

Delays in death reports and their implications for tracking the evolution of COVID-19¹

Emilio Gutierrez,² Adrian Rubli³ and Tiago Tavares⁴

Date submitted: 27 June 2020; Date accepted: 28 June 2020

Understanding the determinants and implications of delays in reporting COVID-19 deaths is important for managing the epidemic. Contrasting England and Mexico, we document that reporting delays in Mexico are larger on average, exhibit higher geographic heterogeneity, and are more responsive to the total number of occurred deaths in a given location-date. We then estimate simple SIR models for each country to illustrate the implications of not accounting for reporting delays. Our results highlight the fact that low and middle-income countries are likely to face additional challenges during the pandemic due to lower quality of real-time information.

Covid Economics 34, 3 July 2020: 116-144

1 The authors acknowledge support from the Asociación Mexicana de Cultura and the ITAM-COVID center. We thank Miguel Messmacher and participants at the ITAM Brown Bag seminar for their helpful comments. Gerardo Sánchez-Izquierdo provided outstanding research assistance. Code and data are available on a GitHub repository: https://github.com/tgstavares/revisions_data_epi. All errors are our own.

2 Instituto Tecnológico Autónomo de México (ITAM), Department of Economics.

3 ITAM, Department of Business Administration.

4 ITAM, Department of Economics and CIE.

Copyright: Emilio Gutierrez, Adrian Rubli and Tiago Tavares

1 Introduction

Tracking the spread of the SARs-CoV-2 virus and subsequently the evolution of the COVID-19 epidemic is important for managing the outbreak (Shea et al., 2020), evaluating the effectiveness of different policy tools to contain it (Kraemer et al., 2020; Chinazzi et al., 2020; Hartl et al., 2020), and for effectively communicating risks and undertaking different policy actions (WHO, 2005; Saliou, 1994), such as enforcing or lifting social distancing measures (WHO, 2020; Greenstone and Nigam, 2020). The effectiveness of surveillance systems is thus critical for the management of pandemics (Olson et al., 2020; Carey et al., 2020; Woolhouse et al., 2015; Brookmeyer, 1991; Krause, 1992). Low surveillance capacity represents not only a threat to the prompt identification of outbreaks in low and middle-income countries, but also for assessing their evolution comparatively across countries (Halliday et al., 2017).

Reporting delays for deaths, defined as the time difference between when a death occurs and when it is registered in the system, have been long recognized in various settings (AbouZahr et al., 2015; Bird, 2015). Nevertheless, in the context of COVID-19, many academics, policy-makers, and media outlets are tracking the pandemic's evolution within and across countries by focusing on death counts (Weinberger et al., 2020), arguing that they are more easily identified and consistently reported than cases (Roser et al., 2020). However, data on death counts may exhibit shortcomings similar to case counts due to reporting delays, and the extent of these issues may also vary across and within countries. This, in turn, may limit policy-makers' ability to effectively communicate the risks associated with individuals' behavior in the midst of the pandemic (Avery et al., 2020).

This paper seeks to characterize reporting delays of COVID-19 deaths across two distinct settings, contrasting the various determinants of these delays and illustrating the implications of these reporting delays for modeling the epidemic. We focus on England and Mexico to contrast delays between a developed and developing country, and since both governments publicly release detailed data that allow us to construct measures of reporting delays.

We document that death counts due to COVID-19 are reported with different delays in England and Mexico. Reporting delays in Mexico are larger, more heterogeneous across space, and most importantly, are more affected to the total number of actual occurred deaths in a particular location on a given day. We then illustrate the implications of not accounting for delays in reporting by

estimating simple SIR models for both countries accounting and not accounting for reporting delays, showing very different predictions for Mexico, consistent with the larger and more heterogeneous delays.

There is a rapidly growing literature touching on various topics related to the COVID-19 pandemic. In particular, our paper speaks directly to two strands of this work. First, recent papers explore how persuasive and/or informative messages correlate with or affect compliance with social distancing measures, which are important for both the economic costs associated with the pandemic and for the evolution of the epidemiological curve (Ajzenman et al., 2020; Allcott et al., 2020; Barrios and Hochberg, 2020; Bursztyn et al., 2020; Grossman et al., 2020; Kushner Gadarian et al., 2020; Painter and Qiu, 2020).

Second, a large literature is attempting to identify the additional restrictions and challenges that low and middle-income countries face in the management of and economic recovery from this pandemic, such as the capacity of the healthcare system, poverty, inequality, and corruption (Gallego et al., 2020; Gottlieb et al., 2020; Loayza, 2020; Monroy-Gómez-Franco, 2020; Ribeiro and Leist, 2020; Walker et al., 2020). By identifying a potential difference in information quality in a middle-income country, we shed light on an additional challenge that policy-makers may face when managing epidemics.

Our main contribution consists in contrasting reporting delays for deaths in two very different settings. We argue (and show supporting evidence in the online appendix) that the difference between Mexico and England in terms of reporting delays is consistent with lower state capacity. To the extent that this is a widespread problem across the developing world, our insights imply that successfully managing the epidemic in these regions will be further complicated by lower quality real-time information.

The rest of the paper is organized as follows. Section 2 presents the data and some descriptives of the evolution of deaths during the COVID-19 epidemic in England and Mexico, as well as the average delays in reporting these deaths. Section 3 decomposes delays into location shifters, date shifters, and the effect of total deaths. Section 4 then illustrates the implications of these reporting delays when modeling the evolution of the epidemic. Lastly, Section 5 concludes.

2 Data and Descriptive Evidence

2.1 Death reports in England and Mexico

We obtain publicly-available data with daily COVID-19 death counts from the England NHS, available from April 1 to June 7, 2020.¹ Each file contains the deaths from the corresponding reporting period, and indicates the date at which these reported deaths occurred. These counts are further disaggregated by NHS trust, which correspond to groups of hospitals and healthcare providers in the public system. Our data covers 198 NHS trusts across seven regions.² It should be noted that these counts do not include deaths that occur outside the hospital system.

For Mexico, we obtain publicly-available data from the Ministry of Health that tracks all patients that were ever suspected of having COVID-19 over time.³ The government uses this database to announce cumulative counts and deaths in a nightly press conference, allowing us to identify each of the newly reported cases and deaths on each date and the deaths' date of occurrence. We observe a few characteristics for each patient, including their municipality of residence. Unfortunately, we cannot observe which healthcare facility they visit and thus reports them to the database. We use these data to construct counts of deaths that are reported each day, as well as the actual date of death for each municipality. Our data includes 593 municipalities.⁴

2.2 Reporting delays in England and Mexico

We present the evolution of death counts over time for each country in Figure 1, distinguishing between aggregates based on the actual date of death and those based on the date of reporting. Figure 1a shows the daily number of total deaths in England by date, while Figure 1b shows the corresponding cumulative deaths. Figures 1c and 1d show analogous plots for Mexico. Over all, we observe that the evolution of death counts over time differs significantly between aggregates by

¹Data are available at <https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-daily-deaths/>. We were unable to find similar data for other countries in the United Kingdom.

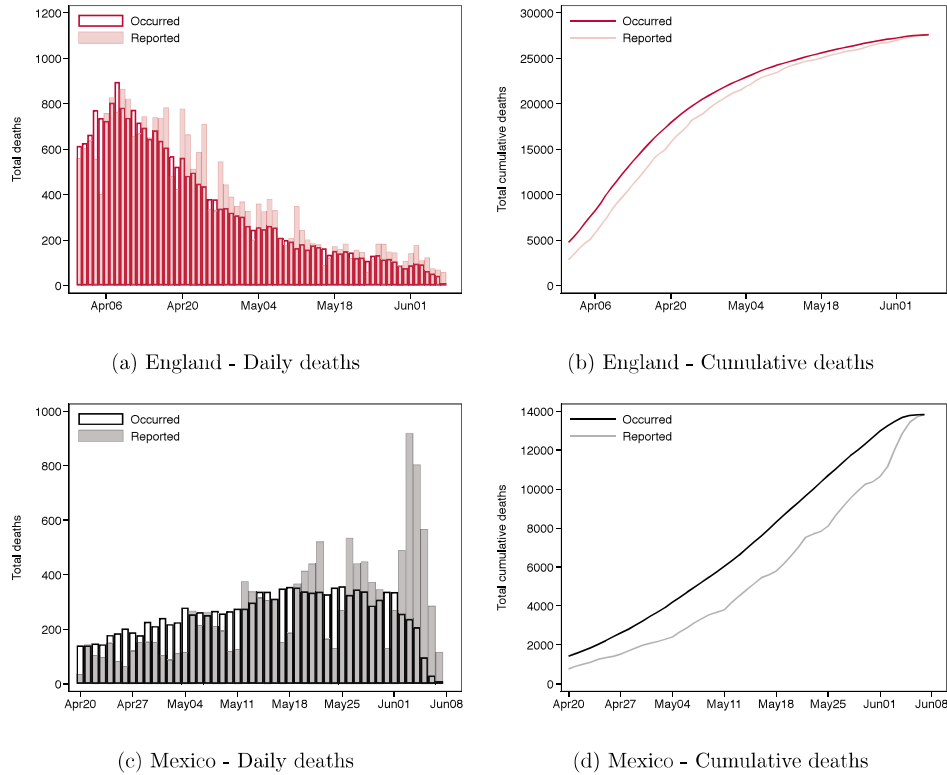
²These regions are East England, London, Midlands, North East and Yorkshire, North West, South East, and South West.

³Data are available at <https://www.gob.mx/salud/documentos/datos-abiertos-bases-historicas-direccion-general-de-epidemiologia>.

⁴Although Mexico has 2,448 municipalities in total, the remaining 1,855 are those that have not reported any COVID-19 cases. These are mostly rural areas.

date of occurrence and date of reporting, especially for Mexico, where delays also appear to be larger.

Figure 1:
COVID-19 Deaths Over Time by Country



Notes: These graphs show the distribution of deaths over time for each country using the full available data (starting April 1 for England and April 20, 2020 for Mexico). We distinguish between deaths that actually occurred on a given date, and deaths that were reported on that date. Figures 1a and 1c show counts of total deaths per day, while Figures 1b and 1d show total cumulative deaths for England and Mexico, respectively.

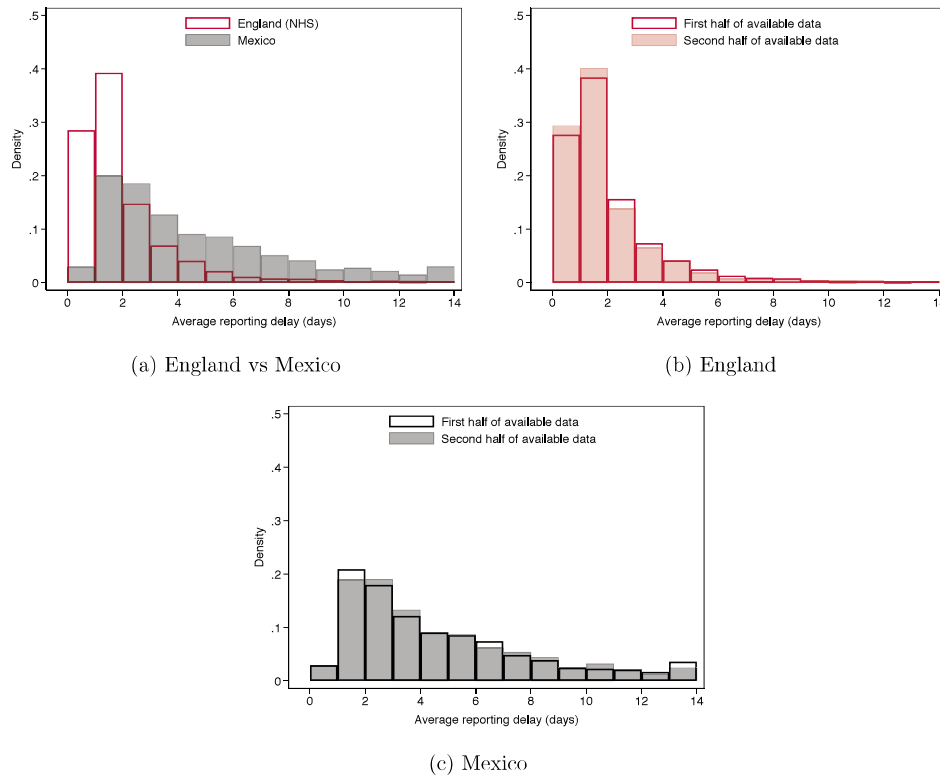
From these datasets, we construct measures of the delays with which deaths are reported by each location in each period, conditional on having at least one observed occurred death. Specifically, we compute the average delay with which deaths that occurred in each location-date pair were reported, conditional on having being reported within k days after their occurrence. Since the data are naturally censored, we drop all deaths reported in the last k days of available reports for each country, regardless of their date of occurrence. The implicit assumption is that the probability of

Covid Economics 34, 3 July 2020: 116-144

observing all deaths that occurred on a given date is one (or close to one) when relying on reports up to k days after that date. We focus on $k = 14$, but present results for $k = 21$ also.

We show the distribution of average reporting delays in Figure 2, where we have restricted the data by setting $k = 14$.⁵ We present a direct comparison between England and Mexico in Figure 2a, and then restrict to data for England in Figure 2b and Mexico in Figure 2c. These last two graphs further split the sample based on the median date of death.

Figure 2:
Average Reporting Delay Over Time by Country



Notes: These graphs show the distribution of the average reporting delay measured in days for each country. In Figures 2b and 2c, the data are further stratified by median date of death for the available span of data. We drop the most recent 14 days of data reports, and delays that are over 14 days. In Figure 2a, the mean for England is 1.74, and 4.29 for Mexico, implying a difference of 2.56, with a 95% CI [2.45,2.66]. The mean for the first half of the data for England in Figure 2b is 1.84, and 1.64 for the second half, implying a difference of 0.20, 95% CI [0.10,0.29]. Lastly, the mean for the first half of the data in Figure 2c is 4.31, 4.28 for the second half, and a difference of 0.03, 95% CI [-0.19,0.25].

⁵See Figure A1 in the online appendix for analogous graphs with $k = 21$.

Figure 2a documents that the average reporting delay is larger in Mexico, with a distribution that is not only mean-shifted but also has a heavier right tail. Figures 2b and 2c further show that the average delay in England decreased over time, while it did not for Mexico. Given the differential timing of the epidemic, one must be cautious when interpreting these differences. Nevertheless, these plots do suggest that trends in delays are not the same between England and Mexico.

There are many potential reasons for why Mexico, a middle-income country, has significantly larger reporting delays than a developed country like England. Although we cannot fully discard alternative explanations, we posit that larger delays are correlated with state capacity, which is lower in Mexico. We present suggestive evidence consistent with this explanation in Figure A2 in the online appendix. We show that the municipalities in Mexico with larger reporting delays are those that have fewer healthcare units per capita, slightly fewer medical staff per capita, and higher patient volumes per healthcare unit. Furthermore, municipalities where a larger share of the population is covered by the public healthcare system, which would indicate a larger presence of the state, are those with significantly lower reporting delays. Over all, this suggests that state capacity plays a key role in decreasing reporting delays.

3 Determinants of reporting delays

There are at least three different determinants of reporting delays for deaths that matter for tracking the evolution of a pandemic. First, there may be spatial differences in delays. Different reporting units may face different staffing and infrastructure constraints that can lead to variation in their reporting capabilities. If this is the case, as the disease spreads geographically, the average delay with which deaths are reported may change, affecting the shape of the curve of reported deaths independently of the pandemic.

Second, there may be system-wide changes in delays over time. If countries improve their reporting systems in real time as the pandemic progresses, then death reports could misrepresent the evolution of the pandemic in different ways over time.

Lastly, there may be decreasing returns in reporting. Hence, as more deaths occur, it may be less likely that these deaths are reported in a timely manner. This too would imply a different shape for the curve from reported deaths when making comparisons.

The data and statistical tools needed to account for the reporting delays implied by these three factors are different and difficult to develop in real time as the pandemic progresses. For geographic differences, large amounts of data are necessary for each location to correctly model location-specific delays and correct for them. Time-specific delays and delays related to the total deaths imply that the corrections should be updated over time. While correcting reports for these factors requires data from which estimates of delays may be inferred, correcting for delays due to changes in death counts also requires information on the total deaths that *actually* occurred in a given moment in addition to those that were reported.

3.1 Framework

In order to characterize and illustrate the differences in the determinants of the delays in death reports in Mexico and England, we assume that reports are a series of Bernoulli trials, so that the number of days it takes for a death to be reported by location l in period t (plus one) follows a geometric distribution with success probability p_{lt} . We further assume that p_{lt} can be parametrized as follows:

$$p_{lt} = p_0 \cdot \exp\left(\sum_{q=1}^Q \alpha_q \mathbb{I}(\text{deaths}_{lt}=q) + \pi_l + \xi_t + \varepsilon_{lt}\right)$$

simply stating that there is a baseline probability p_0 that deaths are reported, which may then be shifted by certain variables as outlined below. This implies then that:

$$\mathbb{E}(\text{delay}_{lt} + 1) = \frac{1}{p_0 \cdot \exp\left(\sum_{q=1}^Q \alpha_q \mathbb{I}(\text{deaths}_{lt}=q) + \pi_l + \xi_t + \varepsilon_{lt}\right)}$$

where we have only used the fact that the mean of a geometric distribution with parameter p is equal to $\frac{1-p}{p}$.

We proceed by log-linearizing this expression. This allows us to decompose the reporting delays into location-specific shifters, period-specific shifters, and parameters that measure the response to the number of occurred deaths through the following ordinary least squares regression:

$$\ln(\overline{\text{delay}}_{lt} + 1) = \sum_{q=1}^Q \alpha_q \mathbb{I}(\text{deaths}_{lt} = q) + \pi_l + \xi_t + \varepsilon_{lt} \quad (1)$$

where $\overline{\text{delay}}_{lt}$ denotes the average delay in reporting deaths that occurred in location l on date t , $\mathbb{I}(\text{deaths}_{lt} = q)$ are indicators equal to one if total occurred deaths in a given location and time fall in a category q , the coefficients α_q measure how log delays respond to total deaths that occurred in a particular location-date, π_l denotes reporting unit fixed effects corresponding to the location-specific shifters (hence, the estimates of these parameters $\hat{\pi}_l$ for each location l recover the estimates of $p_0 + p_l$), ξ_t are the date-specific shifters, and lastly ε_{lt} is an error term that captures any time-varying location-specific shocks to average delays other than the total occurred deaths. We run regressions separately for each country, cluster our standard errors by reporting unit to allow for serial correlation in the error term, and present our results graphically.

3.2 Results

Figure 3 plots the estimated coefficients $\hat{\alpha}_q$ from estimating equation 1 for each country, with Figure 3a using reported data up to 14 days to identify total occurred deaths ($k = 14$) and Figure 3b considering 21 days instead ($k = 21$).⁶ The vertical bars indicate 95% confidence intervals. We use integers of total deaths as our categories, with the last category considering 5 or more deaths.⁷ We take one death as the excluded category, so that our estimated effects are relative to the average delay for reporting units with one death.

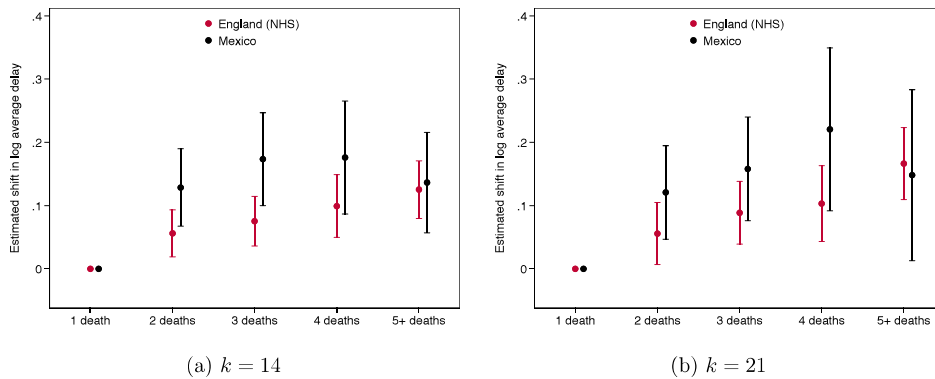
Our point estimates are larger for Mexico across specifications. In Figure 3a, we interpret this to mean that, accounting for location and date effects, the occurrence of two deaths significantly increases the average delay by 0.056 log points in England, which can be approximated as 5.6%, and by 0.129 log points in Mexico on average, or around 12.9%, relative to when there is only one death. For five or more deaths, we find that average delays significantly increase by 0.125 and 0.136 log points for England and Mexico, respectively, relative to the average delay when there is one death only. Importantly, the estimates of how changes in the death toll affect delays are calculated from variation within each reporting unit over time, accounting for system-wide trends in delays. Hence, our results in Figure 3 are not conflating occurred deaths with the general progression or regional variation of death counts and delays.

⁶Table A1 in the online appendix shows the corresponding point estimates.

⁷Figure A3 in the online appendix presents similar results for Mexico using deciles of deaths per capita, by matching population at the municipality level from the 2010 census. We were unable to find a consistent mapping between population and NHS trusts.

Over all, these results show that total death counts matter for average delays, regardless of any spatial and temporal differences, and that they matter much more in Mexico than in England. This relationship is in line with reporting units becoming overwhelmed by higher death tolls, which is exacerbated by settings with low state capacity. Alternatively, a higher death toll may increase the likelihood that at least some deaths require further testing and scrutiny before being reported, which would lead to larger average delays.

Figure 3:
Relationship between Total Deaths and Reporting Delays



Notes: These graphs show the estimated shift in the log average delay in relation to quartiles of total deaths per reporting unit from estimating equation 1. All effects are calculated relative to the mean shift for the first quartile (one death). Figure 3a corresponds to data that exclude the last 14 days of available reports, as well as delays over 14 days (N=5991 for England, N=3660 for Mexico). Figure 3b uses 21 days instead (N=5531 for England, N=2875 for Mexico). The vertical lines indicate 95% confidence intervals from robust standard errors clustered by reporting unit.

Figure 4 shows the estimates of the reporting unit fixed effects. Each coefficient indicates the average shift in delays relative to the excluded location, net of general time trends and accounting for variation in total deaths.⁸ Since there is no commonality in reporting units between England and Mexico, we normalize the median reporting unit for each country in terms of its estimate to zero and order them by size to allow comparability. Figure 4a considers $k = 14$ and 4b $k = 21$.

The results show that the predicted shift in log average delays is more homogenous for England than Mexico, as noted by the larger slope in the coefficients for Mexico. Given that for Mexico each estimate is obtained from a smaller number of observations,⁹ it should not be surprising that these

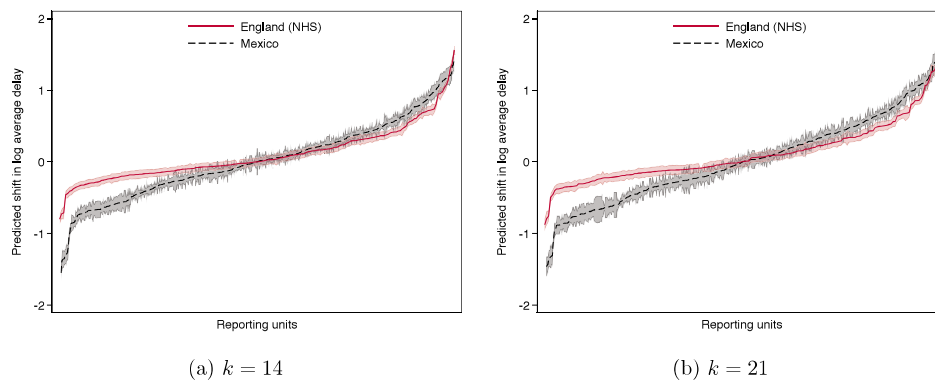
⁸We arbitrarily exclude East Coast Community Healthcare, Beccles Hospital for England, and Aguascalientes municipality in the state of Aguascalientes for Mexico.

⁹For example, for $k = 14$, each reporting unit has on average 27.5 and 6.2 days with occurred deaths in England and Mexico, respectively.

estimates exhibit a higher variance. However, the confidence intervals for Mexico and England do not overlap for a large share of the estimates in both tails of the distribution. This indicates that, indeed, geographic variation in delays is significantly higher in Mexico.

We argue that the larger spatial heterogeneity in Mexico is related to the lower state capacity of this middle-income country, relative to England. Figure A4 in the online appendix shows the correlation between the estimated municipality shifters and various measures of healthcare infrastructure. These plots show that the locations with larger reporting delays, net of time effects and the actual death toll, are those with fewer healthcare units per capita, higher patient loads per healthcare unit, and a smaller share of the population covered by the public healthcare system. Hence, this is consistent with state capacity playing an important role in delays.

Figure 4:
Relationship between Reporting Units and Reporting Delays

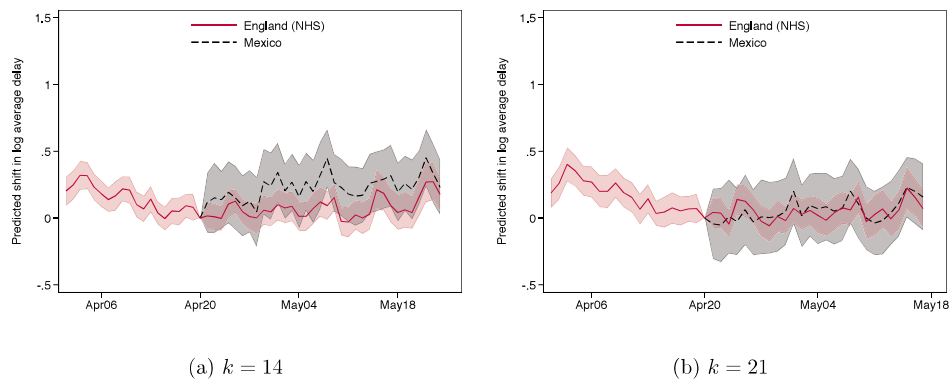


Notes: These graphs show the predicted shift in the log average delay across reporting units from estimating equation 1. The median shift in each series has been normalized to zero to allow comparability. Figure 4a corresponds to data that exclude the last 14 days of available reports, as well as delays over 14 days ($N=5991$ for England, $N=3660$ for Mexico). Figure 4b uses 21 days instead ($N=5531$ for England, $N=2875$ for Mexico). The shaded areas indicate 95% confidence intervals from robust standard errors clustered by reporting unit. Reporting units correspond to 198 NHS trusts in England and 749 municipalities in Mexico in Figure 4a, and 198 trusts and 668 municipalities in Figure 4b.

Finally, we show our estimates of the date effects in Figure 5. The excluded category here is the first calendar day with available data for both countries, April 20. Once again, each plot considers alternative values of k . The point estimates show a decreasing effect over time for England up to April 20. This suggests that NHS trusts improved their reporting over time. For the dates in which we observe data for both countries, the point estimates are mostly flat for both England

and Mexico.¹⁰ Over all, the estimates of the date fixed effects show that England decreased its delays over time, and that, contrary to what the raw data may suggest, there is not much of a relationship between average delays and time for Mexico, once we account for location effects and occurred deaths.

Figure 5:
Relationship between Date and Reporting Delays



Notes: These graphs show the predicted shift in the log average delay over time from estimating equation 1. All date estimates are calculated relative to the mean shift on April 20. Figure 5a corresponds to data that exclude the last 14 days of available reports, as well as delays over 14 days (N=5991 for England, N=3660 for Mexico). Figure 5b uses 21 days instead (N=5531 for England, N=2875 for Mexico). The shaded areas indicate 95% confidence intervals from robust standard errors clustered by reporting unit. Our sample includes 54 days for England and 35 for Mexico in Figure 5a, and 47 days for England and 28 for Mexico in Figure 5b.

Taken together, the results in Figures 3 - 5 suggest that correcting for reporting delays may not be a straight-forward task, since detailed information by location is needed and a single correction factor is unlikely to capture the heterogeneity we document. Moreover, tracking the evolution of COVID-19 from death reports may deliver a biased representation of the epidemic curve that is not comparable across locations.

4 Implications for epidemiological modeling

The results presented in section 3 suggest important challenges for the development of algorithms that can systematically correct for reporting delays, given the heterogeneity we document. We now emphasize the importance of taking delays into account by highlighting how they may affect model

¹⁰Note that there is perhaps a slight increasing trend in the point estimates for Mexico, although the evidence is not strong. This trend is most obvious in Figure 5a.

estimates that are commonly used to support predictions and policy interventions (Zhang et al., 2020; Dehning et al., 2020).

We proceed by contrasting estimates based on reported deaths relative to two alternative ways of counting deaths. First, we consider actual occurred deaths, as reported up to k days after the fact. Second, we consider a hypothetical scenario in which reporting delays are fourteen days shorter than observed, by taking the difference between cumulative occurred deaths at time t and at time $t - 1$ as reported at time $t + 14$ and $t + 13$, respectively. It is important to stress that our objective is not to provide accurate forecasts about the dynamics of the infection in England and Mexico. Instead, we simply illustrate how reporting delays directly impact short-run analyses in modeling the evolution of the COVID-19 pandemic.

To this end, we use a simple homogeneous mixing agent version of the SIR model (Kermack and McKendrick, 1927).¹¹ SIR models are relatively tractable and remain an important tool in epidemiological analysis (Hethcote, 2000), including the current epidemic (Verity et al., 2020; Fox et al., 2020; Kucharski et al., 2020; Giordano et al., 2020). In our model, we allow for time-dependent frequency of contacts to capture the effect of individual behavioral changes or containment policies (Maier and Brockmann, 2020), as this improves fit (Fernández-Villaverde and Jones, 2020). We then evaluate how the main predictions change when the model is estimated using different death counts as outlined above.

4.1 Model setup

Our model is based on Fernández-Villaverde and Jones (2020).¹² The model considers an initial uniform population of mass P_0 .¹³ Assume the time period to be one day. After the outbreak of the epidemic, each individual of the population can be in either one of the following five states at date t : susceptible S_t , infected I_t , resolving R_t , recovered C_t , or dead D_t . Since the dead individuals are

¹¹In the classic SIR model, population is compartmentalized into three states: susceptible, infected, and recovered (also called resistant or removed).

¹²We model behavioral responses exogenously, as in Fernández-Villaverde and Jones (2020). Other recent studies have attempted to endogenize behavior in an optimizing environment where adjustment occurs directly at the contact level (Greenwood et al., 2019; Dasaratha, 2020), or indirectly through decisions of consumption and production (Eichenbaum et al., 2020; Krueger et al., 2020; Acemoglu et al., 2020). For simplicity, we abstract from endogenizing behavior, although our model could be extended accordingly.

¹³Acemoglu et al. (2020) considers heterogeneity of the population across age groups. It would be straight-forward to extend our model in this way as well.

removed from the population, at every point in time the total population P_t equals:

$$P_t = S_t + R_t + I_t + C_t$$

We assume that an infection occurs when a susceptible person enters in contact with an infected person at a rate of $\beta_t I_t / P_t$, where β_t represents the number of random contacts a susceptible person has with the rest of the population and is assumed to change with t to capture behavioral responses to the disease.

The system evolves according to:

$$S_{t+1} = S_t - \beta_t S_t I_t / P_t + C_t$$

$$I_{t+1} = I_t - \beta_t S_t I_t / P_t - \gamma I_t$$

$$R_{t+1} = R_t - \gamma I_t - \theta R_t$$

$$D_{t+1} = D_t + \delta \theta R_t$$

$$C_{t+1} = C_t + (1 - \delta) \theta R_t$$

where the parameter γ captures the rate at which infected agents start recovering and cease to be infectious, and θ is the rate at which recovering agents resolve the disease, where a fraction δ dies while the remaining $(1 - \delta)$ recovers and acquires immunity. The epidemic begins with an initial (exogenous) mass of infections I_0 . Once contagion starts spreading, we allow for time-dependent frequency of contacts, either due to individual behavior changes or public containment policies, according to:

$$\beta_t = \beta^{final} + (\beta^{init} - \beta^{final})e^{-\lambda t}$$

where β^{init} is the initial contact rate across agents that converges at a period rate of λ to a β^{final} rate of contact. Note that the model implies a basic reproduction number of $R^{init} = \beta^{init} / \gamma$ when $t = 0$, which converges to $R^{final} = \beta^{final} / \gamma$ as $t \rightarrow \infty$.

The model is then characterized by seven parameters: $\{\gamma, \theta, \delta, \lambda, \beta^{init}, \beta^{final}, I_0\}$. We take the first three parameters from epidemiological estimates in the literature and estimate the remaining four parameters which we denote by $\Phi \equiv \{\lambda, \beta^{init}, \beta^{final}, I_0\}$. We take $\gamma = 0.2$, meaning individuals

are infectious on average for 5 days; $\theta = 0.1$, meaning that it takes on average 15 days for the infection to resolve; and $\delta = 0.008$, which is the case fatality rate (Bar-On et al., 2020).¹⁴ We fix P_0 to be the initial population of each country and assume that $C_0 = D_0 = R_0 = 0$, that is, at the start of the outbreak there are no fully recovered, dead, or recovering agents. Period $t = 0$ is assumed to be the first day when the official total case count of SARs-CoV-2 reaches more than 150 individuals.

Estimates for Φ are then generated by solving

$$\hat{\Phi} = \arg \min_{\Phi} \left\{ \frac{1}{T} \sum_{t=\bar{t}}^{\bar{t}} (\log D_t - \log D_t(\Phi))^2 \right\} \quad (2)$$

using a global minimizer where the death series $D_t(\Phi)$ is generated by solving numerically the system of equations outlined above for a parameter choice of Φ .

We define D_t^k as the deaths that occurred at time t and were reported up to k days afterward, in line with the data setup used above. With a slight abuse of notation, we define D_t^0 as the deaths that were reported on date t , regardless of when they occurred. Hence, since $D_t^0 \neq D_t^k, \forall k > 0$ due to reporting delays, the model estimates $\hat{\Phi}$ may change whether one uses deaths by reporting date D_t^0 (to $\hat{\Phi}^0$) or deaths by date of occurrence D_t^k (to $\hat{\Phi}^k$). We additionally present estimates of the model in a hypothetical scenario in which delays are fourteen days shorter than observed, by taking the difference between cumulative occurred deaths at time t and at time $t - 1$ as reported at time $t + 14$ and $t + 13$, respectively.

4.2 Model estimates

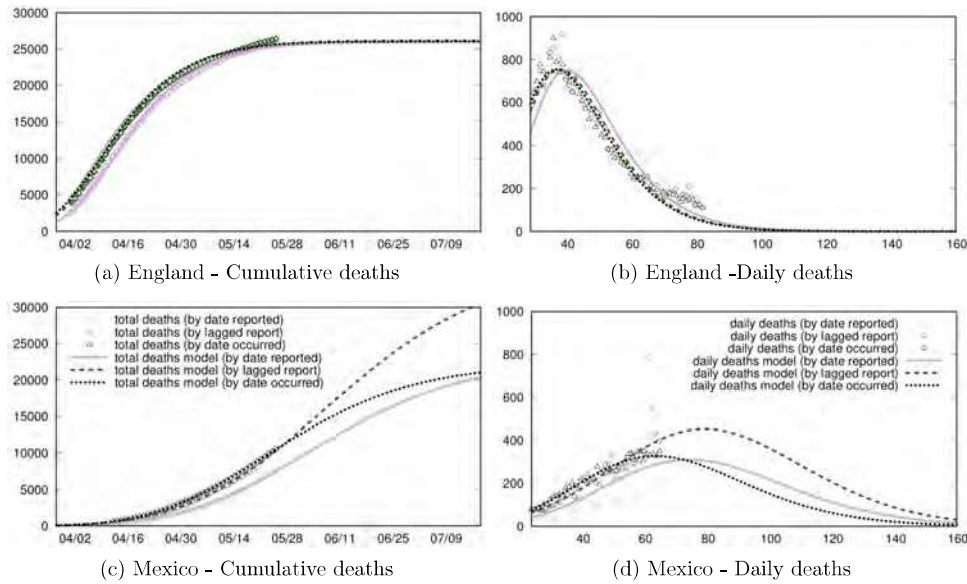
We estimate the model by solving equation 2 using the data for England and Mexico, considering $k = 14$ as before. The dynamics corresponding to the estimation results of this procedure are shown graphically in Figure 6, with the parameter estimates presented in Table 1.

We find that, for England, the differences in predictions when using deaths when reported, when reported with a shorter delay, and when they actually occurred are small, consistent with

¹⁴This estimate is close to evidence presented in a recent sero-epidemiological national survey undertaken by the Spanish government from April 27 to May 11 to measure the incidence of SARs-CoV-2 in Spain. The report is available at https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/ENECOVID_Informe_preliminar_cierre_primera_ronda_13Mayo2020.pdf.

the smaller average delays documented above. In contrast, the differences are striking for Mexico, where we previously found large average delays and considerable heterogeneity. Early estimates for Mexico based on reported deaths would predict a total number of deaths of about 20 thousand (after 120 days from the onset) with a peak in daily deaths of 310. However, if the delay were fourteen days shorter, these estimates change to around 31 thousand total deaths, with a peak of 453 daily deaths (occurring four days later), which implies an increase of about 50%. When using deaths by date of occurrence, the total number of deaths is not that different from the estimates from deaths by date reported, but the shape of the epidemic curve changes: the maximum number of cases is reached eleven days earlier.

Figure 6:
Model Predictions of Deaths Based on Reported vs Occurred Deaths



Notes: These graphs show model predictions of total (Figures 6a and 6b) and daily deaths (Figures 6c and 6d) as a result of estimating equation 2. Figures 6a and 6c correspond to England, while 6b and 6d are for Mexico. Each plot shows the model predictions from using reported deaths (not accounting for delays), the hypothetical scenario with delays that are 14 days shorter, and from using deaths by date of occurrence. To identify occurred deaths, the data exclude the last 14 days of available reports. Markers represent actual data, while lines show model predictions.

Table 1 further shows that the estimation results show a very reasonable fit of the model with respect to the data, with mean sum of square errors ranging from 0.023% to 0.031% for England, and 0.011% to 0.158% for Mexico. The superior fit when using revised data, particularly for Mexico,

accrues from the fact that daily new deaths become less lumpy when considering occurred deaths, as seen for instance in Figure 6d. Moreover, our estimates of the initial number of infected I_0 in Table 1 imply severe under-reporting of total SARs-CoV-2 cases, possibly reflecting a sluggish identification or testing of the first people who contracted the new disease (Li et al., 2020).

Table 1:
Estimates of SIR Model for England and Mexico Accounting vs Not Accounting for Reporting Delays

	I_0	R^{init}	R^{final}	λ	Mean Sum of Square Errors	Total deaths after 120 days	Maximum daily deaths	Days until maximum daily deaths
England								
By date reported	5.9	9.71	0.517	0.095	0.00031	25,985	750	40
By date occurred	340.0	5.95	0.467	0.083	0.00024	26,053	754	37
By date reported (14 days later)	194.3	6.25	0.467	0.082	0.00023	26,158	754	38
Mexico								
By date reported	4015.1	1.77	0.627	0.017	0.00158	20,361	310	75
By date occurred	8344.3	1.77	0.627	0.021	0.00011	21,024	328	64
By date reported (14 days later)	9622.9	1.57	0.470	0.010	0.00030	30,510	453	79

Notes: This table presents the model estimates for England and Mexico corresponding to estimating equation 2. We distinguish between model predictions that use deaths by the date on which they were reported, deaths as they would have been reported if delays were fourteen days shorter, and deaths by the date on which they occurred. The first four columns correspond to the choice parameters defined by Φ . The last three columns show predictions in terms of total deaths after 120 days and the peak of the predicted epidemiological curve (number of daily deaths and days until reached).

Lastly, our results also reveal the importance of taking into consideration behavioral or containment policies when modeling epidemics, as this affects the basic reproduction number. In our estimates for Mexico, for example, the initial magnitudes of this parameter R^{init} range from 1.6 to 1.8, implying an explosive behavior of contagion, but it then swiftly converges to magnitudes that are much smaller, with R^{final} ranging from 0.5 to 0.6.

Over all, our model estimates show that, particularly for Mexico, there are large differences between predictions based on reported deaths and those based on alternative death counts that incorporate the delays. These large differences seem relevant for authorities that use these predic-

Covid Economics 34, 3 July 2020: 116-144

tions to manage the epidemics, especially if we believe that information on model predictions and subsequent policies affect individual behavior, which in turn impacts the spread of the disease.

5 Conclusion

This paper analyzes delays in death reports in two distinct settings. Our data analysis for England and Mexico suggests ample heterogeneity in reporting delays, particularly for Mexico, which we argue is related to lower state capacity. This heterogeneity may then complicate applying systematic corrections to the data, since a single correcting factor is unlikely to adequately capture the full variability of delays. However, failing to accurately account for delays yields drastically different model predictions, which in turn may impact policy-making and policy implementation in undesirable ways.

Ignoring reporting delays and their determinants may lead to biased estimates of demand for healthcare, such as intensive care units (Kissler et al., 2020; Li et al., 2020). Additionally, not considering these delays may give a wrong perception of the severity of the disease to the general public, potentially reducing support for containment policies or a lower adoption rate of individual protective measures. It seems thus imperious that policy-makers recognize early on the role of reporting delays as well as understanding their determinants when formulating policy and communication strategies to fight epidemics.

References

- AbouZahr, C., D. De Savigny, L. Mikkelsen, P. W. Setel, R. Lozano, and A. D. Lopez (2015). Towards universal civil registration and vital statistics systems: the time is now. *The Lancet* 386(10001), 1407–1418.
- Acemoglu, D., V. Chernozhukov, I. Werning, and M. D. Whinston (2020). A multi-risk SIR model with optimally targeted lockdown. Technical report, National Bureau of Economic Research.
- Ajzenman, N., T. Cavalcanti, and D. Da Mata (2020). More than words: Leaders' speech and risky behavior during a pandemic. *Available at SSRN 3582908*.
- Allcott, H., L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Y. Yang (2020). Polarization and public health: Partisan differences in social distancing during the Coronavirus pandemic. *NBER Working Paper* (w26946).
- Avery, C., W. Bossert, A. Clark, G. Ellison, and S. F. Ellison (2020). Policy implications of models of the spread of coronavirus: Perspectives and opportunities for economists. Technical report, National Bureau of Economic Research.
- Bar-On, Y. M., A. Flamholz, R. Phillips, and R. Milo (2020). Science forum: SARS-CoV-2 (COVID-19) by the numbers. *Elife* 9, e57309.
- Barrios, J. M. and Y. Hochberg (2020). Risk perception through the lens of politics in the time of the covid-19 pandemic. Technical report, National Bureau of Economic Research.
- Bird, S. M. (2015). End late registration of fact-of-death in England and Wales. *The Lancet* 385(9980), 1830–1831.
- Brookmeyer, R. (1991). Reconstruction and future trends of the AIDS epidemic in the United States. *Science* 253(5015), 37–42.
- Bursztyjn, L., A. Rao, C. Roth, and D. Yanagizawa-Drott (2020). Misinformation during a pandemic. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2020-44).

- Carey, J. M., V. Chi, D. Flynn, B. Nyhan, and T. Zeitzoff (2020). The effects of corrective information about disease epidemics and outbreaks: Evidence from Zika and yellow fever in Brazil. *Science advances* 6(5), eaaw7449.
- Chinazzi, M., J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. P. y Piontti, K. Mu, L. Rossi, K. Sun, et al. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (Covid-19) outbreak. *Science* 368(6489), 395–400.
- Dasaratha, K. (2020). Virus dynamics with behavioral responses. *arXiv preprint arXiv:2004.14533*.
- Dehning, J., J. Zierenberg, F. P. Spitzner, M. Wibral, J. P. Neto, M. Wilczek, and V. Priesemann (2020). Inferring change points in the spread of Covid-19 reveals the effectiveness of interventions. *Science*.
- Eichenbaum, M. S., S. Rebelo, and M. Trabandt (2020). The macroeconomics of epidemics. Technical report, National Bureau of Economic Research.
- Fernández-Villaverde, J. and C. I. Jones (2020). Estimating and simulating a SIRD model of COVID-19 for many countries, states, and cities. Technical report, National Bureau of Economic Research.
- Fox, G. J., J. M. Trauer, and E. McBryde (2020). Modelling the impact of Covid-19 upon intensive care services in New South Wales. *Med J Aust*.
- Gallego, J. A., M. Prem, and J. F. Vargas (2020). Corruption in the times of pandemia. *Available at SSRN 3600572*.
- Giordano, G., F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, and M. Colaneri (2020). Modelling the Covid-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, 1–6.
- Gottlieb, C., J. Grobovšek, and M. Poschke (2020). Working from home across countries. *Covid Economics* 1(8), 71–91.
- Greenstone, M. and V. Nigam (2020). Does social distancing matter? *University of Chicago, Becker Friedman Institute for Economics Working Paper (2020-26)*.

- Greenwood, J., P. Kircher, C. Santos, and M. Tertilt (2019). An equilibrium model of the African HIV/AIDS epidemic. *Econometrica* 87(4), 1081–1113.
- Grossman, G., S. Kim, J. Rexer, and H. Thirumurthy (2020). Political partisanship influences behavioral responses to governors' recommendations for Covid-19 prevention in the United States. *Available at SSRN 3578695*.
- Halliday, J. E., K. Hampson, N. Hanley, T. Lembo, J. P. Sharp, D. T. Haydon, and S. Cleaveland (2017). Driving improvements in emerging disease surveillance through locally relevant capacity strengthening. *Science* 357(6347), 146–148.
- Hartl, T., K. Wälde, and E. Weber (2020). Measuring the impact of the German public shutdown on the spread of Covid19. *Covid economics, Vetted and real-time papers, CEPR press 1*, 25–32.
- Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM review* 42(4), 599–653.
- Kermack, W. O. and A. G. McKendrick (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* 115(772), 700–721.
- Kissler, S. M., C. Tedijanto, E. Goldstein, Y. H. Grad, and M. Lipsitch (2020). Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* 368(6493), 860–868.
- Kraemer, M. U., C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, L. Du Plessis, N. R. Faria, R. Li, W. P. Hanage, et al. (2020). The effect of human mobility and control measures on the Covid-19 epidemic in China. *Science* 368(6490), 493–497.
- Krause, R. M. (1992). The origin of plagues: old and new. *Science* 257(5073), 1073–1078.
- Krueger, D., H. Uhlig, and T. Xie (2020). Macroeconomic dynamics and reallocation in an epidemic. Technical report, National Bureau of Economic Research.
- Kucharski, A. J., T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R. M. Eggo, F. Sun, M. Jit, J. D. Munday, et al. (2020). Early dynamics of transmission and control of Covid-19: a mathematical modelling study. *The lancet infectious diseases*.

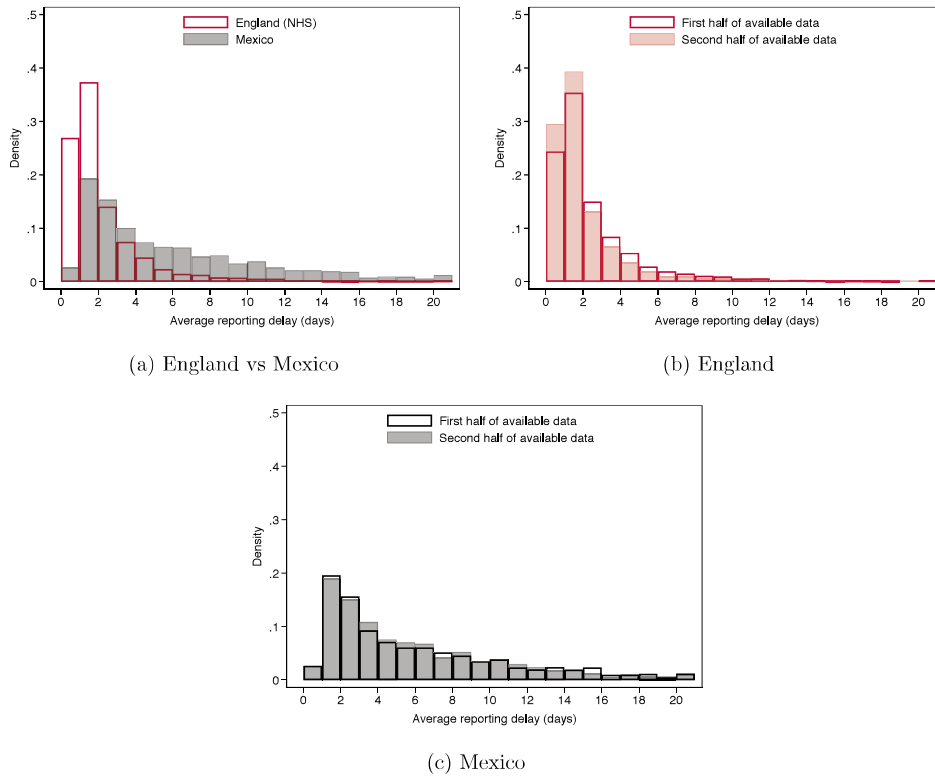
- Kushner Gadarian, S., S. W. Goodman, and T. B. Pepinsky (2020). Partisanship, health behavior, and policy attitudes in the early stages of the Covid-19 pandemic. *Health Behavior, and Policy Attitudes in the Early Stages of the COVID-19 Pandemic (March 27, 2020)*.
- Li, R., S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, and J. Shaman (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* 368(6490), 489–493.
- Li, R., C. Rivers, Q. Tan, M. B. Murray, E. Toner, and M. Lipsitch (2020). The demand for inpatient and ICU beds for Covid-19 in the US: lessons from Chinese cities. *MedRxiv*.
- Loayza, N. V. (2020). Costs and trade-offs in the fight against the Covid-19 pandemic: A developing country perspective.
- Maier, B. F. and D. Brockmann (2020). Effective containment explains subexponential growth in recent confirmed Covid-19 cases in China. *Science* 368(6492), 742–746.
- Monroy-Gómez-Franco, L. (2020). ¿ quién puede trabajar desde casa? Evidencia desde México.
- Olson, D. R., B. A. Lopman, K. J. Konty, R. W. Mathes, V. Papadouka, A. Ternier, J. R. Zucker, L. Simonsen, B. T. Grenfell, and V. E. Pitzer (2020). Surveillance data confirm multiyear predictions of rotavirus dynamics in New York City. *Science advances* 6(9), eaax0586.
- Painter, M. and T. Qiu (2020). Political beliefs affect compliance with covid-19 social distancing orders. *Available at SSRN 3569098*.
- Ribeiro, F. and A. Leist (2020). Who is going to pay the price of Covid-19? Reflections about an unequal Brazil. *International Journal for Equity in Health* 19, 1–3.
- Roser, M., H. Ritchie, E. Ortiz-Ospina, and J. Hasell (2020). Coronavirus Pandemic (COVID-19). *Our World in Data*.
- Saliou, P. (1994). Crisis communication in the event of a flu pandemic. *European journal of epidemiology* 10(4), 515–517.
- Shea, K., M. C. Runge, D. Pannell, W. J. Probert, S.-L. Li, M. Tildesley, and M. Ferrari (2020). Harnessing multiple models for outbreak management. *Science* 368(6491), 577–579.

- Verity, R., L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. Walker, H. Fu, et al. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet infectious diseases*.
- Walker, P. G., C. Whittaker, O. J. Watson, M. Baguelin, P. Winskill, A. Hamlet, B. A. Djafaara, Z. Cucunubá, D. O. Mesa, W. Green, et al. (2020). The impact of Covid-19 and strategies for mitigation and suppression in low-and middle-income countries. *Science*.
- Weinberger, D. M., T. Cohen, F. W. Crawford, F. Mostashari, D. Olson, V. E. Pitzer, N. G. Reich, M. Russi, L. Simonsen, A. Watkins, et al. (2020). Estimating the early death toll of Covid-19 in the United States. *bioRxiv*.
- WHO (2005). WHO outbreak communication guidelines. Technical report, World Health Organization.
- WHO (2020). Coronavirus disease 2019 (?Covid-19?): situation report, 88.
- Woolhouse, M. E., A. Rambaut, and P. Kellam (2015). Lessons from Ebola: Improving infectious disease surveillance to inform outbreak management. *Science translational medicine* 7(307), 307rv5–307rv5.
- Zhang, J., M. Litvinova, Y. Liang, Y. Wang, W. Wang, S. Zhao, Q. Wu, S. Merler, C. Viboud, A. Vespignani, et al. (2020). Changes in contact patterns shape the dynamics of the Covid-19 outbreak in China. *Science*.

Appendix for Online Publication

Supplementary Tables and Figures

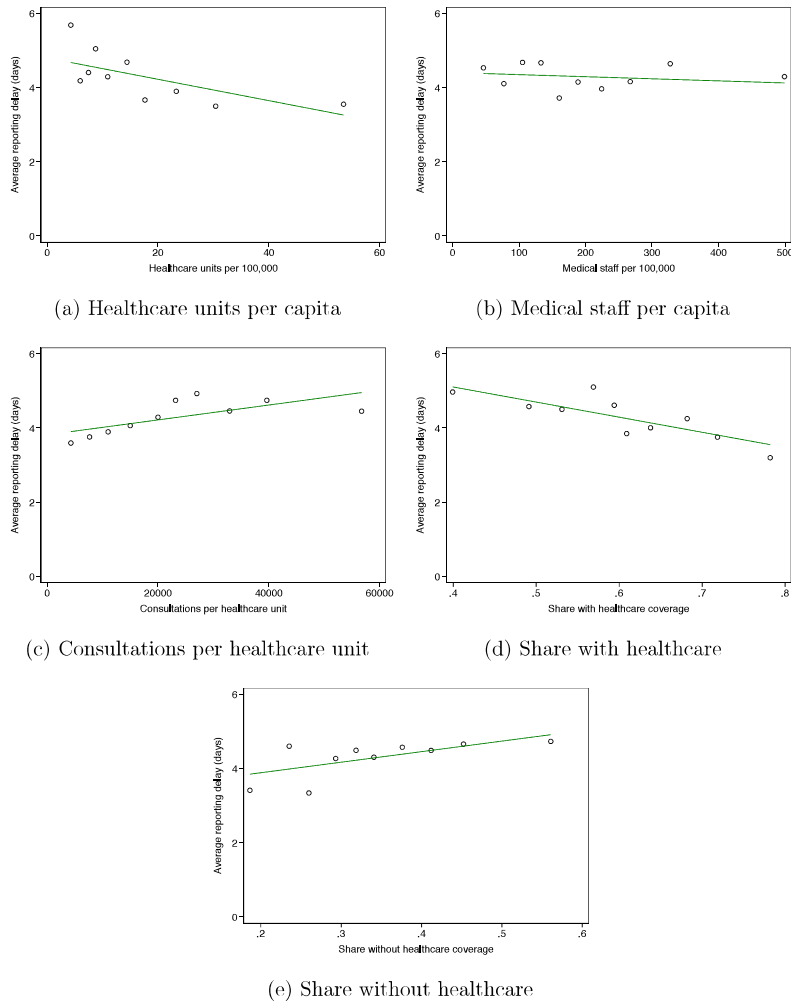
Figure A1:
Average Reporting Delay Over Time by Country



Notes: These graphs show the distribution of the average reporting delay measured in days for each country. In Figures A1b and A1c, the data are further stratified by median date of death for the available span of data. We drop the most recent 21 days of data reports, and delays that are over 21 days. In Figure A1a, the mean for England is 2.10, and 5.56 for Mexico, implying a difference of 3.46, with a 95% CI [3.30,3.62]. The mean for the first half of the data for England in Figure A1b is 2.35, and 1.86 for the second half, implying a difference of 0.49, 95% CI [0.35,0.63]. Lastly, the mean for the first half of the data in Figure A1c is 5.66, 5.47 for the second half, and a difference of 0.19, 95% CI [-0.16,0.54].

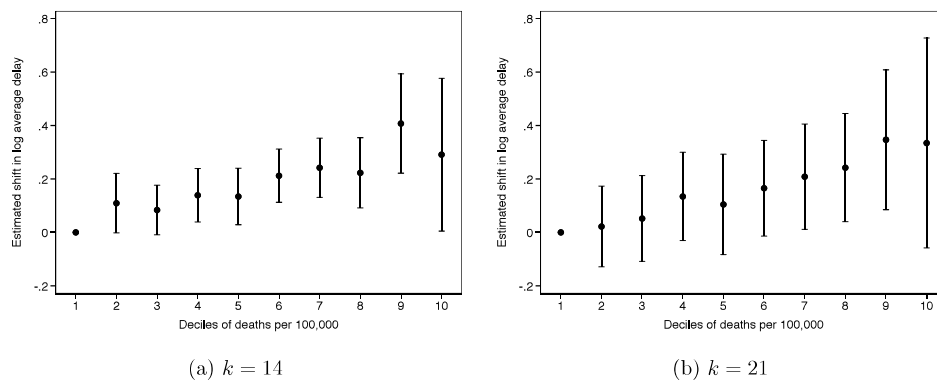
Covid Economics 34, 3 July 2020: 116-144

Figure A2:
Average Reporting Delays and Healthcare Infrastructure in Mexico



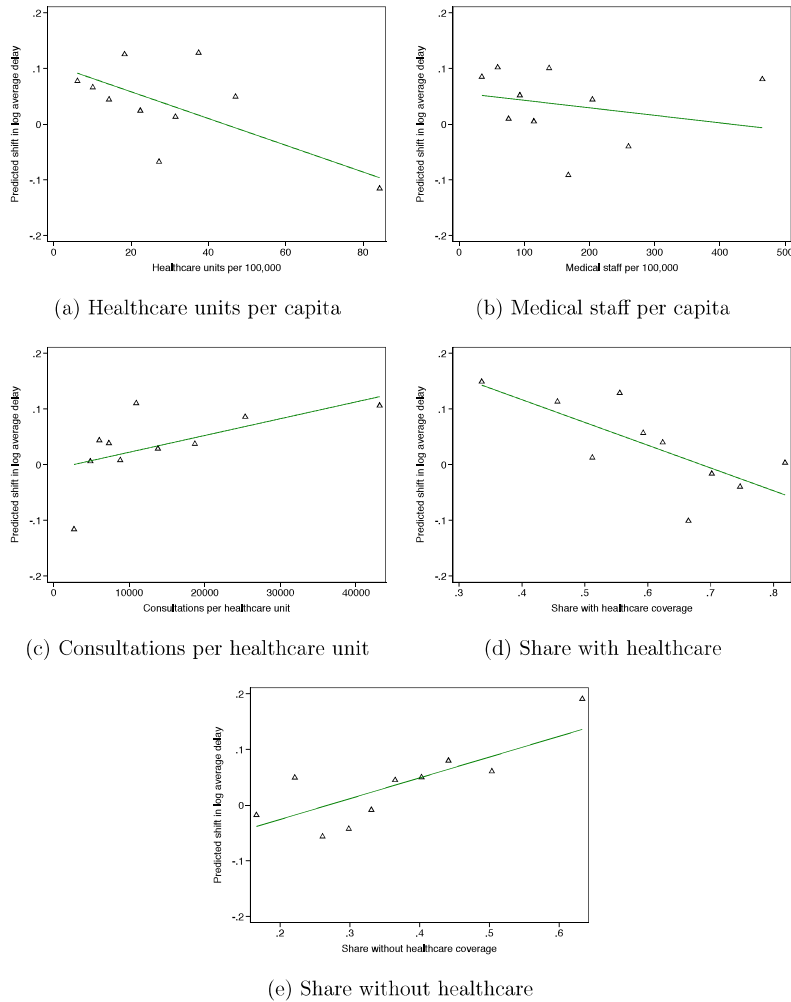
Notes: These graphs show the correlation between average reporting delays measured in days and various measures of healthcare infrastructure in Mexico. Figure A2a shows the number of healthcare units per 100,000 individuals in a municipality as measured in 2016, Figure A2b shows the number of healthcare workers per 100,000 individuals in 2016, Figure A2c shows the number of medical consultations per healthcare unit in 2016, and Figures A2d and A2e show the share of the population in a municipality with public healthcare coverage and without any coverage according to the 2010 census. Each plot shows the average over 10 bins, as well as a line of best fit. To calculate delays, we drop the most recent 14 days of data reports, and delays that are over 14 days.

Figure A3:
Relationship between Total Deaths per Capita and Reporting Delays
in Mexico



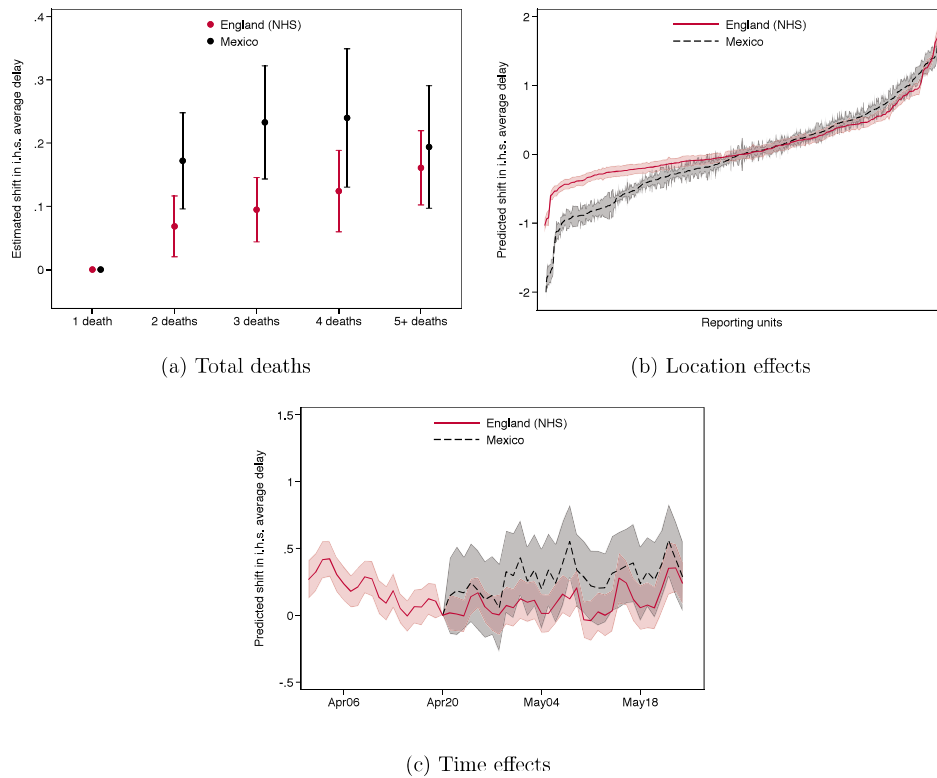
Notes: These graphs show the estimated shift in the log average delay in relation to deciles of total deaths per 100,00 people per municipality from estimating equation 1. All effects are calculated relative to the mean shift for the first decile. Figure A3a corresponds to data that exclude the last 14 days of available reports, as well as delays over 14 days (N=5991 for England, N=3660 for Mexico). Figure A3b uses 21 days instead (N=5531 for England, N=2875 for Mexico). The vertical lines indicate 95% confidence intervals from robust standard errors clustered by municipality.

Figure A4:
Location Fixed Effects and Healthcare Infrastructure in Mexico



Notes: These graphs show the correlation between our estimated location (municipality) fixed effects from estimating equation 1 and various measures of healthcare infrastructure in Mexico. Figure A4a shows the number of healthcare units per 100,000 individuals in a municipality as measured in 2016, Figure A4b shows the number of healthcare workers per 100,000 individuals in 2016, Figure A4c shows the number of medical consultations per healthcare unit in 2016, and Figures A4d and A4e show the share of the population in a municipality with public healthcare coverage and without any coverage according to the 2010 census. Each plot shows the average over 10 bins, as well as a line of best fit. To calculate delays, we drop the most recent 14 days of data reports, and delays that are over 14 days.

Figure A5:
Determinants of Delays using the Inverse Hyperbolic Sine Transformation



Notes: These graphs show estimates from regressions similar to equation 1, using the inverse hyperbolic sine of average delays instead of the log as in Figures 3-5. The vertical lines and shaded areas indicate 95% confidence intervals from robust standard errors clustered by reporting unit. We drop the most recent 14 days of data reports, and delays that are over 14 days.

Table A1:
Point Estimates of Relationship between Total Deaths and Average
Delays

	Data: $k = 14$		Data: $k = 21$	
	(1)	(2)	(3)	(4)
Two deaths	0.056*** (0.019)	0.129*** (0.031)	0.056** (0.025)	0.121*** (0.038)
Three deaths	0.075*** (0.020)	0.173*** (0.037)	0.089*** (0.025)	0.158*** (0.042)
Four deaths	0.099*** (0.025)	0.176*** (0.046)	0.103*** (0.030)	0.220*** (0.066)
Five or more deaths	0.125*** (0.023)	0.136*** (0.040)	0.166*** (0.029)	0.148** (0.069)
Observations	5,991	3,660	5,531	2,875
R-squared	0.394	0.421	0.371	0.483
Sample	England	Mexico	England	Mexico
Mean dep. var.	1.74	4.29	2.10	5.56

Notes: This table presents the point estimates corresponding to categories of total deaths from estimating equation 1, as shown in Figure 3. Robust standard errors clustered by reporting unit are shown in parentheses. We report the mean average reporting delay for each sample. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$