medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

# Risk-driven responses to COVID-19 eliminate the tradeoff between lives and livelihoods

**Authors**:

Hazhir Rahmandad[1*] , Ph.D.

MIT, Room E62-442, 100 Main St., Cambridge, MA 02142

Tse Yang Lim[1], M.S.

MIT, Room E62-441, 100 Main St., Cambridge, MA 02142

Affiliations:

[1]Massachusetts Institute of Technology, Cambridge, MA

*Corresponding Author

Hazhir Rahmandad

hazhir@mit.edu

+1-617-253-8912

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

## Summary

**Background**- Responses to COVID-19 pandemic are conditioned by a perceived tradeoff between saving lives and paying the economic costs of contact-reduction measures. We develop and test the hypothesis that when populations endogenously respond to risk this tradeoff disappears.

**Methods**- We develop a model of SARS-CoV-2 transmission where populations endogenously reduce contacts in response to risk of death. We estimate the model for 118 countries constituting 7.05 billion people and assess the existence of a tradeoff between saving lives and livelihoods.

**Results**- We show that with endogenous responses communities go through three phases – rapid early outbreaks, control through initial response, and a longer period of quasi-equilibrium endemic infection with effective reproduction number ($R_e$) fluctuating around one. Analytical characterization of this phase shows little tradeoff between contact reduction levels (underpinning economic costs) and death rates. Empirically estimating the model, we find no positive correlation ($r$ = -0.241, $p$ = 0.009) between (log) death rates and (normalized) contact levels across nations. While contact reduction levels are broadly similar across countries (5-95 percentile: 0.521-0.867 of pre-pandemic contacts), expected death rates vary greatly, by over two orders of magnitude (5-95 percentile: 0.03-17 deaths per million per day).

**Interpretation**- Whether by choice or by the force of crippling death tolls, most societies will bring down interactions enough to contain SARS-CoV-2's spread. What we control is the severity of death toll required to compel us to act: greater responsiveness to risk can bring down deaths with no excess economic costs.

## Research in Context

### Evidence before this study

We study how endogenous changes in behaviors, in response to risk of death, moderate the tradeoff between deaths and interaction levels in communities exposed to COVID-19 pandemic. We searched for existing models of SARS-CoV-2 transmission and how they incorporate changes in policies and behaviours in response to changing risk levels. Most common are the use of time-based changes in basic reproduction number to capture those behavioural responses, or empirically specified policies that have been enacted in response to the epidemic. These methods, while effective for empirical estimation of historical trajectories, do not inform the tradeoff we are seeking to understand. A smaller subset of prior work does include endogenous responses to risk, mostly in the form of policy switches that activate when risk exceeds some threshold, with notable exceptions that include continuous responses potentially more suitable for the study at hand. To our knowledge these prior formalizations do not focus on teasing out the perceived tradeoff between interaction levels in a community and public health burden of the epidemic. Therefore, the powerful intuition that a strong tradeoff exists between saving lives and livelihoods remains unchallenged.

### Added value of this study

By introducing an endogenous and continuous response function connecting perceived risk levels to contact levels, we examine the existence of a tradeoff between contact levels and death rates across communities. We find this extension leads to a long quasi-equilibrium phase during which effective reproduction number remains around 1. During this phase our model predicts little tradeoff between contact reductions (the driver of economic costs) and death rates. We find empirical support for this extension and its predictions using data from 118 countries constituting more than 7 billion people. While

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

countries show limited variation in their estimated contact reduction levels, deaths vary by two orders of magnitude.

**Implications of all available evidence**

In a pandemic individuals and communities respond to risk, not only by following government mandates, but also through their personal choices protecting their own and others' lives. Formalizing this observation into models of contagion largely eliminates the tradeoff between saving lives and livelihoods in responding to a pandemic. More responsive policies promise to save lives with no additional economic costs compared to weaker response functions.

## Introduction

Is there a tradeoff between protecting lives and protecting livelihoods from the threat of COVID-19? The existence of this tradeoff seems intuitive– the social distancing and lockdown measures needed to control the pandemic's spread have had tremendous economic costs [1,2]. SARS-CoV-2 spreads through human interactions, so slowing down its spread should inevitably hurt economic activities that depend on those interactions.

This apparent tradeoff has dominated the policy discourse around pandemic response, with various nations and communities adopting a range of approaches from stringent lockdowns to hands-off non-responses [3]. Given the uneven distributions of both health risks [4] and economic impacts [3], any resolution to this perceived tradeoff will leave many stakeholders unsatisfied.

Some have sought to better align the two goals by incorporating the economic costs of an uncontrolled outbreak, e.g. from morbidity and mortality [5], or the public health impacts of economic recessions [6-8]. These arguments may reduce the longer-term tradeoffs but do not challenge the basic tension between near-term health and economic outcomes.

In this paper, we go a step further, and propose that fundamentally, little tradeoff exists between health and economic outcomes even during the course of the epidemic. Specifically, we argue theoretically that the degree of contact reduction and by extension, disruption to economic activity that the pandemic necessitates is largely independent of the resultant levels of morbidity and mortality. We find empirical support for this hypothesis estimating a model of endogenous contact reduction in response to COVID-19 pandemic across 118 countries constituting 7.05 billion people.

## Methods

### Model and Predictions

Our propositions build on the observation that behavioral responses to risk perception endogenously bring down contact rates in a community, creating a long quasi-equilibrium phase responsible for much of the public health and economic burden of an epidemic. Prior work has often characterized those responses as a function of time, historical data on policies, or policy switches activated in response to risks [9-12]; and more endogenous formalizations [13,14] have not addressed the tradeoff between contact rates and saving lives. We start with the classic SIR model, and incorporate those behavioral responses as continuous functions of perceived risk.

Populations are divided into three groups: susceptible to the disease (S), infected (I), and removed from circulation due to recovery or death (R). We build our argument starting with the simplest incarnation of the model:

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

$$\frac{dS}{dt} = -r_I;\ \frac{dI}{dt} = r_I - r_R;\ \frac{dR}{dt} = r_R \quad (1)$$

$$r_I = \frac{\beta IS}{N} \quad (2)$$

$$r_R = \frac{I}{\tau} \quad (3)$$

The two main parameters are the infectious contacts per day per index case (β), and the average disease duration (τ). We use $S_f = \frac{S}{N}$ (with $N = S + I + R$) to denote the susceptible fraction of the population.

Basic reproduction number ($R_0$), the expected secondary cases from an index case in a fully susceptible population, would be:

$$R_0 = \beta\tau \quad (4)$$

Because outbreaks in first-affected communities change behaviors in other nations, we call the empirically estimated values for $R_0$ initial reproduction numbers. Effective reproduction number ($R_e$) then accounts for reductions in secondary cases due to reduced susceptible fraction:

$$R_e = \beta\tau S_f \quad (5)$$

With a constant $\beta$ and an $R_0$ notably above one [15] the basic SIR model predicts a rapid COVID-19 outbreak that infects most of a population in a few months, reaching herd immunity when a large fraction of the population has been infected with huge resultant loss of life. In practice, however, few communities have followed this uncontrolled trajectory. Instead, the initial exponential growth phase has been curtailed well before herd immunity, as behavioral and policy responses to perceived risk of death bring down infectious contacts per index case (β) enough to control the epidemic. As a result, few communities have depleted their susceptible populations [16]. Instead, in most countries, the initial outbreak and control phases have been followed by a third phase in which a large fraction of the population remains susceptible and experiences a stream of new cases and deaths, including occasional subsequent waves.

We capture this endogenous change in behaviors and policies by allowing β to go below its initial value ($\beta_0$) in response to perceived risk of death (D):

$$\beta = \beta_0 g(D) \quad (6)$$

Function $g$ reflects contact rates relative to initial levels and offers a proxy for disruption to normal activities, the primary driver of the economic costs of an outbreak. $g(D)$=1 indicates no change in behaviors compared to pre-pandemic, whereas a society in full lockdown may push $g(D)$ down to near zero.

We model perceived risk as a 2nd order Erlang distributed delay of per capita daily death rate ($d_N$) resulting from an infection fatality rate of $f$. An exponential function is used to formulate $g$:

$$d_N = \frac{f r_R}{N} \quad (7)$$

$$D(t) = \int_{k\geq 0} d_N(t-k)\frac{2}{\lambda}k^2 e^{\frac{2k}{\lambda}}dk \quad (8)$$

$$g(D) = e^{-\alpha D} \quad (9)$$

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

α represents the responsiveness of a community to perceived risk and λ represents the time it takes to perceive and respond to risks. α=0 recovers the basic SIR model. Higher α values indicate a community more responsive to the perceived risk of death from the disease.

In estimating the model we note that the time constant for upward vs. downward adjustment (when $D(t)<d_N(t)$ vs. $D(t)>d_N(t)$) of perceived risk and response may differ, e.g. raising alarms may happen faster than forgetting about the risks:

$$\lambda = \begin{matrix} \lambda_u \; if \; \mathrm{D(t)} < \mathrm{d_N(t)} \\ \lambda_d \; if \; \mathrm{D(t)} \geq \mathrm{d_N(t)} \end{matrix} \tag{10}$$

To capture changes in risk-response relationship over time (e.g. due to adherence fatigue), we also allow communities' responsiveness to perceived risk (α) to vary, starting from some initial responsiveness ($\alpha_0$) and moving toward some final value ($\alpha_f$), with parameters $t_0$ and $\theta$ affecting the timing and rapidity of the adjustment:

$$\alpha = \alpha_0 + (\alpha_f - \alpha_0)\frac{1}{1+e^{\frac{-(t-t_0)}{\theta}}} \tag{11}$$

For estimation we predict observed infection rate as a function of observed death rates. Assuming an under-count ratio of $\gamma_I$ for official to true cases and using the subscript "M" for official measures (vs. true values), we can write:

$$r_I = \frac{\beta_0 g(d_N) IS}{N} \Rightarrow \frac{r_{IM}}{\gamma_I} = \frac{\frac{\beta_0 g(d_{NM}) I_M}{\gamma_I} S}{N} \Rightarrow r_{IM} = \frac{\beta_0 g(d_{NM}) I_M S}{N} \xRightarrow{I_M(t)=\sum_{k=1}^{k=\tau} r_{IM}(t-k)}$$

$$r_{IM} = \beta_0 S_f g(d_{NM}) \sum_{k=1}^{k=\tau} r_{IM}(t-k) \tag{12}$$

Note that in this derivation we assume $\gamma_I$ is not changing quickly on the timescale of disease duration, $\tau$, but could change over longer horizons. Susceptible fraction ($S_f$) is not directly observable because the cumulative case counts are significantly under-estimating the true magnitude of the epidemic. Official death data ($d_{NM}$) are more reliable. We therefore estimate total cases by dividing cumulative deaths to-date by country-specific infectious fatality rates ($f$) calculated based on age-specific fatalities [17] and correcting with $\gamma_D$ for country-specific under-reporting and over-estimation of $f$ (details in S2):

$$S_f(t) = 1 - \frac{\gamma_D}{f} \sum_{s=0}^{s=t} d_{NM}(s) \tag{13}$$

Combining equations 8-13 leads to our primary estimation equation:

$$E(r_{IM}(t)) = \beta_0 (1 - \frac{\gamma_D}{f} \sum_{s=0}^{s=t} d_{NM}(s)) e^{-\left(\alpha_0 + (\alpha_f - \alpha_0)\frac{1}{1+e^{\frac{-(t-t_0)}{\theta}}}\right)\left(\int_{s\geq 0} d_{NM}(t-s)\frac{2}{\lambda}s^2 e^{\frac{2s}{\lambda}} ds\right)} \sum_{k=1}^{k=\tau} r_{IM}(t-k) \tag{14}$$

## Quasi-equilibrium phase and its features

The model includes a balancing feedback loop that can control the epidemic before herd immunity is reached: while $R_e>1$, rising deaths raise alarms, bringing down infectious contacts. $R_e<1$ reduces deaths until the diminishing perceived risk leads to relaxation of responses, pushing $R_e$ up. Behavioral risk

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

response thus creates a self-correcting process through which $R_e$ converges to approximately one, resulting in a quasi-equilibrium phase where cases and infections are close to stable (Figure 1).

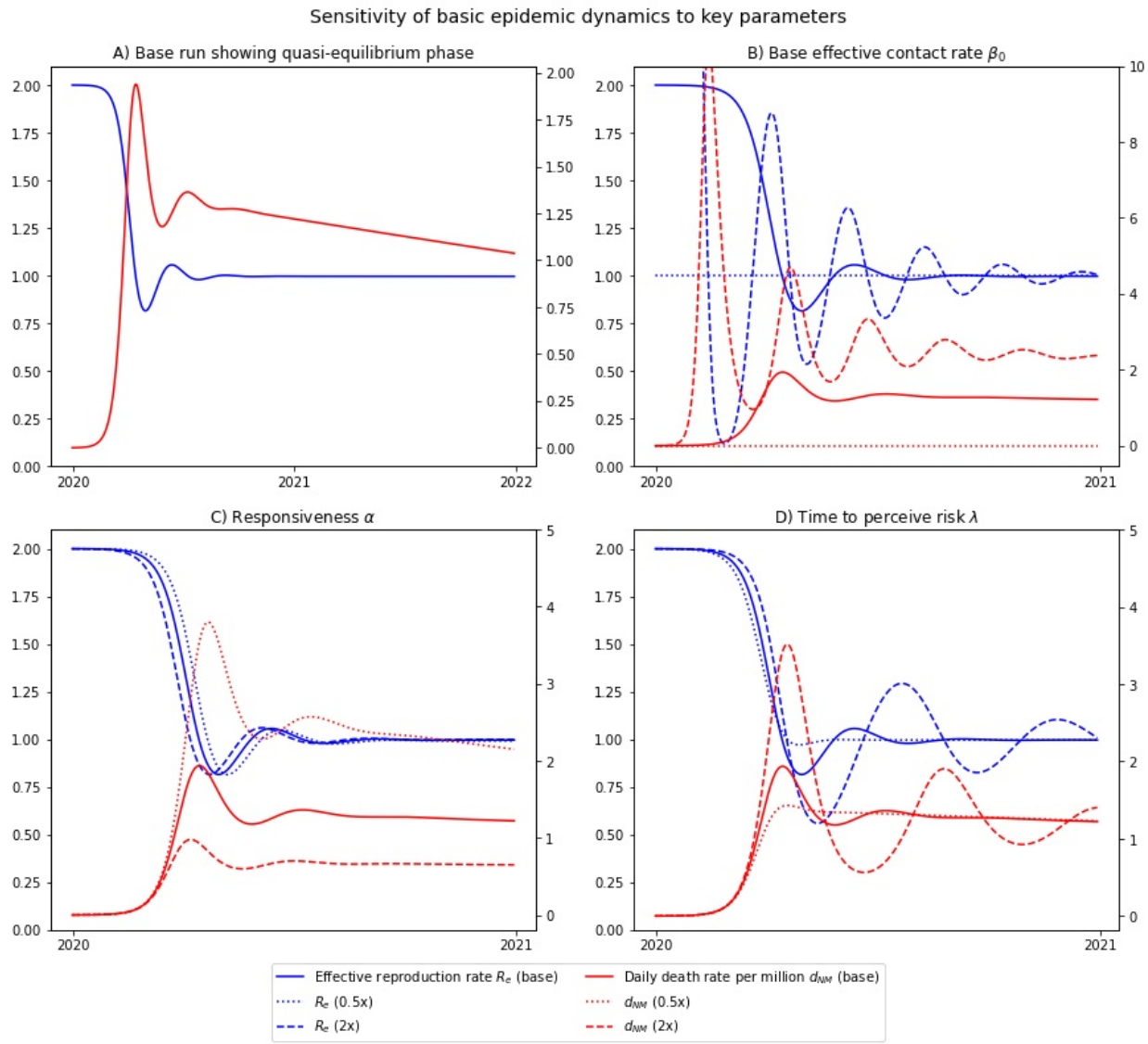


*Figure 1 - $R_e$ (left axes, blue) and daily deaths per million (right, red) in epidemics with endogenous response to risk. A) Illustrative case with $\beta_0$=2, $\alpha$=0.5, and $\lambda$=15 days, f=0.5%absent vaccine quasi-equilibrium extends from mid-2020 to beyond 2022. B-D) sensitivity to variations in parameter values (0.5x and 2x base values).*

This phase is a 'quasi-'equilibrium as an ongoing stream of new cases continues to reduce the susceptible fraction. Moreover, the delays between infection, death, and response may allow transmission and deaths to go higher/lower than equilibrium when rising/falling. Thus, depending on the risk-response delays ($\lambda$), communities converge to quasi-equilibrium expected death rates, $d_{NM}^{eq}$, or oscillate around this value (Figure 1). In that neighborhood, D(t) will on average be fairly stable and close to $d_{NM}$ (equation 8), allowing us to calculate the quasi-equilibrium death rates:

$$d_{NM}^{eq} = \frac{\ln(R_0 S_f)}{\alpha} \tag{16}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

This phase ends when $S_f$ falls below $1/R_0$. As long as communities are somewhat responsive to risk the $d_{NM}$ values remain close to $d_{NM}^{eq}$ and $g$ close to $g^{eq}$ for an extended period, making the quasi-equilibrium phase critical in the overall impact of the pandemic.

We focus on two features of the quasi-equilibrium phase. First, the normalized contact level in this phase, $g^{eq}(D)$ does not depend on the functional form or parameters of risk perception and response or the number of infectious individuals ($I$). It is simply a function of the initial reproduction number ($R_0$) and the susceptible fraction ($S_f$) so that $R_e \approx 1$. Second, the death rate in the quasi-equilibrium phase will depend primarily on α. Thus our model predicts limited variation in normalized contacts across communities, while differences in equilibrium death rates may be much larger. Next, we test these hypotheses by using COVID-19 infections and mortality data to quantify the response functions and the projected quasi-equilibrium death rates across 118 countries.

## Data and Estimation

We use equation 14 to find the posterior distribution of each country's parameters $(\beta_0, \alpha_0, \alpha_f, t_0, \theta, \lambda_u, \lambda_d, \gamma_D)$, using a Markov Chain Monte-Carlo method [18] starting with uninformed priors (details in S1). To accommodate heteroskedasticity and fat tails, we represent observed cases ($r_{IM}$) using a Negative Binomial distribution, which introduces a scale parameter ($\epsilon$). We specify the average disease duration (τ) at 10 days [19], constant across all countries (and find limited sensitivity in the results to plausible alternative values, see S4).

We use a panel of daily data on confirmed cases and deaths for 118 countries comprising 7.05 billion people, covering dates from 31 December 2019 to 02 December 2020. We use 7-day rolling averages on this data to smooth out weekly reporting cycles and better reflect underlying trends. Data come from the OurWorldInData global COVID-19 database [20], which aggregates its data from various official sources.

## Results

This simple model successfully replicates the diverse range of trajectories observed in the epidemic thus far (Figure 2; See S3 for full sample). Over the full sample of 118 countries, $R^2$ for infections against data is 0.936, while the mean absolute errors normalized by mean (MAEN) are 13.2%. For a handful of countries (Argentina, Jordan, Oman) $S_f$ falls below $1/R_0$ in our analysis period so quasi-equilibrium no longer holds; those countries are excluded henceforth.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

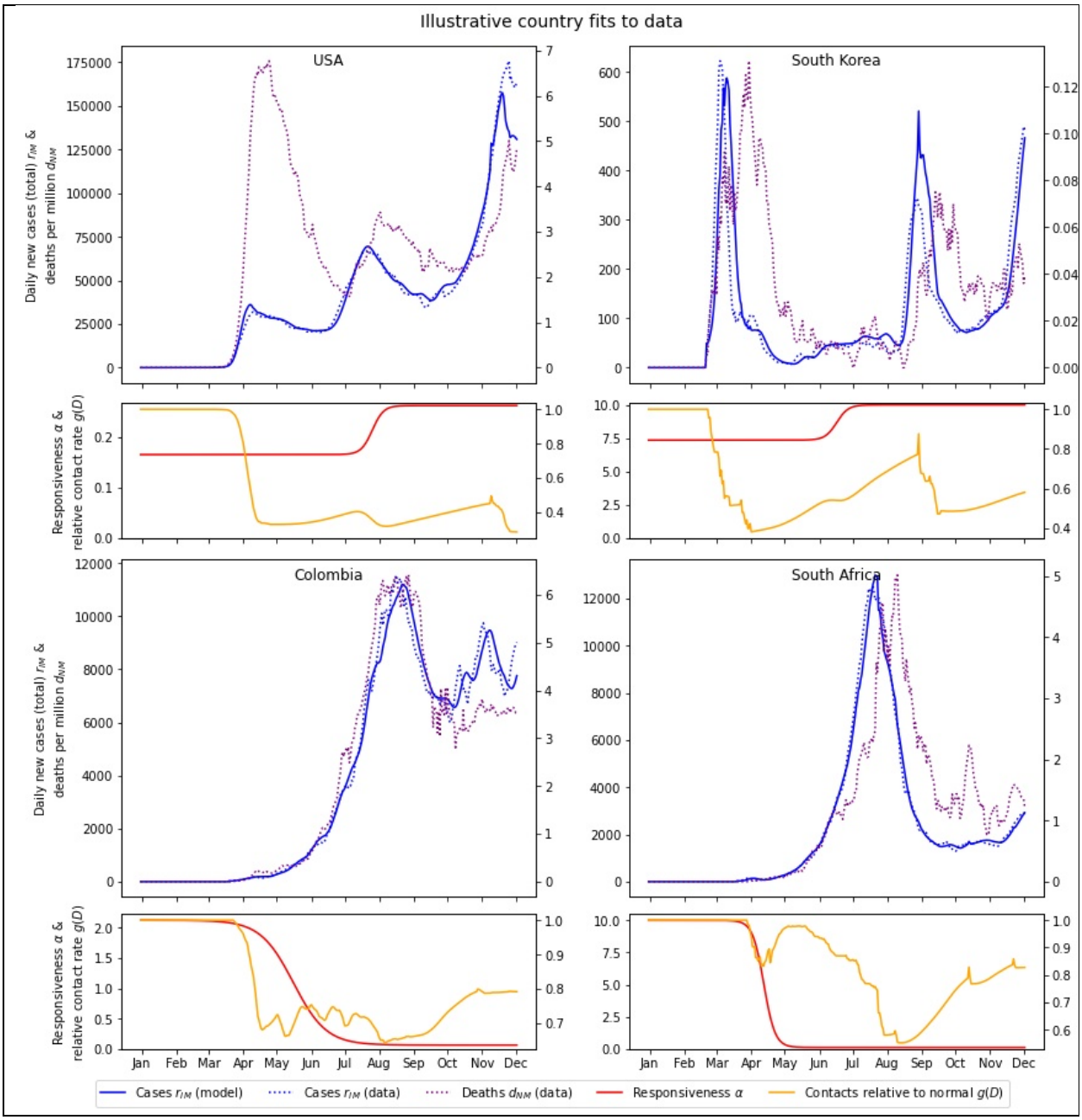


*Figure 2-Sample country-specific trajectories. Daily cases (estimated and actual) are on left axis, and deaths (per million) on the right. Lower panels show alpha (red; left axis) and normalized contacts (yellow; right axis).*

On average, responsiveness decreases over time: On 01 May 2020 the median effect of 1 daily death per million was to reduce contacts to 30.4% of normal, but by 02 December 2020 normalized contacts only drop to 77.7% in the same scenario. There is also much variance in these response factors – in the 1 daily death per million scenario, estimated contact rates by 02 December 2020 range from 0.01%-98.3% ($5^{th}$-$95^{th}$ percentile)of normal. As hypothesized, risk adjustment is asymmetric, with most countries increasing perceived risk much faster (median $\lambda_u$= 8.9 days) than reducing it (median $\lambda_d$= 61 days).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

In most countries, quasi-equilibrium conditions hold as reproduction number remains close to one. Across these 115 countries, the $R_e$ averaged over 05 June 2020-02 December 2020 has a median of 1.10 with a tight range (0.97-1.26, 5th-95th percentile) (Figure 3). In contrast, the median reported deaths ($d_{NM}$) averaged over the same period is 0.73 per million per day with a 5th-95th percentile range spanning 0.02-3.19. While $R_e$ values are clustered around 1 death rates vary by over two orders of magnitude.

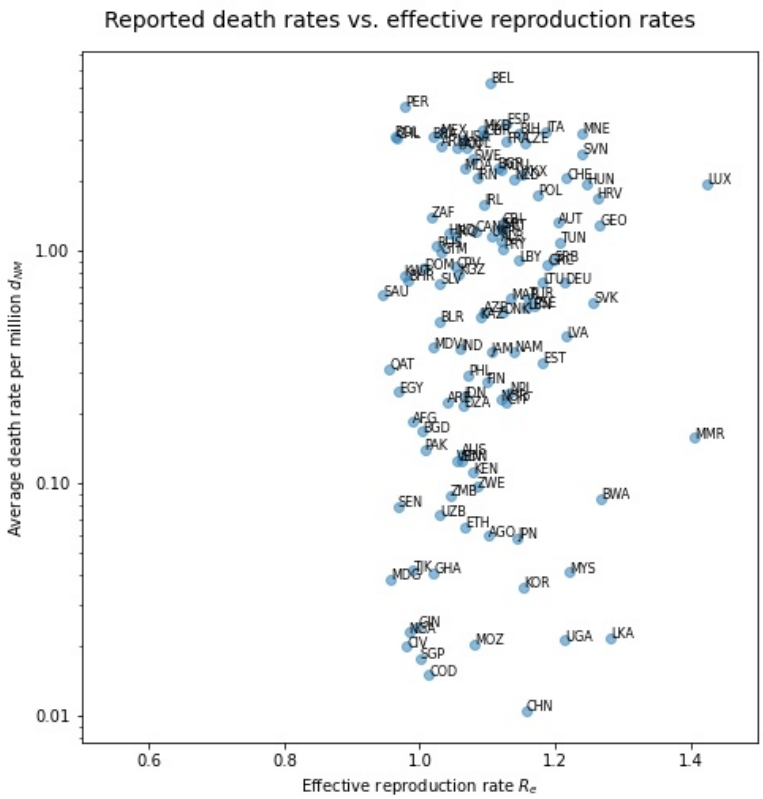


*Figure 3 –Reported death rates (daily; per million) vs. estimated effective reproduction rates ($R_E$), both averaged over the last 180 days.*

Figure 4 shows how the model's predicted quasi-equilibrium deaths compare with actual data. For each day we regress (log-transformed) observed death rates ($d_{NM}$) for all countries against the estimated most likely $\log(d_{NM}^{eq})$ values for that day. We then graph the estimated coefficient for these regressions over time. After a less stable early phase, the coefficient converges to the neighborhood of one and remains there, showing that $d_{NM}^{eq}$ reliably predicts the actual reported deaths for much of the pandemic's history. This result is notable because we did not estimate death rates directly: the projected equilibrium deaths rely on the analytical derivations and the estimated behavioral parameters for risk response, and yet those analytically driven values predict reported deaths with a coefficient close to 1.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

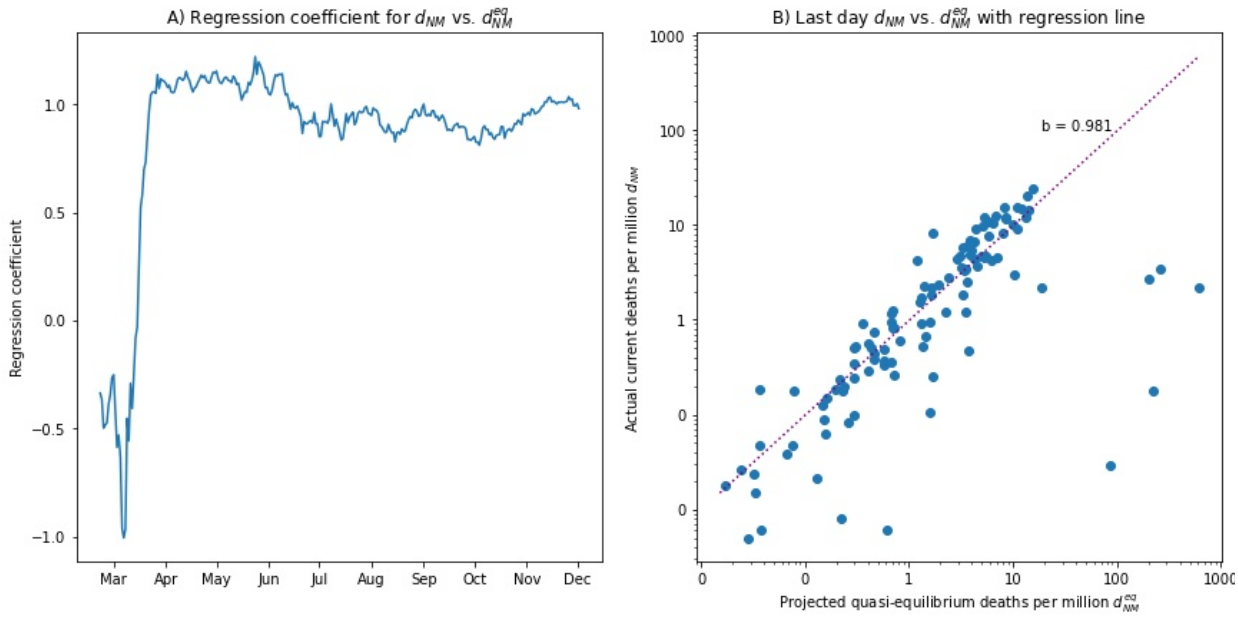


*Figure 4 – Predicting reported deaths based on quasi-equilibrium calculations (using robust regression with no intercept). A) Coefficients for regressing daily death data on analytically derived quasi-equilibrium death rates. B) Scatter-plot for the two variables over the last day of data, also showing the regression line for the last coefficient in A.*

Our key hypothesis informs the relationship between normalized contact levels and equilibrium death rates across nations. Figure 5 presents the scatterplot of the two at the end of estimation period. Quasi-equilibrium daily death rates ($d_{NM}^{eq}$) vary a lot (median: 1.46; 5$^{th}$-95$^{th}$ percentile range of 0.03 to 17). In contrast, normalized contacts in this phase are much less variable with median $g^{eq}$ of 0.690 (0.521-0.867, 5$^{th}$-95$^{th}$ percentile). The estimates of contact rates are highly precise – the median of the interquartile ranges of each country's credible interval, normalized by its estimated value (MNIQR), is 0.057. Credible intervals on projected death rate are wider but still fairly precise (MNIQR = 0.16). Notably, we find no positive correlation between death rate and contact rates (Pearson's $r$ = -0.077, $p$ = 0.411; for log($d_{NM}^{eq}$), $r$ = -0.241, $p$ = 0.009). Overall, both in data (Figure 3) and in estimates (Figure 5), we find no tradeoff between contact levels and death rates across nations; if anything, countries with lower death rates have also benefited from more limited contact reductions.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

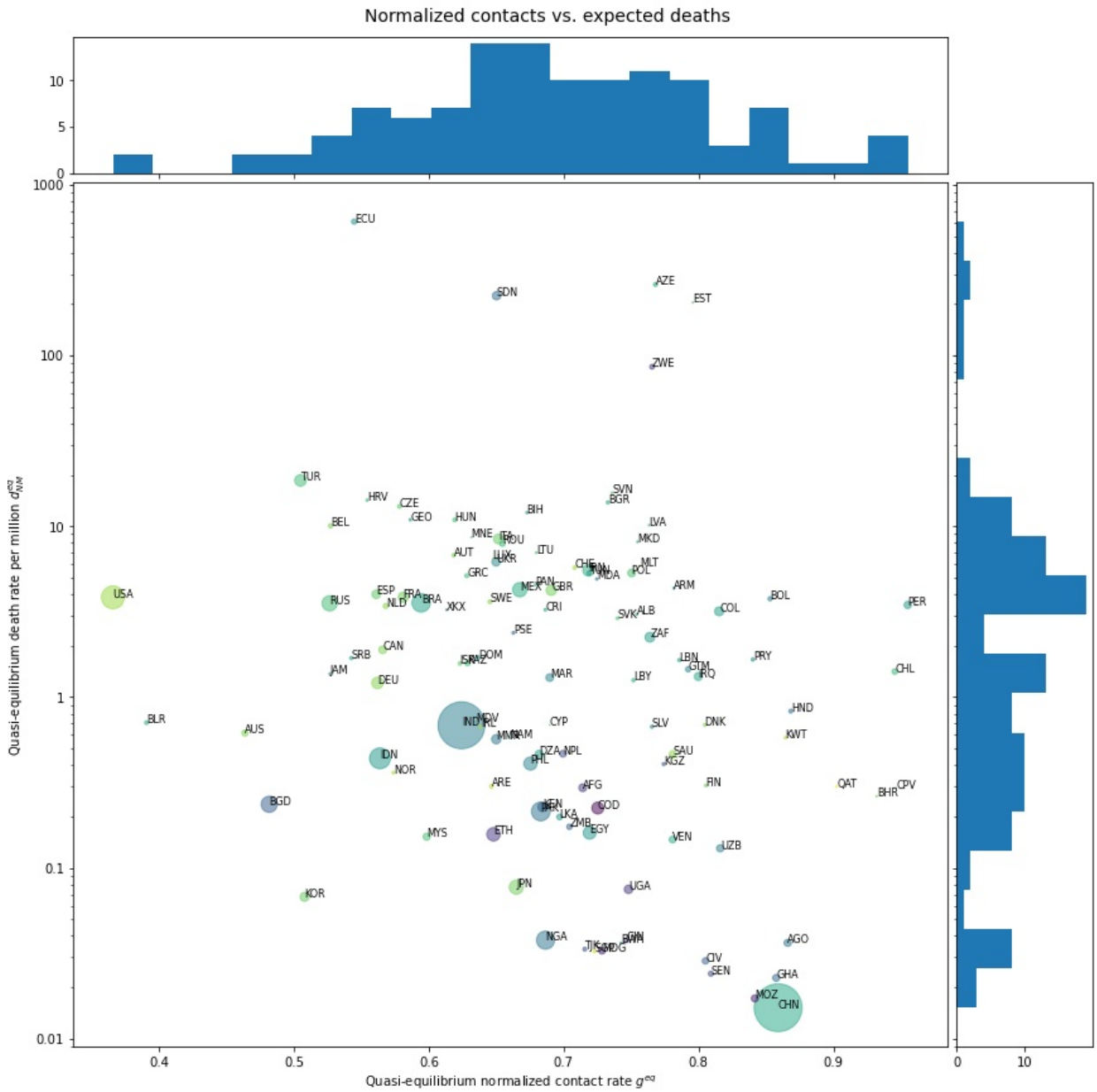


*Figure 5 - Scatterplot and histograms of quasi-equilibrium daily death rates per million (Y-axis) and contact levels normalized to pre-pandemic levels (X-axis) estimated for 02 December 2020. Size of circles scales with population. The few countries with very high (> 100) $d^{eq}_{NM}$ may face large outbreaks and transition out of equilibrium rather than actually experiencing those death rates.*

Normalized deaths are largely stable, but a few countries show notable changes (in both directions). Figure 6 quantifies those trends as daily fractional change in $d^{eq}_{NM}$ at the end of the estimation period. Countries on the left of the graph are at risk of increasing equilibrium death rates in the coming weeks, a warning call to shore up responsiveness to avoid costly, and pointless, rebounds.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

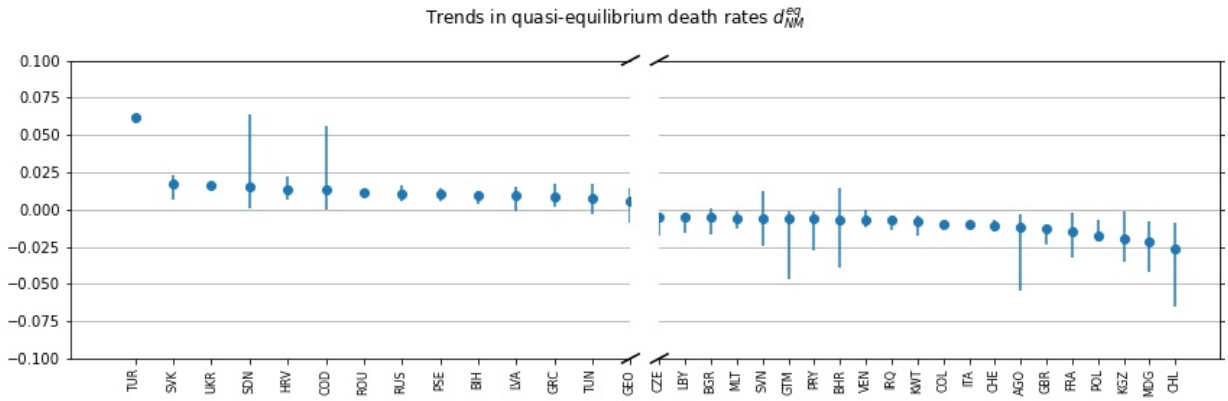


*Figure 6 – Daily fractional change in $d_{NM}^{eq}$ averaged over last 14 days (medians and 90% credible intervals). Graph limited to countries with absolute daily rates of change >0.5%. Turkey is an outlier due to a recent change in its reporting of infections.*

## Discussion

Contrary to the expected tradeoff between protecting lives and livelihoods we found virtually no positive correlation ($r$ = -0.241, $p$ = 0.009) between (log) death rates and (normalized) contact levels in quasi-equilibrium conditions across 115 nations. Furthermore, while contact reduction levels are broadly similar across countries (90% of countries fall between 0.521-0.867 times their initial contact rates), expected death rates vary greatly, by over two orders of magnitude (90% of countries between 0.03-17 deaths per million per day).

What accounts for the counterintuitive disconnect between contact reduction and death rates? Few countries tolerate exponentially growing outbreaks and deaths for long. Whether through formal policy responses or informal behavioral changes, steps will be taken to curb the epidemic, bringing down $R_e$ to the neighborhood of one. Across countries the contact reduction needed for reaching this reproduction number is largely similar, despite some inevitable variation as physical and cultural conditions create some differences in $R_0$ among countries. In contrast, the death rate a country faces in the quasi-equilibrium phase depends primarily on the country's responsiveness to risk. High responsiveness means that even slight upticks in the death rate will trigger a rapid response, bringing down $R_e$ to 1. Low responsiveness requires higher death rates before $R_e$ is pushed down to 1. So far countries have varied by two orders of magnitude in the death rates required to trigger this stabilizing response.

This result offers an important policy insight. Communities could experience vastly lower death rates with comparably stringent contact-reduction measures (i.e. those leading to $R_e$~1), if they first brought down infections to very low levels and responded quickly and vigilantly to keep $R_e$~1 thereafter. Other, more complex models have also noted this (lack of a) tradeoff. For example, Ferguson and colleagues [21] found that different trigger thresholds for re-implementing lockdowns would result in differing mortality levels but with virtually no change in the total time spent in lockdown.

Our simple model was designed primarily for investigating the tradeoffs between contact reduction and deaths. Extensions could more explicitly model under-reporting processes, use testing data to inform the analysis, account for changing fatality rates over time, and couple model parameters across nations in a hierarchical Bayesian framework. Moreover, we do not differentiate between policies used to reduce contact rates (e.g. contact tracing, isolation of vulnerable groups, school closures etc.). Some policies are

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

more efficient in reducing contacts [22,23]. With effective policy sets, some nations may better maximize economic activity given their limited contact 'budget'. Moreover, by keeping infections low, communities can use extensive testing, contact tracing, and isolation to control the epidemic with minimal economic disruption and also offer much better care to those infected, options not available to those struggling with large caseloads. Thus, better economic outcomes may well correlate with fewer deaths, a possibility highlighted both in our analysis and others' [24].

Overall, the tradeoff between lives and livelihoods has dominated policy and popular discourse around COVID-19, but focusing on this tradeoff misses a more fundamental truth. COVID-19 is a deadly disease, and whether by choice or by the force of crippling death tolls, societies will have to bring down interactions enough to contain its spread. What we control is the severity of death toll required to compel us to act. Outbreaks will grow as far as we are willing to let them, and swift reactions will save both lives and livelihoods.

## Data Sharing

No data were collected specifically for this study; all data are from publicly available databases. All data-processing and analysis code, as well as the full model, its associated files, and results files, are available online at https://github.com/tseyanglim/CovidRiskResponse.

## References

1. Coibion O, Gorodnichenko Y, Weber M. The cost of the covid-19 crisis: Lockdowns, macroeconomic expectations, and consumer spending: National Bureau of Economic Research, 2020.
2. McKee M, Stuckler D. If the world fails to protect the economy, COVID-19 will damage health not just now but also in the future. *Nature Medicine* 2020; **26**(5): 640-2.
3. Loayza NV. Costs and Trade-Offs in the Fight Against the COVID-19 Pandemic: A Developing Country Perspective. World Bank; 2020.
4. Hooper MW, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. *Jama* 2020.
5. Cutler DM, Summers LH. The COVID-19 pandemic and the $16 trillion virus. *Jama* 2020; **324**(15): 1495-6.
6. Webb L. Covid-19 lockdown: a perfect storm for older people's mental health. *Journal of psychiatric and mental health nursing* 2020.
7. Bhuiyan AI, Sakib N, Pakpour AH, Griffiths MD, Mamun MA. COVID-19-related suicides in Bangladesh due to lockdown and economic factors: case study evidence from media reports. *International Journal of Mental Health and Addiction* 2020.
8. Douglas M, Katikireddi SV, Taulbut M, McKee M, McCartney G. Mitigating the wider health effects of covid-19 pandemic response. *BMJ* 2020; **369**: m1557.
9. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* 2020; **368**(6493): 860-8.
10. Flaxman S, Mishra S, Gandy A, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 2020.
11. Walker PGT, Whittaker C, Watson OJ, et al. The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science* 2020.
12. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science* 2020.
13. Struben J. The coronavirus disease (COVID-19) pandemic: simulation-based assessment of outbreak responses and postpeak strategies. *Syst Dyn Rev* 2020.
14. Ghaffarzadegan N, Rahmandad H. Simulation-based estimation of the early spread of COVID-19 in Iran: actual versus confirmed cases. *Syst Dyn Rev* 2020; **36**(1): 101-29.
15. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* 2020; **382**(13): 1199-207.
16. Rahmandad H, Lim TY, Sterman J. Estimating COVID-19 under-reporting across 86 nations: implications for projections and control. *Available at SSRN 3635047* 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

17. Verity R, Okell LC, Dorigatti I, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet infectious diseases* 2020.
18. Vrugt JA, Ter Braak C, Diks C, Robinson BA, Hyman JM, Higdon D. Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *International Journal of Nonlinear Sciences and Numerical Simulation* 2009; **10**(3): 273-90.
19. He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med* 2020; **26**(5): 672-5.
20. Roser M, Ritchie H, Ortiz-Ospina E, Hasell J. Coronavirus pandemic (COVID-19). *Our World in Data* 2020.
21. Ferguson N, Laydon D, Nedjati-Gilani G, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. *Imperial College London* 2020; **10**: 77482.
22. Han E, Tan MMJ, Turk E, et al. Lessons learnt from easing COVID-19 restrictions: an analysis of countries and regions in Asia Pacific and Europe. *Lancet* 2020; **396**(10261): 1525-34.
23. Benzell SG, Collis A, Nicolaides C. Rationing social contact during the COVID-19 pandemic: Transmission risk and social benefits of US locations. *Proc Natl Acad Sci U S A* 2020; **117**(26): 14642-4.
24. Hasell J. Which countries have protected both health and the economy in the pandemic. *Our World in Data* 2020; **1**.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Online Appendix to Accompany:

# Risk-driven responses to COVID-19 eliminate the tradeoff between lives and livelihoods

**Authors**: Hazhir Rahmandad[1*], Tse Yang Lim[1]

Affiliations:

[1]Massachusetts Institute of Technology, Cambridge, MA

*Correspondence to: hazhir@mit.edu

## Contents

S1 – Estimation method ..... 16
S2 – Data processing ..... 16
IFR calculation ..... 17
S3 – Full results ..... 17
S4 – Sensitivity results ..... 20
Disease Duration ..... 20
Estimation without behavioural response ..... 23
S5 – Model equations listing ..... 25
References ..... 26

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

## S1 – Estimation method

The estimation equation for our model (equation 14 in the main paper) is replicated below:

$$E\big(r_{IM}(t)\big) =$$

$$\beta_0\left(1-\frac{\gamma_D}{f}\sum_{s=0}^{s=t} d_{NM}(s)\right)e^{-\left(\alpha_0+(\alpha_f-\alpha_0)\frac{1}{1+e^{\frac{-(t-t_0)}{\theta}}}\right)\left(\int_{s\geq 0} d_{NM}(t-s)\frac{2}{\lambda}s^2 e^{\frac{2s}{\lambda}}ds\right)}\sum_{k=1}^{k=\tau} r_{IM}(t-k)$$

This equation predicts the expected number of new infections for day t as a function of past infections (to calculate stock of infectious individuals) and past death rates (to calculate the response function). We assume the observed infections follow a negative binomial distribution with the given mean from this equation and a scale parameter, $\epsilon$, that is estimated to replicate observed error distributions. The negative binomial distribution provides the flexibility to account for heteroscedasticity, over-dispersion, and fat tails, allowing a robust estimation despite substantial randomness in the data-generating process.

The model includes the following unknown parameters: $\boldsymbol{\omega} = [\beta_0, \alpha_0, \alpha_f, t_0, \theta, \lambda_u, \lambda_d, \gamma_D, \epsilon]$ – 4 $(\alpha_0, \alpha_f, t_0, \theta)$ quantify responsiveness and how it changes over time, 2 $(\lambda_u, \lambda_d)$ specify the risk perception delays, $\beta_0$ estimates the initial rate of infectious contacts per day per index case, and $\gamma_D$ estimates the potential under-valuation of Infection Fatality Rates (*f* values) for each country . Finally, $\epsilon$ informs the shape of negative binomial distribution for each country. Besides the f values (see IFR calculation below for details) the model includes two other given parameters, the duration of disease (10 days) and the order of the perception delay (2).

More compactly, the model yields a (log)likelihood for observing the daily reported cases ($r_{IM}(t)$), given a vector of unknown model parameters $\boldsymbol{\omega}$ and the reported daily cases and death rate for each country prior to the current date ($d_{NM}(t-)$) and $r_{IM}(t-)$): $LL_{NegBin}(r_{IM}(t)|\boldsymbol{\omega}, r_{IM}(t-), d_{NM}(t-))$

The estimation process seeks to identify the values for $\boldsymbol{\omega}$ that maximise this likelihood or are likely to be observed given that peak.

We estimate the model separately for each country. Separating countries significantly speeds up estimation and makes it feasible to conduct the full analysis within days on a 48-core server. For each country, we estimate the parameter vector $\boldsymbol{\omega}$ using the Powell direction search method implemented in Vensim™ DSS simulation software, restarting the optimization at 20 random points in the feasible parameter space. From the resulting optimum, we use MCMC to explore the payoff landscape to identify the high-likelihood region of parameter space around the peak. The MCMC algorithm used is designed for exploring high-dimensional parameter spaces; for more details see [1]. We draw a total of 500000 samples for each country, of which the first 300000 are discarded (the burn-in period); by the end of the burn-in, the chains are well-mixed and stable (Gelman-Rubin PSRF statistic < 1.1). We use the remaining 200000 samples to derive credible intervals for parameter and outcome estimates.

## S2 – Data processing

Data on daily confirmed cases and deaths come from the OurWorldInData (OWID) global COVID-19 database [2], which draws on the Johns Hopkins University CSSE COVID dashboard [3]. The CSSE dashboard in turn aggregates its data primarily from official sources such as the US Centres for Disease Control and Prevention (CDC), the European CDC, the World Health Organization, and national health ministries, updating at least daily.

We use OWID's 7-day rolling averages for new cases ('new_cases_smoothed') and deaths per million population ('new_deaths_smoothed_per_million'). COVID-19 case and death reporting data show strong weekly cycles in many countries, as well as occasional anomalous spikes due to e.g. irregularities in test reporting or redefinitions by government statistical agencies; using the rolling average data smooths out these cycles, which we are not attempting to model here, to better reflect underlying trends.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Our analysis includes all countries in the dataset with at least 10000 cumulative cases reported, and at least 20 days of data. We exclude countries with fewer than 10000 cumulative cases to avoid skewing the results with outliers. The minimum datapoint requirement helps ensure robust estimation. In total, 118 countries meet these criteria as of 02 December 2020.

For countries included, we utilise data starting from the date when they exceed 100 cumulative cases reported. Excluding early data entails a tradeoff. Excluding it makes estimating the true basic reproduction number ($R_0$) more difficult – as discussed in the main text, after forceful outbreaks in the first countries, most others adopted various precautions that brought down $R_0$ below its pre-pandemic level. Furthermore, excluding the early data may cut out the initial dynamics of infection. As a result, our estimated values for initial reproduction number are likely underestimates of basic reproduction number, and thus the $g$ estimates may tend to be larger than the true changes in the contact rates compared to pre-pandemic levels. On the other hand, many of the early cases reported in most countries were due to travellers, and often identified and isolated early on. The data during this earliest 'importation' stage therefore do not accurately reflect community transmission dynamics we are modelling. Rapid changes in the testing coverage also impact our ability to use assume ascertainment rates are stable in the $\tau$ time horizon as needed in our derivations (see equation 12 in the main text). We selected the 100 case cut-off to balance reasonably estimating $R_0$ with correctly reflecting transmission dynamics rather than travel networks, which are out of scope for this model.

Data are downloaded and processed with Python 3 code, using Pandas and NumPy packages. For the full data processing code, see https://github.com/tseyanglim/CovidRiskResponse.

### IFR calculation

Most countries' reported case counts substantially under-estimate the true magnitude of the epidemic [4]. To estimate the remaining susceptible fraction ($S_f$) for each country over time, we therefore rely on reported deaths, which while still variable are more reliable, multiplying cumulative reported deaths by an estimated country-specific under-reporting ratio ($\gamma_D$) and using country-specific infection fatality rates (IFR) to calculate cumulative infections.

Age strongly influences IFR, with older patients far more likely to die of COVID-19 [5]. We therefore calculate country-specific IFRs based on each country's age structure.

We use data from the World Bank's World Development Indicators [6] on the age distribution of each country's population in 10-year age strata to calculate an age-weighted average of the IFRs for COVID patients by 10-year age group estimated in [5]. The resulting demography-adjusted country-specific IFRs range from 0.14% (Uganda) to 1.51% (Japan), with a mean of 0.54% and median of 0.44% (Lebanon). For the handful of countries for which up-to-date demographic data are unavailable, we use a baseline IFR of 0.50%.

We incorporate an estimated multiplier for actual to calculated deaths, $\gamma_D$, to account for potential undercounts of death and reductions in IFR compared to early values estimated from the methods we used. In estimation we restrict this multiplier to be between 1 and 4.

## S3 – Full results

Figure S1 shows fits to data for simulated infection rates for all 118 countries. Blue lines show model-generated daily infection rates, while red lines show 7-day rolling average infection rates from OWID. The correspondence between model and data is very close for most countries, with a few outliers bringing down the quality of fit a bit; yet over the full sample of 118 countries, $R^2$ for infections against data is 0.936, while the mean absolute errors normalized by mean (MAEN) are 13.2%. The quality of fit should not come as a total surprise: the model uses infection rates from the past 10 days to predict current-day infections, and thus to the extent that infections are auto-correlated, the estimation process can use this anchor to offer close approximations for the number of new cases. However, the behavioural response function does add significant value in terms of quality of fit, which we demonstrate in the sensitivity analysis section by comparing results against estimates that do not account for behaviour responses.

Table S1 summarises estimated parameter values across the 115 countries which met the quasi-equilibrium condition ($S_f > 1/R_0$). 3 of the 118 total countries estimated (ARG: Argentina, JOR: Jordan, OMN: Oman) no longer meet this

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

condition as of 02 December 2020, and were excluded from further analyses. For the full table of country-by-country parameter estimates, see https://github.com/tseyanglim/CovidRiskResponse.

Figure S2 shows 90% credible intervals estimated for each country for the two main outcome measures, quasi-equilibrium normalized contact rate ($g^{eq}$) and quasi-equilibrium daily death rate per million ($d_{NM}^{eq}$).

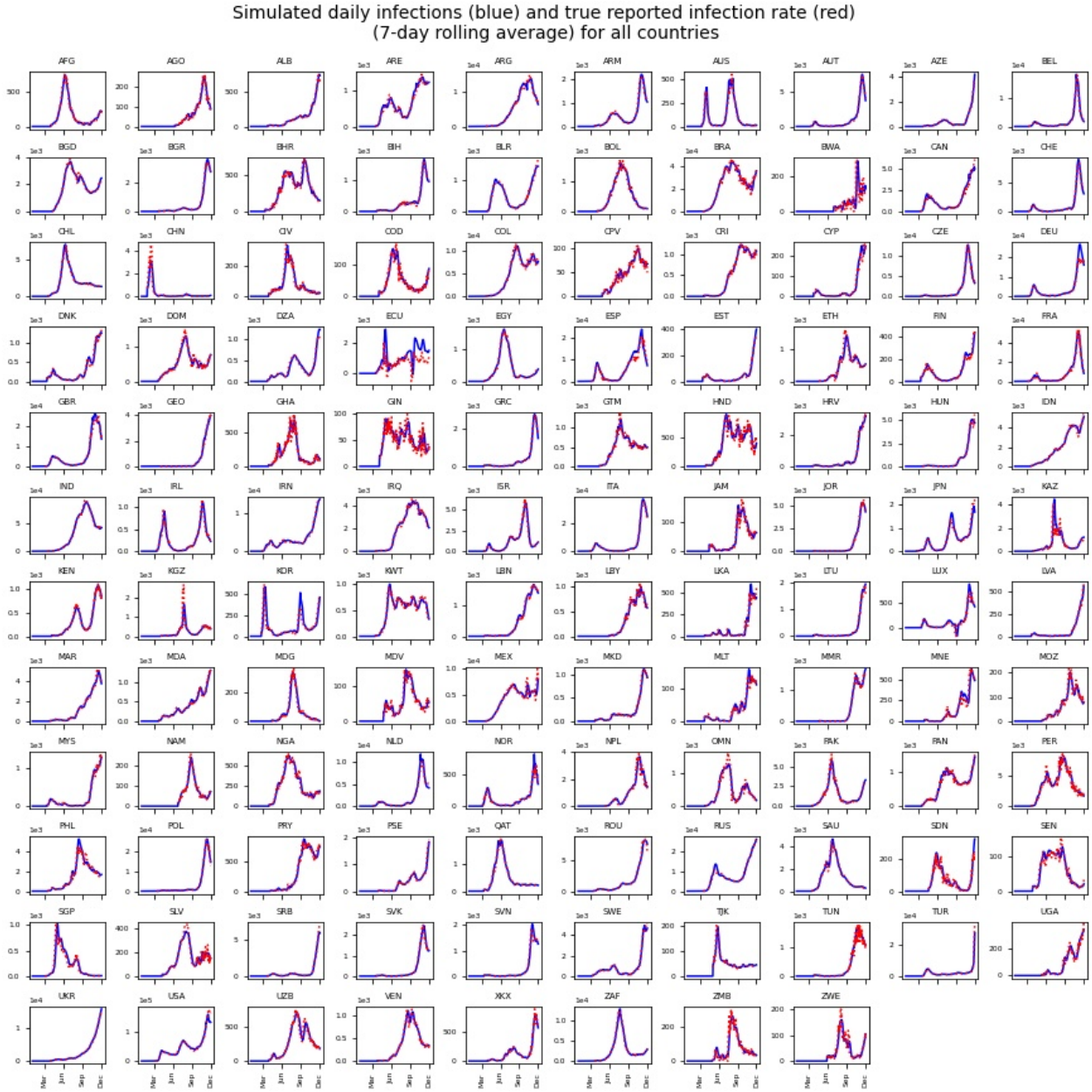


*Figure S1 - Comparison of simulated infection rates with data across all 118 countries*

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

*Table S1 - Summary statistics of estimated parameter values*

| Parameter | Symbol | Mean | StDev | Median | Med. IQR |
|---|---|---|---|---|---|
| Reference effective contact rate | $\beta 0$ | 0.17 | 0.034 | 0.16 | 0.008 |
| Initial responsiveness | $\alpha 0$ | 3.87 | 4.01 | 1.91 | 0.97 |
| Final responsiveness | $\alpha f$ | 1.46 | 2.79 | 0.19 | 0.064 |
| Responsiveness inflection point | t0 | 197 | 84 | 211 | 13 |
| Responsiveness scaling factor | $\theta$ | 16.1 | 22.0 | 5.1 | 5.4 |
| Time to upgrade risk | $\lambda U$ | 12.2 | 9.8 | 8.9 | 3.9 |
| Time to downgrade risk | $\lambda D$ | 58 | 38 | 61 | 11 |
| Death underreporting multiplier | $\gamma D$ | 2.52 | 1.29 | 2.35 | 1.66 |
| Likelihood scaling factor | $\epsilon$ | 0.058 | 0.150 | 0.022 | 0.004 |

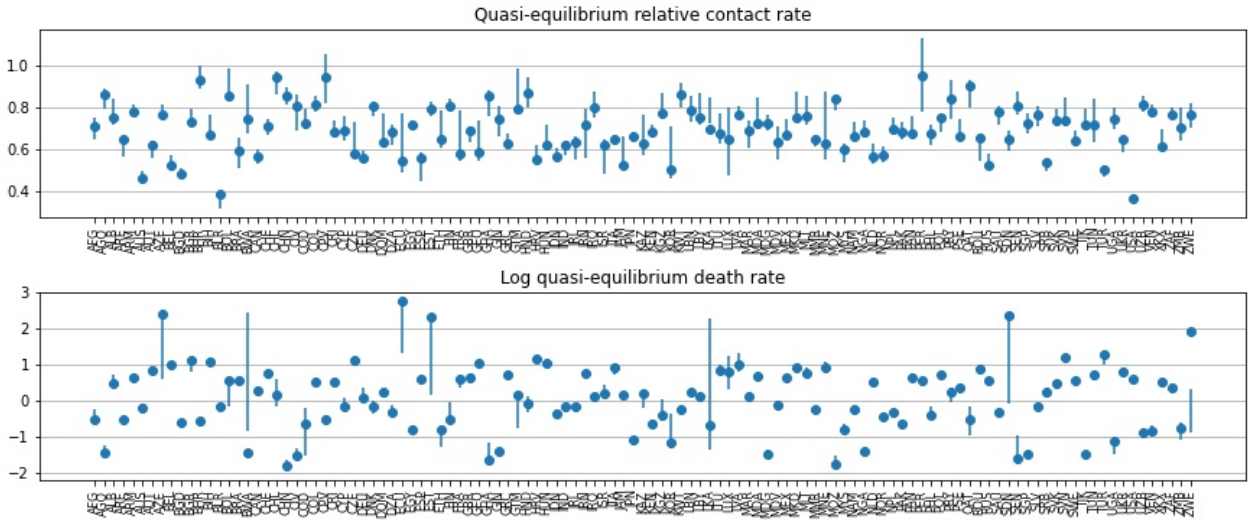


*Figure S2 - 90% credible intervals for estimates of normalized contact rate and log10 daily deaths per million in quasi-equilibrium.*

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

## S4 – Sensitivity results

### Disease Duration

We specify the average disease duration (τ) at 10 days, constant across all countries. This duration is consistent with prior findigns [7,8]. To test for sensitivity to this parameter, we re-ran model estimation and analysis with τ = 8 and 14 days.

Figure S3 and Figure S4 show the main result for τ = 8 and 14 days respectively. The primary insight has not changed – expected deaths and normalized contact rates in quasi-equilibrium conditions have no positive correlation (for τ = 8 and 14 days respectively, Pearson's $r$ = -0.044, $p$ = 0.647 and $r$ = 0.099, $p$ = 0.303; for $\log(d_{NM}^{eq})$, $r$ = -0.130, $p$ = 0.177 and $r$ = -0.238, $p$ = 0.012), with if anything a slight negative correlation as per the main result.

The model is still able to fit the data reasonably well with changes in disease duration. Table S2 and Table S3 summarise estimated parameter values with τ = 8 and 14 days respectively. On average, reducing τ to 8 days results in a 10.0% absolute change in estimated parameter values, while increasing it to 14 days results in a 7.4% absolute change. The fit between simulated infections and data deteriorates slightly at τ = 14 days ($R^2$ = 0.930, MAEN = 14.7%) compared to baseline ($R^2$ = 0.936, MAEN = 13.2%). Fit improves slightly at τ = 8 days ($R^2$ = 0.950, MAEN = 11.5%). The primary driver of infection rates is the number of currently infected people, which is calculated exogenously from the data based on the specified disease duration. As such, some inverse relationship between quality of fit and disease duration is to be expected, as shorter durations allow autocorrelation in the [smoothed] infection data to exert a stronger influence on the accuracy of model predictions.

These results indicate that both overall model performance, and more importantly, the main results of this analysis, are robust to alternative specifications of average disease duration within a broadly reasonable range.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

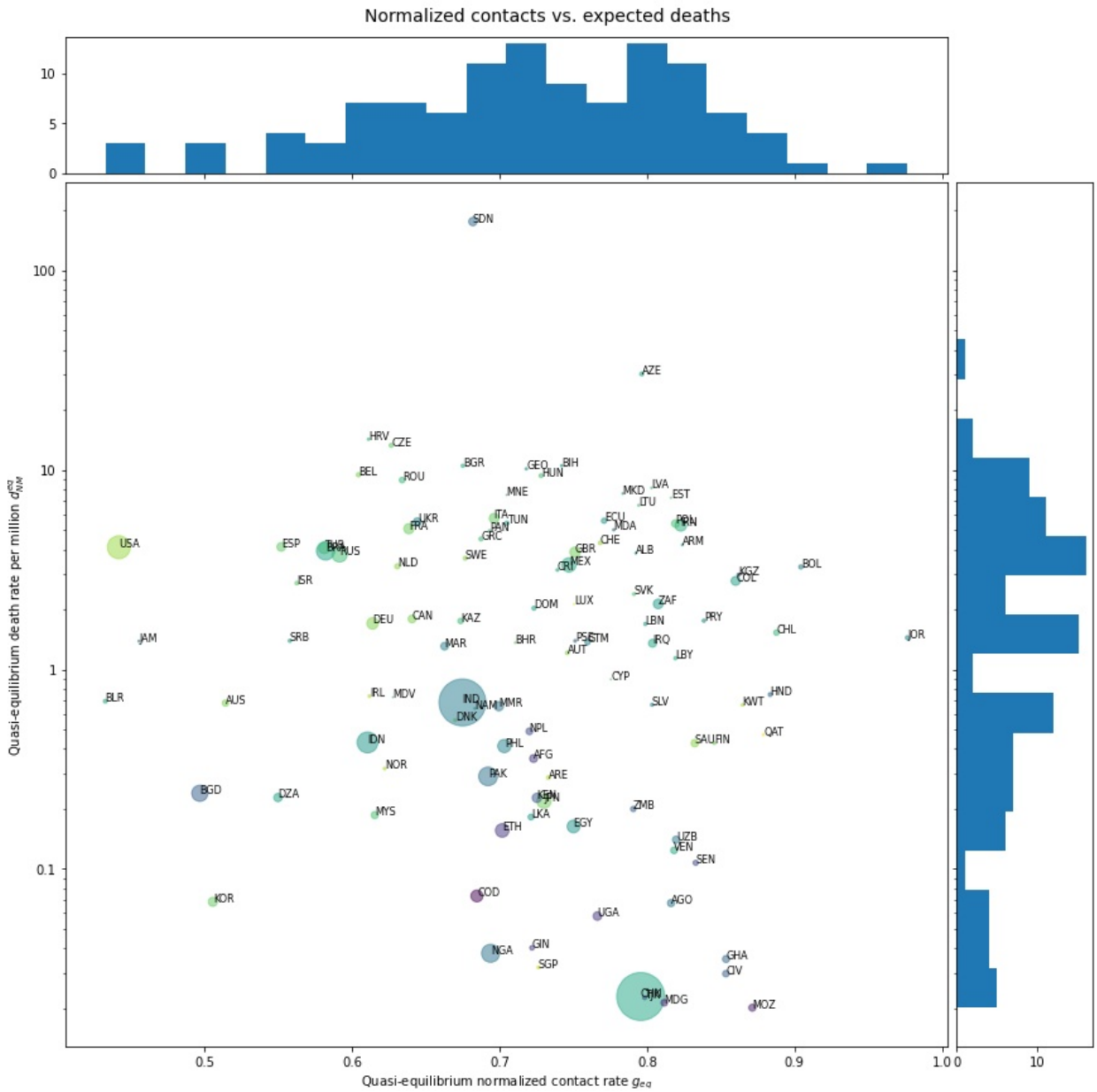


*Figure S3 - Normalized contacts vs. expected deaths for $\tau$ = 8 days*

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

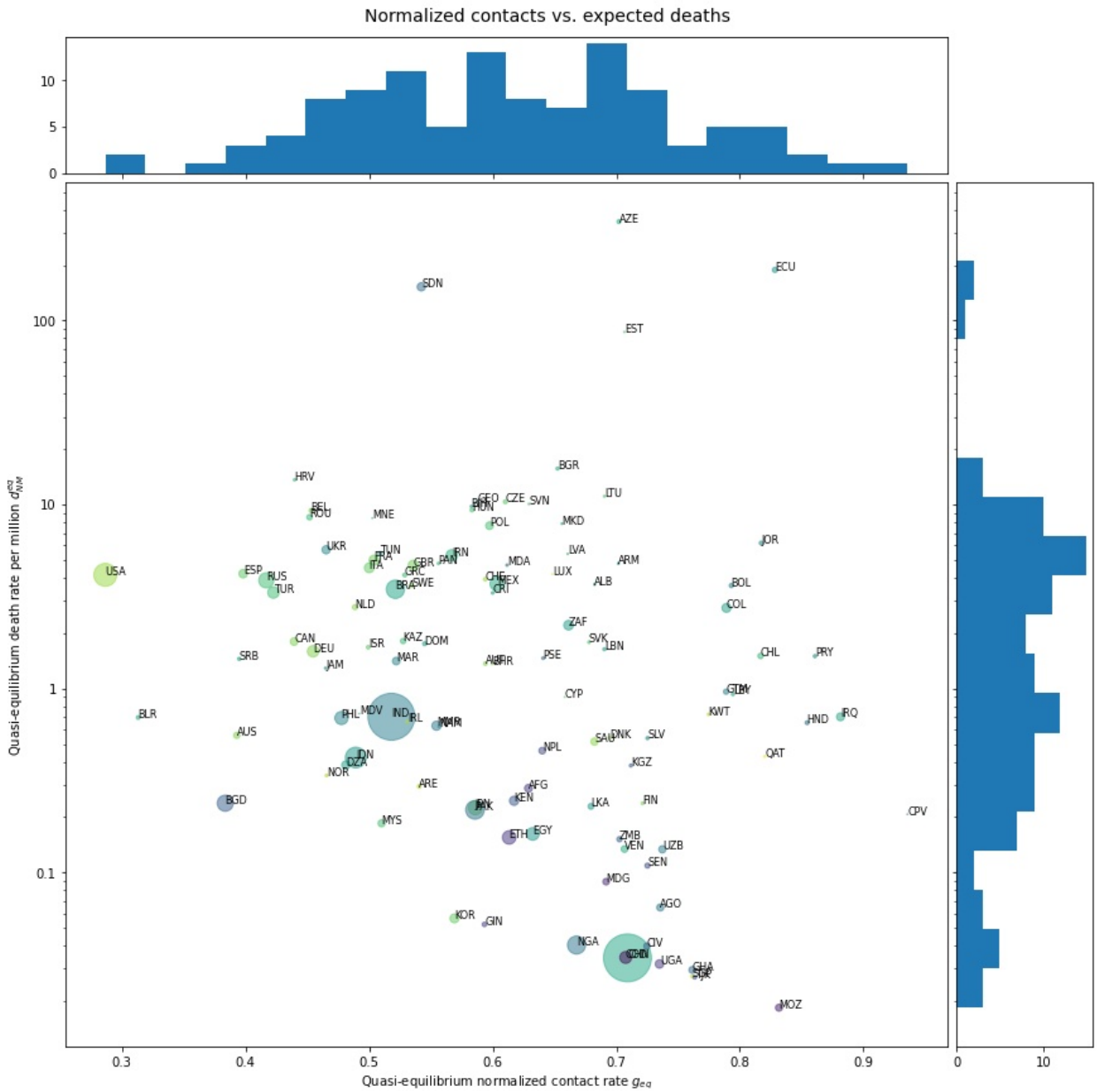


*Figure S4 - Normalized contacts vs. expected deaths for τ = 14 days*

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

*Table S2 - Summary of parameter estimates for $\tau$ = 8 days, with change in mean & median estimates*

| Parameter | Symbol | Mean | Change | Median | Change |
|---|---|---|---|---|---|
| Reference effective contact rate | $\beta 0$ | 0.20 | 0.187 | 0.19 | 0.195 |
| Initial responsiveness | $\alpha 0$ | 4.30 | 0.111 | 2.12 | 0.109 |
| Final responsiveness | $\alpha f$ | 1.30 | -0.114 | 0.21 | 0.098 |
| Responsiveness inflection point | $t0$ | 190 | -0.031 | 197 | -0.068 |
| Responsiveness scaling factor | $\theta$ | 14.1 | -0.127 | 5.0 | -0.012 |
| Time to upgrade risk | $\lambda U$ | 11.6 | -0.052 | 7.4 | -0.175 |
| Time to downgrade risk | $\lambda D$ | 60 | 0.036 | 63 | 0.030 |
| Death underreporting multiplier | $\gamma D$ | 2.42 | -0.039 | 2.21 | -0.059 |
| Likelihood scaling factor | $\epsilon$ | 0.046 | -0.202 | 0.017 | -0.205 |

*Table S3 - Summary of parameter estimates for $\tau$ = 14 days, with change in mean & median estimates*

| Parameter | Symbol | Mean | Change | Median | Change |
|---|---|---|---|---|---|
| Reference effective contact rate | $\beta 0$ | 0.14 | -0.16 | 0.13 | -0.174 |
| Initial responsiveness | $\alpha 0$ | 3.95 | 0.02 | 1.70 | -0.109 |
| Final responsiveness | $\alpha f$ | 1.65 | 0.13 | 0.35 | 0.848 |
| Responsiveness inflection point | $t0$ | 201 | 0.02 | 214 | 0.016 |
| Responsiveness scaling factor | $\theta$ | 13.3 | -0.18 | 5.0 | -0.013 |
| Time to upgrade risk | $\lambda U$ | 12.7 | 0.04 | 10.3 | 0.160 |
| Time to downgrade risk | $\lambda D$ | 59 | 0.02 | 61 | 0.007 |
| Death underreporting multiplier | $\gamma D$ | 2.66 | 0.06 | 2.85 | 0.214 |
| Likelihood scaling factor | $\epsilon$ | 0.060 | 0.04 | 0.030 | 0.386 |

### Estimation without behavioural response

We estimate the model with the endogenous behavioural response deactivated, i.e. $\alpha = 0$. In the absence of behavioural response, the fit of simulated infections to data deteriorates by 38% ($R^2$ = 0.914, MAEN = 18.2%), as expected, indicating that the behavioural response mechanism does improve the quality of fit. As the primary driver of infection rates is the number of currently infected people, which is calculated exogenously from the data, overall model fit remains notably good (Figure S5).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

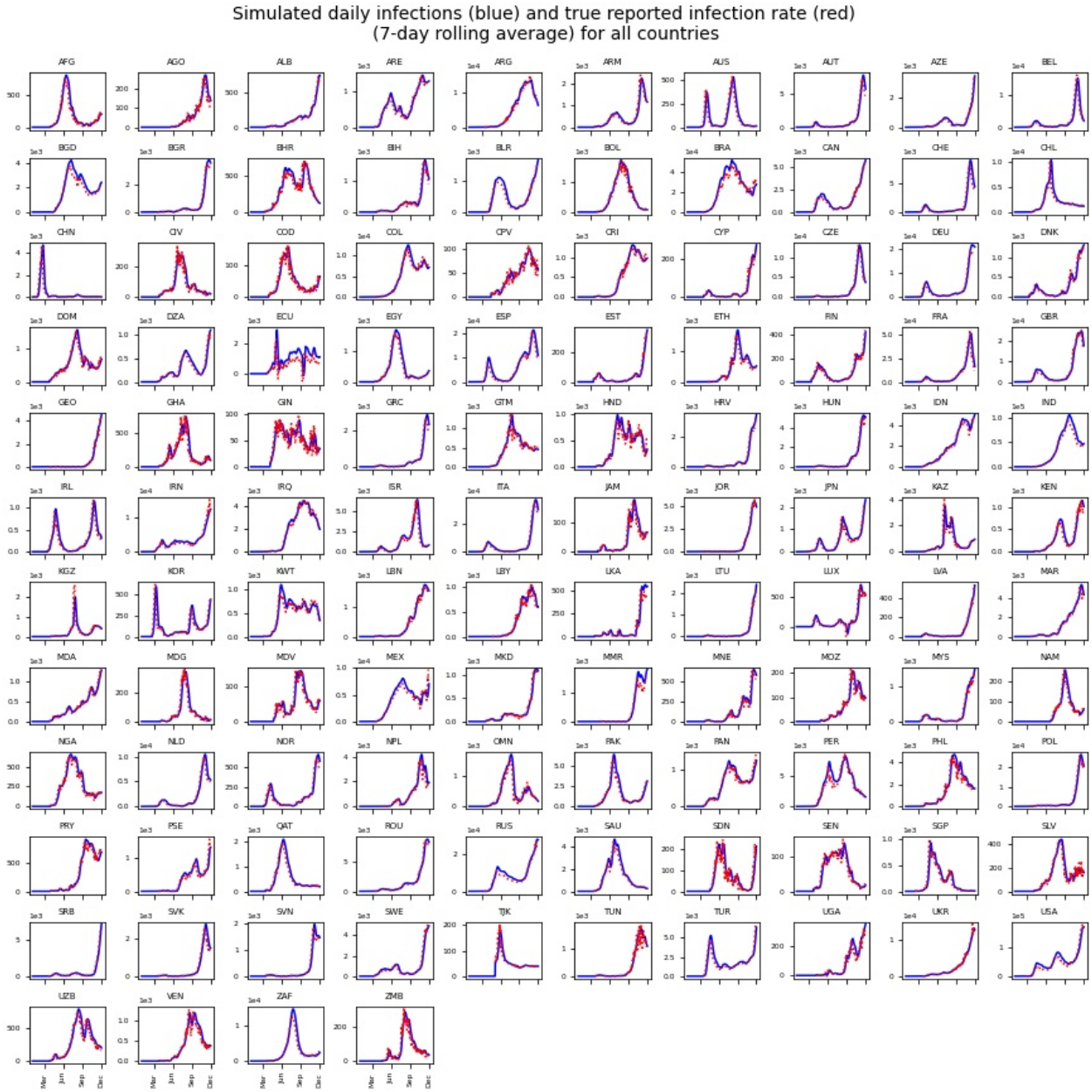


*Figure S5 - Comparison of simulated infection rates with data in the absence of behavioural risk response.*

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

## S5 – Model equations listing

1) AdjIFR[Rgn] = GET VDF CONSTANTS('InputConstants.vdf', 'AdjIFR[Rgn]', 1)
2) alp[Rgn] = 0.1 This parameter is 1 over the number of failures in negative binomial before experiment is stopped. A value between 0 and 1 (excluding zero) is legitimate calibraiton parameters here.
3) alpha[Rgn] = alpha 0[Rgn] + 1 / ( 1 + exp ( timesens[Rgn] ) ) * ( alpha f[Rgn] - alpha 0[Rgn] )
4) alpha 0[Rgn] = 1
5) alpha f[Rgn] = 2
6) BaseIFR = 0.005
7) beta[Rgn] = 0.1
8) CumulativeDpm[Rgn] = INTEG( DeathsOverTime[Rgn] , 0)
9) DataFlowOverTime[Rgn] = if then else ( new cases[Rgn] = :NA:, :NA:, new cases[Rgn] )
10) DataIncluded[Rgn] = 1
11) DataStartTimeCases[Rgn] = INITIAL( GET DATA FIRST TIME ( new cases[Rgn] ) )
12) DataStartTimeDeaths[Rgn] = INITIAL( GET DATA FIRST TIME ( new dpm[Rgn] ) )
13) DeathReportingRatio[Rgn] = 500
14) DeathsOverTime[Rgn] = if then else ( Time < DataStartTimeDeaths[Rgn] , 0, new dpm interpolated[Rgn] )
15) DeathsOverTimeRaw[Rgn] = if then else ( new dpm[Rgn] = :NA:, :NA:, new dpm[Rgn] )
16) Di[Rgn] = DataFlowOverTime[Rgn]
17) DiseaseDuration = 10
18) dn[Rgn] = SMOOTH N ( DeathsOverTime[Rgn] , if then else ( dn[Rgn] < DeathsOverTime[Rgn] , PMean[Rgn] , PMeanRelax[Rgn] ) ,0, PMeanOrder )
19) eps = 0.01
20) eqDeath[Rgn] = ZIDZ ( ln ( beta[Rgn] * DiseaseDuration * SFrac[Rgn] ) , alpha[Rgn] )
21) FINAL TIME = 334 The final time for the simulation.
22) g death[Rgn] = exp ( - alpha[Rgn] * dn[Rgn] )
23) IFR[Rgn] = INITIAL( if then else ( AdjIFR[Rgn] = -1, BaseIFR , AdjIFR[Rgn] ) ) Note: -1 is placeholder value for missing data in InputConstants.vdf
24) inf exp[Rgn] = beta[Rgn] * roll[Rgn] * g death[Rgn] * SFrac[Rgn]
25) InfShift[Shft] := - Shft
26) INITIAL TIME = 0 initial time for the simulation.
27) Mu[Rgn] = Max ( eps , inf exp[Rgn] )
28) NBL1[Rgn] = if then else ( DataFlowOverTime[Rgn] = 0, - ln ( 1 + alp[Rgn] * Mu[Rgn] ) / alp[Rgn] , 0) This is the part of negative binomial distribution calcualted when outcomes are zero.
29) NBL2[Rgn] = if then else ( DataFlowOverTime[Rgn] > 0, GAMMA LN ( Di[Rgn] + 1 / alp[Rgn] ) - GAMMA LN ( 1 / alp[Rgn] ) - GAMMA LN ( Di[Rgn] + 1) - ( Di[Rgn] + 1 / alp[Rgn] ) * ln ( 1 + alp[Rgn] * Mu[Rgn] ) + Di[Rgn] * ( ln ( alp[Rgn] ) + ln ( Mu[Rgn] ) ) , 0) This is the second piece in the loglikelihood for negative binomial which only applies to non-zero data points.
30) NBL3[Rgn] = if then else ( Di[Rgn] > 0, - GAMMA LN ( Di[Rgn] + 1) - ( Di[Rgn] + 1 / alp[Rgn] ) * ln ( 1 + alp[Rgn] * Mu[Rgn] ) + Di[Rgn] * ( ln ( alp[Rgn] ) + ln ( Mu[Rgn] ) ) , 0)
31) NBLLFlow[Rgn] = ( NBL1[Rgn] + NBL2[Rgn] ) * DataIncluded[Rgn]
32) new cases[Rgn] :RAW:
33) new dpm[Rgn] :RAW:
34) new dpm interpolated[Rgn] := new dpm[Rgn]
35) PMean[Rgn] = 5
36) PMeanOrder = 2
37) PMeanRelax[Rgn] = 20
38) Pssn : (p1-p100)
39) Re[Rgn] = beta[Rgn] * g death[Rgn] * DiseaseDuration * SFrac[Rgn]
40) Rgn : AFG, AGO, ALB, ARE, ARG, ARM, AUS, AUT, AZE, BEL, BGD, BGR, BHR, BIH, BLR, BOL, BRA, BWA, CAN, CHE, CHL, CHN, CIV, COD, COL, CPV, CRI, CYP, CZE, DEU, DNK, DOM, DZA, ECU, EGY, ESP, EST, ETH, FIN, FRA, GBR, GEO, GHA, GIN, GRC, GTM, HND, HRV, HUN, IDN,

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247924; this version posted December 14, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

IND, IRL, IRN, IRQ, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KGZ, KOR, KWT, LBN, LBY, LKA, LTU, LUX, LVA, MAR, MDA, MDG, MDV, MEX, MKD, MLT, MMR, MNE, MOZ, MYS, NAM, NGA, NLD, NOR, NPL, OMN, PAK, PAN, PER, PHL, POL, PRY, PSE, QAT, ROU, RUS, SAU, SDN, SEN, SGP, SLV, SRB, SVK, SVN, SWE, TJK, TUN, TUR, UGA, UKR, USA, UZB, VEN, XKX, ZAF, ZMB, ZWE

41) roll[Rgn] = if then else ( Time < DataStartTimeCases[Rgn] , 0, sum ( SelectRoll[Shft!] * ShiftedInfection[Rgn,Shft!] ) )
42) SAVEPER = TIME STEP [0,?] The frequency with which output is stored.
43) SelectRoll[Shft] = if then else ( Shft > DiseaseDuration , 0, 1)
44) Series : Infection
45) SFrac[Rgn] = Max ( 1e-06, 1 - ( CumulativeDpm[Rgn] * DeathReportingRatio[Rgn] / IFR[Rgn] ) / 1e+06)
46) Shft : (S1-S20)
47) ShiftedInfection[Rgn,Shft] := TIME SHIFT ( new cases[Rgn] , InfShift[Shft] )
48) t0[Rgn] = 20
49) theta[Rgn] = 1
50) TIME STEP = 1 [0,?] The time step for the simulation.
51) timesens[Rgn] = MIN ( 50, - ( Time - t0[Rgn] ) / theta[Rgn] )

## References

1. Vrugt JA, Ter Braak C, Diks C, Robinson BA, Hyman JM, Higdon D. Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *International Journal of Nonlinear Sciences and Numerical Simulation* 2009; **10**(3): 273-90.
2. Roser M, Ritchie H, Ortiz-Ospina E, Hasell J. Coronavirus pandemic (COVID-19). *Our World in Data* 2020.
3. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020; **20**(5): 533-4.
4. Rahmandad H, Lim TY, Sterman J. Estimating COVID-19 under-reporting across 86 nations: implications for projections and control. *Available at SSRN 3635047* 2020.
5. Verity R, Okell LC, Dorigatti I, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet infectious diseases* 2020.
6. World-Bank. World development indicators. 2014.
7. He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med* 2020; **26**(5): 672-5.
8. Wolfel R, Corman VM, Guggemos W, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature* 2020; **581**(7809): 465-9.