this premise. They show that there is a higher prevalence of antibodies among people living in crowded conditions, low-income households, and ethnic minorities; and while the prevalence of antibodies does not differ significantly across men and women, it is significantly higher among the population in age range 20 to 59. To shed light on the extent to which these factors play a role, we complement the cross-country and patient-level analyses with cross-city analysis of COVID-19 hospital deaths in Brazil, the hardest-hit developing country in terms of the number of COVID-19 deaths at the time of writing.

Figure 3 reports the estimated coefficients from the following city-level multivariate regression:

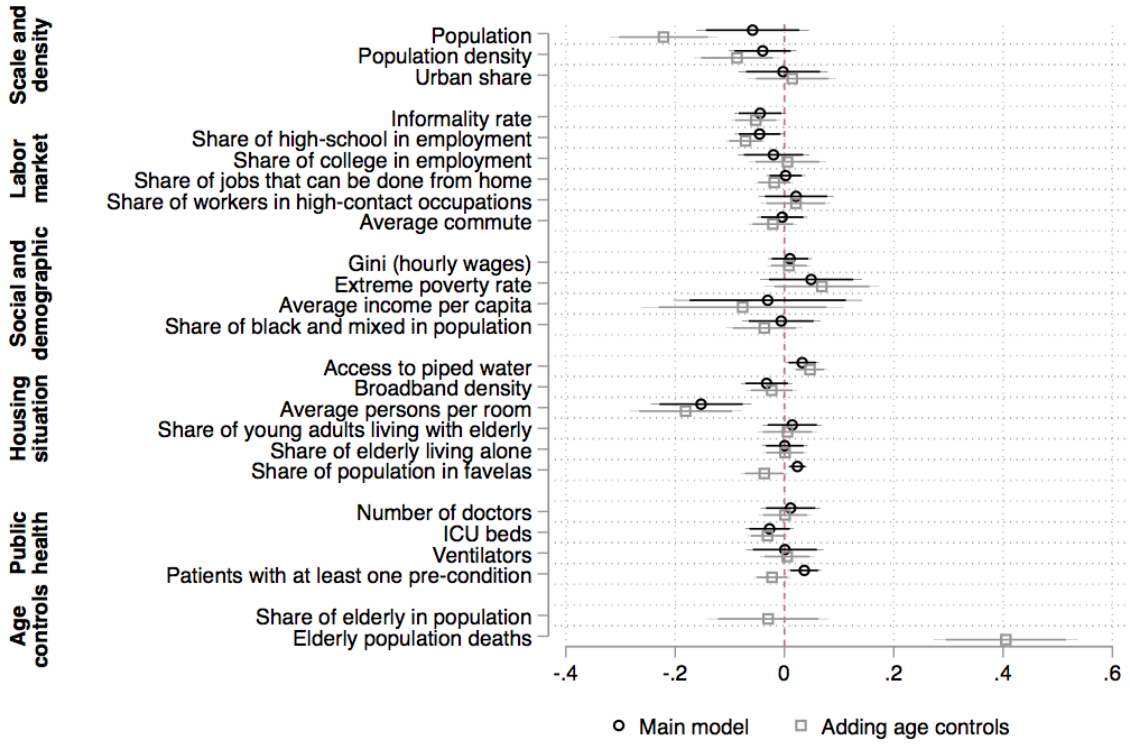
$$s_{j,elderly} - s_{j,non-elderly} = \gamma_0 + \gamma_{\mathbf{j}}^{'}\mathbf{X_j} + \varepsilon_j \tag{3}$$

in which the dependent variable is the difference in COVID-19 death shares between the elderly (aged 60 or older) and the non-elderly population (younger than 60) in city j, $\mathbf{X_j}$ is a vector of regressors measured at the city level before the pandemic, and ε_j is the error term. The explanatory variables come primarily from Chauvin (2020), and are city characteristics that the literature has linked – theoretically or empirically – to the spread of COVID-19 at the local level. Appendix Table C8 also reports estimates of equation 3, showing that the results remain generally stable as new covariates are added to the regression. Definitions and sources are detailed in appendixes A and B.

We report estimates for two models. The first includes the main regressors of interest. The second adds two age-related controls: the share of individuals aged 60 or older in the population, and the total number of deaths of individuals aged 60 or older. Larger gaps in the shares of excess deaths can be driven by either higher mortality among the elderly population, lower mortality among younger cohorts, or both. The first control allows us to account for the fact that some cities may see a relatively larger share of excess deaths among the elderly because their population is relatively older. The second allows us to isolate the variation coming from differences in mortality among younger cohorts. Specifically, if the observed effects are driven by mortality in the older group, this coefficient should be different from zero, and other estimates may also change. While the coefficient on the share of the elderly in the population is not statistically different from zero, the point estimate for the total elderly deaths from COVID-19 is positive, large, and statistically significant. This implies that, if the the estimate of a coefficient changes after including this control, the

original coefficient is at least partially driven by cross-city differences in the mortality rates of the population aged 60 or older.

Figure 3: Correlates of the age-share difference in COVID-19 deaths across Brazilian cities



Notes: The figure reports coefficient estimates and confidence intervals of a multivariate regression at the city level, restricted to cities with population of at least 5,000 (N=1,385). The regression also includes geographic and climate controls (maximum annual temperature, average annual precipitation, and distance to Sao Paulo), state fixed effects, and a constant. The dependent variable is the difference between the shares in COVID-19 deaths of the elderly (aged 60 or older) and of the non-elderly (younger than 60). The bars represent confidence intervals at 90% (thick bars) and at 95% (thin bars), both from robust standard errors clustered at the state level. All regressors are standardized.

Urban scale and the labor market

We first look at the role of urban scale. City size is one of the strongest correlates of the overall impact of COVID-19 at the local level in Brazil (Chauvin, 2020), and large cities played a key role in the expansion of the epidemic by exporting the virus to the rest of the country (Candido et al., 2020). Population density also has a positive though smaller effect on both cases and deaths, whereas the share of inhabitants that live in urbanized areas within the city perimeter is strongly and positively related to the number of cases, but less so to the number of deaths. Our results suggest that this scale effect does not have

a systematic age bias. When we look at the effects on the age-share difference in COVID-19 deaths in our main specification we do not find a statistically significant correlation with the overall city population, density, or the share of people living in urbanized areas. When we introduce the control for elderly deaths, the coefficients on both population and density turn negative and statistically significant. This implies that there are more non-elderly COVID-19 deaths in larger cities, but that this doesn't affect the age-share difference because these cities also have more elderly deaths in similar proportion.

A second set of regressors captures labor market conditions in the city. The literature has pointed out that higher levels of informality may be connected with a faster spread of COVID-19 in developing countries, particularly due to the larger role that the informal sector plays in their economies (Alon et al., 2020; Busso et al., 2020). Because younger cohorts are more likely to be economically active, places with more informality could see a larger mortality gap between elderly and non-elderly individuals. Our results suggest that this is indeed the case. After controlling for the other covariates, informality rates and the share of high-school educated workers in employment have a negative and significant correlation with the age-share difference. This holds with and without the control for the total number of elderly deaths, implying that the smaller difference is driven by a higher mortality among individuals younger than 60 in cities with more informal labor markets and less-skilled labor forces. In contrast, the share of college graduates in employment, and measures of the share of workers in occupations that can be performed from home (Dingel and Neiman, 2020) and that involve high levels of physical contact (Mongey et al., 2020) have, on average, close to zero correlation with the age-share difference.

Another potential work-related source of exposure to COVID-19 infections that disproportionately affects the working-age population is commuting. The length of work-related trips could be positively linked to infection rates, as commuters frequently need to use public transportation, which has been linked to higher infection rates in other countries (Harris, 2020; McLaren, 2020). We fail to find a statistically significant correlation between the average commuting time in the city and the age-share difference in COVID-19 deaths.

Socioeconomic and housing conditions

A third group of risk factors that we consider include social and demographic characteristics of the local population. In the US, low-income populations have had a lower propensity to stay at home during the pandemic (Brown and Ravallion, 2020; Papageorge et al., 2020; Weill et al., 2020), and this population has disproportionately had more infections and deaths (Benitez et al., 2020; Borjas, 2020; Desmet and Wacziarg, 2020; McLaren, 2020). Both ethnic minorities and people from low-income households tend to be younger in Brazil

(Appendix Figure C14), suggesting that socioeconomic differences could partially account for the age-share difference in COVID-19 mortality. However, we find no statistically significant connections between the age-share difference and measures of income, poverty, inequality, and the presence of large ethnic minority populations across cities.

The literature has also highlighted the potential role of housing and neighborhood conditions in the spread of COVID-19 in developing countries. Our analysis includes a fourth set of regressors to assess the role that these mechanisms may play. We first consider access to piped water, a service that facilitate hygiene practices that help prevent the spread of the virus (Brown et al., 2020). We find that cities with higher access tend to have a larger age-share difference in COVID-19 deaths. This holds in both specifications reported on Figure 3, implying that the wider gap reflects fewer non-elderly deaths in places with more access to water. In contrast, we do not find a statistically-significant connection between the age-share difference and access to high-speed internet – which in principle facilitates remote work and education and thus stay-at-home behavior (Chiou and Tucker, 2020), particularly among younger cohorts.

Next, we look at residential overcrowding. We find this to be one of the strongest correlates of the COVID-19 deaths' age-share difference. Researchers have suggested that the virus may spread faster in the developing world because of the higher prevalence of large, often multi-family, multi-generational households living together in cramped conditions (Brotherhood et al., 2020a). Our results suggest that this is also related to larger shares of non-elderly individuals in COVID-19 deaths. Holding all other covariates constant, a one- standard-deviation increase in overcrowding – as measured by the average number of persons per room – is associated with a 15-percentage-point smaller elderly-non elderly shares difference in COVID-19 deaths across Brazilian cities. This is driven by larger non-elderly mortality, as evidenced by the results showing an even more pronounced effect when controlling for total non-elderly deaths. Nonetheless, we fail to find a statistically significant link between the age-share difference in deaths and measures of multi-generational co-habitation – including the share of elderly individuals living alone or the share of young adults living with the elderly, factors that researchers have argued could increase infections and deaths among the older group (Walker et al., 2020).

Another dimension of housing conditions that we also find to be connected with the age-share difference in COVID-19 deaths is the presence of slums. Brotherhood et al. (2020a) model COVID-19 dynamics in alternative scenarios, with and without favelas, and find that overall mortality increases with the presence of these informal neighborhoods. First, we find a positive correlation in the main model between the age-shares difference in COVID-19 deaths and the share of the local population that lives in favelas. When we

include the age controls in the second model, however, the point estimate flips sign. These results suggest that, in cities with a significant presence of these informal neighborhoods, elderly individuals tend to die at a higher rate – relative to the non-elderly – than elsewhere in the country.

Public health conditions

To complete our analysis we look at the effects of local pre-pandemic public health conditions on the age-share difference in COVID-19 deaths. First we consider differences in installed healthcare capacity. As previously discussed, mortality rates, and, specifically, the rates of recovery of different age groups, can be affected by the ability of each city's healthcare facilities to handle increased rates of infection. To measure this effect we collect data on the local number of hospital beds, ICU beds, and ventilators in the period right before the pandemic. After controlling for the other covariates, we do not find any link between these variables and the age-share difference in COVID-19 deaths across Brazilian cities.

Last, we consider the role of differences across cities in the prevalence of preexisting health conditions associated with severe complications from COVID-19. We examine this issue by using hospitalization data from the Brazilian Ministry of Health (see Appendix A for details). The prevalence of these conditions, on average, increases sharply with age (Appendix Figure C15), implying that cities with higher incidence of preexisting conditions are likely to have relatively worse COVID-19 outcomes for the elderly. We find this to be the case. Holding other covariates constant, we find that the share of the city's patients hospitalized in 2019 that had at least one preexisting condition is positively associated with the age-share difference in COVID-19 deaths. Introducing age controls in the second model flips the sign of the coefficient, indicating that higher incidence of preexisting conditions is also associated with higher mortality among non-elderly individuals, although to a lesser degree than is the case for the elderly.

5 Conclusions

Developing countries have a younger age profile of COVID-19 deaths than high-income countries. This paper studies the drivers of this difference using cross-country, patient-level, and cross-city analyses. Our cross-country results, based on a sample of 13 developing countries and 18 high-income countries, show that the difference is not explained by the relative age distributions of susceptible populations, or by differences in testing coverage. Instead, the difference stems from lower rates of recovery among younger cohorts in

developing countries, which in turn are associated with a higher prevalence of preexisting conditions linked to severe COVID-19 symptoms. The results from our patient-level analysis with data from Canada, Colombia, Mexico, and the US are consistent with this finding. They show, in addition, that non-elderly adults with a positive COVID-19 diagnosis have less access to ICUs in the lower-income countries, also contributing to lower recovery rates. A statistically significant gap in COVID-19 deaths between lower- and high-income countries remains among middle-aged adults even after controlling for these factors. Our city-level analysis, which uses data from Brazil, provides indirect evidence that differences in infection rates also play a role. This analysis shows that higher mortality in younger cohorts is associated with higher levels of informal employment, less access to piped water, and more residential overcrowding – all of which have been linked in the literature to faster disease spread. Further research is needed to provide a more complete understanding of these and other potential mechanisms.

References

Acemoglu, D., Chernozhukov, V., Werning, I., and Whinston, M. D. (2020). A multi-risk SIR model with optimally targeted lockdown. *NBER Working Paper 27102*.

Alfaro, L., Becerra, O., and Eslava, M. (2020). EMEs and COVID-19: Shutting down in a world of informal and tiny firms. *NBER Working Paper 27360*.

Alon, T. M., Kim, M., Lagakos, D., and Van Vuren, M. (2020). How should policy responses to the COVID-19 pandemic differ in the developing world? *NBER Working Paper 27273*.

Baqui, P., Bica, I., Marra, V., Ercole, A., and van der Schaar, M. (2020). Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study. *The Lancet Global Health*, 8(8):e1018–e1026.

Barnett-Howell, Z. and Mobarak, A. M. (2020). The benefits and costs of social distancing in rich and poor countries. *arXiv preprint*, (arXiv:2004.04867):1–14.

Benitez, J. A., Courtemanche, C. J., and Yelowitz, A. (2020). Racial and ethnic disparities in COVID-19: Evidence from six large cities. *NBER Working Paper 27592*.

Borjas, G. J. (2020). Demographic determinants of testing Incidence and COVID-19 infections in New York City neighborhoods. *NBER Working Paper 26952*.

Brotherhood, L., Cavalcanti, T., Da Mata, D., and Santos, C. (2020a). Slums and Pandemics. *Mimeo*.

Brotherhood, L., Kircher, P., Santos, C., and Tertilt, M. (2020b). An economic model of the COVID-19 epidemic: The importance of testing and age-specific policies. *IZA Institute of Labor Economics - Discussion Paper Series*, (13265):1–71.

Brown, C., Ravallion, M., and Van de Walle, D. (2020). Can the world's poor protect themselves from the new coronavirus? *NBER Working Paper 27200*.

Brown, C. S. and Ravallion, M. (2020). Inequality and the coronavirus: Socioeconomic covariates of behavioral responses and viral outcomes across US counties. *NBER Working Paper 27549*.

Busso, M., Camacho, J., Messina, J., and Montenegro, G. (2020). The challenge of protecting informal households during the COVID-19 pandemic: Evidence from Latin America. *IADB Discussion Paper 780*, (June).

Canabarro, A., Tenório, E., Martins, R., Martins, L., Brito, S., and Chaves, R. (2020). Data-driven study of the COVID-19 pandemic via age-structured modelling and prediction of the health system failure in Brazil amid diverse intervention strategies. *PLoS ONE*, 15(7 July):1–13.

Candido, D. S., Claro, I. M., de Jesus, J. G., Souza, W. M., Moreira, F. R., Dellicour, S., Mellan, T. A., du Plessis, L., Pereira, R. H., Sales, F. C., Manuli, E. R., Thézé, J., Almeida, L., Menezes, M. T., Voloch, C. M., Fumagalli, M. J., Coletti, T. M., da Silva, C. A., Ramundo, M. S., Amorim, M. R., Hoeltgebaum, H. H., Mishra, S., Gill, M. S., Carvalho, L. M., Buss, L. F., Prete, C. A., Ashworth, J., Nakaya, H. I., Peixoto, P. S., Brady, O. J., Nicholls, S. M., Tanuri, A., Rossi, Á. D., Braga, C. K., Gerber, A. L., de C Guimarães, A. P., Gaburo, N., Alencar, C. S., Ferreira, A. C., Lima, C. X., Levi, J. E., Granato, C., Ferreira, G. M., Francisco, R. S., Granja, F., Garcia, M. T., Moretti, M. L., Perroud, M. W., Castiñeiras, T. M., Lazari, C. S., Hill, S. C., de Souza Santos, A. A., Simeoni, C. L., Forato, J., Sposito, A. C., Schreiber, A. Z., Santos, M. N., de Sá, C. Z., Souza, R. P., Resende-Moreira, L. C., Teixeira, M. M., Hubner, J., Leme, P. A., Moreira, R. G., Nogueira, M. L., Ferguson, N. M., Costa, S. F., Proenca-Modena, J. L., Vasconcelos, A. T. R., Bhatt, S., Lemey, P., Wu, C. H., Rambaut, A., Loman, N. J., Aguiar, R. S., Pybus, O. G., Sabino, E. C., and Faria, N. R. (2020). Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science*, 369(6508):1255–1260.

Chauvin, J. P. (2020). Why does COVID-19 affect some cities more than others? Evidence from Brazil. *Unpublished manuscript, Inter-American Development Bank*.

Chiou, L. and Tucker, C. E. (2020). Social distancing, internet access and inequality. *NBER Working Paper 26982*.

Clark, A., Jit, M., Warren-Gash, C., Guthrie, B., Wang, H. H., Mercer, S. W., Sanderson, C., McKee, M., Troeger, C., Ong, K. L., Checchi, F., Perel, P., Joseph, S., Gibbs, H. P., Banerjee, A., and Eggo, R. M. (2020). Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. *The Lancet. Global health*, (20):1–15.

Davies, N. G., Klepac, P., Liu, Y., Prem, K., Jit, M., Pearson, C. A., Quilty, B. J., Kucharski, A. J., Gibbs, H., Clifford, S., Gimma, A., van Zandvoort, K., Munday, J. D., Diamond, C., Edmunds, W. J., Houben, R. M., Hellewell, J., Russell, T. W., Abbott, S., Funk, S., Bosse, N. I., Sun, Y. F., Flasche, S., Rosello, A., Jarvis, C. I., and Eggo, R. M. (2020). Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine*, 26(8):1205–1211.

Demombynes, G. (2020). COVID-19 age-mortality curves are flatter in developing countries. *World Bank Policy Research Working Paper*, (9313).

Desmet, K. and Wacziarg, R. (2020). Understanding spatial variation in COVID-19 across the United States. *NBER Working Paper 27329*.

Dingel, J. I. and Neiman, B. (2020). How many jobs can be done at home? *Journal of Public Economics*, (189).

Dowd, J. B., Andriano, L., Brazel, D. M., Rotondi, V., Block, P., Ding, X., Liu, Y., and Mills, M. C. (2020). Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proceedings of the National Academy of Sciences*.

Dudel, C., Riffe, T., Acosta, E., Van Raalte, A. A., Strozza, C., and Myrskylä, M. (2020). Monitoring trends and differences in COVID-19 case fatality rates using decomposition methods: Contributions of age structure and age-specific fatality. *medRxiv Preprint*, pages 1–17.

Goldstein, J. R. and Lee, R. D. (2020). Demographic perspectives on mortality of COVID-19 and other epidemics. *NBER Working Paper 27043*, page 18.

Hallal, P., Hartwig, F., Horta, B., Silveira, M., Struchiner, C., Vidaletti, L. P., Neumann, N., Pellanda, L. C., Dellagostin, O. A., Burattini, M. N., Menezes, A. M., Barros, F. C., Barros, A. J., and Victora, C. G. (2020). Remarkable variability in SARS-CoV-2 antibodies across Brazilian regions: Report on two successive nationwide serological household surveys. *medRxiv*, page 2020.05.30.20117531.

Harris, J. E. (2020). The subways seeded the massive coronavirus epidemic in New York city. *NBER Working Paper 27021*.

Hausmann, R. and Schetter, U. (2020). Horrible trade-offs in a pandemic: Lockdowns, transfers, fiscal space, and compliance. *CID Faculty Working Paper No. 382, Harvard University*.

IBGE (2016). *Arranjos Populacionais e Concentrações Urbanas do Brasil*. IBGE, Rio de Janeiro, second edition.

Ioannidis, J. P. A., Axfors, C., and Contopoulos-Ioannidis, D. G. (2020). Population-level COVID-19 mortality risk for non-elderly individuals overall and for non-elderly individuals without underlying diseases in pandemic epicenters. *medRxiv*.

Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772):700–721.

Klein, S. L., Dhakal, S., Ursin, R. L., Deshpande, S., Sandberg, K., and Mauvais-Jarvis, F. (2020). Biological sex impacts COVID-19 outcomes. *PLoS Pathogens*, 16(6):1–5.

Laxminarayan, R., Wahl, B., Dudala, S. R., Gopal, K., Mohan, C., Neelima, S., Reddy, K. J., Radhakrishnan, J., and Lewnard, J. A. (2020). Epidemiology and transmission dynamics of COVID-19 in two Indian states. *Science*.

Levin, A. T., Cochran, K. B., and Walsh, S. P. (2020). Assessing the age specificity of infection fatality rates for COVID-19: Meta-analysis & public policy implications. *NBER Working Paper 27597*.

Lyra, W., Do Nascimento, J. D., Belkhiria, J., De Almeida, L., Chrispim, P. P., and De Andrade, I. (2020). COVID-19 pandemics modeling with determinist SEIR, social distancing, and age stratification. The effect of vertical confinement and release in Brazil. *PLoS ONE*, 15(9)(e0237627).

Manski, C. F. and Molinari, F. (2020). Estimating the COVID-19 infection rate: Anatomy of an inference problem. *NBER Working Paper 27023*.

McLaren, J. (2020). Racial disparity in COVID-19 deaths: Seeking economic roots with census data. *NBER Working Paper 27407*.

Menezes-Filho, N. and Politi, R. (2020). Estimating the causal effects of private health insurance in Brazil: Evidence from a regression kink design. *Social Science and Medicine*, 264(510282):113258.

Mohamed, M. O., Gale, C. P., Kontopantelis, E., Doran, T., de Belder, M., Asaria, M., Luscher, T., Wu, J., Rashid, M., Stephenson, C., Denwood, T., Roebuck, C., Deanfield, J., and Mamas, M. A. (2020). Sex-differences in mortality rates and underlying conditions for COVID-19 deaths in England and Wales. *Mayo Clinic Proceedings*.

Mongey, S., Pilossoph, L., and Weinberg, A. (2020). Which workers bear the burden of social distancing policies? *Becker Friedman Institute Working Paper 2020-51*.

Papageorge, N., Zahn, M., Belot, M., Van den Broek-Altenburg, E., Choi, S., Jamison, J., and Tripodi, E. (2020). Socio-demographic factors associated with self-protecting behavior during the COVID-19 pandemic. *NBER Working Paper 27378*.

Philip, M., Ray, D., and Subramanian, S. (2020). Decoding India's low COVID-19 case fatality rate. *NBER Working Paper 27696*.

Roser, M., Ritchie, H., Ortiz-Ospina, E., and Hasell, J. (2020). Coronavirus Disease (COVID-19). *Our World in Data. Available at https://ourworldindata.org/coronavirus*.

Stock, J. (2020). Data gaps and the policy response to the novel coronavirus. *Nber Working Paper 26902*.

Walker, P. G. T., Whittaker, C., Watson, O. J., Baguelin, M., Winskill, P., Hamlet, A., Djafaara, B. A., Cucunubá, Z., Olivera Mesa, D., Green, W., Thompson, H., Nayagam, S., Ainslie, K. E. C., Bhatia, S., Bhatt, S., Boonyasiri, A., Boyd, O., Brazeau, N. F., Cattarino, L., Cuomo-Dannenburg, G., Dighe, A., Donnelly, C. A., Dorigatti, I., van Elsland, S. L., FitzJohn, R., Fu, H., Gaythorpe, K. A., Geidelberg, L., Grassly, N., Haw, D., Hayes, S., Hinsley, W., Imai, N., Jorgensen, D., Knock, E., Laydon, D., Mishra, S., Nedjati-Gilani, G., Okell, L. C., Unwin, H. J., Verity, R., Vollmer, M., Walters, C. E., Wang, H., Wang, Y., Xi, X., Lalloo, D. G., Ferguson, N. M., and Ghani, A. C. (2020). The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science*, 35(June).

Weill, J. A., Stigler, M., Deschenes, O., and Springborn, M. R. (2020). Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proceedings of the National Academy of Sciences*, 117(33):19658–19660.

Wiemers, E. E., Abrahams, S., AlFakhri, M., Hotz, V. J., Schoeni, R. F., and Seltzer, J. A. (2020). Disparities in vulnerability to severe complications from COVID-19 in the United States. *NBER Working Paper 27294*.

A Further data and methods details

A.1 Measure of high-risk of developing severe COVID-19 complications

The data to calculate age-specific shares of the population at high risk of developing severe COVID-19 complications used in the cross-country analysis come from Clark et al. (2020) and estimates they have made publicly available. These authors estimate for 5-year age groups in 188 countries the total population at high risk, defined as those that are likely to require hospitalization if infected by the virus. The estimates are based on infection-hospitalization ratios observed in China, which are adjusted to account for country-specific underlying conditions and other risk factors. The preconditions considered in the analysis include the following categories: cardiovascular disease, chronic kidney disease, chronic respiratory disease, chronic liver disease, diabetes, cancers with direct immunosuppression, cancers without direct immunosuppression but with possible immunosuppression caused by treatment, HIV/AIDS, tuberculosis, chronic neurological disorders, and sickle cell disorders.

A.2 Triple interaction estimates for gender and ethnic minorities

We estimate triple interaction models to explore how the result vary by gender and ethnic minority status in the patient-level data. In the cross-country analysis we use OLS to estimate the following equation:

$$Y_{a,s,c} = \alpha_0 + \sum^{3} \alpha_{1a} \left(a_{a,c} \times d_c \times m_{a,s,c}\right) + \sum^{3} \alpha_{2a} \left(a_{a,c} \times d_c\right) + \sum^{3} \alpha_{3a} \left(a_{a,c} \times m_{a,s,c}\right) + \sum^{3} \alpha_4 \left(m_{a,s,c} \times d_c\right) + \sum^{3} \alpha_5 a_{a,c} + \alpha_6 d_c + \alpha_7 m_{a,s,c} + \alpha_{\mathbf{c}}^{'} \mathbf{C_c} + \varepsilon_{a,s,c}$$

where the dependent variable is the share of each age group a in the total COVID-19 deaths corresponding to sex s in country c, $m_{a,s,c}$ is an indicator that takes the value 1 when the shares corresponds to men within each age group, and the value 0 when it corresponds to women, and all other variables are defined as in equation 1. The estimates of this regression with different sets of controls are reported in Appendix Table C2.

We use a similar equation to estimate, using a Probit model, patient-level conditional probabilities of dying or receiving healthcare by sex and by ethnic minority status. Specifi-

cally, we estimate:

$$Y_{i,a,g,c} = \delta_0 + \sum^{4} \delta_{1a} \left(a_{i,a,c} \times d_c \times g_{i,a,g,c}\right) + \sum^{4} \delta_{2a} \left(a_{i,a,c} \times d_c\right) + \sum^{4} \delta_{3a} \left(a_{i,a,c} \times g_{i,a,g,c}\right) + \sum^{4} \delta_4 \left(g_{i,a,g,c} \times d_c\right) + \sum^{4} \delta_5 a_{i,a,c} + \delta_6 d_c + \delta_7 g_{i,a,g,c} + \delta_{\mathbf{c}}^{'} \mathbf{C_c} + \varepsilon_{i,a,g,c}$$

where $Y_{i,a,g,c}$ is an indicator that individual i from age group a and minority group g in country c has either died or received a specific type of healthcare service (hospitalization or ICU), the term $g_{i,a,g,c}$ corresponds to a male dummy in the gender specification and to an ethnic minority dummy in the ethnicity specification (as defined in Appendix B, Table 2), and all other variables are defined as in equation 2. The results of both specifications are reported in Appendix figure C13.

A.3 Estimation of predictive margins

Figures 1 and 2 in the main text plot, as well as figures C7 through C13 and tables C1, C6 and C7 in the Appendix report the predictive margins for each age group included in their respective regressions. They are all computed using the *margins* command in Stata. In the case of Figure 1 and Appendix Table C1, this corresponds to the average predicted shares of each age group in their country's total COVID-19 deaths by country type (developing or high-income). It is computed using the actual values of all regressors in the data and the estimated coefficients of the corresponding OLS regression on table 1 plus all the coefficients on the control variables and the constant (not reported in the table). In the case of Figure 2 and appendix figures C7 through C13 and tables C6 and C7, the reported margins correspond to the predicted conditional probabilities of dying, being hospitalized, or entering an ICU for each age group by country type. The computations are based on the estimated coefficients of Probit regressions and the actual value of the regressors in the data. The standard errors of the predicted shares and probabilities are calculated using the Delta method.

A.4 Patient-level analysis at similar testing coverage and positivity rates

To explore the sensitivity of our patient-level results to differences in testing across countries we produce estimates of our main specification using data from dates in which the countries in the sample had similar testing characteristics using data from Roser et al. (2020). First, we construct a data set in which the sample for each country is censored at a given date, selected so that they have the closest levels of tests per capita between them. This corresponds to

5.58 for Canada (March 29), 5.81 for Colombia (May 28), 5.88 for Mexico (July 17), and 5.92 for USA (April 6). The estimates using this data set are reported in Appendix Figure C7. Second, we repeat the procedure but focusing on dates in which the positivity rates are as close as possible across countries. This corresponds to 6.6% in Canada (May 6), 12.5% in Colombia (June 30), 12.5% in Mexico (April 17), and 12.1% in USA (May 7). The positivity rates match is less precise, because Canada's rate was consistently below that of all other countries across the period. Moreover, the Mexico data corresponds to an early stage of the country's outbreak, when there were only 486 reported deaths, making age-specific estimates noisy. The country's positivity rate increased monotonically since that date until the time of writing. The estimates using this data set are reported in Appendix Figure C8.

A.5 Brazilian cities definition and COVID-19 deaths data

The local administrative boundary typically associated with cities in Brazil is the municipality. However, municipalities are oftentimes integrated to larger metropolitan areas through commuting and other economic links, being effectively part of the same local labor and housing markets. To account for this, we use a definition that considers all municipalities within the same commuting zone as a single city. We use the commuting zones ("arranjos populacionais") definition provided by the Brazilian Institute of Geography and Statistics (IBGE, 2016). All municipalities that are not part of one of these commuting zones are considered a city on their own.

The COVID-19 deaths data from Brazil comes from the SIVEP-Gripe data set —also used in Baqui et al. (2020)—. This data is at the patient level, and includes all cases diagnosed with any type of Severe Acute Respiratory Syndrome (SARS), whose symptoms included fever accompanied by cough or sore throat, along with dyspnea or oxygen saturation lower than 95% or respiratory distress, or who died due to Severe Acute Respiratory Syndrome (SARS) regardless of hospitalization. The total number of deaths of patients for whom the final diagnostic was COVID-19 coincides with the official cumulative deaths count for the disease on that date (Roser et al., 2020). We aggregate deaths by age group at the city level using the city definition described above. See Data Appendix B for further details.

A.6 Prevalence of preexisting conditions in Brazilian cities

To study the role of preexisting conditions in the elderly–non-elderly gap in COVID-19 death shares across Brazilian cities we construct measures of the local prevalence of these morbidities. We first map the list of conditions provided by Clark et al. (2020) to the classification used in the DATASUS system from the Brazilian Ministry of Health. We then

use 2019 hospitalization-level data provided by this system to produce age-specific counts of the number of patients with at least one precondition in each city, and normalize it by the total number of patients, both measured in 2019. All data sources are detailed in Appendix B.

A.7 Data sources and availability

All of our source data are publicly available, and the original sources are listed in Appendix Table 4. In addition, the consolidated data, along with all the replication files for the results of this paper, will be made publicly available in a Github repository upon publication.

B Data appendix

Table 1: Country-level variables sources and description

Variable	Sources	Description / comments
Share of age group in national COVID-19 deaths	Multiple data sources (see Table 4)	Share of each age group in the total of deaths from COVID-19 in the country to date
GDP per capita	World Development Indicators	Gross Domestic Product divided by population
Population	World Development Indicators	Total population
Shares of age group in total population	World Population Prospects (United Nations)	Estimates of the share of each age group in the total population of the country in 2019.
Total tests performed	Our World in Data	Logarithm of the total number of tests performed in the country to date
Positivity rate	Our World in Data	Share of positive test over the total number of tests
School closure delay	Our World in Data	Logarithm of the number of days between the date of the first confirmed case until schools were closed
Stay at home delay	Our World in Data	Logarithm of the number of days between the date of the first confirmed case until stay-at-home orders were issued
Total cases	Our World in Data	Cumulative confirmed COVID-19 cases
Total deaths	Our World in Data	Cumulative confirmed COVID-19 deaths
Share in all cause deaths	World Population Prospects	Estimates of the share of each age group in the total number of deaths for any cause in a year
Share in high-risk of severe complications	Clark et al. (2020)	Share of individuals of an age group in high risk of severe COVID-19 complications due to preconditions.
Participation female	ILOSTAT database	Percentage of female population ages 15-64 participating in the labor force
Participation male-female	ILOSTAT database	Difference male-female in labor force participation

Notes: All data sources are publicly available. The data from the World Population Prospects can be found at https://population.un.org/wpp. The variables from Our World in Data (Roser et al., 2020) can be found at https://github.com/owid/covid-19-data/tree/master/public/data. The Clark et al. (2020) estimates are available at https://cmmid.github.io/topics/covid19/Global_risk_factors.html. The World Development Indicators and the ILOSTAT data can be accessed through the World Bank Open Data Platform, at https://data.worldbank.org.

Table 2: Patient-level variables sources and description

Variable	Availability	Description / comments
Developing country	Colombia, Mexico, Philippines, Canada, USA, Czech Republic, and Korea	Indicator that takes the value one if a country was classified as low or middle income by the World Bank in 2019.
Age	Colombia, Mexico, Philippines, Canada, USA, Czech Republic, and Korea	Age declared by patient at the time of test
Male	Colombia, Mexico, Philippines, Canada, USA, Czech Republic, and Korea	Indicator that takes the value one if the patient is male
Dead	Colombia, Mexico, Philippines, Canada, USA, Czech Republic, and Korea	Indicator that patient died from COVID-19 as registered by local authorities
Hospitalized	Colombia, Mexico, Philippines, Canada and USA	Indicator that patient was hospitalized for COVID-19
ICU	Colombia, Mexico, Canada and USA	Indicator that patient was admitted to the ICU for COVID-19
Ethnic Minority	Colombia, Mexico and USA	For Mexico indicates whether the patient speaks an indigenous language, for Colombia it indicates whether patient is black or Indigenous, and for the US if the patient is hispanic, black or Native American.
Preconditions	Mexico and USA	Indicator of whether the patient presents at least one of the following conditions: diabetes, hypertension, severe obesity, cardiovascular disease, chronic renal disease, asthma, emphysema, COPD or smoking.

Notes: Data sources are the same as those described in Table 4.

Table 3: Brazilian city-level variables description

Variable	Sources	Description / comments
Age shares gap	Ministry of Health (SIVEP-Gripe)	Difference in the mortality rates between individuals aged 60 or older and individuals aged less than 60 according to the SIVEP-Gripe data.
Informality rate	Census 2010 (IBGE)	An informal worker is someone who during the period of reference worked without a signed work card, or was self-employed. The informality rate is the share of informal workers in the labor force.
Share of high-school in employment	Census 2010 (IBGE)	Share of employed population with high-school but not college degree
Share of college in employment	Census 2010 (IBGE)	Share of employed population with at least college degree
Average commute	Census 2010 (IBGE)	Average commuting time of workers, estimated based on midpoints of the time intervals available in the census (in logs).
Share of jobs that can be done from home	Census 2010 (IBGE)	Share of workers in occupations that can be done from home according to the classification by Dingel and Neiman (2020).
Share of workers in high contact occupations	Census 2010 (IBGE)	Share of workers with jobs that involve high levels of physical contact according to the classification by Mongey et al. (2020).
Share of black and mixed in population	Census 2010 (IBGE)	Share of self-identified black and mulatto individuals in total population.
Extreme poverty rate	Census 2010 (IBGE)	Share of families with a daily income less than 1.90 US dollars per day in 2011 PPP.
Gini (hourly wages)	Census 2010 (IBGE)	Gini coefficient of the hourly wages of city workers.
Average income per capita	Census 2010 (IBGE)	City-level average of the household income per capita capita, calculating dividing total household income from all sources by the number of people in the household.
Urban share	Census 2010 (IBGE)	Proportion of individuals living in urban areas according to the census definition.
Population	Census 2010 (IBGE)	Number of persons living in the city (in logs).
Population density	Census 2010 (IBGE)	Ratio between total population and the area of the city in square kilometers (in logs).
Share of 60+ in population	Census 2010 (IBGE)	Share of individuals aged 60 or more in population.
Access to drinking water	Census 2010 (IBGE)	Share of households with access to piped water.
Broadband density	National Telecommunications Agency (ANATEL)	Number of accesses to fixed broadband per 100,000 households in February 2020.
Share of young adults living with elderly	Census 2010 (IBGE)	Share of population aged 20 to 44 living in a household with at least one person aged 65 or older.
Share of elderly living alone	Census 2010 (IBGE)	Share of population aged 65 or older living in a single-person household.
Average persons per room	Census 2010 (IBGE)	Cross-households average of the number of persons - number of rooms household ratio.
Share of population in favelas	Census 2010 (IBGE)	Share of household located in "abnormal agglomerations" by the census definition.
Number of doctors	Health ministry (DATASUS)	Number of doctors in local hospitals in February 2020 (inverse hyperbolic sine).
ICU beds	Health ministry (DATASUS)	Number of beds in Intensive Care Units in local hospitals in February 2020 (inverse hyperbolic sine).
Ventilators	Health ministry (DATASUS)	Number of ventilators in local hospitals in February 2020 (inverse hyperbolic sine).
Patients with at least in population	Health ministry (DATASUS)	Number of patients with at least one of the morbidities associated with severe COVID-19 complications identified in Clark et al. (2020).
Maximum yearly temperature	University of East Anglia Climatic Research Unit (CRU)	Highest registered temperature between 1900-2019, interpolated to the city level.
Average yearly precipitation	University of East Anglia Climatic Research Unit (CRU)	Average yearly precipitation 1900-2019, interpolated to the city level.
Distance to Sao Paulo	Census 2010 (IBGE)	Distance (in KM) of the shortest path between the city's centroid and Sao Paulo's centroid.
Elderly population COVID-19 deaths	Ministry of Health (SIVEP-Gripe)	Total number of confirmed COVID-19 deaths of individuals aged 60 or older (in logs).

Notes: All data sources are publicly available. The SIVEP-Gripe data can be obtained from http://plataforma.saude.gov.br/coronavirus/dados-abertos. The 2010 census data is available at https://downloads.ibge.gov.br. The ANATEL data can be found at https://www.anatel.gov.br/paineis/acessos. The DATASUS microdata can be obtained from http://www2.datasus.gov.br/DATASUS/index.php. The CRU data is available at http://www.cru.uea.ac.uk/data. All city-level CENSUS, SIVEP-Gripe and DATASUS variables are constructed from the microdata, using provided sample weights in the case of the CENSUS.

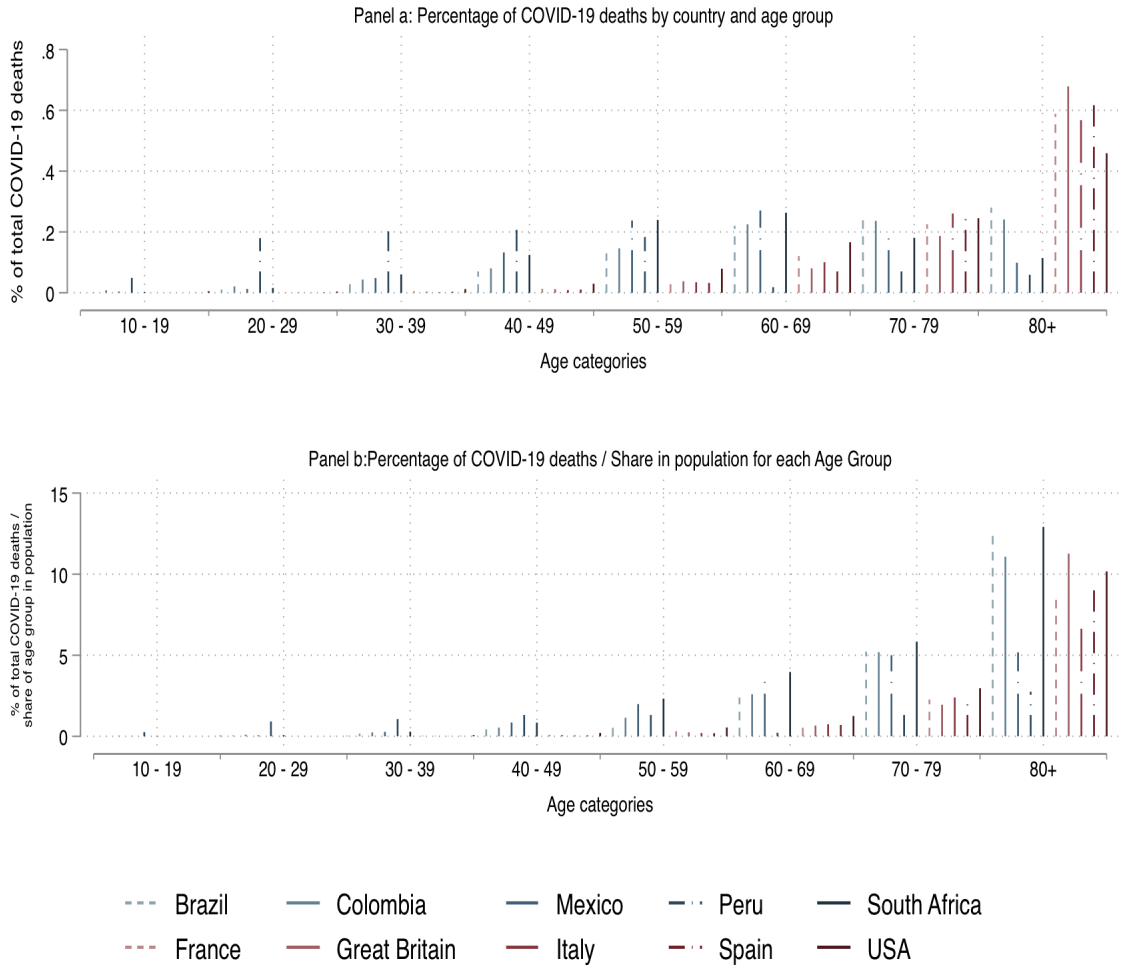
Table 4: International data sources on COVID-19 mortality

Country	Updated	Sources	URL
		Panel A: High-income countries	
Australia	7/11/20	Australian Government Department of Health	https://www.health.gov.au/news/health-alerts/novel-coronavirus-2019-ncov-health-alert/coronavirus-covid-19-current-situation-and-case-numbers#total-cases-and-deaths-by-state-and-territory
Belgium	6/7/20	Belgian institute for health	https://epistat.wiv-isp.be/covid/
Canada	7/16/20	Public Health Agency of Canada	https://www150.statcan.gc.ca/t1/tbl1/fr/tv.action?pid=1310078101
Czech Republic	7/12/20	Health Ministry Czech Republic	https://onemocneni-aktualne.mzcr.cz/api/v2/covid-19
France	8/18/20	Health Ministy of France	https://www.gouvernement.fr/info-coronavirus/carte-et-donnees
Germany	6/9/20	Robert Koch Institut	https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/dd4580c810204019a7b8eb3e0b329dd6_0?page=3123
Hungary	7/3/20	Hungarian Government	https://koronavirus.gov.hu/elhunytak
Israel	5/20/20	Independent news-sourced data repository	https://github.com/dennisyar/covid19_isr
Italy	8/18/20	Statista	https://www.statista.com/topics/6061/coronavirus-covid-19-in-italy/
Japan	7/3/20	Ministry of Health, Labor and Welfare via Tokyo Keizai	https://toyokeizai.net/sp/visual/tko/covid19/en.html
Netherlands	6/7/20	National Institute for Public Health and the Environment	https://www.rivm.nl/coronavirus-covid-19/grafieken
Portugal	7/12/20	General Directorate of Health	https://covid19.min-saude.pt/relatorio-de-situacao/
South Korea	6/28/20	Korea Centers for Disease Control & Prevention	https://www.kaggle.com/kimjihoo/coronavirusdataset
Spain	8/18/20	Statista	https://es.statista.com/estadisticas/1125974/covid-19-porcentaje-de-fallecimientos-por-edad-y-genero-en-espana/
Sweden	6/7/20	Public Health Agency of Sweden	https://www.folkhalsomyndigheten.se/smittskydd-beredskap/utbrott/aktuella-utbrott/covid-19/statistik-och-analyser/bekraftade-fall-i-sverige/
Switzerland	6/10/20	Federal Office of Public Health	https://www.bag.admin.ch/bag/en/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/situation-schweiz-und-international.html#-1199962081
United Kingdom	8/18/20	Office for National Statistics	https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/weeklyprovisionalfiguresondeathsregisteredinenglandandwales
USA	6/27/20	Centers for Disease Control and Prevention	https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data/vbim-akqf
		Panel B: Developing countries	
Afghanistan	7/3/20	Ministry of Public Health	http://covid.moph-dw.org
Argentina	5/5/20	Health Ministry of Argentina	https://www.argentina.gob.ar/salud/coronavirus-COVID-19/sala-situacion
Brazil	6/11/20	Ministry of Health (SIVEP-Gripe)	http://plataforma.saude.gov.br/coronavirus/dados-abertos
Colombia	7/16/20	National Health Insitutite of Colombia	https://www.datos.gov.co/Salud-y-Protecci-n-Social/Casos-positivos-de-COVID-19-en-Colombia/gt2j-8ykr/data
Ecuador	5/7/20	Ministry of Public Health	https://www.coronavirusecuador.com/data
Kenya	7/8/20	Ministry of Health	https://www.health.go.ke/#1591180376422-52af4c1e-256b
Malaysia	7/3/20	Ministry of Health of Malaysia	http://covid-19.moh.gov.my
Mexico	7/16/20	General Directorate of Epidemiology of Mexico	https://www.gob.mx/salud/documentos/datos-abiertos-152127
Pakistan	7/3/20	Government of Pakistan	http://covid.gov.pk/stats/pakistan
Peru	8/18/20	Health Ministy of Peru	https://www.datosabiertos.gob.pe/group/datos-abiertos-de-covid-19
Phillipines	7/9/20	DOH Epidemiology Bureau of the Phillipines	https://www.doh.gov.ph/2019-nCoV
Romania	6/12/20	Independent news-source data repository	https://covid-19.shinyapps.io/romania/#facts
South Africa	8/18/20	Data Science for Social Impact research group at the University of Pretoria	https://github.com/dsfsi/covid19za

C Additional figures and tables

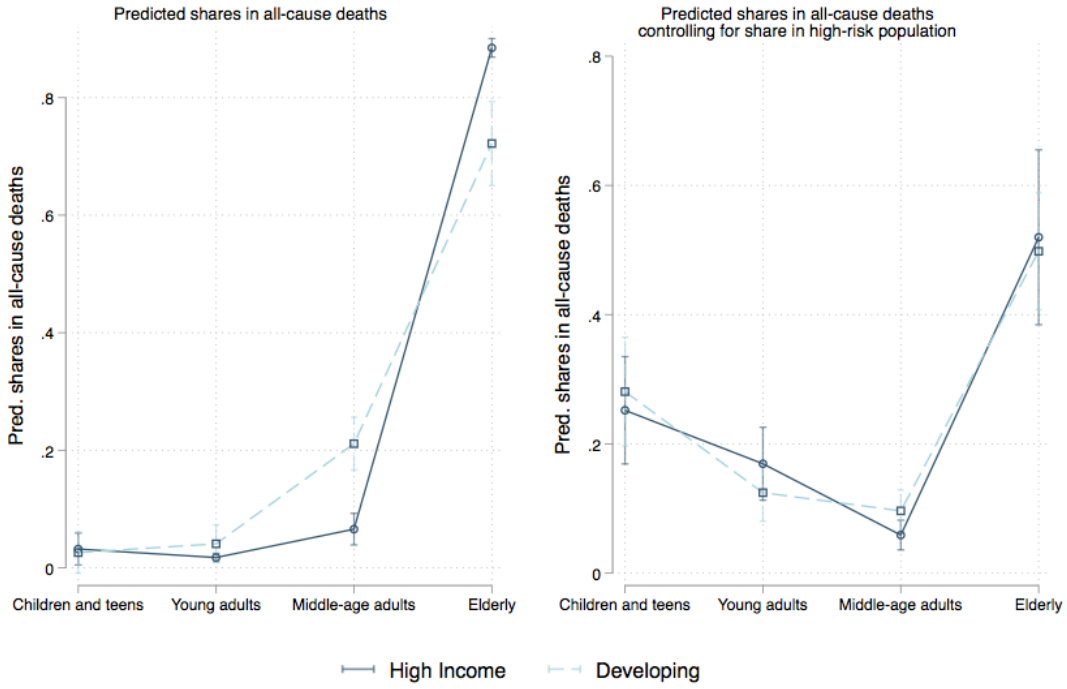
C1 Additional figures

Figure C1: Distributions of age-group shares in COVID-19 deaths among highly affected countries



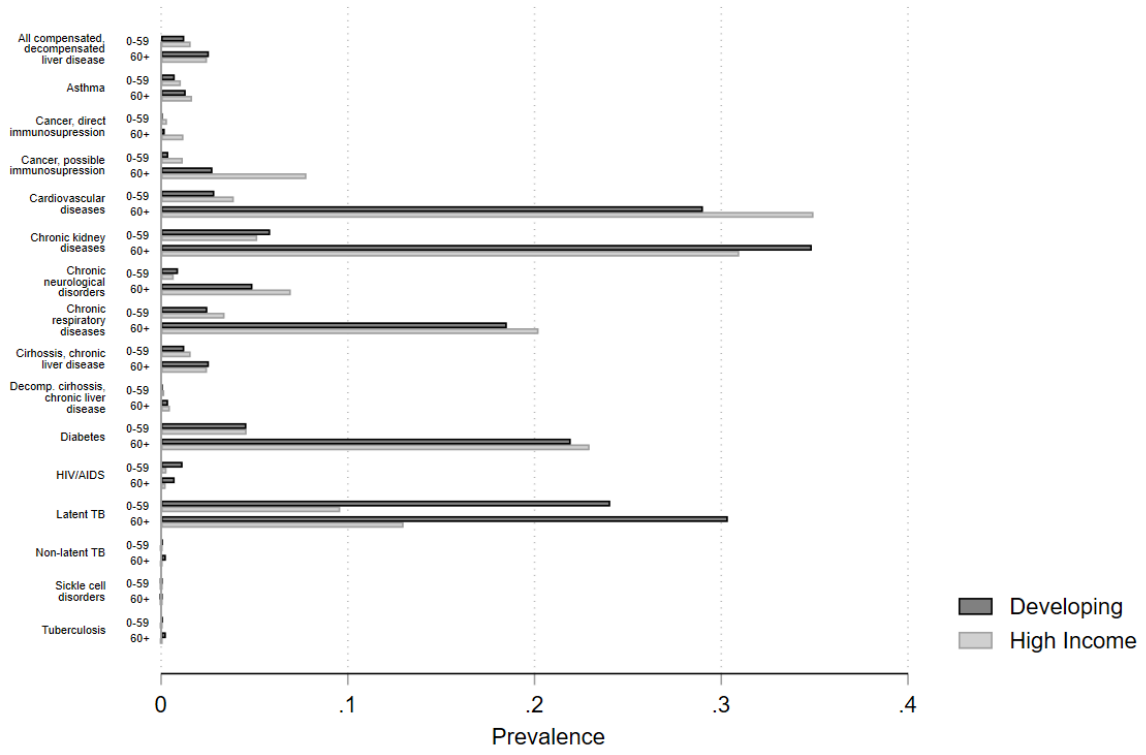
Notes: Panel (a) shows the share of each age group in total deaths attributed to COVID-19 by country, expressed in percentages. Panel (b) shows these shares normalized by the shares of each age group in the country's population. Although India and Russia are among the countries with the highest number of cases in the world at the time of writing, they were unable to obtain age-specific mortality rates for these countries. Data sources are listed in Table 4.

Figure C2: Predicted age-group shares in all-cause deaths in developing and high-income countries



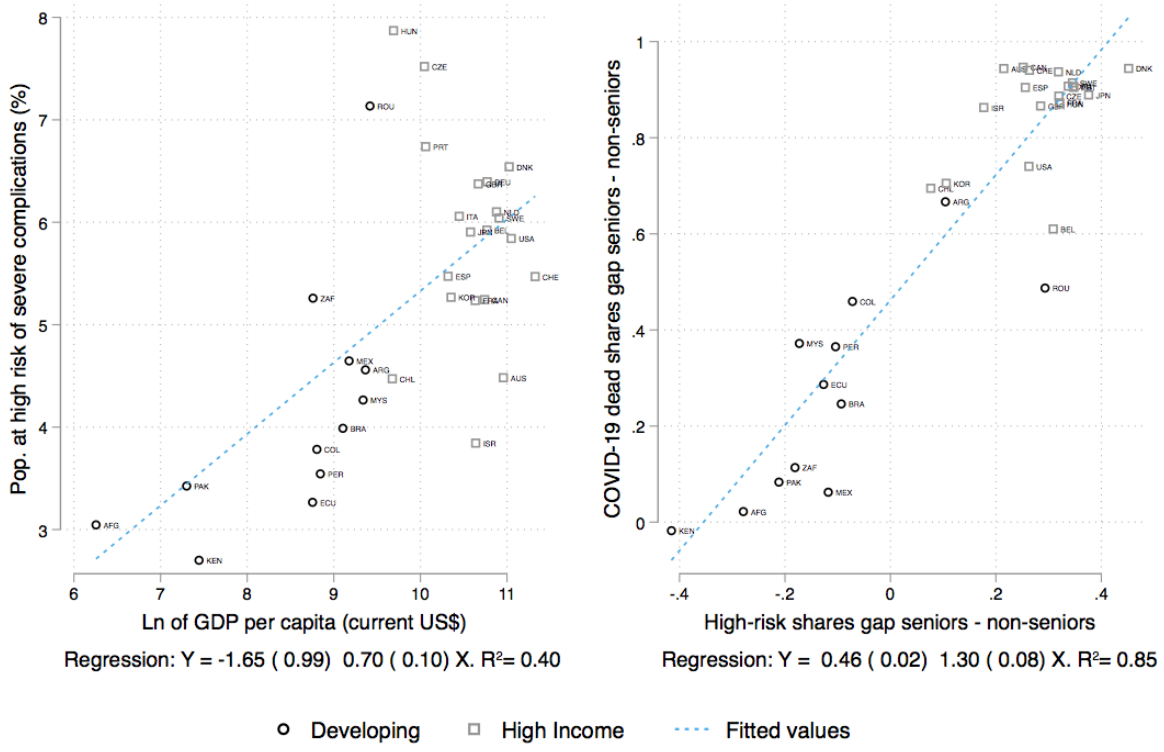
Notes: This figure reports estimates of equation 1 using as dependent variable the age-specific shares in all-cause deaths. The reported 95% confidence intervals are constructed using the Delta method. See Appendixes A and B for further details.

Figure C3: Prevalence of pre-existing conditions associated with severe COVID-19 complications by age group across countries in sample



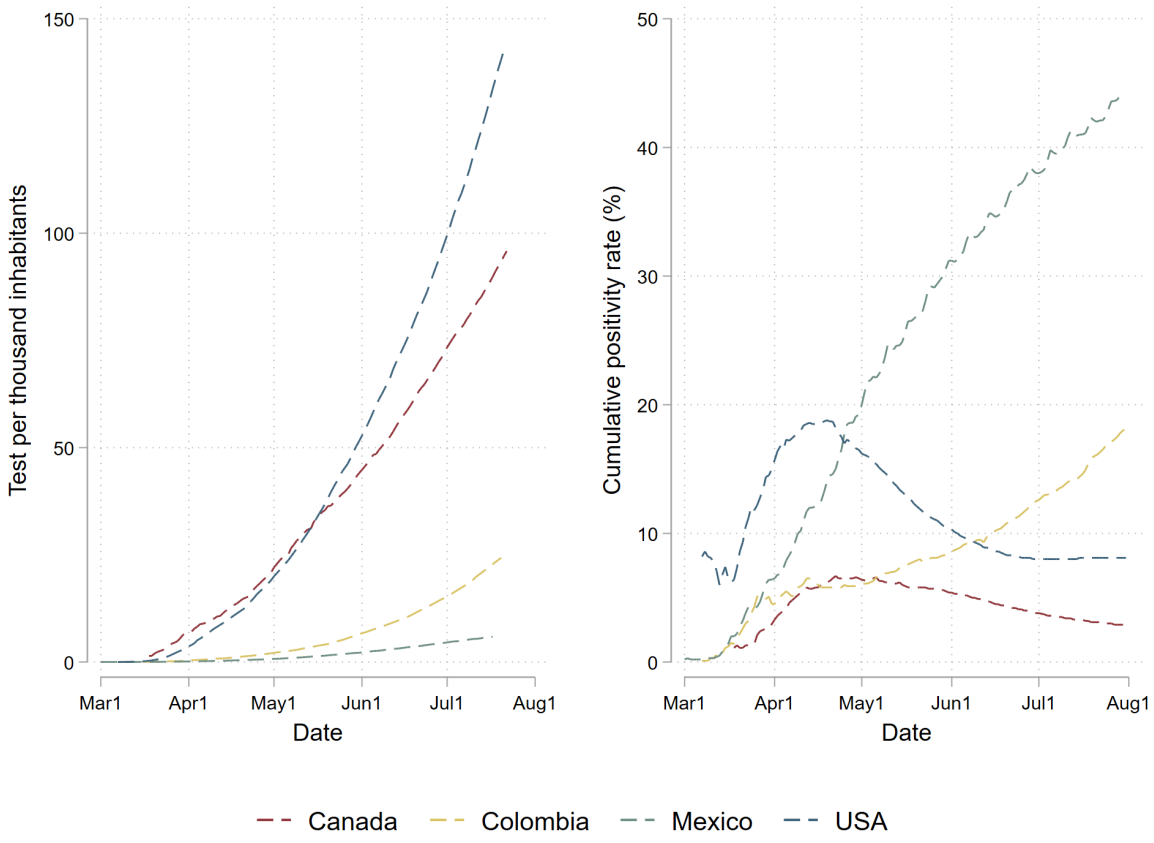
Notes: The bars represent cross-country population-weighted averages of the country-age specific prevalence of pre-existing conditions associated with severe COVID-19 complications, based on data from Clark et al. (2020).

Figure C4: High-risk of developing severe COVID-19 complications due to pre-existing conditions, income levels, and the age gap in mortality



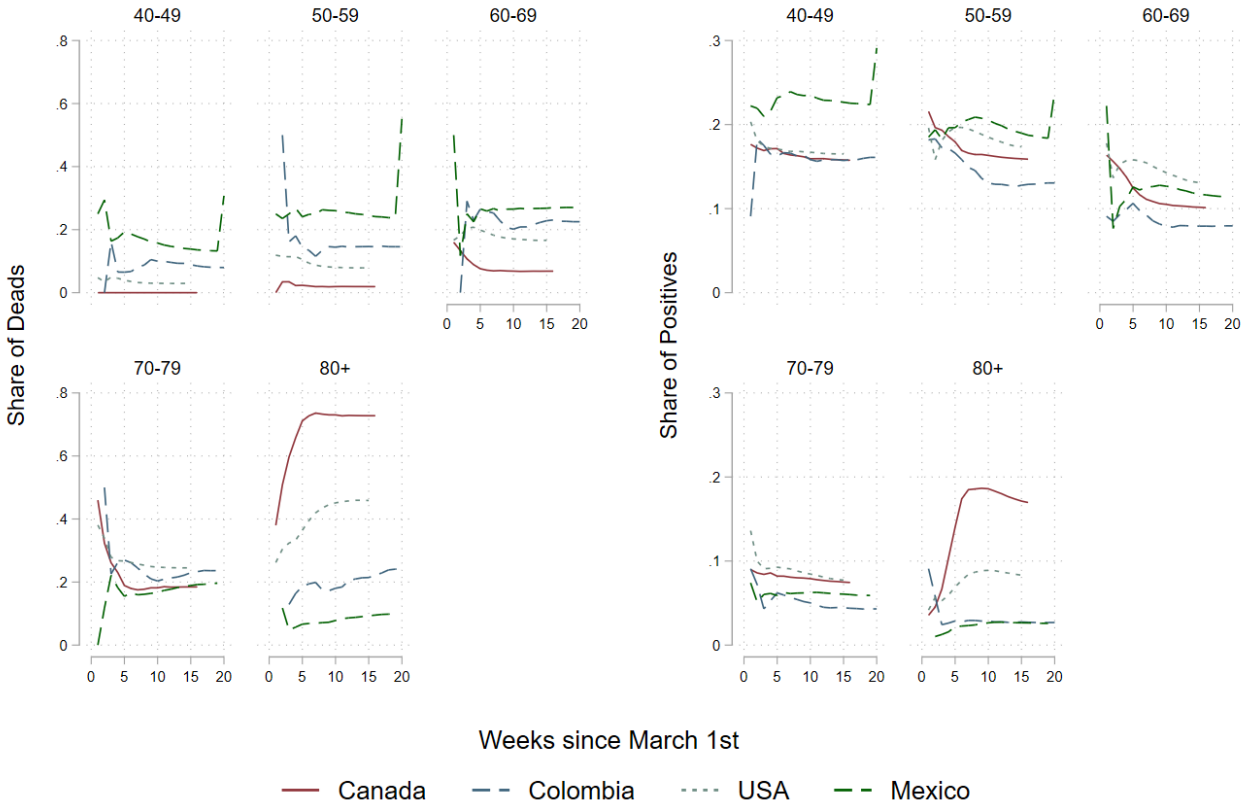
Notes: Country-level regressions. Data sources described in table 1.

Figure C5: Tests per capita and positivity rates in Canada, Colombia, Mexico and the US over time



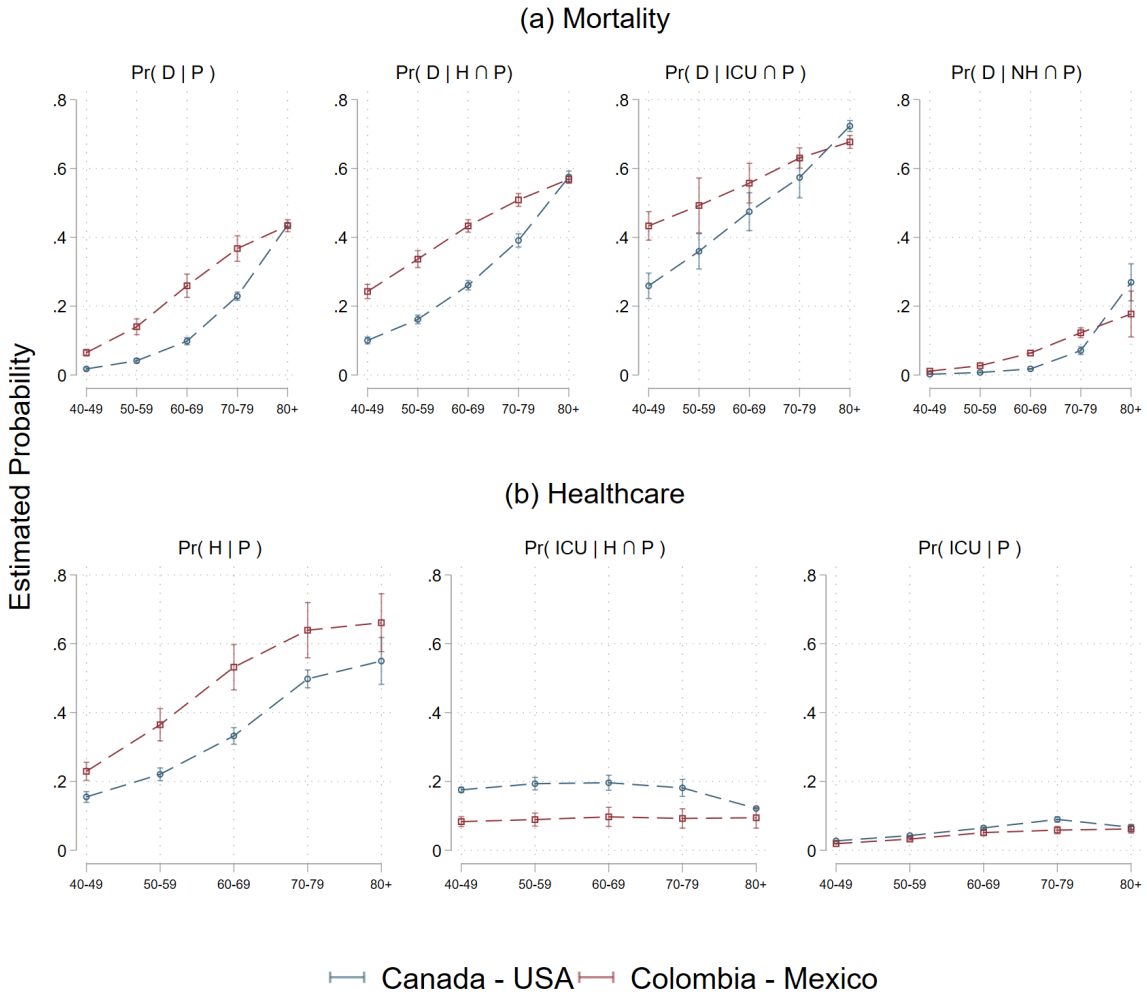
Source: Our World in Data Roser et al. (2020).

Figure C6: Evolution over time of the share in COVID-19 positive diagnoses and deaths of each age group across countries



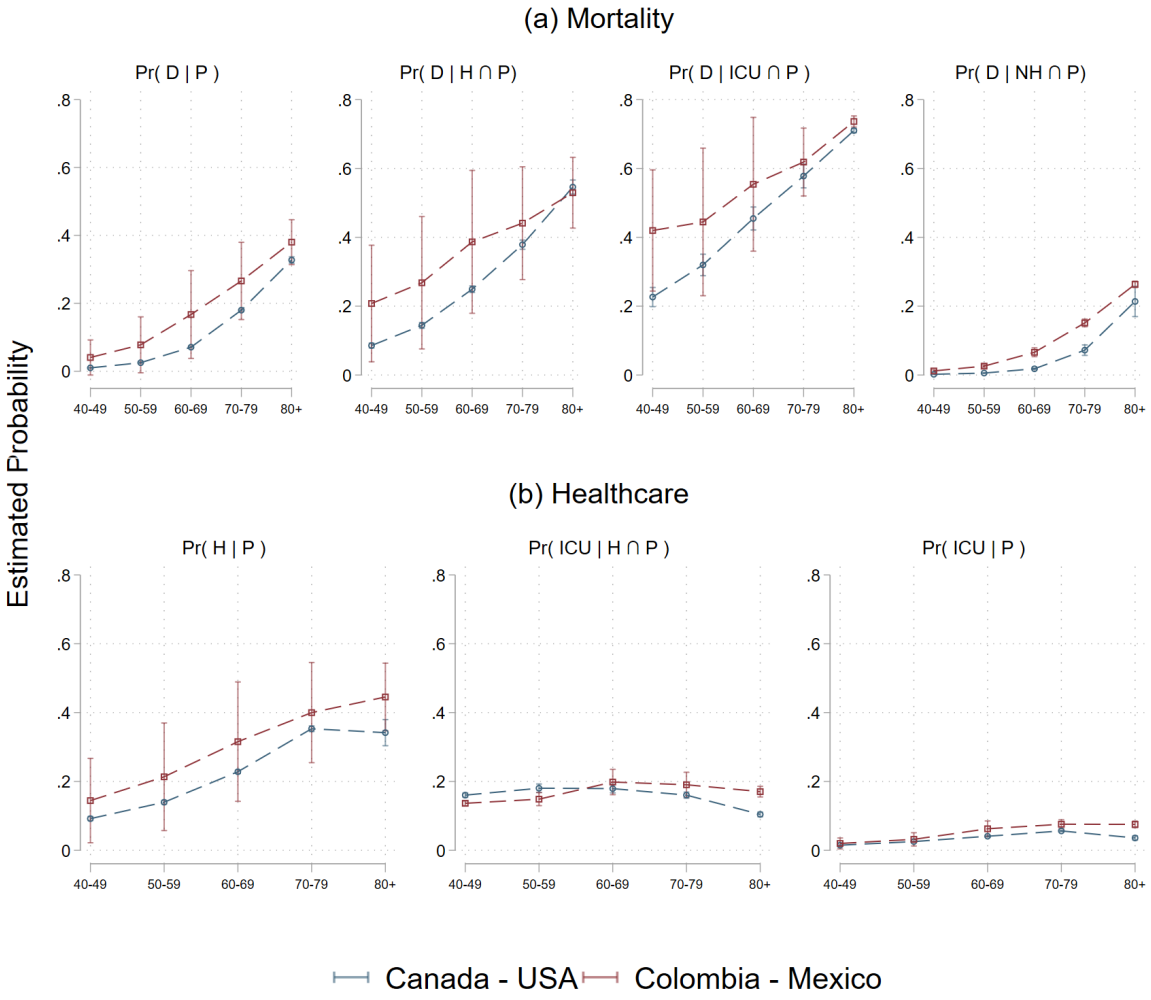
Note: Data from different sources, detailed in Table 1.

Figure C7: Differences between developing and high-income countries in the share of COVID-19 deaths by age group, at similar levels of testing per capita across countries



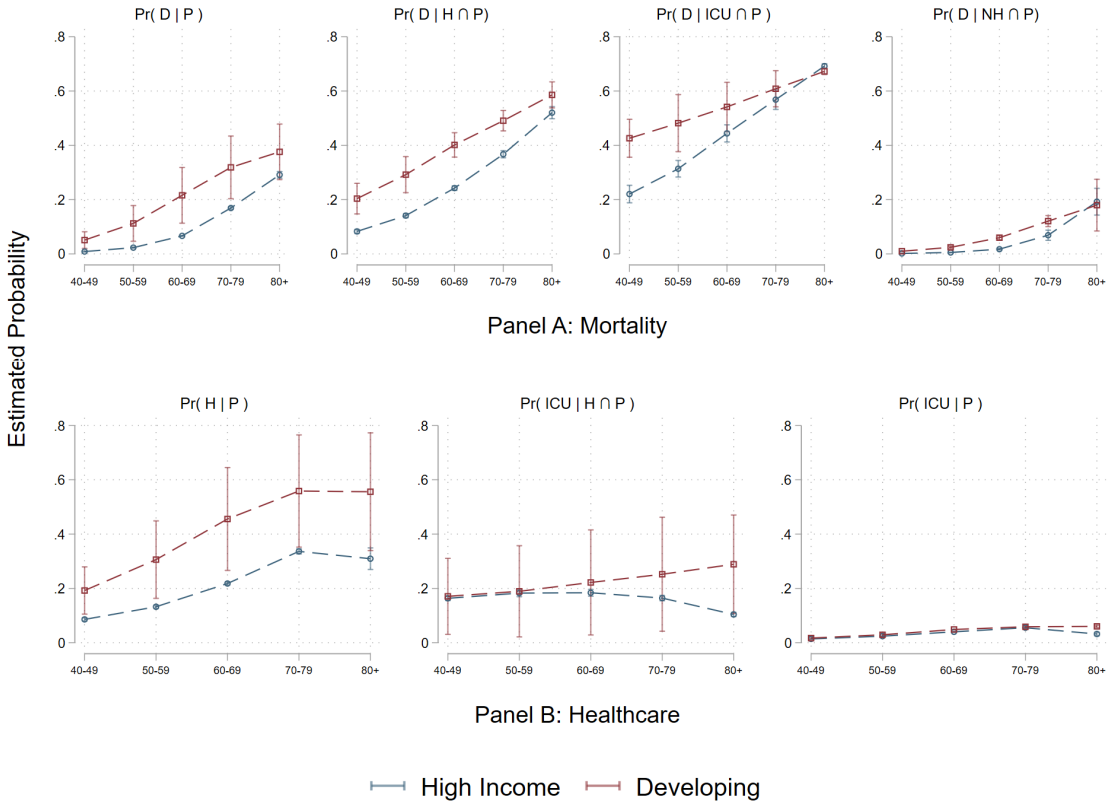
Notes: The figure shows the estimated conditional probabilities of dying, being hospitalized, or entering an ICU by country income level with a 95% CI. Regressions estimated at individual level using a Probit model. Robust standard errors clustered at the country level.

Figure C8: Differences between developing and high-income countries in the share of COVID-19 deaths by age group, at similar positivity rates across countries



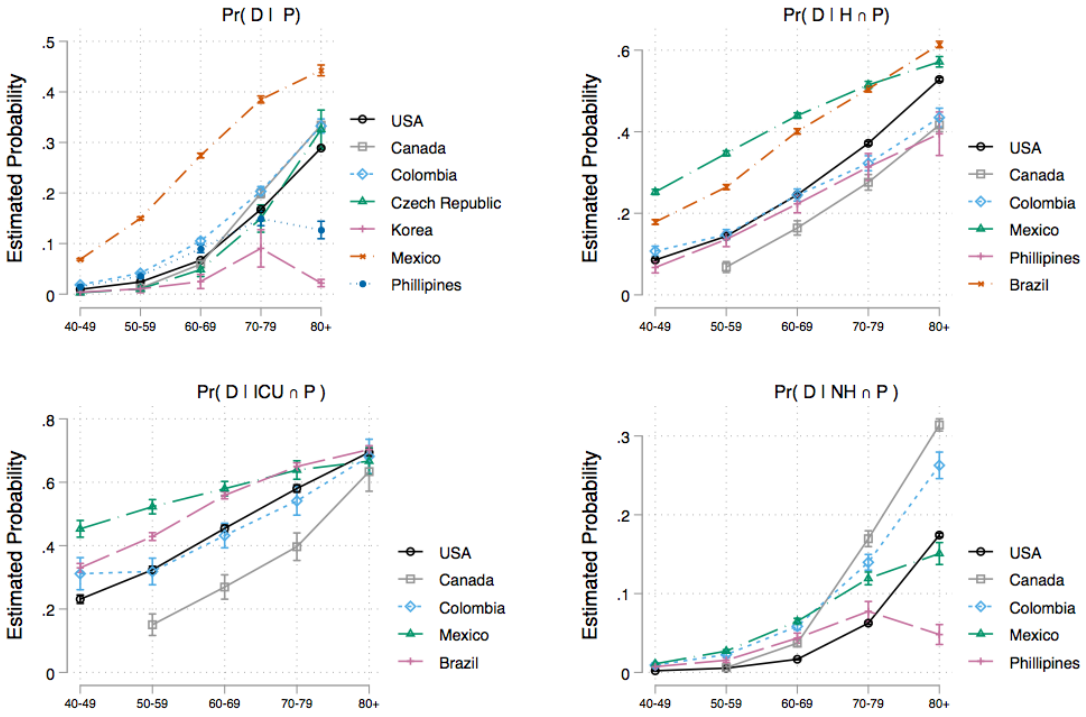
Notes: The figure shows the estimated conditional probabilities of dying, being hospitalized, or entering an ICU by country income level with a 95% CI. Regressions estimated at individual level using a Probit model. Robust standard errors clustered at the country level.

Figure C9: Differences between developing and high-income countries in the age-specific conditional probabilities of dying from COVID-19 and accessing healthcare when infected, using eight countries with available information



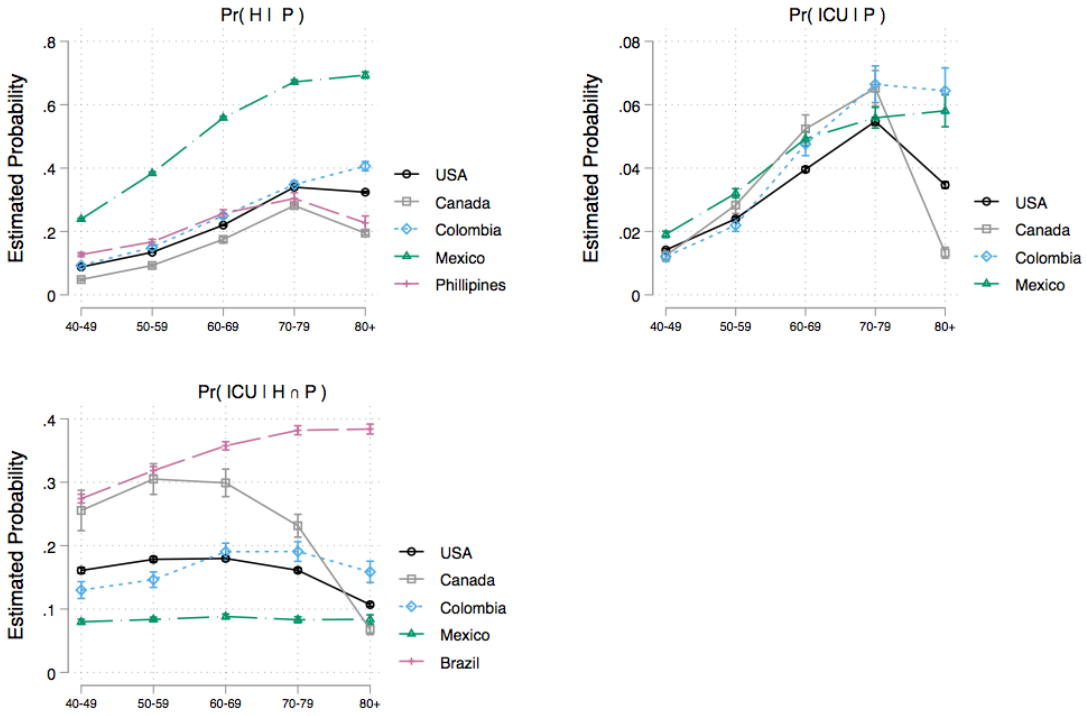
Notes: The figure shows the estimated conditional probabilities of dying, being hospitalized, or entering an ICU by country income level with a 95% CI. Regressions estimated at individual level using a Probit model. Robust standard errors clustered at the country level. In contrast with the main figure in the paper, the countries included in each graph vary, depending on the data available. The countries included in each figure are the following: All except Brazil in $\Pr(D \mid P)$; Brazil, Canada, Colombia, Mexico, Phillipines and USA in $\Pr(D \mid H \cap P)$; Brazil, Canada, Colombia, Mexico and USA in $\Pr(D \mid ICU \cap P)$ and $\Pr(ICU \mid H \cap P)$; Canada, Colombia, Mexico, Phillipines, and USA in $\Pr(D \mid NH \cap P)$ and $\Pr(H \mid P)$; and Canada, Colombia, Mexico and USA in $\Pr(ICU \mid P)$.

Figure C10: Age-specific conditional probabilities of dying from COVID-19 when infected for eight countries



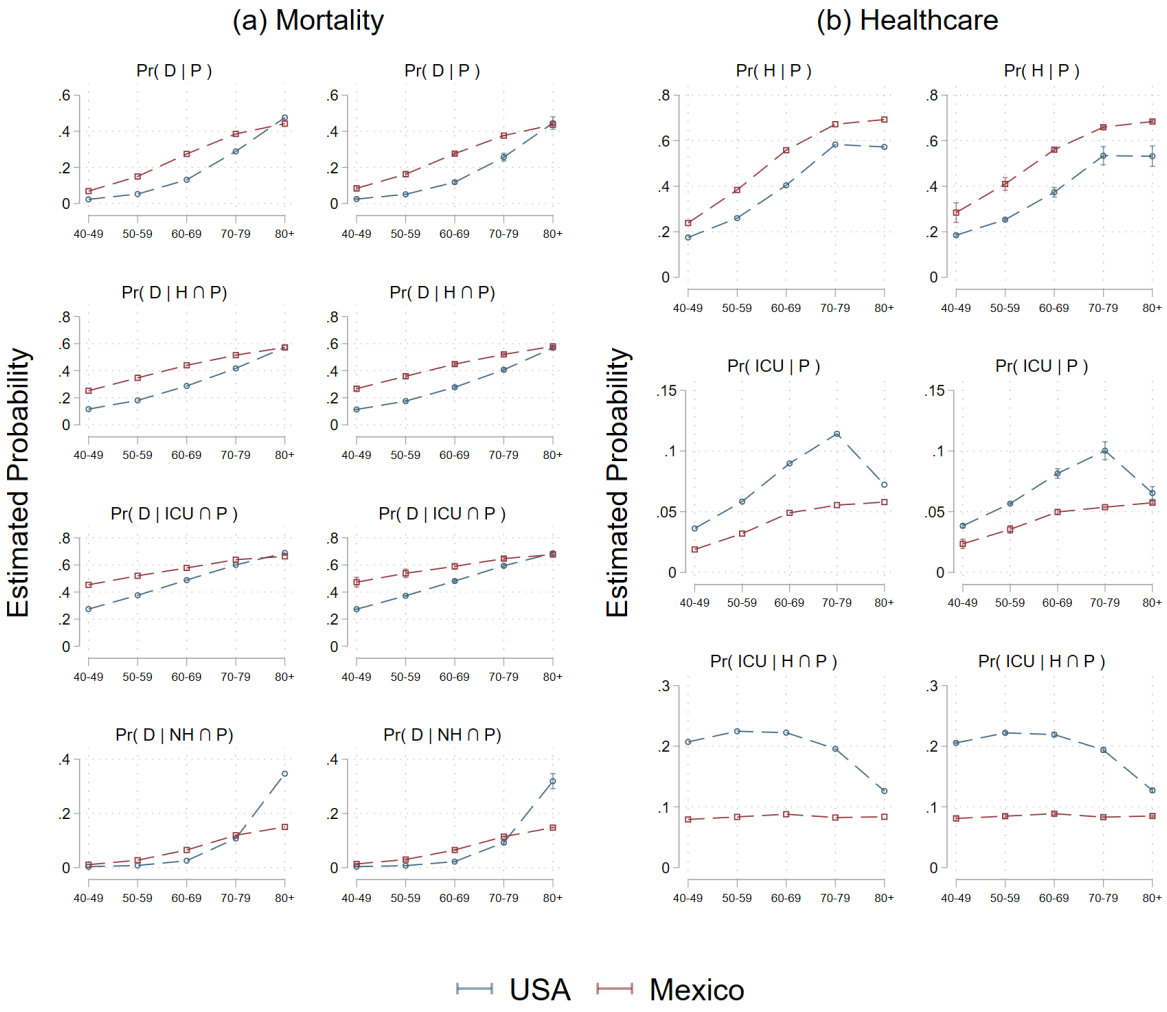
Notes: The figure shows the estimated conditional probabilities of dying by country with a 95% CI. The estimates are obtained in pairwise regressions comparing each country to the U.S., using equation 2 with indicators for each country. Regressions estimated at individual level using a Probit model. The countries included in each graph vary, depending on the data available. Robust standard errors clustered at the country level.

Figure C11: Age-specific conditional probabilities of accessing healthcare if infected with COVID-19 for eight countries



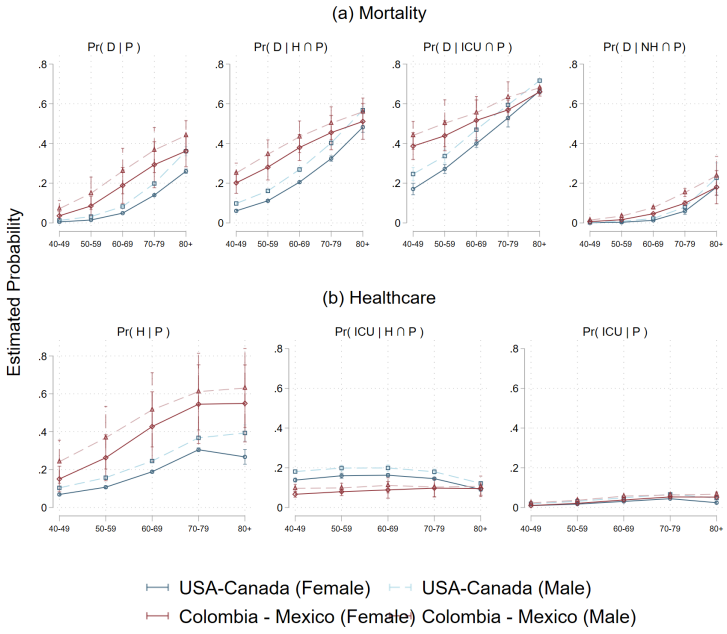
Notes: The figure shows the estimated conditional probabilities of accessing healthcare if infected with COVID-19 by country with a 95% CI. The estimates are obtained in pairwise regressions comparing each country to the U.S., using equation 2 with indicators for each country. Regressions estimated at individual level using a Probit model. The countries included in each graph vary, depending on the data available. Robust standard errors clustered at the country level.

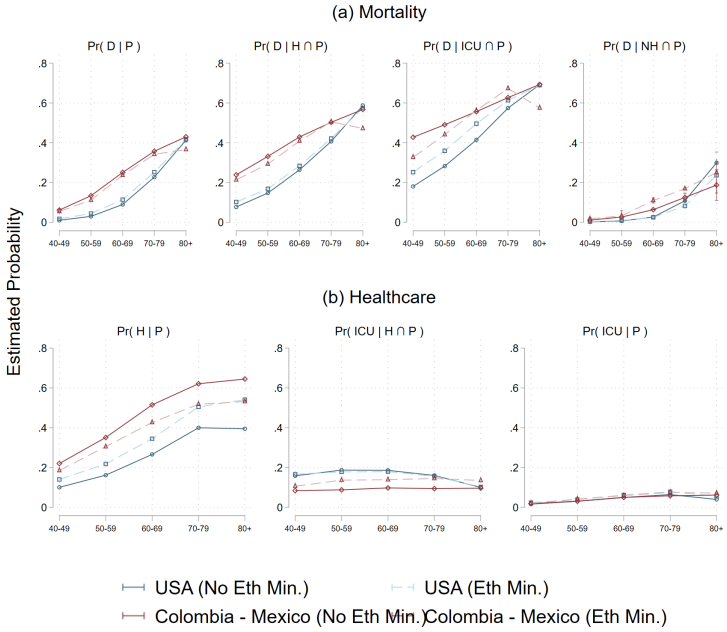
Figure C12: Differences between Mexico and USA in the share of COVID-19 deaths by age group with preexisting conditions and ethnic minority controls



Notes: The figure shows the estimated conditional probabilities of dying, being hospitalized, or entering an ICU by country income level with a 95% CI. Regressions estimated at individual level using a Probit model. Robust standard errors clustered at the country level.

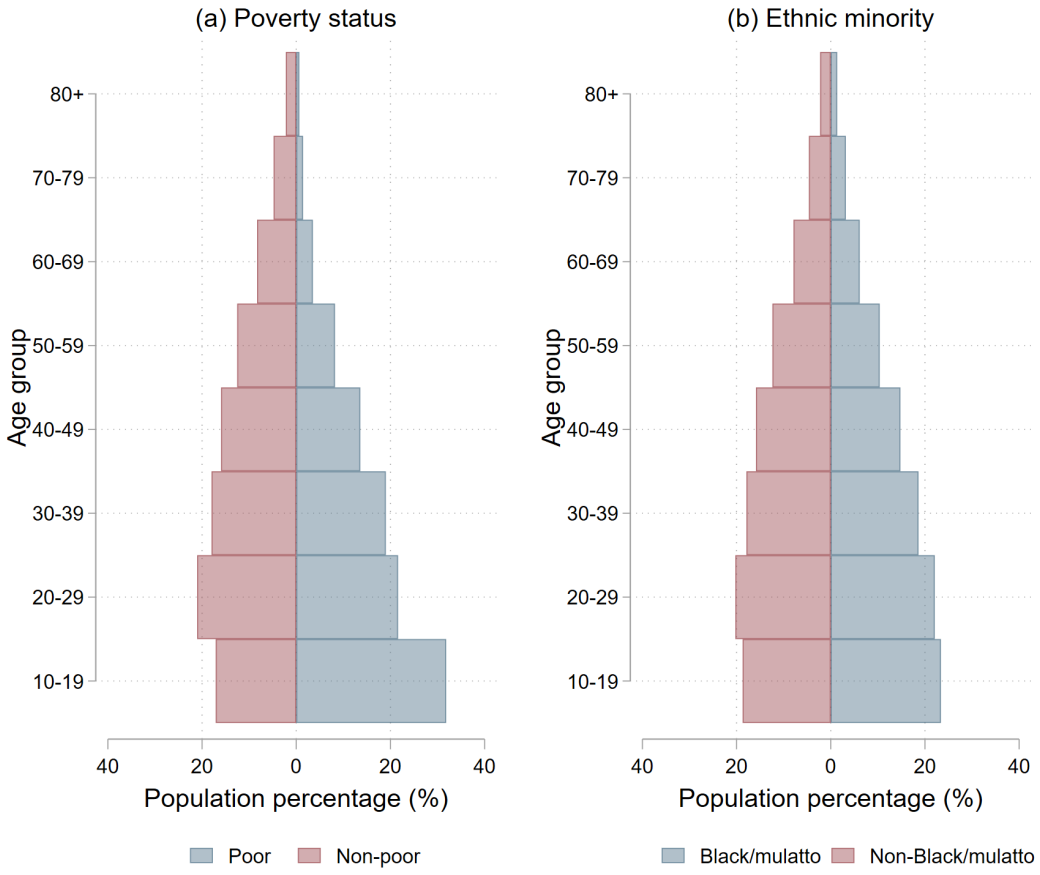
Figure C13: Heterogeneity by gender and ethnic minority status in the age-specific differences between developing and high-income countries in the probabilities of dying from COVID-19 and accessing healthcare when diagnosed positive





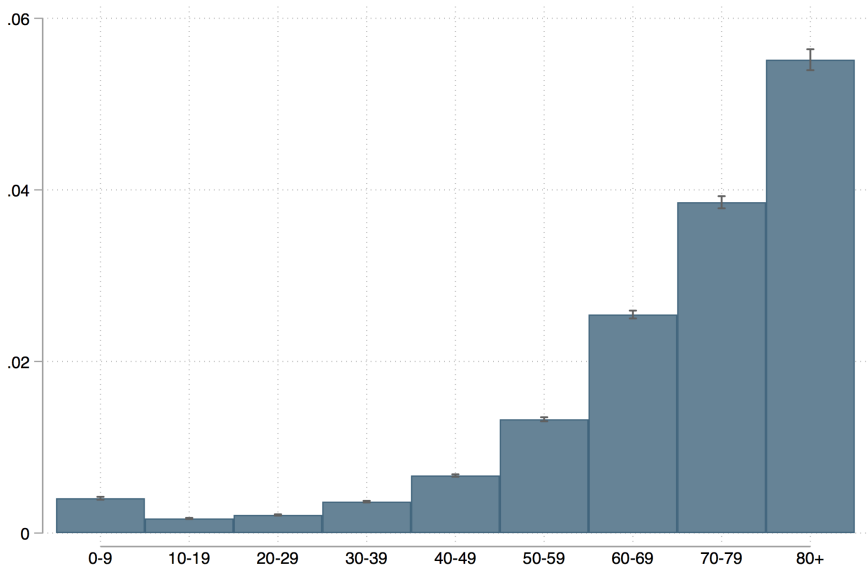
Notes: The figure shows the estimated conditional probabilities of dying, being hospitalized, or entering an ICU by country income level and demographic group, with a 95% CI. Regressions estimated at individual level using a Probit model. Robust standard errors clustered at the country level.

Figure C14: Age pyramids by poverty and by ethnic minority status in Brazil, 2010



Notes: The figures are computed directly from the microdata of the 2010 census, using the sample weights provided by the IBGE.

Figure C15: Share of the population in each age group with preexisting conditions associated with severe COVID-19 complications in Brazil, 2019



Notes: The figure depicts the average and 95% confidence interval across cities of the percentage of the population in each age group with at least one preexisting condition associated with severe COVID-19 complications. The estimates are computed directly from the DATASUS microdata.

C2 Additional tables

Table C1: Cross-country estimates of age-group shares in their country's COVID-19 deaths by country type (13 developing countries and 18 high-income countries)

	Developing	High-income
Ages 0-19	0.006***	0.000
	(0.002)	(0.000)
Ages 20-39	0.060***	0.006***
	(0.009)	(0.001)
Ages 40-59	0.281***	0.053***
	(0.028)	(0.006)
Ages 60+	0.653***	0.942***
	(0.036)	(0.007)

Notes: The table reports estimates of the average shares of each age group in their country's COVID-19 deaths by country type (developing and high-income) from a cross-country OLS regression. All estimates are calculated without controls, and their standard errors are reported in parentheses. The regressions includes a constant. See Appendix A for further details. *** p<0.01, ** p<0.05, * p<0.1.

Table C2: Differences between developing and high-income countries in the share of COVID-19 deaths by age group and gender

	Dependent variable: Share of age-gender group in national COVID-19 deaths						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Developing country indicator	-0.26***	-0.23***	-0.23***	-0.23***	-0.23***	-0.15**	-0.03
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)
Male indicator	-0.03***	-0.02**	-0.02**	-0.02**	-0.02**	0.00	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Developing country × male	-0.00	-0.01	-0.01	-0.01	-0.01	0.02	-0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Age groups indicators (reference group is age 60+)							
Ages 0-19	-0.95***	-0.94***	-0.94***	-0.94***	-0.94***	-0.42***	-0.07
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.11)	(0.10)
Ages 20-39	-0.95***	-0.94***	-0.94***	-0.94***	-0.94***	-0.41***	-0.13
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.11)	(0.10)
Ages 40-59	-0.91***	-0.91***	-0.91***	-0.91***	-0.91***	-0.40***	-0.25***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.11)	(0.08)
Age groups × Developing country							
Ages 0-19	0.27***	0.22***	0.22***	0.22***	0.22***	0.13**	0.03
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)
Ages 20-39	0.32***	0.28***	0.28***	0.28***	0.28***	0.19**	-0.01
	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.06)
Ages 40-59	0.46***	0.45***	0.45***	0.45***	0.45***	0.28**	0.08
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)	(0.10)
Age groups × male							
Ages 0-19	0.03***	0.02**	0.02**	0.02**	0.02**	0.00	-0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ages 20-39	0.03***	0.02**	0.02**	0.02**	0.02**	-0.00	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ages 40-59	0.05***	0.04***	0.04***	0.04***	0.04***	-0.01	-0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Age groups × Developing country × male							
Ages 0-19	0.00	0.01	0.01	0.01	0.01	-0.03	-0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Ages 20-39	0.00	0.01	0.01	0.01	0.01	-0.05	-0.02
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Ages 40-59	0.01	0.02	0.02	0.02	0.02	-0.01	0.03
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)
Observations	168	168	168	168	168	168	168
R^2	0.97	0.97	0.97	0.97	0.97	0.98	0.98
Country characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shares of age group in total population	No	Yes	Yes	Yes	Yes	Yes	Yes
Testing controls	No	No	Yes	Yes	Yes	Yes	Yes
Policy controls	No	No	No	Yes	Yes	Yes	Yes
Labor force participation of males and females	No	No	No	No	Yes	Yes	Yes
Share in all-cause deaths	No	No	No	No	No	Yes	Yes
Share in high-risk of severe complications	No	No	No	No	No	No	Yes

Notes: Regressions at the country - age group - gender level. Relative to the main regression (Table 1), the following countries are not included due to lack data on age-specific mortality *by sex* at the time of update included in our data set: Afghanistan, Ecuador, France, Italy, Japan, Romania, South Africa, Spain and Sweden. Robust standard errors clustered at the country level in parentheses. All regressions include a constant. *** p<0.01, ** p<0.05, * p<0.1.

Table C3: Differences between developing and high-income countries in the share of COVID-19 deaths by age group with alternative sample definitions

	Dependent variable: Share of age group in national COVID-19 deaths				
	Full sample (1)	Full sample (2)	No data flags (3)	Most pop in cities < 1m (4)	Most pop in cities < 500k (5)
Developing country indicator	-0.12**	-0.12**	-0.14***	-0.13**	-0.15***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Age groups indicators (*reference group is age 60+*)					
Ages 0-19	-0.33**	-0.33**	-0.31**	-0.32**	-0.40***
	(0.12)	(0.12)	(0.13)	(0.13)	(0.13)
Ages 20-39	-0.36***	-0.36***	-0.33***	-0.34***	-0.40***
	(0.11)	(0.11)	(0.11)	(0.12)	(0.13)
Ages 40-59	-0.41***	-0.41***	-0.37***	-0.39***	-0.42***
	(0.08)	(0.08)	(0.09)	(0.09)	(0.12)
Age groups × Developing country					
Ages 0-19	0.12**	0.12**	0.16***	0.14**	0.17***
	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
Ages 20-39	0.12*	0.12*	0.15**	0.13*	0.17**
	(0.07)	(0.07)	(0.06)	(0.07)	(0.06)
Ages 40-59	0.25***	0.25***	0.27***	0.25***	0.28**
	(0.08)	(0.08)	(0.09)	(0.09)	(0.10)
Observations	116	116	92	100	80
R^2	0.99	0.99	0.98	0.98	0.99
Date of data update	No	Yes	No	No	No
Country characteristics	Yes	Yes	Yes	Yes	Yes
Shares of age group in total population	Yes	Yes	Yes	Yes	Yes
Testing controls	Yes	Yes	Yes	Yes	Yes
Policy controls	Yes	Yes	Yes	Yes	Yes
Share in all-cause deaths	Yes	Yes	Yes	Yes	Yes
Share in high-risk of severe complications	Yes	Yes	Yes	Yes	Yes

Notes: Regressions at the country-age group level. Robust standard errors clustered at the country level in parentheses. All regressions include a constant. *** p<0.01, ** p<0.05, * p<0.1.

Table C4: Age-group-specific differences in mortality outcomes, Probit regression estimates

	Mortality outcomes				Healthcare outcomes		
	$\Pr(D \mid P)$	$\Pr(D \mid H \cap P)$	$\Pr(D \mid ICU \cap P)$	$\Pr(D \mid NH \cap P)$	$\Pr(H \mid P)$	$\Pr(ICU \mid H \cap P)$	$\Pr(ICU \mid P)$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Developing country indicator	0.292***	0.0425	-0.0652***	0.0469	0.717**	-0.0168	0.288***
	(0.107)	(0.113)	(0.0153)	(0.185)	(0.288)	(0.125)	(0.0511)
Age group effects							
Ages 40-49	-1.830***	-1.436***	-1.271***	-2.010***	-0.875***	0.278***	-0.350***
	(0.0514)	(0.00418)	(0.0429)	(0.121)	(0.0408)	(0.0408)	(0.0448)
Ages 50-59	-1.452***	-1.126***	-0.986***	-1.683***	-0.625***	0.353***	-0.129**
	(0.0371)	(0.0114)	(0.0314)	(0.0901)	(0.0460)	(0.0494)	(0.0547)
Ages 60-69	-0.966***	-0.749***	-0.642***	-1.244***	-0.288***	0.357***	0.0958
	(0.0242)	(0.0158)	(0.0284)	(0.0602)	(0.0507)	(0.0476)	(0.0588)
Ages 70-79	-0.422***	-0.390***	-0.329***	-0.623***	0.0684	0.281***	0.250***
	(0.00923)	(0.0116)	(0.0343)	(0.0190)	(0.0474)	(0.0436)	(0.0568)
Interactions with developing country indicator							
Ages 40-49	0.475***	0.617***	0.649***	0.514**	-0.196*	-0.370***	-0.210**
	(0.0656)	(0.0229)	(0.112)	(0.217)	(0.119)	(0.0774)	(0.0820)
Ages 50-59	0.524***	0.570***	0.502***	0.547***	-0.0675	-0.410***	-0.204**
	(0.0777)	(0.0140)	(0.149)	(0.207)	(0.0847)	(0.0741)	(0.0846)
Ages 60-69	0.468***	0.434***	0.306**	0.523***	0.00193	-0.352***	-0.196***
	(0.0762)	(0.0185)	(0.131)	(0.189)	(0.0606)	(0.0490)	(0.0626)
Ages 70-79	0.240***	0.254***	0.162	0.300**	-0.0872*	-0.286***	-0.260***
	(0.0577)	(0.0117)	(0.108)	(0.122)	(0.0481)	(0.0452)	(0.0580)
Observations	1,368,174	298,495	41,558	1,070,054	1,368,174	298,495	1,368,174
$Pseudo\ R^2$	0.175	0.0795	0.0597	0.229	0.0785	0.0161	0.0232

Notes: Regressions at the patient level. Sample is constant across regressions, and includes observations from Canada, Colombia, Mexico, and USA. All regressions include a constant. Robust standard errors clustered at the country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table C5: Descriptive statistics for countries with patient-level information

	USA (1)	CAN (2)	KOR (3)	CZE (4)	MEX (5)	COL (6)	PHL (7)	BRA (8)
Panel A: Country-level data								
Testing								
Testing per 100k	142.7	95.82	28.68	59.62	5.88	24.30	10.14	11.93
Positivity rate	6.40%	1.90%	0.30%	1.40%	43.20%	12.50%	5.30%	47.80%
Age-group shares in individuals at high risk due to pre-existing conditions								
Children and teens	0.4%	0.4%	0.3%	0.2%	0.8%	0.9%	1.1%	0.8%
Young adults	9.2%	9.6%	9.9%	7.4%	15.8%	16.9%	19.8%	16.7%
Middle-age adults	27.2%	27.4%	34.4%	26.4%	39.3%	35.7%	40.0%	37.1%
Elderly	63.1%	62.6%	55.3%	66.0%	44.1%	46.4%	39.1%	45.3%
Share of individuals at high-risk in population (all age groups)	5.8%	5.2%	5.3%	7.5%	4.6%	3.8%	3.9%	4.0%
All cause mortality								
Children and teens	0.1%	0.0%	0.0%	0.0%	0.1%	0.1%	0.2%	0.0%
Young adults	0.2%	0.1%	0.1%	0.1%	0.2%	0.2%	0.2%	0.2%
Middle-age adults	0.7%	0.3%	0.2%	0.3%	0.5%	0.4%	0.7%	0.5%
Elderly	4.1%	2.9%	2.4%	3.8%	3.6%	3.0%	3.7%	3.2%
Overall all-cause mortality rate (all age groups)	1.2%	0.8%	0.7%	1.0%	1.1%	0.9%	1.2%	1.0%
Panel B: Patient-level data								
Pre-existing conditions and minorities								
Share of patients with pre-existing conditions	70.97%	N/A	N/A	N/A	45.57%	N/A	N/A	N/A
Share of patients that are ethnic minorities	65.19%	N/A	N/A	N/A	1.07%	9.51%	N/A	N/A
Share of non-minorities with pre-existing conditions	74.43%	N/A	N/A	N/A	45.60%	N/A	N/A	N/A
Share of minorities with pre-existing conditions	70.01%	N/A	N/A	N/A	51.51%	N/A	N/A	N/A
Avg. days between symptoms and death	N/A		17.50	N/A	11.83	12.87	10.59	13.36

Table C6: Patient-level estimates of conditional probabilities of dying from COVID-19 by age group and country type

	Developing	High-income
Panel A: Conditional on testing positive (3 developing, 4 high-income countries)		
Ages 40-49	0.0511***	0.00896***
	(0.0156)	(0.000742)
Ages 50-59	0.113***	0.0234***
	(0.0335)	(0.000941)
Ages 60-69	0.216***	0.0666***
	(0.0523)	(0.000628)
Ages 70-79	0.319***	0.169***
	(0.0589)	(0.00241)
Ages 80+	0.390***	0.294***
	(0.0426)	(0.00604)
Panel B: Conditional on testing positive and being hospitalized (4 developing, 2 high-income countries)		
Ages 40-49	0.204***	0.0830***
	(0.0287)	(0.00375)
Ages 50-59	0.292***	0.141***
	(0.0339)	(0.00396)
Ages 60-69	0.402***	0.243***
	(0.0229)	(0.00423)
Ages 70-79	0.491***	0.367***
	(0.0192)	(0.00658)
Ages 80+	0.586***	0.520***
	(0.0245)	(0.0113)
Panel C: Conditional on testing positive and not being hospitalized (3 developing, 2 high-income countries)		
Ages 40-49	0.0101***	0.00205***
	(0.000701)	(0.000174)
Ages 50-59	0.0245***	0.00548***
	(0.00218)	(5.07e-05)
Ages 60-69	0.0600***	0.0177***
	(0.00385)	(0.00140)
Ages 70-79	0.121***	0.0690***
	(0.0103)	(0.00958)
Ages 80+	0.195***	0.195***
	(0.0428)	(0.0249)
Panel D: Conditional on testing positive and being put in an ICU (2 developing, 2 high-income countries)		
Ages 40-49	0.426***	0.221***
	(0.0356)	(0.0164)
Ages 50-59	0.482***	0.314***
	(0.0536)	(0.0155)
Ages 60-69	0.542***	0.444***
	(0.0462)	(0.0161)
Ages 70-79	0.608***	0.568***
	(0.0340)	(0.0184)
Ages 80+	0.672***	0.692***
	(0.00579)	(0.00438)

Notes: The table reports estimated conditional probabilities of dying from COVID-19 by age group and country type from patient-level Probit regressions. All estimates are calculated without controls, and their standard errors are reported in parentheses. All estimating regressions include a constant. See Appendix A for further details. *** p<0.01, ** p<0.05, * p<0.1.

Table C7: Patient-level estimates of conditional probabilities of being hospitalized or entering an ICU if diagnosed with COVID-19 by age group and country type

	Developing	High-income
Panel A: Probability of being hospitalized conditional on testing positive (3 developing, 2 high-income countries)		
Ages 40-49	0.192***	0.0860***
	(0.0444)	(0.00308)
Ages 50-59	0.306***	0.132***
	(0.0728)	(0.00311)
Ages 60-69	0.456***	0.218***
	(0.0966)	(0.00295)
Ages 70-79	0.559***	0.337***
	(0.105)	(0.00482)
Ages 80+	0.576***	0.312***
	(0.102)	(0.0209)
Panel B. Probability of being put in an ICU conditional on testing positive (2 developing, 2 high-income countries)		
Ages 40-49	0.0172***	0.0141***
	(0.00227)	(0.000143)
Ages 50-59	0.0294***	0.0242***
	(0.00324)	(0.000336)
Ages 60-69	0.0488***	0.0401***
	(0.000540)	(0.000867)
Ages 70-79	0.0588***	0.0553***
	(0.00343)	(0.000896)
Ages 80+	0.0603***	0.0325***
	(0.00235)	(0.00355)
Panel C: Probability of entering an ICU conditional on testing positive and being hospitalized (3 developing, 2 high-income countries)		
Ages 40-49	0.171**	0.164***
	(0.0714)	(0.00424)
Ages 50-59	0.189**	0.183***
	(0.0857)	(0.00676)
Ages 60-69	0.222**	0.184***
	(0.0987)	(0.00634)
Ages 70-79	0.252**	0.164***
	(0.107)	(0.00493)
Ages 80+	0.289***	0.104***
	(0.0925)	(0.00403)

Notes: The table reports estimated conditional probabilities of being hospitalized or entering an ICU if diagnosed with COVID-19 by age group and country type from patient-level Probit regressions. All estimates are calculated without controls, and their standard errors are reported in parentheses. All estimating regressions include a constant. See Appendix A for further details. *** p<0.01, ** p<0.05, * p<0.1.

Table C8: Correlates of the age-share difference in COVID-19 deaths across Brazilian cities

	Dependent variable: Gap in the elderly and non-elderly shares in COVID-19 deaths							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	-0.07***	-0.07***	-0.07***	-0.08**	-0.09**	-0.06	-0.06	-0.22***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)
Population density	-0.02	-0.02	-0.02	-0.01	-0.01	-0.04	-0.04	-0.09**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Urban share	-0.02	-0.02	-0.02	0.01	0.01	-0.01	-0.00	0.01
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Maximum yearly temperature			-0.01	-0.01	-0.01	-0.01	-0.01	-0.04
			(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.04)
Average yearly precipitation			-0.04	-0.05	-0.05	-0.02	-0.02	-0.09*
			(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Distance to Sao Paulo			-0.03	-0.04	-0.03	-0.08	-0.08	-0.03
			(0.04)	(0.04)	(0.05)	(0.05)	(0.06)	(0.04)
Informality rate				-0.04*	-0.04	-0.04*	-0.04*	-0.05**
				(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Share of high-school in employment				-0.03	-0.03	-0.05**	-0.05*	-0.07***
				(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Share of college in employment				0.01	-0.00	-0.02	-0.02	0.01
				(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Share of jobs that can be done from home				0.00	-0.00	0.01	0.00	-0.02
				(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Share of workers in high-contact occupations				-0.01	-0.01	0.02	0.02	0.02
				(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Average commute				0.00	0.01	0.00	-0.00	-0.02
				(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Gini (hourly wages)					0.01	0.01	0.01	0.01
					(0.02)	(0.02)	(0.02)	(0.02)
Extreme poverty rate					0.05	0.06	0.05	0.07
					(0.05)	(0.05)	(0.05)	(0.05)
Average income per capita					0.06	-0.02	-0.03	-0.08
					(0.07)	(0.08)	(0.08)	(0.09)
Share of black and mixed in population					-0.02	-0.00	-0.01	-0.04
					(0.04)	(0.03)	(0.03)	(0.03)
Access to piped water						0.03**	0.03**	0.05***
						(0.01)	(0.01)	(0.02)
Broadband density						-0.03	-0.03	-0.02
						(0.02)	(0.02)	(0.02)
Average persons per room						-0.16***	-0.15***	-0.18***
						(0.04)	(0.04)	(0.05)
Share of young adults living with elderly						0.01	0.01	0.01
						(0.03)	(0.03)	(0.03)
Share of elderly living alone						0.00	0.00	0.00
						(0.02)	(0.02)	(0.02)
Share of population in favelas						0.03***	0.02***	-0.04*
						(0.01)	(0.01)	(0.02)
Number of doctors							0.01	0.00
							(0.03)	(0.02)
ICU beds							-0.03	-0.03*
							(0.02)	(0.02)
Ventilators							0.00	0.00
							(0.03)	(0.02)
Patients with at least one pre-condition							0.04**	-0.02
							(0.01)	(0.02)
Share of elderly in population								-0.03
								(0.05)
Elderly population COVID-19 deaths								0.41***
								(0.06)
Observations	1,412	1,412	1,389	1,389	1,389	1,386	1,386	1,385
R^2	0.06	0.06	0.06	0.06	0.06	0.08	0.08	0.20

Notes: Regressions at the city level, restricted to cities with population of at least 5,000. The dependent variable is calculated subtracting the share of non-elderly individuals (younger than 60) from the share of elderly individuals (aged 60 or older) total COVID-19 dates confirmed until June 16, 2020. Robust standard errors clustered at the state level in parentheses. All regressions include state fixed effects and a constant. *** p<0.01, ** p<0.05, * p<0.1