

# COVID-19 severity: A new approach to quantifying global cases and deaths

Daniel L. Millimet<sup>1</sup>  | Christopher F. Parmeter<sup>2</sup> 

<sup>1</sup>Department of Economics, Southern Methodist University, Dallas, Texas, USA

<sup>2</sup>Department of Economics, Miami Herbert Business School, University of Miami, Miami, Florida, USA

## Correspondence

Christopher F. Parmeter, Department of Economics, Miami Herbert Business School, University of Miami, Miami, Florida, USA.

Email: [cparmeter@bus.miami.edu](mailto:cparmeter@bus.miami.edu)

## Abstract

As the COVID-19 pandemic has progressed, so too has the recognition that cases and deaths have been underreported, perhaps vastly so. Here, we present an econometric strategy to estimate the true number of COVID-19 cases and deaths for 61 and 56 countries, respectively, from 1 January 2020 to 3 November 2020. Specifically, we estimate a ‘structural’ model based on the SIR epidemiological model extended to incorporate underreporting. The results indicate significant underreporting by magnitudes that align with existing research and conjectures by public health experts. While our approach requires some strong assumptions, these assumptions are very different from the equally strong assumptions required by other approaches addressing underreporting in the assessment of the extent of the pandemic. Thus, we view our approach as a complement to existing methods.

## KEY WORDS

COVID-19, nonclassical measurement error, stochastic frontier analysis

*If we stop testing right now, we'd have very few cases, if any.*

—President Donald J. Trump on COVID-19 (20 June 2020)

*Political meddling, disorganization and years of neglect of public-health data management mean the country is flying blind.*

—Nature on COVID-19 (25 August 2020)

*Many of the infected do not get sick. Many who do are never seen by any health system.*  
—The Economist on COVID-19 (26 September 2020)

## 1 | INTRODUCTION

Efforts by government, media and public health officials to track the severity and public health impact of COVID-19 in the United States and around the globe have encountered numerous difficulties. As a result, the *official* number of COVID-19 cases and deaths likely undercounts the *true* number of cases and deaths, perhaps significantly so. Manski and Molinari (2021, p. 181) state: ‘It is well appreciated that accurate characterization of the time path of the coronavirus pandemic has been hampered by a serious problem of missing data’. To estimate the magnitude of this missing data problem, we use data from roughly 60 countries over the period from 1 January 2020 to 3 November 2020, a ‘structural’ model based on the Susceptible-Infected-Removed (SIR) epidemiological model extended to incorporate missing data, and the econometric methodology explored in Millimet and Parmeter (2021) to account for one-sided (nonclassical) measurement error.

In addition to yielding insights into the determinants of cases and deaths, our approach also generates estimates for the true number of cases and deaths due to COVID-19 across the countries in our sample. We estimate the true number of cases to be two to nine times higher than official reports. In terms of deaths, we estimate the true number to be 1.6 to 3.0 times higher than official reports. Over our entire sample, the case fatality rate—observed deaths divided by observed cases—is 2.6%. The infectious fatality rate—estimated deaths divided by estimated cases—ranges from 0.9% to 2.1%. The fact that our estimates of the infectious fatality rate are below the case fatality rate results from our estimate of more missing cases than deaths.

Accurate information on the count of cases and deaths is critical for several reasons. First, underreporting of cases and deaths inhibits scientific efforts to understand the SARS-CoV-2 virus that causes COVID-19. Correct estimation of rates of transmission, factors affecting transmission, fatality rates, etc. are made all the more difficult without knowledge of the true number of cases and fatalities. As noted in Korolev (2021, p. 83), ‘... estimates of  $R_0$  based on the confirmed cases data under the assumption that all cases are reported may be biased downward’, where  $R_0$  is the reproduction number, or the expected number of secondary cases spawning from each infected individual. Second, inaccurate information, particularly underreporting, may inhibit public and political support for warranted responses such as expansions of hospitals, investments in medical equipment, mask mandates and lockdowns (Depalo, 2021; Hortaçsu et al., 2021). Knowledge of undocumented cases is also necessary to assess the risk of infection by the healthy population. As a result, Manski and Molinari (2021), Stock (2020) and others have called for random testing to determine the size of the population of asymptomatic individuals.<sup>1</sup> Third, underreporting of cases can be fatal. Missing cases precludes complete contact tracing, which in turn leads to greater spread and death (Fetzer & Graeber, 2020). Finally, undercounting may lead to sub-optimal government fiscal responses, such as business loans and expanded unemployment insurance, and decisions regarding school closures.

Owing to the importance of accurate information, official figures on cases and deaths have come under increasing scrutiny in the United States (US) and abroad. As recently as June 2020 at

<sup>1</sup>This is consistent with the approach in Iceland, where the government had randomly tested the equivalent of 26,762 per million inhabitants by 21 March 2020, a number much higher than anywhere else in the world at the time. See <https://www.government.is/news/article/newsid=f96a270c-66e8-11ea-945f-005056bc4d74>.

least 28 states were not following US Centers for Disease Control and Prevention (CDC) reporting guidelines. While CDC guidelines are voluntary, 'states not reporting probable cases likely undercount the number of people infected and make it difficult for officials to get the true picture of where the nation stands in the midst of a pandemic that has rocked almost every aspect of life'.<sup>2</sup> Antibody tests performed in the US in late Spring 2020 provide evidence of such underreporting, suggesting that 'a good rough estimate now' of the number of missed cases 'is 10 to 1'.<sup>3</sup> However, a mere few days later the CDC reported that the true number of cases 'could be up to 24 times higher than reported'.<sup>4</sup> Anecdotal evidence of underreporting is also rampant. In Houston some hospitals stopped updating their data once they reached full capacity, 'rattling policymakers and residents'.<sup>5</sup> Government officials in Florida came under fire in May 2020 when the data scientist responsible for publishing the state's data on COVID-19 cases and deaths was terminated.<sup>6</sup> And, US deaths officially attributed to Alzheimer's disease and dementia show 'how the coronavirus pandemic has exacted a higher fatality toll than official numbers have shown'.<sup>7</sup>

Outside the US, data concerns are no less common. Many 'experts ... fear substantial under-reporting' globally.<sup>8</sup> In China, antibody tests performed in Wuhan suggest that the true number of infected was 'almost 10 times the official number of confirmed cases'.<sup>9</sup> Russia, initially claiming its low death toll was a 'miracle', subsequently revised its official number of deaths in April 2020 upward by a factor of two.<sup>10</sup> However, some speculate that 'about 70% of coronavirus-related deaths have not been reported in Moscow and about 80 percent in the country's regions'.<sup>11</sup> While the official number of cases in Brazil recently topped six million, the reported cases are viewed as a vast undercount 'because there is underreporting of a magnitude of five to 10 times'.<sup>12</sup> On 18 July 2020 the president of Iran suggested that 25 million residents have been infected, 'nearly a third of the population and massively higher than the official number of COVID-19 cases'.<sup>13</sup> A study in Italy documented evidence that residents in Italy were infected as far back as September 2019 even though the first official case was not recorded until February 2020.<sup>14</sup>

In total, the World Health Organization (WHO) confirmed in May 2021 that the 'at least 6-8 million people may have died due to the COVID-19 pandemic so far' despite an official WHO death toll of roughly 3.4 million.<sup>15</sup> Even the range of 6-8 million is described as 'an estimate on a cautionary note'.<sup>16</sup>

These measurement issues are perhaps not surprising as undercounting COVID-19 cases and deaths can arise due to several reasons. First, a shortage of information, a lack of tests and testing inaccuracies during the early stages of the pandemic may have led to many cases and/or

<sup>2</sup>See <https://www.cnn.com/2020/06/09/health/us-coronavirus-tuesday/index.html>.

<sup>3</sup>See <https://www.cnn.com/2020/06/25/health/us-coronavirus-thursday/index.html>.

<sup>4</sup>See <https://www.cnn.com/2020/06/29/health/us-coronavirus-monday/index.html>.

<sup>5</sup>See <https://www.houstonchronicle.com/news/houston-texas/houston/article/Houston-hospitals-hit-100-base-ICU-capacity-15372256.php>.

<sup>6</sup>See <https://www.cnn.com/2020/05/20/us/florida-georgia-covid-19-test-data/index.html>.

<sup>7</sup>See <https://www.wsj.com/articles/coronavirus-pandemic-led-to-surge-in-alzheimers-deaths-11593345601>.

<sup>8</sup>See <https://www.reuters.com/article/us-health-coronavirus-global-cases/global-coronavirus-cases-surpass-3-5-million-amid-underreporting-fears-idUSKBN22G00Z>.

<sup>9</sup><https://www.cnn.com/2020/12/29/asia/china-coronavirus-seroprevalence-study-intl-hnk/index.html>.

<sup>10</sup>See <https://www.nytimes.com/2020/05/29/world/coronavirus-update.html>.

<sup>11</sup>See <https://www.nytimes.com/2020/05/11/world/europe/coronavirus-deaths-moscow.html>.

<sup>12</sup>See <https://www.cnbc.com/2020/06/20/brazil-passes-1-million-coronavirus-cases-with-no-end-in-sight.html>.

<sup>13</sup>See <https://www.reuters.com/article/us-health-coronavirus-iran-idUSKCN24K0E3>.

<sup>14</sup>See <https://www.reuters.com/article/us-health-coronavirus-italy-timing/coronavirus-emerged-in-italy-earlier-than-thought-italian-study-shows-idUSKBN27V0KF>.

<sup>15</sup><https://www.reuters.com/world/india/total-death-toll-covid-19-could-be-least-6-8-million-who-2021-05-21/>.

<sup>16</sup><https://www.reuters.com/world/india/total-death-toll-covid-19-could-be-least-6-8-million-who-2021-05-21/>.

deaths being misdiagnosed or missed altogether (Manski & Molinari, 2021; Reese et al., 2021). Evidence suggests that ‘underreporting is a problem faced by health authorities in many countries, often due to a lack of capacity and resources’.<sup>17</sup> Second, a significant number of the infected are asymptomatic, thereby evading official counts in the absence of universal or randomized testing protocols (e.g. Reese et al., 2021). Third, there is concern in many countries, including the US, that political corruption and institutions play a role in underreporting (Badman et al., 2021). Greer et al. (2020, p. 1413) state that ‘there will be no way to understand the different responses to COVID-19 and their effects without understanding policy and politics’.

Fourth, cases and deaths are not counted if they occur outside public health surveillance systems. Individuals may remain outside such systems due to lack of access, fear or personal choice (e.g., McCulloh et al., 2020, p. 2). For example, this is suspected of leading to significant underreporting in India, where ‘deaths … have always been counted poorly … [as] the vast majority of deaths, especially in rural villages, take place at home and routinely go unregistered … leading experts to estimate that only between 20% and 30% of all deaths in India are properly medically certified’.<sup>18</sup> Fifth, even when individuals are symptomatic and captured by public health surveillance systems, these are local in nature. Data must then be passed upstream to be added to official tallies. As discussed in Dubrow (2021, p. 1), ‘countries with different data infrastructures have different capacities to collect information and turn them into data’ making inconsistencies in upstream reporting an ‘overlooked source’ of errors in administrative data on COVID-19.

Finally, reporting of deaths is inconsistent over time and across countries (Flaxman et al., 2020). In the US, for example, doctors and coroners have broad discretion in the determination of cause of death which can lead to inconsistencies.<sup>19</sup> That such discretion has led to underreporting is borne out by a few instances of *ad hoc* corrections that have occurred. For instance, in early April, New York City experienced a one day increase in COVID-19 deaths of over 3,700 as they retroactively counted past deaths that were likely from COVID-19.<sup>20</sup> New Jersey reported a 1 day jump of over 1,900 deaths due to COVID-19 towards the end of June, again due to retroactive changes in how deaths were classified.

To overcome this missing data problem, researchers have pursued several approaches to date which we discuss in detail in the next section. In this paper, we propose an alternative method to assess undercounting based on an extended version of the SIR epidemiological model and the econometric methodology discussed in Millimet and Parmeter (2021). Specifically, Millimet and Parmeter (2021) propose using stochastic frontier analysis (SFA) to estimate models where the outcome is believed to suffer from one-sided or skewed measurement error. Given the overwhelming evidence suggesting that both COVID-19 cases and deaths are undercounted, SFA offers a convenient framework to assess the true extent of the pandemic. See Badunenko et al. (2012) for a recent investigation of SFA.

Estimating stochastic frontier models on weekly data on cases (deaths) from 61 (56) countries under varying distributional assumptions concerning the underreporting, we reach several striking conclusions. First, we find evidence of significant underreporting, with our preferred estimates suggesting that the true case (death) count to be roughly nine (two) times higher than official reports. Second, we obtain an infectious fatality rate less than 1%, well below the observed

<sup>17</sup><https://www.cnn.com/2020/12/29/asia/china-coronavirus-seroprevalence-study-intl-hnk/index.html>.

<sup>18</sup><https://www.bloomberg.com/news/articles/2021-04-22/even-record-death-toll-may-hide-extent-of-india-s-covid-crisis>.

<sup>19</sup>See <https://www.scientificamerican.com/article/how-covid-19-deaths-are-counted1/>.

<sup>20</sup>See <https://abcnews.go.com/Health/coronavirus-updates-us-navy-battles-growing-outbreak-hospital/story?id=70134122>.

case fatality rate. Third, consistent with conjectures in the epidemiological literature, we find no evidence that the extent of underreporting is waning. Third, we document a robust positive (negative) association between testing and the number of cases (deaths). Finally, we generally find a negative association between the stringency of non-pharmaceutical interventions (NPIs) and cases and deaths. However, this association varies over time and is sensitive to assumptions concerning the nature of the underreporting.

We are fully aware that no statistical approach to estimating the true number of cases and deaths is beyond reproach. As Manski and Molinari (2021, p. 192) state, instead of relying on strong assumptions, ‘a more satisfactory approach to increase knowledge of the infection rate is to obtain better data’. In lieu of better data, SFA is a convenient tool that provides a valuable complement to existing research. While one can certainly criticize the reliance of SFA on distributional assumptions, we provide a range of estimates by employing various assumptions on the underreporting process. This range is broadly consistent with other estimates of underreporting of cases and deaths. Thus, even with the adherence to parametric distributional assumptions, our findings are both timely and representative. Finally, as detailed in the next section, alternatives to SFA either rely on alternative sets of strong assumptions or rely on weaker assumptions at the expense of point identification. Moreover, these alternatives are often country or time specific. Our analysis should be viewed as adding to the totality of scientific evidence using assumptions that are likely orthogonal to those employed by alternative methodologies.

The remainder of the paper proceeds as follows. Section 2 provides a brief literature review. Section 3 outlines the model and econometric framework and discusses the data. Section 4 presents the results. Section 5 concludes.

## 2 | LITERATURE REVIEW

While there is ample evidence of underreporting of COVID-19 cases and deaths appearing in various media outlets, the scientific research is limited. The most common approach is to focus on fatalities and quantify the amount of missing data via excess mortality. Measures of excess mortality are computed by comparing deaths in 2020 to the number of deaths over the same time period in previous years. This count of ‘excess’ deaths is then compared to official death tolls from COVID-19. Although straightforward, this approach is not without criticism. First, and most importantly, data on all-cause mortality are not available for most countries, let alone data on deaths by category (such as pneumonia and influenza). Moreover, even when such data are available, reporting standards are inconsistent across countries. The *Financial Times* has been most diligent about reporting excess mortality, but these data are only available for 21 countries.<sup>21</sup>

Second, conclusions based on excess mortality potentially confound true deaths with other changes as a result of the pandemic. For example, Barnes et al. (2020) and Oguzoglu (2020) report a reduction in traffic accidents and fatalities across Louisiana and 31 metropolitan areas in Turkey, respectively, due to implementation of various social distancing measures. Dang and Trinh (2020) documents a significant reduction of air pollution in Vietnam due to the pandemic. Reductions in elective health procedures or changes in usage of preventative medical services may also affect mortality during the pandemic but not be directly attributable to COVID-19. Third, excess mortality is not particularly useful in predicting the true number of cases absent knowledge of the true fatality rate. Finally, as Cerqua et al. (2021, p. 2) state, computation based on excess

<sup>21</sup>See <https://www.ft.com/content/6bd88b7d-3386-4543-b2e9-0d5c6fac846c>.

mortality ‘does not employ any covariates, nor indeed any model, and may be sensitive to outliers and other data issues’.

Despite these shortcomings, excess mortality serves as a useful benchmark for gauging the extent of undercounting of deaths. For instance, Weinberger et al. (2020) use excess deaths related to pneumonia and influenza to estimate the true death toll from COVID-19 in the US early in the pandemic. Their approach fits a Poisson regression over 5 years of weekly data and then predicts deaths that should have arisen in a given period. Comparing predicted and reported deaths, the authors find roughly 2.6 predicted deaths per officially reported death; this is referred to as the multiplication factor (MF) for deaths. Similarly, the National Statistical Agency of Italy used data on excess deaths through 31 March and reported a MF of 1.8 (Backhaus, 2020). More recently, the CDC reported a MF of 1.4 based on excess mortality in the US.<sup>22</sup>

Overcoming the missing data problem for the number of infections has a lengthy history in the infectious disease and epidemiological literatures; see Gibbons et al. (2014) for a recent discussion. Approaches are often cast in terms of quantifying underestimation (UE) of a particular disease and decomposing UE into underreporting (UR) and under-ascertainment (UA). UE is defined as the number of cases missed by official counts and arises from both UR and UA. UR is defined as cases that are missed despite individuals seeking treatment. This could arise either unintentionally (due to misdiagnosis) or intentionally (due to political corruption). UA is defined as cases that are missed because infected individuals remain outside surveillance systems either due to being asymptomatic or not seeking treatment. The objective is to estimate UR and UA, and hence UE, to calculate a MF for cases (i.e. the ratio between the true and reported number of cases).

The standard suite of methods and study designs which can be used to determine the extent of UR and UA include: community-based studies (CBS)<sup>23</sup>; serological surveys (which measure sero-incidence or sero-prevalence); returning traveller studies (RTS); and capture–recapture studies (CRS). Upon estimation of UE, researchers can derive country-specific MFs. Typically, researchers rely on simulation methods or national data coupled with CBS or CRS. For example, during the H1N1 outbreak, Reed et al. (2009) estimated MFs on the order of 79, implying a median estimate of nearly three million H1N1 infections even though at the time of their study there were only 43,677 laboratory confirmed cases. Additionally, there were 302 deaths from H1N1 reported at the time of the study by Reed et al. (2009), but their MFs for UR of hospitalizations, coupled with a 6% death rate, suggest that there were over 800 deaths from H1N1. Reese et al. (2021) follow the same strategy as in Reed et al. (2009) to estimate the number of COVID-19 cases in the US through the end of September. The authors estimate roughly 52.9 million cases. Given a reported count of 6.9 million cases, this yields an MF of 7.7.

The studies most closely related to ours include Li et al. (2020), Flaxman et al. (2020), Wu et al. (2020), McCulloh et al. (2020), and Hortaçsu et al. (2021). Li et al. (2020) study undocumented COVID-19 infections in China for the weeks bracketing the closure of Wuhan. They use an iterated filter-ensemble adjustment Kalman filter framework to estimate the trajectories of four key variables: the susceptible population, the exposed population and both documented and undocumented infections. Their results suggest that 86% of infections went undocumented (a MF over

<sup>22</sup>See <https://www.ft.com/content/752384ef-fc13-3e46-a76d-c4ef5dc5444b>.

<sup>23</sup>Gibbons et al. (2014, p. 147) state: ‘CBS can take many forms but generally involve active searching within the community for disease episodes, pathogen carriage or infection, with questionnaire-based data acquisition often accompanied by biological sampling. Active searching can be conducted face-to-face, by telephone, internet or post, with several possible study designs e.g. based on probability samples, prospective or retrospective cohorts, population cross-sections, involving representative samples of the whole population or certain interest or high-risk groups only’.

7). Flaxman et al. (2020) chart the course of the COVID-19 epidemic by back-calculating infections from observed deaths across 11 European countries. They introduce a Bayesian mechanistic model to connect the infection cycle to observed deaths. The model is then used to infer the total population infected and the reproduction number,  $R_0$ , over time in each country. For the countries considered, the authors estimate 12 to 15 million COVID-19 cases whereas the official count was roughly one million (a MF of at least 12). Consistent with our approach, Flaxman et al. (2020, p. 258) conclude that ‘... there are orders-of-magnitude fewer infections detected ... than true infections, most likely due to mild and asymptomatic infections as well as limited testing capacity and changes in testing policy’. Wu et al. (2020) use a semi-Bayesian probabilistic bias analysis on US state-level data from early in the pandemic (through 18 April 2020). The analysis assumes prior distributions for parameters capturing incomplete testing and inaccurate test results. Allowing for uncertainty in these prior distributions, the authors find a MF ranging from three to 20. McCulloh et al. (2020) use Monte Carlo methods to infer true cases from data on observed deaths, clinical estimates of the infection fatality ratios, and estimates of time-to-death distributions. Results for New York State indicate that fewer than one-third of cases were reported through April 2020.

Finally, Hortaçsu et al. (2021) devise a creative strategy to estimate a MF for cases in the US based on data on travel patterns. As with our approach, their strategy entails a number of strong assumptions. Moreover, their approach is perhaps limited in its generalizability to other countries and later periods of the pandemic due to data demands and the nature of the identifying assumptions. Specifically, their approach assumes that the pandemic originates in an epicentre and spreads to other regions via travellers. Given data on travel patterns, combined with assumptions on selection into travel, the transmission rate, and the infection rate in the epicentre, the authors estimate the expected number of infections outside the epicentre. The ratio of this expectation and the official number of cases yields the MF. For the US very early in the pandemic (through 16 March 2020), they obtain a MF between six and 24.

Two other studies merit discussion. Depalo (2021) and Manski and Molinari (2021) diverge from the preceding studies in that they consider what can be learned about COVID-19 given missing data under weaker assumptions. As a result, both papers focus on partial identification of the incidence and severity of COVID-19. Depalo (2021) combines incomplete administrative data on excess mortality in Lombardia, Italy with assumptions on excess mortality in unobserved municipalities to partially identify the true number of COVID-19 deaths. The author also uses incomplete data on testing to partially identify the true number of COVID-19 cases. Manski and Molinari (2021) partially identify the infection rate and the infectious fatality rate accounting for uncertainty due to the lack of universal or random testing and the inaccuracy of tests. Unfortunately, with the information available, it is not possible to partially identify MFs from these studies.

Beyond the specific issue of underreporting of COVID-19 cases and deaths, our approach also yields estimates of the association between NPIs (along with other country-specific attributes such as testing) and COVID-19 cases and deaths. Thus, our study also relates to existing analyses of the efficacy of NPIs. Chudik et al. (2020) extend the SIR model to allow for different degrees of compulsory and voluntary social distancing. Their model is used to study the impact of social distancing on both the spread of COVID-19 and unemployment across Chinese provinces. Note-worthy, the authors specifically account for underreporting using an assumed MF of 2, which they justify on the basis of data on the number of asymptomatic individuals aboard the Diamond Princess cruise ship. Their model assumes that underreporting follows a Log-Normal distribution, which is then used in a nonlinear equation to discern the impact of social distancing on unemployment. Chernozhukov et al. (2021) and Flaxman et al. (2020), similar to Chudik et al. (2020),

find that NPIs (stay-at-home orders and business closures, in particular) substantially reduced transmission. Additionally, Chernozhukov et al. (2021) estimate that a mask mandate would have reduced the growth rate of infections by 10 percentage points and the number of deaths by at least 19% early in the pandemic. Lastly, Askatas et al. (2020) exploit both between and within country variation in the type, timing, and level of six different NPIs across 135 countries. The authors find that a unit increase in the intensity of public events cancellations leads to a 25% decrease in the number of new infections of COVID-19.

### 3 | EMPIRICS

#### 3.1 | An epidemiological model with underreporting

We begin with the classic SIR model (Kermack & McKendrick, 1927). As stated in Avery et al. (2020, p. 80), the SIR model ‘serves as the basis of much of modern epidemiology of infectious disease, both theoretical and empirical’ However, we extend the model to incorporate incorrect reporting. Broadly, the SIR model posits that the population of country  $i$  at any time  $t$  is divided into three distinct compartments: susceptible,  $S^*$ , infected,  $I^*$ , and removed,  $R^*$ . Thus,

$$S_{it}^* + I_{it}^* + R_{it}^* = N_{it}, \quad (1)$$

where  $N_{it}$  is the overall population.

The change in the number of susceptibles at time  $t$  is given by

$$\frac{dS_{it}^*}{dt} = -\beta_{it} I_{it}^* S_{it}^*, \quad (2)$$

where  $\beta_{it}$  is the number of contacts that occur in country  $i$  in period  $t$  between infected and susceptible individuals. As the pool of infected individuals increases, the number of susceptible individuals decreases; newly infected individuals are reclassified from  $S_{it}^*$  to  $I_{it}^*$ .

The change in the number of infected depends on the number of susceptibles that become infected and the number of infected that are removed. Removal can happen either through recovery or death, leading some to refer to this as a Susceptible-Infected-Recovered-Dead (SIRD) model (e.g. Chernozhukov et al., 2021). The change in the number of infected at time  $t$  is given by

$$\frac{dI_{it}^*}{dt} = \beta_{it} I_{it}^* S_{it}^* - (\gamma_{1it} + \gamma_{2it}) I_{it}^*, \quad (3)$$

where  $\gamma_{1it}$  is the recovery rate and  $\gamma_{2it}$  is the fatality rate. The reproduction number,  $R_0$ , is  $R_{0it} = \beta_{it}/(\gamma_{1it} + \gamma_{2it})$ ; Atkeson (2020) refers to  $R_0$  as the normalized transmission rate. Our modelling of the reproduction rate as country- and time-specific follows from the fact that individual and government behaviour evolves over the course of the pandemic (Avery et al., 2020). Finally, the change in the number of recovered and dead at time  $t$  is given by

$$\frac{dR_{1it}^*}{dt} = \gamma_{1it} I_{it}^* \quad (4)$$

$$\frac{dR_{2it}^*}{dt} = \gamma_{2it} I_{it}^*, \quad (5)$$

respectively.

Equation (3) gives the change in the infected population. The number of new cases at time  $t$ ,  $Y_{it}^*$ , is given by the first term,

$$Y_{it}^* \equiv \beta_{it} I_{it}^* S_{it}^*. \quad (6)$$

In the presence of underreporting, we model the relationship between the true number of new cases,  $Y^*$ , and the reported number of cases,  $Y$ , as

$$Y_{it}^* = Y_{it} \exp(u_{it}^Y), \quad (7)$$

where  $Y_{it}^*/Y_{it}$  is the MF for cases. If  $u_{it}^Y \in \mathcal{R}^+$ , then  $Y_{it} < Y_{it}^*$ , MF is greater than one, and new cases at time  $t$  are underreported. Modelling underreporting in this way is similar to Atkeson (2020); however, Atkeson (2020) assumes  $u_{it}^Y$  is constant over time. Moreover, Atkeson (2020) does not take a stance on the sign of  $u_{it}^Y$ . As discussed above, there is widespread belief that reported cases of COVID-19 are underreported. Nonetheless, our analysis can be viewed as a test of this (as discussed in Section 3.2).

Substituting (7) into (6), taking logarithms, and rearranging yields

$$\ln(Y_{it}) \equiv \ln(\beta_{it}) + \ln(I_{it}^*) + \ln(S_{it}^*) - u_{it}^Y. \quad (8)$$

While  $S_{it}^*$  and  $I_{it}^*$  are unobserved, Equation (1) implies the following relationship

$$\begin{aligned} \ln(N_{it}) &= \ln(S_{it}^* + I_{it}^* + R_{it}^*) \\ &= \ln\left[S_{it}^* \left(1 + \frac{I_{it}^* + R_{it}^*}{S_{it}^*}\right)\right] \\ &= \ln(S_{it}^*) + \ln\left(1 + \frac{I_{it}^* + R_{it}^*}{S_{it}^*}\right) + \ln(I_{it}^*) - \ln(I_{it}^*) \\ &= \ln(S_{it}^*) + \ln(I_{it}^*) + \ln\left(\frac{N_{it}}{S_{it}^* I_{it}^*}\right) \\ &= \ln(S_{it}^*) + \ln(I_{it}^*) + \tilde{v}_{it}, \end{aligned} \quad (9)$$

where  $\tilde{v}_{it} \in \mathcal{R}$ .<sup>24</sup> Substituting (13) into (8) yields

$$\ln(Y_{it}) \equiv \ln(\beta_{it}) + \ln(N_{it}) - u_{it}^Y + v_{it}, \quad (10)$$

where  $v_{it} = -\tilde{v}_{it}$ .

Finally, we assume that  $\beta_{it}$  can be modelled partially in a deterministic fashion. Specifically, let

$$\ln(\beta_{it}) = X_{it}\beta_X + v_{it}^\beta, \quad (11)$$

<sup>24</sup>Setting  $R_{it}^* = 0$  for simplicity and noting that  $S_{it}^* >> 0$ , then  $N_{it} > (<)S_{it}^* I_{it}^*$  if  $I_{it}^* < (>)1$ . Because  $I_{it}^*$  (as well as  $S_{it}^*$  and  $N_{it}$ ) can be measured in, say, population in millions, it may be greater than or less one. Thus,  $\tilde{v}_{it}$  is a two-sided error term.

where  $v_{it}^\beta \in \mathcal{R}$ .  $X$  includes geographic, demographic and political characteristics of country  $i$ , as well as variables measuring the extent of COVID-19 testing being performed and the strength of NPIs taken by the country prior to  $t$ . We also include region-specific time trends in  $\beta_{it}$  and interactions between testing and NPIs with a linear time trend to allow the efficacy of each to vary over time. We do so to avoid the critique levied in Avery et al. (2020, p. 89) that ‘tightly parameterized models lack the flexibility to respond to qualitative changes in disease behavior that are inconsistent with earlier apparent patterns’ such as ‘repeated changes in the rate of spread due to changing regulations, changing public perception, and “quarantine fatigue”’. That said, it is important to recognize that the analysis conducted in this study pre-dates the rise of new variants of the disease, namely Delta and Omicron. Future work analysing data that spans different variants of the virus should account for possible structural breaks in the data-generating process.

Our setup is similar in spirit to the random walk approach used in Arroyo-Marioli et al. (2020). While Arroyo-Marioli et al. (2020) model  $\beta_{it}$  in levels as a random walk with an additive error, we model  $\beta_{it}$  in levels as a deterministic function with a proportional error. Thus, taking logarithms in our model produces an additive error. Incorporating this, our estimating equation becomes

$$\ln(Y_{it}) \equiv X_{it}\beta_X + \beta_N \ln(N_{it}) - u_{it}^Y + v_{it}^Y, \quad (12)$$

where  $v_{it}^Y \equiv v_{it} + v_{it}^\beta$  and we introduce  $\beta_N$  rather than impose the restriction that the coefficient on  $\ln(N_{it})$  is one. Note, if  $S_{it}^*, I_{it}^* >> 0$ , then  $v_{it}$  will be skewed in the positive direction. Assuming  $v_{it}^\beta$  is symmetrically distributed around zero, this will make our estimates of underreporting a lower bound.

Turning to deaths, Equation (5) gives the change in deaths. Thus, the number of new deaths at time  $t$ ,  $D_{it}^*$ , is given by

$$D_{it}^* \equiv \gamma_{2it} I_{it}^*. \quad (13)$$

In the presence of underreporting, we model the relationship between the true number of new deaths,  $D^*$ , and the reported number of new deaths,  $D$ , as

$$D_{it}^* = D_{it} \exp(u_{it}^D), \quad (14)$$

where  $D_{it}^*/D_{it}$  is the MF for deaths. If  $u_{it}^D \in \mathcal{R}^+$ , then  $D_{it} < D_{it}^*$ , the MF is greater than one, and new deaths at time  $t$  are underreported. Substituting (14) into (13), taking logarithms, and rearranging yields

$$\ln(D_{it}) \equiv \ln(\gamma_{2it}) + \ln(I_{it}^*) - u_{it}^D. \quad (15)$$

As before, assume that  $\gamma_{2it}$  can be modelled partially in a deterministic fashion. Specifically, let

$$\ln(\gamma_{2it}) = X_{it}\gamma_X + v_{it}^\gamma, \quad (16)$$

where  $v_{it}^\gamma \in \mathcal{R}$ . Substituting this into (15) and replacing  $I^*$  with  $N$  produces our estimating equation

$$\ln(D_{it}) \equiv X_{it}\gamma_X + \gamma_N \ln(N_{it}) - u_{it}^D + v_{it}^D, \quad (17)$$

where  $v_{it}^D \equiv v_{it}^\gamma + \ln(I_{it}^*) - \gamma_N \ln(N_{it}) = v_{it}^\gamma - \tilde{v}_{it} - \ln(S_{it}^*) - (1 - \gamma_N) \ln(N_{it})$ .

### 3.2 | Estimation

Our objective is to estimate Equations (12) and (17). In addition to obtaining coefficient estimates, we also wish to obtain estimates of  $u_{it}^Y$  and  $u_{it}^D$  to estimate the actual number of cases,  $Y_{it}^*$ , and deaths,  $D_{it}^*$ . This requires estimation of the coefficients and then decomposition of the residuals into its constituent parts. This is not trivial given the structure of the model. Fortunately, the estimating equations above have a parallel structure to SFA if one is willing to assume that the errors,  $u_{it}^Y$  and  $u_{it}^D$ , are one-sided (Aigner et al., 1977). In the setting where  $u$  is interpreted as one-sided measurement error, Millimet and Parmeter (2021) show via simulations and several empirical replications that SFA can be a valuable tool to address one-sided measurement error.

Prior to continuing, it is important to make one point. As should be clear from Sections 1 and 2, the assumption that cases and deaths are underreported is not particularly controversial. Nonetheless, it is worth emphasizing that estimation of a stochastic frontier models does not *impose* underreporting on the data. Instead, underreporting is a ‘testable’ assumption in the sense that identification in the stochastic frontier model requires the composite error term to be negatively skewed. If reporting errors are not one sided (or at least strongly skewed towards negative values), but instead classical or non-existent, then the residuals will not be negatively skewed and identification will fail. Following Kumbhakar et al. (2015), we find statistically meaningful evidence in favour of negative skewness of the ordinary least squares (OLS) residuals (results available upon request).

To proceed with estimation, we make distributional assumptions for  $v_{it}^q$  and  $u_{it}^q$ ,  $q \in \{Y, D\}$ , and estimate the models by maximum likelihood where the dependent variables are transformed using the inverse hyperbolic sine transformation (IHS) to account for zeros (Askitas et al., 2020; Bellemare & Wichman, 2020; Goolsbee & Syverson, 2020). Importantly, the coefficient estimates are consistent regardless of the distributional assumptions (Papadopoulos & Parmeter, 2021). However, estimates of actual cases and deaths may be sensitive to the distributional assumptions. Estimation is performed in Stata using `frontier` or `xtfrontier`.

Since  $u^q$  and  $v^q$  are not observed independently, estimation of the model requires knowledge of the composite error structure of the model,  $\epsilon_{it}^q \equiv v_{it}^q - u_{it}^q$  so that maximum likelihood can be deployed. Finding the composite density,  $f(\epsilon^q)$ , is an exercise in convolution if  $v_{it}^q$  and  $u_{it}^q$  are assumed independent, which is found as

$$f(\epsilon^q) = \int_0^\infty f(u^q) \cdot f(\epsilon^q + u^q) du^q. \quad (18)$$

The classic set of assumptions is that  $v_{it}^q \stackrel{\text{iid}}{\sim} N(0, \sigma_{vq}^2)$  and  $u_{it}^q \stackrel{\text{iid}}{\sim} N^+(0, \sigma_{uq}^2)$ , known colloquially as the Normal-Half Normal setup. In the cross-sectional setting, Aigner et al. (1977) demonstrate that  $f(\epsilon_{it}^q) = \frac{2}{\sigma_q} \phi\left(\frac{\epsilon_{it}^q}{\sigma_q}\right) \Phi\left(-\frac{\lambda \epsilon_{it}^q}{\sigma_q}\right)$ , where  $\sigma_q^2 = \sigma_{vq}^2 + \sigma_{uq}^2$ ,  $\lambda = \sigma_{uq}/\sigma_{vq}$ , and  $\phi(\cdot)$  and  $\Phi(\cdot)$  are the standard Normal density and distribution functions respectively. This gives rise to the following log-likelihood function

$$\ln L = \sum_{i=1}^N \sum_{t=1}^{T_i} \left\{ -\ln \sigma_q + \ln \Phi \left[ -\frac{\lambda \epsilon_{it}^q}{\sigma_q} \right] - \frac{(\epsilon_{it}^q)^2}{2\sigma_q^2} \right\}, \quad q \in \{Y, D\}. \quad (19)$$

Beyond the double summation, this is identical to the likelihood function for estimation of the Skew Normal distribution (Azzalini, 1985, 2014).

To assess sensitivity of our findings, we consider three alternative sets of distributional assumptions. First, we assume that  $v_{it}^q \stackrel{\text{iid}}{\sim} N(0, \sigma_{vq}^2)$  and  $u_{it}^q \stackrel{\text{iid}}{\sim} \exp(\lambda^q)$  (Meeusen & van den Broeck, 1977). The Exponential distribution has several diagnostic advantages over the Half Normal as it allows for a greater range of both skewness and kurtosis of the composite error,  $\varepsilon_{it}^q$ , providing greater flexibility to match its shape (Papadopoulos & Parmeter, 2021). Both Aigner et al. (1977) and Meeusen and van den Broeck (1977) show that  $f(\varepsilon_{it}^q) = \frac{1}{\sigma_u} \exp\left[\frac{\varepsilon_{it}^q}{\sigma_u} + \frac{\sigma_v^2}{2\sigma_u^2}\right] \Phi\left(-\frac{\varepsilon_{it}^q}{\sigma_v} - \frac{\sigma_v}{\sigma_u}\right)$ . The log-likelihood function is given by

$$\ln L = \sum_{i=1}^N \sum_{t=1}^{T_i} \left\{ -\ln \sigma_{uq} + \frac{\sigma_{vq}^2}{2\sigma_{uq}^2} + \ln \Phi\left(-\frac{\varepsilon_{it}^q}{\sigma_v} - \frac{\sigma_v}{\sigma_u}\right) + \frac{\varepsilon_{it}^q}{\sigma_{uq}} \right\}, \quad q \in \{Y, D\}. \quad (20)$$

Second, we relax the independence assumption by estimating a stochastic frontier model with time invariant underreporting,  $u_i^q$  (Pitt & Lee, 1981). Specifically,  $\varepsilon_{it}^q = v_{it}^q - u_i^q$ , where  $v_{it}^q \stackrel{\text{iid}}{\sim} N(0, \sigma_{vq}^2)$  and  $u_i^q \stackrel{\text{iid}}{\sim} N^+(0, \sigma_{uq}^2)$ . The composite error density of  $\varepsilon_i^q = (\varepsilon_{i1}^q, \dots, \varepsilon_{iT}^q)$  is

$$f(\varepsilon_i^q) = \frac{2\sigma_{uq}}{\sigma_{vq}^2 + T\sigma_{uq}^2} \exp\left(-\frac{\varepsilon_i^q' A \varepsilon_i^q}{2\sigma_{vq}^2}\right) \Phi\left[-\frac{\sigma_{uq}}{\sigma_{vq}(\sigma_{vq}^2 + T\sigma_{uq}^2)} \sum_{t=1}^T \varepsilon_{it}^q\right], \quad (21)$$

where  $A = I_T - (\sigma_{uq}^2 / \sigma_{vq}^2 + T\sigma_{uq}^2) l_T l_T'$  with  $l_T'$  a  $T \times 1$  vector of ones and  $I_T$  a  $T \times T$  identity matrix. The corresponding log-likelihood function is

$$\begin{aligned} \ln L = & -\frac{N(T-1)}{2} \ln \sigma_{vq}^2 - \frac{N}{2} l_N(\sigma_{vq}^2 + T\sigma_{uq}^2) \\ & - \frac{1}{2\sigma_{vq}^2} \sum_{i=1}^N \varepsilon_i^q' A \varepsilon_i^q + \sum_{i=1}^N \ln \Phi\left[-\frac{\sigma_{uq}}{\sigma_{vq}(\sigma_{vq}^2 + T\sigma_{uq}^2)} \sum_{t=1}^T \varepsilon_{it}^q\right]. \end{aligned} \quad (22)$$

The assumption of time invariant underreporting is restrictive, but relaxes the assumption of independent underreporting within a country. Moreover, it may be a reasonable approximation since the dependent variables are transformed using the IHS. As such, time invariant underreporting implies a constant percentage of underreporting. For cases, this may be plausible given that a constant fraction of infected are likely to be asymptomatic. It is also consistent with the claim in Avery et al. (2020, p. 89) that ‘observation error can rise as larger swaths of the population are infected and contact tracing becomes less reliable’. For deaths, it may be plausible as well given that reporting practices for cause of death are likely fixed during the initial waves of the pandemic.

That said, it is also possible that the proportion of missing data may vary over time. Murray (2020, p. 110) states: ‘For diseases with more common mild and asymptomatic cases … the completeness of the data can vary in complex ways over time’. Thus, in our final specification, we relax the assumption of time invariant underreporting and estimate a time-varying decay model. In this case,  $\varepsilon_{it}^q = v_{it}^q - u_{it}^q$ , where  $v_{it}^q \stackrel{\text{iid}}{\sim} N(0, \sigma_{vq}^2)$ ,  $u_{it}^q = \exp[-\eta^q(t - T_i)] u_i^q$ , and  $u_i^q \stackrel{\text{iid}}{\sim} N^+(0, \sigma_{uq}^2)$ , which follows Battese and Coelli (1992) who provide the derivation of  $f(\varepsilon_{it}^q)$ . The sign of  $\eta$  determines whether the extent of underreporting is increasing or decreasing over time;  $\eta < 0$  implies that it is increasing over time while  $\eta > 0$  implies the reverse. The log-likelihood function is given by

$$\ln L = \sum_{i=1}^N \left\{ \ln \Phi \left( \frac{\mu_{i*}^q}{\sigma_*^q} \right) + \frac{1}{2} \ln (\sigma_*^q)^2 - \frac{1}{2} \left[ \frac{1}{\sigma_{v^q}^2} \sum_{t=1}^{T_i} (\varepsilon_{it}^q)^2 - \left( \frac{\mu_{i*}^q}{\sigma_*^q} \right)^2 \right] - T_i \ln (\sigma_{v^q}) - \ln (\sigma_{u^q}) \right\},$$

$q \in \{Y, D\}$

(23)

where

$$\eta_{it}^q = \exp [-\eta^q (t - T_i)]$$

$$\begin{aligned}\mu_{i*}^q &= \frac{-\sigma_{u^q}^2 \sum_{t=1}^{T_i} \eta_{it}^q \varepsilon_{it}^q}{\sigma_{v^q}^2 + \sigma_{u^q}^2 \sum_{t=1}^{T_i} (\eta_{it}^q)^2} \\ \sigma_*^q &= \frac{\sigma_{v^q}^2 \sigma_{u^q}^2}{\sigma_{v^q}^2 + \sigma_{u^q}^2 \sum_{t=1}^{T_i} (\eta_{it}^q)^2}.\end{aligned}$$

The log-likelihood of the time invariant model is obtained by maximizing (23) imposing the constraint  $\eta^q = 0$ . Note, the likelihood function in Equation (19) is identical to (23) with  $\eta^q = 0$  and  $T_i = 1 \forall i$ .

One of the main novelties of SFA is that beyond calculation of average underreporting (for a given distributional assumption on  $u^q$ ), the degree of underreporting can be estimated for each observation (Jondrow et al., 1982). The intuition for this is that since  $\varepsilon_{it}^q$  can be estimated, it contains information on  $u_{it}^q$ , which, if used correctly, can provide an estimation strategy; the natural starting point for this is the conditional density,  $f(u^q | \varepsilon^q)$ . Again, for a given set of distributional assumptions, this conditional density can be derived.

A variety of options exist to estimate/predict  $u_{it}^q$  from  $f(u^q | \varepsilon^q)$ , but the most popular is through the calculation of the conditional expectation of the one-sided error given the composite residual. In the case of the time invariant and time-varying decay models, this is given by

$$E[u_i^q | \varepsilon_i^q] = \mu_{i*}^q + \sigma_*^q \left[ \frac{\phi(-\omega_{i*}^q)}{1 - \Phi(-\omega_{i*}^q)} \right],$$
(24)

where  $\omega_{i*}^q = \mu_{i*}^q / \sigma_*^q$ .

There certainly exist alternative approaches to estimate both the model itself and measure the degree of underreporting through a panel setup; see Parmeter and Kumbhakar (2014) for a thorough review. However, our approach is consistent with several recent COVID-19 studies. Li and Linton (2021) estimate linear regression models for the logarithm of daily cases and deaths to forecast future trajectories. Chernozhukov et al. (2021) estimate linear regression models for the growth rate of weekly cases and deaths in the US to assess the impact of NPIs and behavioural responses. Weinberger et al. (2020) examine pneumonia and influenza deaths in order to assess excess mortality using a Poisson regression framework, treating deaths as a count variable. Teixeira da Silva and Tsigaris (2020) estimate cross-country linear regression models to explore determinants of the case fatality rate.

Although continuing in this line of analysis, our approach is not meant to displace alternative, more traditional epidemiological approaches, nor alternative structural approaches such as in Hortaçsu et al. (2021), but to complement them. Our approach offers numerous advantages. First,

it provides estimates under alternative assumptions than existing methods designed to correct for underreporting. Second, our method is estimable across a large range of countries. Finally, our approach could prove especially useful early in pandemics when epidemiological approaches based on community-based or serological surveys take time for the spread of the diseases to reveal itself.

### 3.3 | Data

Data on official reports of daily COVID-19 cases and deaths are obtained from Our World in Data.<sup>25</sup> We utilize data from 1 January 2020 to 3 November 2020 (i.e. the first 44 weeks of 2020). We restrict our sample to countries observed in every week. This leads to a sample of 63 countries which are listed in Table 1. In our analysis, we further restrict the sample along two dimensions. In the models for both cases and deaths, we restrict the sample to weeks following the first week with at least five newly reported cases. In the models for deaths, we further restrict the sample to weeks following the first week with at least 10 newly reported deaths. These sample restrictions are not strictly necessary. However, the stochastic frontier models can have a difficult time converging otherwise. That said, it is also consistent with the manner in which cases and deaths are reported in the popular press, where countries' observed paths are discussed after the country passes an initial milestone. Nonetheless, one might view our estimates of underreporting as underestimates given that they fail to capture missed cases and deaths prior to the start of our estimation sample.

Our dependent variables are new cases and deaths aggregated to the weekly level. Control variables include population, population density, median age, the per cent of the population aged 65 and above, diabetes prevalence, and per capita gross domestic product. These variables come from the same source as the COVID-19 data. Sá (2020) and Olmo and Sanso-Navarro (2021) use similar controls to discern determinants of reported cases and deaths in England and Wales and cases in New York City respectively. In addition, we include Transparency International's 2019 Corruption Perception Index as a country-level measure of corruption.<sup>26</sup> Next, we include lagged weekly tests obtained from the same source as the COVID-19 data. Data on new tests are measured in different units across countries (people tested, samples tested, tests conducted, or unclear). To account for this, we include dummy variables for the unit of measurement as well as interactions between these dummy variables and the continuous measure of new tests. As all testing variables are lagged 1 week—to allow for lags in completing the test—the possibility of feedback from shocks to new cases or deaths to testing is limited. In Section 4.2 we also allow for underreporting to depend directly on testing. Finally, we include a time-varying index of the stringency of NPIs undertaken in each country. The index is taken from the Oxford COVID-19 Government Response Tracker (OxCGR) which varies between 0 and 100 and over time (Hale et al., 2020).<sup>27</sup> Given that the estimated incubation period for COVID-19 has been reported to be between 14 and 21 days, and deaths likely occur after a further delay, we use lags of the NPI stringency index. When modelling new cases, we consider the average of the index computed over the third and fourth lags. When modelling new deaths, we use the average of the index computed over the fifth and sixth lags. Excluding countries with missing covariate

<sup>25</sup>The data are available at <https://github.com/owid/covid-19-data/tree/master/public/data>.

<sup>26</sup>See <https://www.transparency.org/en/cpi>.

<sup>27</sup>The data are available at <https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker>. We use the containment and health index.

TABLE 1 Countries included in sample

<i>Europe and Central Asia</i>	<i>East Asia and the Pacific</i>	<i>Latin America and the Caribbean</i>
Austria	Australia	Brazil
Azerbaijan	Cambodia	Dominican Republic
Belarus	China	Ecuador
Belgium	Indonesia	Mexico
Croatia	Japan	
Czech Republic	Malaysia	<i>Middle East and Africa</i>
Denmark	New Zealand	Algeria
Estonia	Philippines	Bahrain
Finland	Singapore	Egypt
France	South Korea	Iran
Georgia	Thailand	Iraq
Germany	Vietnam	Israel
Greece		Kuwait
Iceland	<i>South Asia</i>	Lebanon
Ireland	Afghanistan	Nigeria
Italy	India	Oman
Lithuania	Nepal	Qatar
Luxembourg	Pakistan	United Arab Emirates
Netherlands	Sri Lanka	
Norway		<i>North America</i>
Romania		Canada
Russia		United States
Spain		
Sweden		
Switzerland		
United Kingdom		

data reduces our sample to 61 countries when assessing cases and 56 countries when analysing deaths.

All models include region fixed effects (East Asia and the Pacific, Europe and Central Asia, Latin America and the Caribbean, Middle East and Africa, North America, and South Asia) and region-specific linear time trends (see Table 1). The time trends are measured relative to time in the sample (i.e. weeks after first week with reported cases or deaths above the threshold), not calendar time. We also interact NPI stringency and testing with a linear (calendar) time trend to

allow for changes in the effectiveness of each over time. Finally, we include quadratics for the continuous country characteristics to allow for nonlinear effects.

## 4 | RESULTS

### 4.1 | COVID-19 cases and deaths

#### 4.1.1 | Stochastic frontier models assuming independent underreporting

Table 2 presents select coefficient estimates from our specifications assuming independent under-reporting (full results are available upon request). Columns 1 and 2 assume the one-sided error is distributed Half Normal; columns 3 and 4 assume an Exponential distribution. The dependent variable in columns 1 and 3 is the IHS transformation of weekly cases. The dependent variable in columns 2 and 4 is the IHS transformation of weekly deaths. In the table we also report the mean of the observation-specific marginal effects of the lagged containment index (along with its standard error and  $p$ -value associated with the two-sided test that this effect is zero). The marginal effects of a covariate, say for weekly cases, is given by  $\partial E[IHS(Y_{it})|X_{it}, \ln N_{it}]/\partial X_{it}$ . Due to the inclusion of quadratic and interaction terms, these marginal effects vary across country and over time. We report the overall average marginal effect for the lagged containment index in the table. Lastly, the final two rows in the table display the cumulative total number of observed and estimated cases and deaths over the sample period. In total, roughly 37.66 million cases have been reported along with about 0.98 million deaths.

Figures 1 and 2 plot the average of the country-specific marginal effects over time for the number of tests performed and NPIs, respectively, to facilitate interpretation of the coefficients. Figure 1 shows a positive association between testing and cases since early April. There is a consistently negative association between testing and deaths (when tests are measured in clearly defined units). Moreover, as the marginal effects represent elasticities (i.e. the per cent change in cases given a 1% increase in testing), the associations are significant in magnitude. Although one should be cautious giving this a causal interpretation, it is nonetheless reassuring. It is also comforting that the distributional assumption concerning the underreporting has little impact on the estimates.

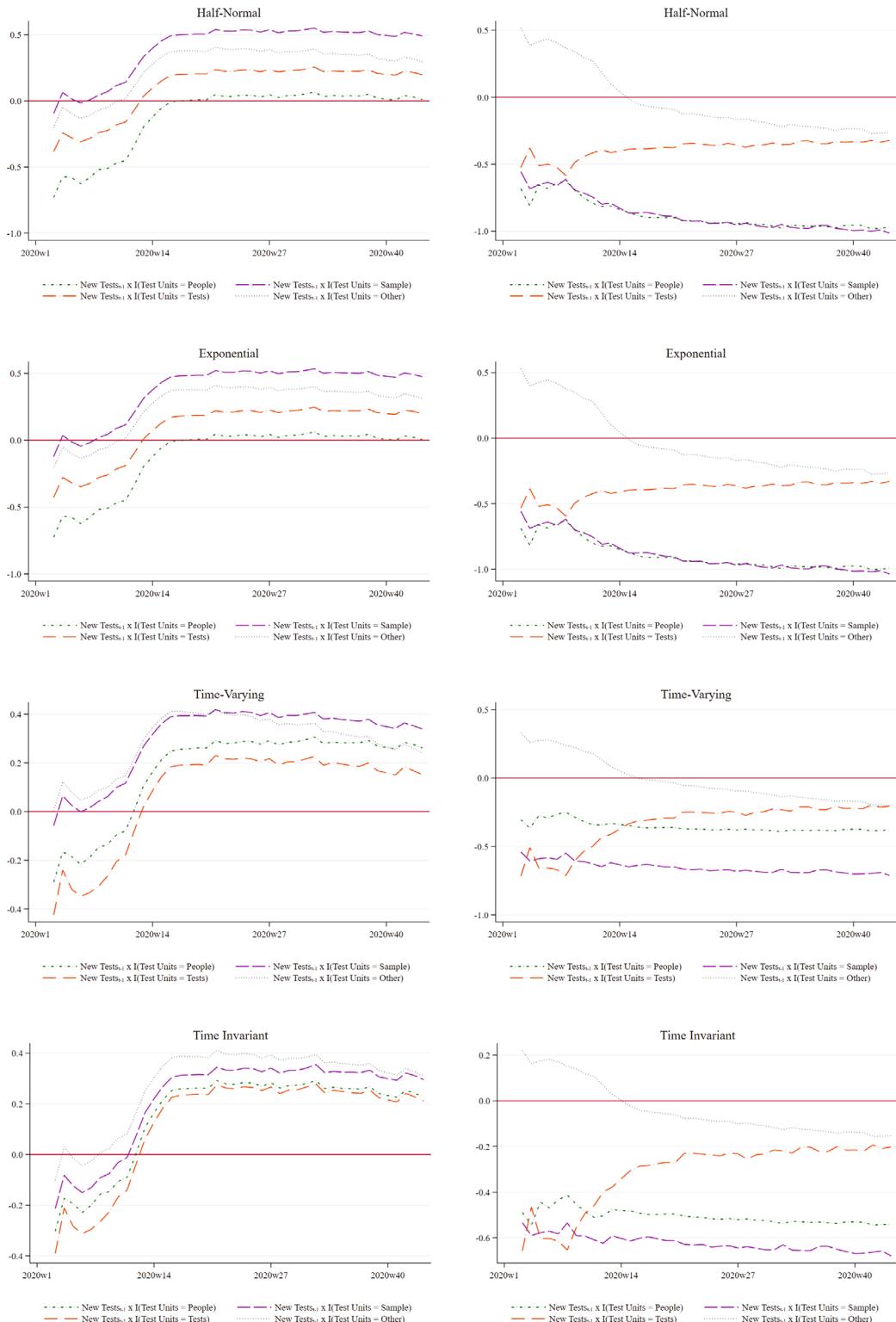
Figure 2 reveals a volatile association between NPIs and cases and deaths, although the time pattern is similar across the two outcomes and invariant to the distributional assumptions considered. For both cases and deaths, the association between NPIs and cases and deaths is most negative during Summer 2020. From April to July, NPIs are associated with fewer cases and deaths. However, perhaps due to pandemic fatigue, this association falls to zero and even becomes positive for cases in Fall 2020. To interpret the magnitude of the association, recall that the containment index measuring NPIs ranges from 0 to 100 with a mean close to 50 and a standard deviation of roughly 25. Thus, a one standard deviation increase in this index is associated with a decline in deaths of roughly 40% ( $\exp(-0.02 \cdot 25) - 1 = -0.4$ ) during the summer. This is substantial, yet consistent with Chernozhukov et al. (2021). It does, however, contrast with Goolsbee and Syverson (2020) who find that most of the decline in economic activity in the US is attributable to behavioural responses due to fear, not policy changes. Nonetheless, as stated previously, one should be cautious giving our results a causal interpretation.

Turning to the issue of underreporting for cases, we estimate a MF of 2.9 (2.0) when the under-reporting is assumed to be distributed Half Normal (Exponential). Given the ranges discussed in

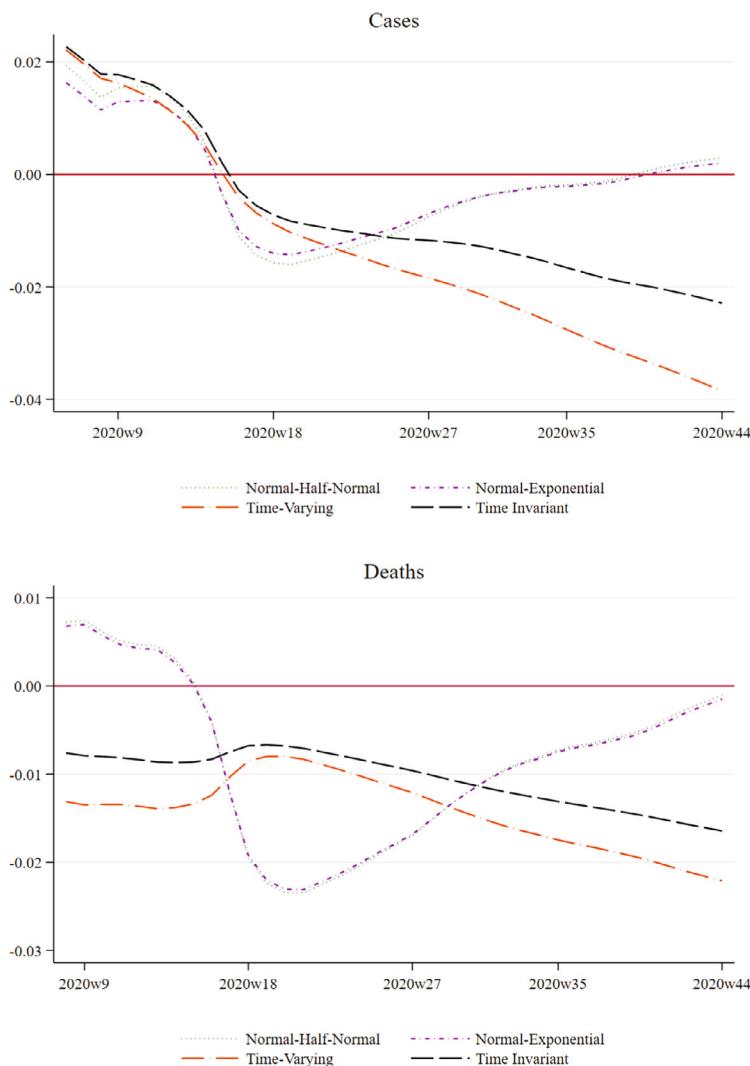
TABLE 2 Stochastic frontier results assuming independence

	Half normal		Exponential	
	Cases	Deaths	Cases	Deaths
	(1)	(2)	(3)	(4)
Lagged Containment Index	0.020 (0.016)	0.006 (0.015)	0.017 (0.016)	0.006 (0.016)
(Lagged Containment Index) <sup>2</sup>	-0.000* (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
(Lagged Containment Index) × (Linear Time Trend)	0.001 (0.000)	0.001 (0.000)	0.000 (0.000)	0.001 (0.000)
New Tests <sub>t-1</sub>	-0.185 (0.128)	0.523** (0.225)	-0.193 (0.139)	0.537** (0.215)
(New Tests <sub>t-1</sub> ) <sup>2</sup>	0.045*** (0.010)	-0.035** (0.016)	0.045*** (0.011)	-0.035** (0.015)
(New Tests <sub>t-1</sub> ) × (Linear Time Trend)	-0.008*** (0.002)	-0.004* (0.002)	-0.007*** (0.002)	-0.004* (0.002)
(New Tests <sub>t-1</sub> ) × I(Test Units = People)	-0.535*** (0.194)	-1.204*** (0.270)	-0.523** (0.209)	-1.222*** (0.262)
(New Tests <sub>t-1</sub> ) × I(Test Units = People) <sup>2</sup>	0.033* (0.017)	0.096*** (0.023)	0.033* (0.018)	0.098*** (0.021)
(New Tests <sub>t-1</sub> ) × I(Test Units = People)) × (Linear Time Trend)	0.003 (0.003)	0.003 (0.003)	0.002 (0.003)	0.002 (0.003)
(New Tests <sub>t-1</sub> ) × I(Test Units = Samples)	0.101 (0.193)	-1.072*** (0.282)	0.080 (0.208)	-1.088*** (0.270)
(New Tests <sub>t-1</sub> ) × I(Test Units = Samples) <sup>2</sup>	-0.009 (0.020)	0.104*** (0.022)	-0.005 (0.020)	0.105*** (0.021)
(New Tests <sub>t-1</sub> ) × I(Test Units = Samples)) × (Linear Time Trend)	0.003 (0.003)	0.001 (0.004)	0.002 (0.003)	0.001 (0.004)
(New Tests <sub>t-1</sub> ) × I(Test Units = Tests)	-0.188 (0.164)	-1.048*** (0.288)	-0.226 (0.172)	-1.070*** (0.266)
(New Tests <sub>t-1</sub> ) × I(Test Units = Tests) <sup>2</sup>	-0.005 (0.012)	0.075*** (0.021)	-0.003 (0.013)	0.076*** (0.020)
(New Tests <sub>t-1</sub> ) × I(Test Units = Tests)) × (Linear Time Trend)	0.003* (0.002)	0.004* (0.002)	0.003** (0.002)	0.004* (0.002)
N	2195	1712	2195	1712
Marginal Effect: Lagged Containment Index	-0.001 (0.006)	-0.009 (0.006)	-0.001 (0.006)	-0.009 (0.007)
	<i>p</i> = 0.934	<i>p</i> = 0.156	<i>p</i> = 0.909	<i>p</i> = 0.162
Observed (Cumulative)	37,663,438	980,123	37,663,438	980,123
Predicted (Cumulative)	109,767,009	2,108,317	73,746,440	1,534,761

Standard errors in parentheses. Half normal model assumes  $u_{it} \sim N^+(0, \sigma_u^2)$ . Exponential model assumes  $u_{it} \sim \exp(\lambda)$ . I(·) is the indicator function, equal to one if the argument is true and zero otherwise. Lagged containment index is the average of the third and fourth lags in the models for cases; fifth and sixth lags for deaths. Number of tests is transformed using the inverse hyperbolic sine. Other covariates included in all models: log population, population density, median age, percent of population aged 65+, percent of population with diabetes, log per capita Gross Domestic Product, corruption index, the quadratic of each of the preceding variables, region fixed effects, and region-specific time trends. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



**FIGURE 1** Marginal effects of testing on cases and deaths. Marginal effects on cases are in the left column. Marginal effects on deaths are in the right column [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 2** Marginal effects of non-pharmaceutical interventions on cases and deaths. Marginal effects on cases are in the top panel. Marginal effects on deaths are in the bottom panel [Colour figure can be viewed at [wileyonlinelibrary.com](https://wileyonlinelibrary.com)]

Section 2, this is on the low end. Hortaçsu et al. (2021) estimate a MF ranging from 6 to 24, Li et al. (2020) estimate a MF of at least 7, Flaxman et al. (2020) estimate a MF of at least 12, and Reese et al. (2021) estimate an MF of nearly 8. However, our MF is consistent with the argument put forth in Chudik et al. (2020). Needing to choose a MF to calibrate their model, the authors note that of the 3,711 passengers and crew on board the Diamond Princess cruise ship, 712 tested positive for COVID-19. Of these 712 individuals, 331 were asymptomatic at the time of testing; Chudik et al. (2020) report 311, but this number is erroneous. Under the assumption that the testing is accurate, this suggests that the 331 asymptomatic individuals likely would not have been tested—and hence not officially counted—had they not been aboard the Diamond Princess. Thus, the MF is  $712/381 \approx 1.9$ . Sutton et al. (2020) provide an alternative estimate derived from the testing of all pregnant women delivering infants at the New York-Presbyterian Allen Hospital and

Columbia University Irving Medical Center. Of the 215 women who delivered, 214 were tested for COVID-19 despite only four mothers displaying symptoms. In addition to these four, another 29 were asymptomatic but tested positive. This suggests a MF of  $33/4 \approx 8$ , although this is likely not a random sample of the population.

We find lower levels of underreporting for deaths. We estimate a MF of 2.2 (1.6) when the underreporting is assumed to be distributed Half Normal (Exponential). The smaller MF for deaths is not surprising, however, as the underreporting of COVID-19 cases is most likely due to asymptomatic carriers who never interact with the medical system, whereas underreported deaths are predominantly due to unintentional or intentional misdiagnosis. Flaxman et al. (2020, p. 258) state that data on deaths are ‘... likely to be far more reliable than case data’. Moreover, our estimated MF is in line with estimates discussed in Section 2, ranging from 1.4 to 2.6.

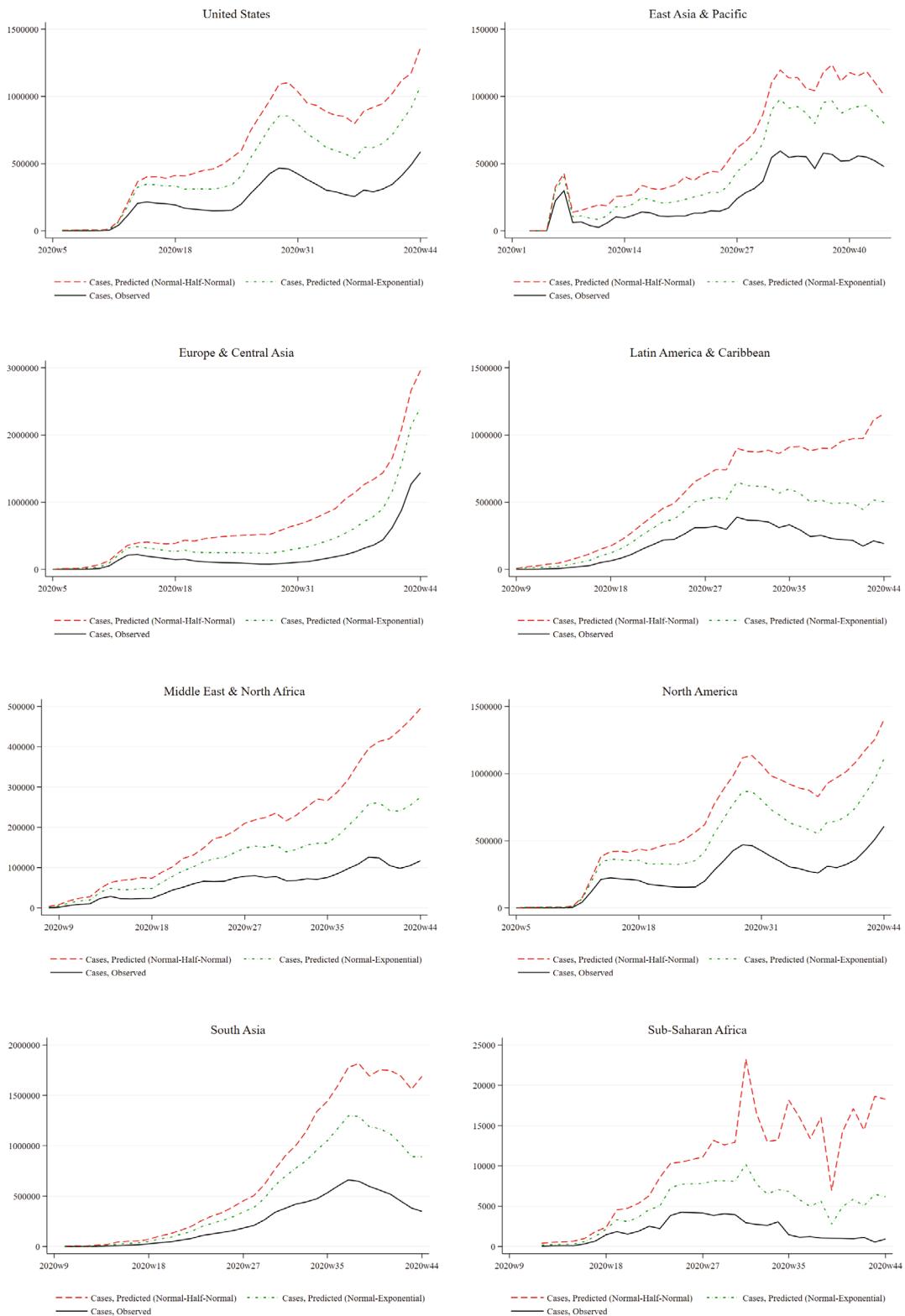
For the US we find that through 3 November the total number of cases was 9,291,234 while the predicted number of cases under the Half Normal specification is 24,178,170 (a MF of 2.6) and that for the Exponential specification is 17,874,482 (a MF of 1.9). For deaths, again through 3 November, reported deaths are 231,545, while the Half Normal specification predicts deaths at 485,320 (a MF of 2.1) while the Exponential specification predicts deaths to be 359,449 (a MF of 1.6).

To visualize the estimated levels of underreporting, Figures 3 and 4 show graphically the observed and estimated weekly cases and deaths, respectively, for different geographic regions. These figures are obtained directly from the model estimates in Table 2. A few interesting findings stand out. First, the Exponential model estimates less underreporting for all regions for both cases and deaths. Second, in an absolute sense, underreporting of cases has not waned over time. For the US, Europe and Central Asia, and North America, the absolute amount of underreported appears to have stabilized in Fall 2020. However, the absolute amount of underreporting continues to rise in other regions. This is consistent with idea that contact tracing may get exponentially more difficult as more of the population becomes infected (Avery et al., 2020).

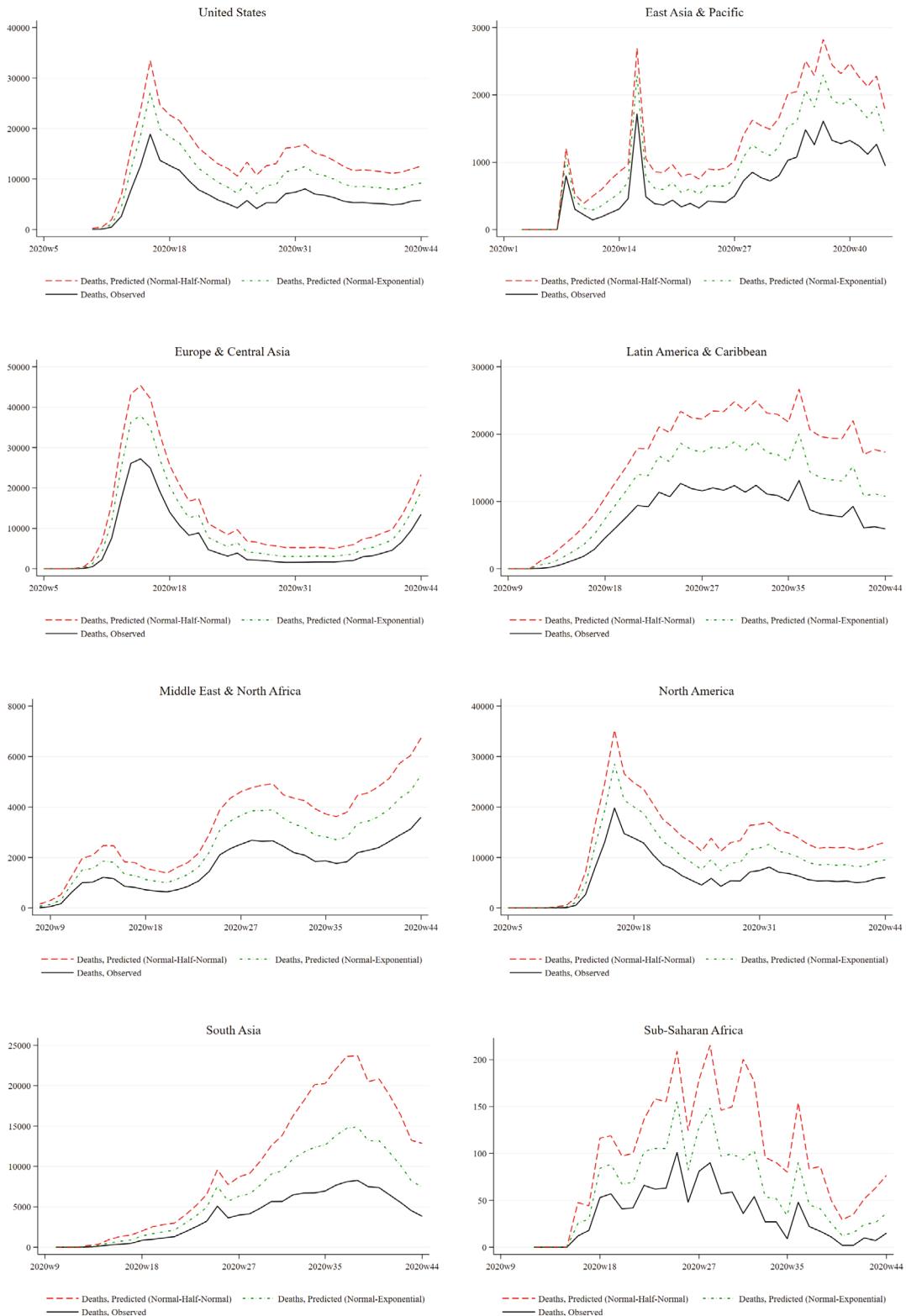
Third, whereas the observed data show a decline in new cases for some regions at some points in time, the Half Normal model suggests that some of this is illusory. For example, the decline in observed cases in Latin America and the Caribbean starting in late summer is attributable to an increase in underreporting. Even in Europe and Central Asia during the summer, the modest decline in reported cases is potentially an artefact of underreporting. Finally, while underreporting of deaths does not seem to be disappearing either, there is some evidence that it has declined over time in certain regions. For example, for both the United States, Europe and Central Asia, and North America, underreporting appears most severe early in the pandemic. However, for other regions, most notably Latin America and the Caribbean, South Asia, and East Asia and the Pacific, underreporting seems to have become worse in Fall 2020.

#### 4.1.2 | Stochastic frontier models allowing for dependence

Table 3 is analogous to Table 2 except the estimates are from our specifications allowing for dependence over time in the underreporting. As the assumption of independent reporting errors over time seems perhaps too strong due to, say, constraints within the health care system or persistence in sub-populations operating outside public health surveillance systems, allowing for serial dependence is desirable. Columns 1 and 2 assume the underreporting follows the time-varying decay structure; columns 3 and 4 assume time invariant underreporting. The average of the



**FIGURE 3** Comparison of observed and predicted weekly cases assuming independent underreporting [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 4** Comparison of observed and predicted weekly deaths assuming independent underreporting  
[Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

TABLE 3 Stochastic frontier results allowing for dependence

	Time-varying measurement error		Time invariant measurement error	
	Cases	Deaths	Cases	Deaths
	(1)	(2)	(3)	(4)
Lagged Containment Index	0.031*** (0.007)	-0.008 (0.007)	0.028*** (0.007)	-0.004 (0.007)
(Lagged Containment Index) <sup>2</sup>	-0.000 (0.000)	0.000 (0.000)	-0.000* (0.000)	0.000 (0.000)
(Lagged Containment Index) × (Linear Time Trend)	-0.001*** (0.000)	-0.001** (0.000)	-0.001*** (0.000)	-0.000* (0.000)
(New Tests <sub>t-1</sub> )	0.031 (0.243)	0.334 (0.378)	-0.076 (0.244)	0.223 (0.380)
(New Tests <sub>t-1</sub> ) <sup>2</sup>	0.034 (0.024)	-0.017 (0.034)	0.036 (0.025)	-0.015 (0.035)
(New Tests <sub>t-1</sub> ) × (Linear Time Trend)	-0.010*** (0.003)	-0.005 (0.004)	-0.007** (0.003)	-0.002 (0.004)
(New Tests <sub>t-1</sub> ) × I(Test Units = People)	-0.313 (0.269)	-0.631 (0.401)	-0.225 (0.270)	-0.728* (0.406)
(New Tests <sub>t-1</sub> × I(Test Units = People)) <sup>2</sup>	0.014 (0.026)	0.059* (0.036)	0.018 (0.027)	0.072* (0.037)
(New Tests <sub>t-1</sub> × I(Test Units = People)) × (Linear Time Trend)	0.006* (0.003)	0.005 (0.004)	0.001 (0.003)	0.000 (0.004)
(New Tests <sub>t-1</sub> ) × I(Test Units = Samples)	-0.081 (0.257)	-0.856** (0.392)	-0.135 (0.256)	-0.755* (0.392)
(New Tests <sub>t-1</sub> × I(Test Units = Samples)) <sup>2</sup>	-0.006 (0.025)	0.077** (0.035)	0.005 (0.026)	0.078** (0.036)
(New Tests <sub>t-1</sub> × I(Test Units = Samples)) × (Linear Time Trend)	0.004 (0.003)	0.003 (0.004)	0.002 (0.003)	-0.000 (0.004)
(New Tests <sub>t-1</sub> ) × I(Test Units = Tests)	-0.433* (0.249)	-1.024*** (0.385)	-0.299 (0.251)	-0.864** (0.386)
(New Tests <sub>t-1</sub> × I(Test Units = Tests)) <sup>2</sup>	0.017 (0.025)	0.073** (0.034)	0.014 (0.025)	0.067* (0.035)
(New Tests <sub>t-1</sub> × I(Test Units = Tests)) × (Linear Time Trend)	0.004 (0.003)	0.005 (0.004)	0.001 (0.003)	0.001 (0.004)
N	2195	1712	2195	1712
Marginal Effect: Lagged Containment Index	-0.015*** (0.003)	-0.014*** (0.003)	-0.007*** (0.003)	-0.010*** (0.003)
	<i>p</i> = 0.000	<i>p</i> = 0.000	<i>p</i> = 0.006	<i>p</i> = 0.001
Observed (Cumulative)	37,663,438	980,123	37,663,438	980,123
Predicted (Cumulative)	337,360,283	2,954,841	130,411,499	1,849,022

Standard errors in parentheses. Time invariant model assumes  $u_i \sim N^+(0, \sigma_u^2)$ . The time-varying model assumes  $u_{it} = \exp\{-\eta(t - T_i)\} u_i$ ,  $u_i \sim N^+(0, \sigma_u^2)$ .  $I(\cdot)$  is the indicator function, equal to one if the argument is true and zero otherwise. Lagged containment index is the average of the third and fourth lags in the models for cases; fifth and sixth lags for deaths. Number of tests is transformed using the inverse hyperbolic sine. Other covariates included in all models: log population, population density, median age, per cent of population aged 65+, per cent of population with diabetes, log per capita Gross Domestic Product, corruption index, the quadratic of each of the preceding variables, region fixed effects and region-specific time trends. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

country-specific marginal effects over time for the number of tests performed and NPIs are shown in Figures 1 and 2 respectively.

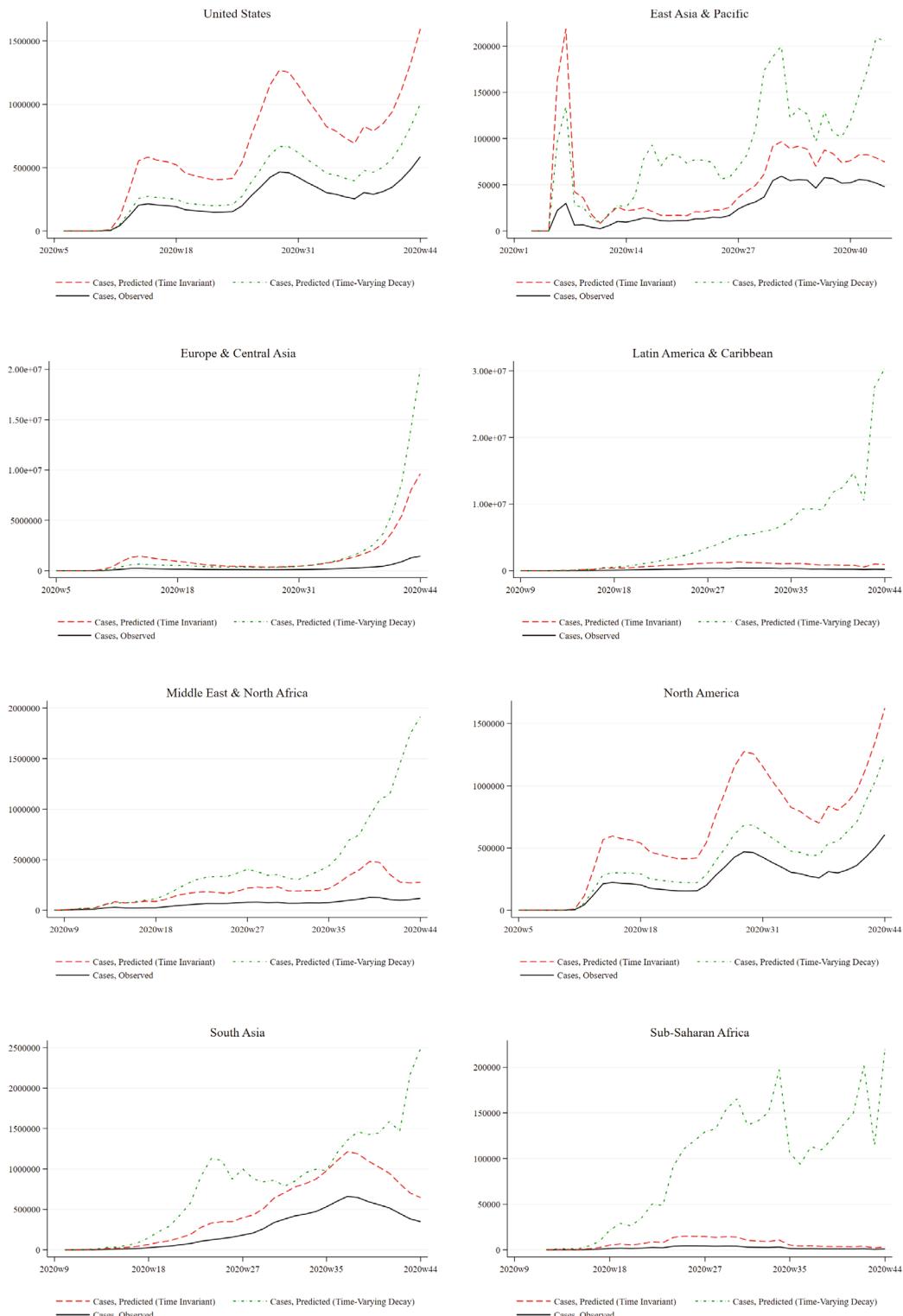
Figure 1 shows a positive association between testing and cases since early April in both the time-varying decay and time invariant specifications as in the previous models. If anything, the specifications allowing for dependence find a stronger association between testing and the number of cases as the marginal effects are strictly positive regardless of the units in which the testing data are reported. That said, there is little qualitative difference relative to the prior results obtained assuming independent underreporting. For deaths there is also little difference compared to the prior results assuming independent underreporting. Thus, we continue to find a negative association of large magnitude between testing and deaths (when tests are measured in clearly defined units).

Figure 2 shows the marginal effects of NPIs on cases and deaths. Here, the results do differ from the prior specifications assuming independent underreporting. Specifically, we find a stronger, negative association between NPIs and both cases and deaths when allowing for dependence in underreporting. In addition, while the association between NPIs and cases remains positive in the early weeks of the pandemic (as in the prior specifications), the association with deaths is now negative over the entire course of the pandemic to date. Overall, then, the results suggest a stronger, negative association between stringent NPIs and deaths once underreporting is allowed to be temporally dependent.

Turning to the issue of underreporting for cases, we estimate a MF of 9.0 (3.5) in the time-varying decay (time invariant) specifications. These MFs are notably higher than in the specifications assuming independent underreporting. The MFs for the US, however, are roughly similar to the independent error setting. That said, these higher MFs are more in line with alternative estimates presented in Flaxman et al. (2020), Hortaçsu et al. (2021), Li et al. (2020), Reese et al. (2021), and Sutton et al. (2020). For deaths we also find higher levels of underreporting than in our previous specifications (although still lower than for cases). We estimate a MF of 3.0 (1.9) in the time-varying decay (time invariant) specifications. These figures continue to be in line with the figures discussed in Section 2, ranging from 1.4 to 2.6.

To visualize the estimated levels of underreporting, Figures 5 and 6 show graphically the observed and estimated weekly cases and deaths, respectively, for different geographic regions. As before, these figures are obtained directly from the model estimates in Table 3. A few noteworthy findings emerge. First, while the time-varying decay model predicts more underreporting overall, there is variation across regions. For example, Figure 5 shows that the time invariant specification yields greater underreporting for the US and North America, but not in the other regions. Figure 6 shows that the time-varying decay specification yields greater underreporting in Europe and Central Asia, Latin America and the Caribbean, Middle East and North Africa, and Sub-Saharan Africa. There is little difference between the two specifications for the remaining regions, with the possible exception of South Asia. For this region, the two specifications yield similar results until late summer, after which the time-varying decay model predicts greater underreporting.

Second, the time-varying decay model indicates that underreporting for both cases and deaths has increased over time. For both outcomes, the estimate of  $\eta$  is negative and statistically significant ( $p < 0.01$ ). For cases,  $\hat{\eta} = -0.027$  (standard error = 0.003), while  $\hat{\eta} = -0.029$  (standard error = 0.003) for deaths. This is particularly noticeable for Latin America and the Caribbean, South Asia, East Asia and the Pacific, and Middle East and North Africa. As stated previously, this is consistent with the notion that contact tracing may become exceedingly more difficult as the infection spreads (Avery et al., 2020). Finally, both specifications indicate that Europe



**FIGURE 5** Comparison of observed and predicted weekly cases under the assumption of dependent underreporting [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

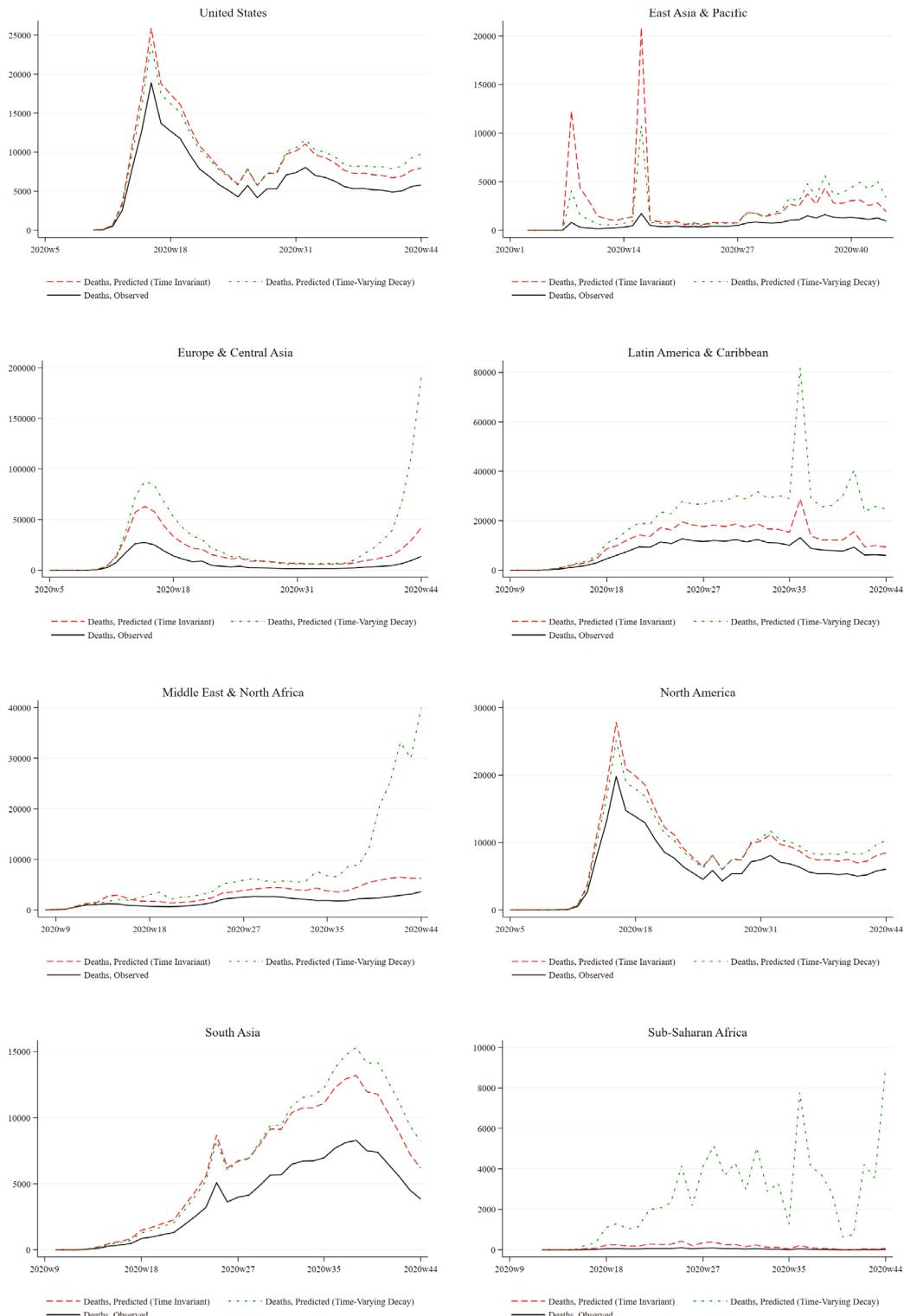


FIGURE 6 Comparison of observed and predicted weekly deaths under the assumption of dependent underreporting [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

and Central Asia are embroiled in a substantial second wave. While both specifications paint a similar picture for the US in terms of cases, there is not a corresponding rise in deaths in the US.

## 4.2 | Alternative stochastic frontier models

To assess the sensitivity of our results to modelling choices, we estimate two additional specifications. First, we augment the set of controls to address possible spillovers from neighbouring countries. Such cross-sectional dependence is typically viewed as important when modelling COVID-19 at the subnational level (e.g. Olmo & Sanso-Navarro, 2021). Specifically, we add controls for the (equally weighted) average value of population, population density, median age, the per cent of the population aged 65 and above, diabetes prevalence, per capita gross domestic product, the corruption perception index and the index of the stringency of NPIs computed over all other countries located in the same region (see Section 3.3 and Table 1). In addition, we add quadratic terms for these spatially lagged covariates, as well as an interaction between spatially lagged NPIs and a linear time trend.

With this expanded set of controls, we re-estimate the models assuming independent underreporting, distributed either Half Normal or Exponential, as well as the time-varying decay models allowing for dependence. The results are available upon request; however, we note that the time invariant stochastic frontier models failed to converge. For cases, the augmented models yield slightly lower MFs of 2.5 (1.8) when the underreporting is assumed to be distributed Half Normal (Exponential). However, the augmented time-varying decay model yields a significantly higher MF of 63. For deaths, the MFs in the augmented models when the underreporting is assumed to be distributed Half Normal (Exponential) are virtually unchanged. The augmented time-varying decay model again yields a much higher MF of 8.6. Thus, allowing for spillovers in the determinants of cases and deaths suggests, if anything, greater levels of underreporting.

Second, the specifications estimated to this point have assumed that the covariates directly affect the true number of cases and deaths, while the extent of underreporting is a random draw from a particular distribution that is identical across countries. Here, we relax this assumption and allow the underreporting to be drawn from a Truncated Normal distribution with a country- and time-specific mean. Formally,  $u_{it}^q \stackrel{\text{iid}}{\sim} N^+(\mu_{it}^q, \sigma_{u^q}^2)$ ,  $q \in \{Y, D\}$ . We then model the pre-truncation mean,  $\mu_{it}^q$ , as a function of the corruption perception index and testing controls that had previously been included in the set of covariates. Note, while  $u_{it}^q$  are i.i.d draws, serial correlation in the pre-truncation mean,  $\mu_{it}^q$ , implies that underreporting will be serially correlated.

Again, the results suggest similar levels of underreporting (results are available upon request). For cases, the model yields an MF of 3.3, between the prior estimates in Tables 2 and 3. For deaths, the MF is 4.7, higher than in the previous results. Thus, allowing for the degree of underreporting to depend directly on corruption and testing has little effect on the estimated underreporting of cases, but suggests greater levels of underreporting of deaths.

## 4.3 | Country-specific multiplication factors, infection rates, case fatality rates, and infectious fatality rates

The model estimates can also be used to assess heterogeneity in pandemic experiences across countries. To start, Table 4 displays country-specific MFs for cases and deaths from each

TABLE 4 Country-specific multiplication factors

Country	Cases				Deaths			
	Half-normal		Time-exponential	Time-varying	Half-normal		Time-exponential	Time-varying
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Afghanistan	2.846	1.731	106.986	1.749	2.214	1.583	2.856	1.664
Algeria	5.188	2.498	8.397	3.039	3.409	1.938	4.208	7.599
Australia	2.029	1.584	1.704	1.360	1.579	1.351	1.218	1.218
Austria	2.642	1.787	6.025	4.220	2.128	1.549	5.061	2.252
Azerbaijan	2.843	1.914	1.550	1.414	3.376	1.941	2.988	6.371
Bahrain	3.336	2.163	1.439	2.091	2.492	1.687	1.897	1.616
Belarus	2.465	1.785	1.963	1.514	2.913	1.800	4.504	4.285
Belgium	1.944	1.585	1.460	2.169	1.682	1.397	2.967	1.310
Brazil	2.070	1.640	2.116	1.259	2.072	1.545	2.175	1.286
Cambodia	7.006	3.293	71.152	7.550				
Canada	4.247	2.311	6.843	1.414	2.292	1.608	1.600	2.114
China	3.001	1.817	6.473	7.332	1.589	1.357	7.377	15.410
Croatia	2.228	1.680	28.549	6.668	1.943	1.487	1.472	3.060
Czech Republic	2.295	1.713	1.807	5.685	1.993	1.489	3.733	4.942
Denmark	4.496	2.371	9.436	5.131	2.296	1.605	11.659	3.936
Dominican Republic	2.865	1.958	1.451	1.374	2.754	1.761	2.408	1.538
Ecuador	4.841	2.470	99.026	12.068	2.334	1.615	11.456	3.814
Egypt	2.968	1.735	12.432	1.273	1.866	1.464	3.589	1.249
Estonia	3.106	1.937	10.947	2.549	2.337	1.687	4.930	2.964
Finland	3.441	2.081	4.544	2.939	1.908	1.484	7.726	2.549
France	3.492	2.143	7.926	9.219	2.005	1.512	6.777	3.563
Georgia	2.090	1.590	3.451	5.880	1.618	1.376	1.404	1.343
Germany	4.235	2.221	12.718	6.969	2.049	1.530	5.730	1.981
Greece	2.760	1.867	1.467	1.783	2.120	1.545	1.249	1.343
Iceland	2.857	1.867	7.098	2.997				
India	2.854	2.008	1.569	1.826	2.669	1.739	1.744	1.592
Indonesia	1.787	1.555	1.215	1.166	1.709	1.411	1.176	1.149
Iran	2.424	1.822	4.025	1.511	1.777	1.440	1.218	1.219
Iraq	2.974	2.081	4.454	1.568	2.000	1.518	1.283	1.299
Ireland	2.810	1.864	3.188	2.047	1.774	1.430	2.604	1.468
Israel	4.304	2.471	1.471	7.191	3.175	1.880	12.046	10.611
Italy	4.340	2.342	21.337	11.324	1.900	1.477	1.909	1.620

(Continues)

TABLE 4 (Continued)

Country	Cases				Deaths			
	Half-normal	Exponential	Time-varying	Time-invariant	Half-normal	Exponential	Time-varying	Time-invariant
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Japan	2.387	1.753	1.478	1.411	2.491	1.666	1.854	1.511
Kuwait	2.554	1.843	1.719	1.263	2.094	1.548	2.565	1.232
Lebanon	1.807	1.538	1.710	1.300	2.351	1.640	2.338	3.206
Lithuania	1.982	1.564	1.521	1.249	1.952	1.495	1.646	1.325
Luxembourg	2.260	1.672	1.789	1.270	1.804	1.461	2.460	1.313
Malaysia	2.525	1.714	12.951	2.279	2.221	1.588	10.283	4.926
Mexico	8.117	3.164	179.189	16.781	2.269	1.612	2.849	1.776
Nepal	1.859	1.573	1.482	1.295	2.163	1.572	2.703	1.429
Netherlands	1.981	1.587	1.657	1.221	1.773	1.431	1.288	1.177
New Zealand	2.295	1.626	2.341	1.603				
Nigeria	5.353	2.629	49.543	3.531	2.876	1.794	76.622	4.349
Norway	10.164	3.848	176.037	19.166	3.008	1.843	22.919	17.060
Oman	3.164	1.987	3.424	3.009	2.172	1.577	1.649	1.755
Pakistan	4.583	2.332	31.946	4.445	2.244	1.592	2.030	2.275
Philippines	2.408	1.822	1.272	1.685	2.365	1.636	5.529	4.348
Qatar	4.250	2.266	8.001	5.039	2.840	1.779	32.713	6.550
Romania	2.973	1.958	1.250	3.263	2.088	1.540	1.072	1.295
Russia	3.186	2.063	5.842	4.088	2.945	1.818	3.207	5.706
Singapore	2.760	1.852	23.178	1.552				
South Korea	2.836	1.807	8.942	4.411	2.043	1.526	1.473	2.037
Spain	2.136	1.662	1.100	1.148	1.709	1.412	1.215	1.310
Sri Lanka	3.973	2.127	21.326	2.180				
Sweden	2.427	1.790	2.203	1.220	1.928	1.490	4.071	1.679
Switzerland	1.869	1.525	4.780	2.375	1.771	1.429	5.032	1.470
Thailand	4.043	2.141	8.783	5.232	3.698	2.160	5.896	11.612
United Arab Emirates	5.320	2.589	55.226	8.187	3.263	1.893	307.806	6.642
United Kingdom	3.959	2.264	14.934	11.819	1.795	1.434	10.593	2.713
United States	2.602	1.924	1.483	2.721	2.096	1.552	1.387	1.372
Vietnam	5.275	2.823	4.220	11.792	2.494	1.754	38.478	7.990

Multiplication factor is defined as the total number of estimated cases or deaths divided by the total number of reported cases or deaths.

specification over the full sample period. A few key findings stand out. First, rankings of data accuracy across models are highly correlated. The Spearman rank correlations for cases (deaths) all exceed 0.50 (0.43). The MFs from the time-varying decay model are the most divergent from the other specifications; the Spearman rank correlations among the other three specifications all exceed 0.67 (0.64) for cases (deaths). Second, reporting accuracy for the US is roughly middle of the pack. In each of the eight models (two outcomes times four specifications) at least 27 countries have higher MFs than the US. Third, the time-varying decay model produces the greatest heterogeneity in MFs across countries, with some countries having very high MFs.

Figures A1 and A2 in the Appendix plot the weekly MFs for cases and deaths, respectively, for the US and the regions of the world. For cases, the MFs are quite high in the very early stages of the pandemic, as one might expect, for all regions except East Asia and the Pacific and Sub-Saharan Africa. However, the MFs quickly fall and remain relatively constant over the remainder of the sample period. For deaths, a similar pattern emerges, but it is less extreme and less common. Specifically, the MFs are no longer significantly higher early in pandemic for Europe and Central Asia.

Next, we tabulate cumulative infection rates for the countries in our sample. Column 1 in Table 5 shows the observed cumulative infection rate by country, defined as the cumulative number of cases divided by population. The remaining columns report estimated infection rates, defined as the cumulative estimated number of cases divided by population. Note, the ratio of the estimated infection rate to the observed infection rate yields the MFs reported in Table 4. The estimated infection rates for the US and Italy lie within the bounds reported in Manski and Molinari (2021) for New York State, Illinois, and Italy. For the US, our estimated cumulative infection rate ranges from 4.2% to 7.6%, in contrast to the observed rate of 2.8%. Again, this is consistent with the broad findings in the literature that cases in the US are severely undercounted. We also point out that any individual estimate should not be relied on too heavily as for instance both Ecuador and Mexico have predicted cumulative infection rates over 100%. Although not realistic, our estimation procedure does not restrict total cases to be less than the total population. To incorporate such a restriction into the estimation would require perhaps directly modelling the cumulative infection rate and combining a stochastic frontier model with a fractional logit/probit setup. This is beyond the scope of the current paper. Alternatively, one could predict cases in these instances using an appropriate prediction interval (Simar & Wilson, 2010).

Finally, Figure 7 plots the case fatality rate (CFR) and the infectious fatality rate (IFR) over time for the US and the regions of the world. We define the CFR as the number of observed deaths 1-week ahead divided by this week's observed cases. The IFR is defined as the number of estimated deaths 1-week ahead divided by this week's estimated cases. Before turning to the figure, we note that over our entire sample, the CFR is 2.6%, while the IFR ranges from 0.9% to 2.1%, with the 0.9% coming from the time-varying decay model which follows from the fact that this model produces the highest estimated caseload. The fact that our estimates of the IFR are below the CFR results from our estimate of more missing cases than deaths. Manski and Molinari (2021) reach a similar conclusion; however, they assume that deaths are reported correctly. Regardless, one important finding from this is that the COVID-19 virus is less lethal than what is commonly reported in the news media (based on the CFR).

Turning to the figure, we see that the CFR and IFRs peak very early in the pandemic for the US, and for most of the world, outside of later spikes in Sub-Saharan Africa and East Asia and the Pacific. Since the late summer, there is little difference between the CFR and the estimated IFRs for Europe and East Asia, North America, and the Middle East and North Africa. The steady decline of the CFR/IFRs is likely due to increased experience handling the virus, improvements

TABLE 5 Observed and predicted country-specific cumulative infection rates

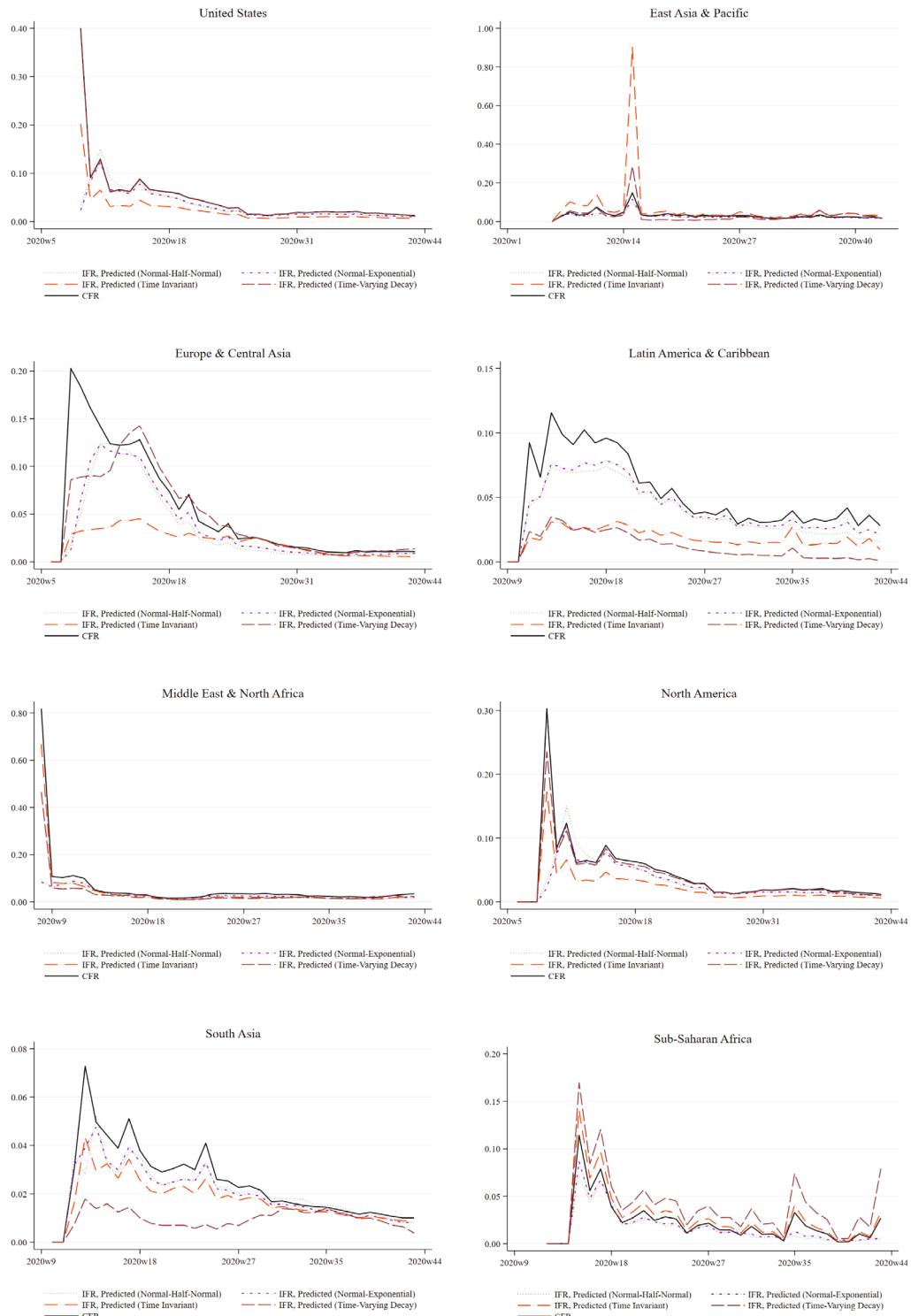
Country	Observed (1)	Predicted			Time varying (5)
		Half- normal (2)	Exponential (3)	Time invariant (4)	
Afghanistan	0.107	0.305	0.186	0.187	11.467
Algeria	0.134	0.693	0.334	0.406	1.122
Australia	0.108	0.220	0.171	0.147	0.184
Austria	1.269	3.352	2.267	5.355	7.645
Azerbaijan	0.563	1.599	1.076	0.795	0.872
Bahrain	4.827	16.104	10.441	10.095	6.946
Belarus	1.063	2.619	1.896	1.609	2.085
Belgium	3.846	7.476	6.098	8.342	5.615
Brazil	2.613	5.410	4.285	3.291	5.529
Cambodia	0.002	0.012	0.006	0.013	0.123
Canada	0.637	2.704	1.471	0.900	4.356
China	0.005	0.015	0.009	0.036	0.032
Croatia	1.283	2.858	2.156	8.553	36.621
Czech Republic	3.277	7.520	5.614	18.629	5.921
Denmark	0.833	3.744	1.975	4.273	7.859
Dominican Republic	1.176	3.370	2.303	1.616	1.707
Ecuador	1.100	5.325	2.717	13.275	108.925
Egypt	0.105	0.313	0.183	0.134	1.311
Estonia	0.380	1.181	0.737	0.969	4.163
Finland	0.296	1.018	0.616	0.870	1.345
France	2.249	7.852	4.818	20.733	17.825
Georgia	1.067	2.230	1.697	6.275	3.684
Germany	0.669	2.833	1.485	4.661	8.506
Greece	0.404	1.114	0.754	0.720	0.592
Iceland	1.445	4.128	2.698	4.331	10.257
India	0.599	1.710	1.203	1.094	0.940
Indonesia	0.152	0.271	0.236	0.177	0.184
Iran	0.749	1.814	1.364	1.132	3.014
Iraq	1.190	3.540	2.477	1.866	5.301
Ireland	1.271	3.571	2.369	2.601	4.051
Israel	3.651	15.712	9.022	26.250	5.372
Italy	1.210	5.253	2.835	13.708	25.829

(Continues)

TABLE 5 (Continued)

Country	Observed (1)	Predicted			
		Half-normal (2)	Exponential (3)	Time invariant (4)	Time varying (5)
Japan	0.081	0.193	0.142	0.114	0.120
Kuwait	2.981	7.612	5.494	3.765	5.124
Lebanon	1.226	2.216	1.886	1.594	2.097
Lithuania	0.616	1.221	0.963	0.769	0.937
Luxembourg	3.493	7.895	5.841	4.435	6.250
Malaysia	0.103	0.260	0.176	0.235	1.334
Mexico	0.724	5.875	2.290	12.145	129.688
Nepal	0.606	1.126	0.953	0.784	0.898
Netherlands	2.146	4.250	3.405	2.620	3.556
New Zealand	0.033	0.076	0.054	0.053	0.078
Nigeria	0.031	0.164	0.080	0.108	1.515
Norway	0.381	3.869	1.465	7.295	67.002
Oman	2.274	7.197	4.520	6.843	7.787
Pakistan	0.152	0.698	0.355	0.677	4.863
Philippines	0.352	0.847	0.641	0.593	0.447
Qatar	4.613	19.608	10.455	23.246	36.910
Romania	1.303	3.875	2.552	4.252	1.628
Russia	1.134	3.613	2.339	4.636	6.625
Singapore	0.991	2.737	1.836	1.539	22.980
South Korea	0.052	0.148	0.094	0.231	0.467
Spain	2.658	5.679	4.417	3.051	2.924
Sri Lanka	0.053	0.210	0.113	0.115	1.129
Sweden	1.231	2.988	2.204	1.503	2.712
Switzerland	2.029	3.791	3.094	4.818	9.697
Thailand	0.005	0.022	0.012	0.028	0.048
United Arab Emirates	1.366	7.268	3.538	11.185	75.452
United Kingdom	1.552	6.146	3.515	18.347	23.184
United States	2.807	7.305	5.400	7.639	4.163
Vietnam	0.001	0.006	0.003	0.014	0.005

Infection rate is defined as the total number of cases—observed in column (1) or predicted in columns (2)–(5)—divided by population (multiplied by 100).



**FIGURE 7** Comparison of the case fatality rate and infectious fatality rates obtained under different underreporting assumptions. Case fatality rate is observed deaths one-week ahead divided by observed cases. Infectious fatality rate is predicted death one-week ahead divided by predicted cases [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

in therapeutics and heightened public awareness. Finally, it is noteworthy that there is apparent convergence between the CFR and the IFRs as we enter the later stages of the pandemic. The one exception is the Latin American and Caribbean region; the CFR and IFRs have yet to converge. Thus, for this region, the lethality of the disease appears to be much lower in magnitude than suggested by the CFR even in Fall 2020. We did investigate whether this lack of convergence in the CFR and IFRs is driven by Brazil. Omitting Brazil from the analysis does not change this finding; if anything, the divergence at the end of the sample period is even greater.

#### 4.4 | Country- and time-specific infection rates, fatality rates and recovery rates

We use the model estimates in one final manner. As discussed in Section 3.1, the reproduction number,  $R_{0it}$ , is equal to  $\beta_{it}/(\gamma_{1it} + \gamma_{2it})$ , where  $\beta_{it}$  is the infection rate,  $\gamma_{1it}$  is the recovery rate, and  $\gamma_{2it}$  is the fatality rate. Using Equations (11) and (16) and assuming that each has a log-Normal distribution,  $\beta_{it}$  and  $\gamma_{2it}$  are consistently estimated by

$$\hat{\beta}_{it} = \exp\left(X\hat{\beta} + \frac{1}{2}\sigma_{v^\beta}^2\right) \quad (25)$$

$$\hat{\gamma}_{2it} = \exp\left(X\hat{\gamma} + \frac{1}{2}\sigma_{v^\gamma}^2\right), \quad (26)$$

where  $\sigma_{v^\beta}^2$  and  $\sigma_{v^\gamma}^2$  are the variances of  $v^\beta$  and  $v^\gamma$ , respectively. In practice,  $\sigma_{v^\beta}^2$  and  $\sigma_{v^\gamma}^2$  are not identified. Instead, we use  $\sigma_{v^\beta}^2$  and  $\sigma_{v^\gamma}^2$  (see Equations 12 and 17).

These are plotted in Figures A3 and A4 for select countries. For the US, the infection rate has increased at the start of the pandemic, slowed and perhaps even dipped during the summer, and recently has begun to increase again. Estimates of the infection rate at the end of our sample range from roughly 0.7% to 1.3%. Other countries shown also have seen an increase in the infection rate this fall with the exception of South Korea. The fatality rates, on the other hand, have declined since the summer in the majority of the countries shown; Brazil, and to a lesser extent Italy, are the exceptions.

As we do not have data on recoveries from COVID-19, we cannot directly estimate  $\gamma_{1it}$  and, hence, cannot estimate  $R_{0it}$ . However, to provide some information on the recovery rate we use available estimates for  $R_{0it}$ , combined with our estimates of  $\beta_{it}$  and  $\gamma_{2it}$ , to obtain estimates of  $\gamma_{1it}$ .<sup>28</sup> Specifically, estimates of  $\gamma_{1it}$  are given by

$$\hat{\gamma}_{1it} = \frac{\beta_{it} - \gamma_{2it}R_{0it}}{R_{0it}}. \quad (27)$$

As the estimates of  $R_{0it}$  do not account for underreporting, the results should be viewed cautiously.

These are plotted in Figures A5. For the US, the recovery rate increased from the spring to the summer and has remained relatively constant since. Estimates of the infection rate at the end of our sample range from roughly 0.6% to 1.1%. The other countries examined also show similar rises in the recovery rate entering Fall 2020. South Korea is the lone exception. However, this is consistent with the low infection rate in South Korea (see Figure A3) that also shows no signs of increasing entering Fall 2020.

<sup>28</sup>Estimates are provided at <http://metrics.covid19-analysis.org/>. They are obtained using the EpiEstim method implemented using the EpiEstim R package.

## 5 | CONCLUSION

Not since the Spanish Flu of 1918-1919 has the world faced a public health crisis of the magnitude of COVID-19. Meeting this challenge is made more difficult due to the lack of complete information. Understanding the science behind the SARS-CoV-2 virus and the COVID-19 disease and devising optimal policy responses requires accurate data. Unfortunately, missing data are rampant. The reasons for this are numerous and diverse, ranging from the apparently high rate of asymptomatic cases to inadequate surveillance systems to political corruption.

In this paper, we devise a method to overcome this missing data problem. A ‘structural’ model based on the SIR framework extended to allow for underreporting yields estimating equations for cases and deaths that conveniently resembles the classic stochastic frontier model. This connection has not been made to our knowledge and represents a tractable approach to incorporate missing data into the standard SIR model.

Admittedly, the stochastic frontier model relies on strong distributional assumptions. It is crucial to note that while these assumptions affect the magnitude of our conclusions, they do not change our fundamental conclusions. Above all else, this is our main takeaway. Furthermore, as shown in Badunenko et al. (2012), the setting we are operating in is perhaps the most robust to discrepancies in the distributional assumptions: a large degree of underreporting relative to the stochastic noise present overall. Moreover, the assumptions of the stochastic frontier model are orthogonal to the strong assumptions made in other studies seeking to overcome missing data to assess the pandemic. Thus, we view our approach as adding to the totality of our knowledge. Seen in this light, we find it reassuring that our range of predictions align with the limited existing research seeking to understand both the extent of the COVID-19 pandemic, as well as the negative association between NPIs and testing and COVID-19 severity.

Specifically, our approach yields a multiplication factor for the true number of cases between two and nine; a multiplication factor of the true number of deaths between 1.6 and 3.0. Our approach yields an infectious fatality rate ranging from 0.9% to 2.1%, in contrast to a case fatality rate of 2.6%. The fact that our estimates of the infectious fatality rate are below the case fatality rate results from our estimate of more missing cases than deaths.

Our analysis here suggests a number of avenues for future research that should prove fruitful in modelling pandemics with missing data. First, as noted, our estimated cumulative infection rates exceed 100% for some countries in our sample in some specifications. Adjusting the functional form to constrain the number of infections by the total population of the country (or some fraction thereof), while not immediately obvious, would be valuable. While this might entail placing bounds on the degree of undercounting, making the model more similar to existing approaches and less similar to a standard stochastic frontier setting, this could potentially be done using the doubled-bounded efficiency estimator in Almanidis et al. (2014) or a Tobit-type estimator which truncates the degree of undercounting based on the population of the country. Or, as mentioned previously, directly modelling the cumulative infection rate, rather than new cases, and constraining the rate to be in the unit interval as in fractional logit/probit models.

Second, our estimation procedure has focused on cases and deaths separately. Joint estimation might exploit information in cases to improve estimation of actual deaths. However, doing so risks mis-specification in one equation leading to bias in the other. Finally, future research might consider modelling underreporting of cases and deaths as a deterministic function of a set of observable covariates, thereby reducing reliance on distributional assumptions. If the covariates that determine the quality of the data—such as testing and reporting protocols, inclusivity of

surveillance systems, rigor of contact tracing, dissemination of information, etc.—are measurable, such a model may be informative.

We conclude by re-iterating our message at the outset: no model addressing missing data is beyond reproach as each must rely on assumptions, not all of which are testable. While the stochastic frontier SIR model we propose requires assumptions on the nature of underreporting of cases and deaths, so too do other methods that have recently received critical attention. Rather than relying on any single model, and hence one particular set of assumptions, scientists and policymakers ought to aggregate information across a wide range of methodologies. Stochastic frontier modelling should be part of this set.

## ACKNOWLEDGEMENTS

All data and Stata code used in this paper are available upon request. We are grateful for comments from James Carpenter, Hao Dong, Alecos Papadopoulos and three anonymous referees.

## ORCID

*Daniel L. Millimet*  <https://orcid.org/0000-0001-5178-5993>

*Christopher F. Parmeter*  <https://orcid.org/0000-0001-6123-0107>

## REFERENCES

- Aigner, D.J., Lovell, C.A.K. & Schmidt, P. (1977) Formulation and estimation of stochastic frontier production functions. *Journal of Econometrics*, 6(1), 21–37.
- Almanidis, P., Qian, J. & Sickles, R.C. (2014) Stochastic frontier models with bounded inefficiency. In: Sickles, R.C. & Horrace, W.C. (Eds.) *Festschrift in honor of Peter Schmidt econometric methods and applications*. New York: Springer, pp. 47–82.
- Arroyo-Marioli, F., Bullano, F., Kučinskas, S. & Rondón-Moreno, C. (2020) Tracking R of COVID-19: A new real-time estimation using the Kalman filter. medRxiv. Available from: <https://www.medrxiv.org/content/early/2020/04/29/2020.04.15.20066431>
- Askitas, N., Tatsiramos, K. & Verheyden, B. (2020), Lockdown strategies, mobility patterns and COVID-19. IZA Discussion Paper No. 13293.
- Atkeson, A. (2020) How deadly is COVID-19? understanding the difficulties with estimation of its fatality rate. NBER Working Paper No. 26965.
- Avery, C., Bossert, W., Clark, A., Ellison, G. & Fisher Ellison, S. (2020) An economist's guide to epidemiology models of infectious disease. *Journal of Economic Perspectives*, 34, 79–104.
- Azzalini, A. (1985) A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12(2), 171–178.
- Azzalini, A. (2014) *The skew-normal and related families*. Cambridge: Cambridge University Press.
- Backhaus, A. (2020) Common pitfalls in the interpretation of COVID-19 data and statistics. *Intereconomics*, 55, 162–166.
- Badman, R.P., Wu, Y., Inukai, K. & Akaishi, R. (2021) Blessing or curse of democracy? Current evidence from the Covid-19 pandemic. *arXiv*. Available from: <https://arxiv.org/abs/2105.10865v1>
- Badunenko, O., Henderson, D.J. & Kumbhakar, S.C. (2012) When, where and how to perform efficiency analysis. *Journal of the Royal Statistical Society, Series A*, 175(4), 863–892.
- Barnes, S.R., Beland, L.-R., Huh, J., Kim, D., Barnes, S.R., Beland, L.-P., Huh, J. & Kim, D. (2022) COVID-19 lockdown and traffic accidents: Lessons from the pandemic. *Contemporary Economic Policy*, 40(2), 349–368. Available from: <https://doi.org/10.1111/coep.12562>
- Battese, G.E. & Coelli, T.J. (1992) Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of Productivity Analysis*, 3, 153–169.
- Bellemare, M.F. & Wichman, C.J. (2020) Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, 82, 50–61.

- Cerqua, A., Di Stefano, R., Letta, M. & Miccoli, S. (2021) Local mortality estimates during the COVID-19 pandemic in Italy. *Journal of Population Economics*, Forthcoming.
- Chernozhukov, V., Kasahara, H. & Schrimpf, P. (2021) Causal impact of masks, policies, behavior on early COVID-19 pandemic in the U.S. *Journal of Econometrics*, 220, 23–62.
- Chudik, A., Pesaran, M.H. & Rebucci, A. (2020) Voluntary and mandatory social distancing: Evidence on Covid-19 exposure rates from Chinese provinces and select countries. NBER Working paper 27039.
- Dang, H.-A. & Trinh, T.-A. (2020) The beneficial impacts of covid-19 lockdowns on air pollution: Evidence from Vietnam. IZA Discussion Paper No. 13651.
- Depalo, D. (2021) True COVID-19 mortality rates from administrative data. *Journal of Population Economics*, 34, 253–274.
- Dubrow, J.K. (2021) Local data and upstream reporting as sources of error in the administrative data undercount of Covid 19 *International Journal of Social Research Methodology*, Forthcoming.
- Fetzer, T., Graeber, T., Fetzer, T. & Graeber, T. (2021) Measuring the scientific effectiveness of contact tracing: Evidence from a natural experiment. *Proceedings of the National Academy of Sciences*, 118(33). Available from: <https://doi.org/10.1073/pnas.2100814118>
- Flaxman, S., Mishra, S., Gandy, A., Unwin, H.J.T., Mellan, T.A., Coupland, H. et al. (2020) Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 584, 257–261.
- Gibbons, C.L., Mangen, M.-J.J., Plass, D., Havelaar, A.H., Brooke, R.J., Kramarz, P. et al. (2014) Measuring under-reporting and under-ascertainment in infectious disease datasets: a comparison of methods. *BMC Public Health*, 14(1), 147–164.
- Goolsbee, A., Syverson, C., Goolsbee, A. & Syverson, C. (2021) Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of Public Economics*, 193, 104311. Available from: <https://doi.org/10.1016/j.jpubeco.2020.104311>
- Greer, S.L., King, E.J., Massard da Fonseca, E. & Peralta-Santos, A. (2020) The comparative politics of COVID-19: the need to understand government responses. *Global Public Health*, 15, 1413–1416.
- Hale, T., Webster, S., Petherick, A., Phillips, T. & Kira, B. (2020) Oxford COVID-19 Government Response Tracker. Blavatnik School of Government. Available from: <http://www.bsg.ox.ac.uk/covidtracker>
- Hortaçsu, A., Liu, J. & Schieg, T. (2021) Estimating the fraction of unreported infections in epidemics with a known epicenter: an application to COVID-19. *Journal of Econometrics*, 220, 106–129.
- Jondrow, J., Lovell, C.A.K., Materov, I.S. & Schmidt, P. (1982) On the estimation of technical efficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19(2/3), 233–238.
- Kermack, W.O. & McKendrick, A.G. (1927) A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A*, 115, 700–721.
- Korolev, I. (2021) Identification and estimation of the SEIRD epidemic model for COVID-19. *Journal of Econometrics*, 220, 63–85.
- Kumbhakar, S.C., Amundsveen, R., Kvile, H.M. & Lien, G. (2015) Scale economies, technical change and efficiency in Norwegian electricity distribution, 1998–2010. *Journal of Productivity Analysis*, 43(3), 295–305.
- Li, S. & Linton, O. (2021) When will the COVID-19 pandemic peak. *Journal of Econometrics*, 130–157.
- Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W. et al. (2020) Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, 368(6490), 489–493.
- Manski, C.F. & Molinari, F. (2021) Estimating the COVID-19 infection rate: Anatomy of an inference problem. *Journal of Econometrics*, 220, 181–192.
- McCulloh, I., Kiernan, K. & Kent, T. (2020) Inferring true COVID19 infection rates from deaths. *Frontiers in Big Data*, 3, 1–10.
- Meeusen, W. & van den Broeck, J. (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18(2), 435–444.
- Millimet, D.L. & Parmeter, C.F. (2022) Accounting for skewed or one-sided measurement error in the dependent variable. *Political Analysis*, 30(1), 66–88. Available from: <https://doi.org/10.1017/pan.2020.45>
- Murray, E. J. (2020) Epidemiology's time of need: COVID-19 calls for epidemic-related economics. *Journal of Economic Perspectives*, 34, 105–120.
- Oguzoglu, U. (2020) Covid-19 lockdowns and decline in traffic related deaths and injuries. IZA Discussion Paper No. 13278.

- Olmo, J. & Sanso-Navarro, M. (2021) Modeling the spread of COVID-19 in New York City. *Papers in Regional Science*, Forthcoming.
- Papadopoulos, A. & Parmeter, C.F. (2021) Type II failure and specification testing in the stochastic frontier model. *European Journal of Operational Research*, 293, 990–1001.
- Parmeter, C.F. & Kumbhakar, S.C. (2014) Efficiency analysis: a primer on recent advances. *Foundations and Trends in Econometrics*, 7(3-4), 191–385.
- Pitt, M.M. & Lee, L.-F. (1981) The measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics*, 9(1), 43–64.
- Reed, C., Angulo, F.J., Swerdlow, D.L., Lipsitch, M., Meltzer, M.I., Jernigan, D. et al. (2009) Estimates of the prevalence of pandemic (H1N1) 2009, United States, April–July 2009. *Emerging Infectious Diseases*, 15(12), 2004–2007.
- Reese, H., Iuliano, A.D., Patel, N.N., Garg, S., Kim, L., Silk, B.J. et al. (2021) Estimated incidence of COVID-19 illness and hospitalization – United States, February – September, 2020. *Clinical Infectious Diseases*, 72, e1010–e1017.
- Sá, F. (2020) Socioeconomic determinants of COVID-19 infections and mortality: Evidence from England and Wales. IZA Policy Paper No. 159.
- Simar, L. & Wilson, P.W. (2010) Inferences from cross-sectional, stochastic frontier models. *Econometric Reviews*, 29(1), 62–98.
- Stock, J.H. (2020) Random testing is urgently needed. Available from: <http://www.igmchicago.org/covid-19-random-testing-is-urgently-needed/>
- Sutton, D., Fuchs, K., D'Alton, M. & Goffman, D. (2020) Universal screening for SARS-CoV-2 in women admitted for delivery. *New England Journal of Medicine*, 382(22), 2163–2164.
- Teixeira da Silva, J. & Tsigaris, P. (2020) Policy determinants of COVID-19 pandemic induced fatality rate across nations. *Public Health*, 187, 140–142.
- Weinberger, D., Cohen, T., Crawford, F., Mostashari, F., Olson, D., Pitzer, V.E. et al. (2020) Estimating the early death toll of COVID-19 in the United States. *medRxiv*. Available from: <https://www.medrxiv.org/content/early/2020/04/29/2020.04.15.20066431>
- Wu, S.L., Mertens, A.N., Crider, Y.S., Nguyen, A., Pokpongkiat, N.N., Djajadi, S. et al. (2020) Substantial underestimation of SARS-CoV-2 infection in the United States. *Nature Communications*, 11, 4507–4516.
- Goolsbee, A., Syverson, C., Goolsbee, A. & Syverson, C. (2021) Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of Public Economics*, 193, 104311. Available from: <https://doi.org/10.1016/j.jpubeco.2020.104311>

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Millimet, D.L. & Parmeter, C.F. (2022) COVID-19 severity: A new approach to quantifying global cases and deaths. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(3), 1178–1215. Available from: <https://doi.org/10.1111/rssa.12826>