medRxiv preprint doi: https://doi.org/10.1101/2021.02.17.21251898; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 4.0 International license .

Title page

Lukas Boesch, University of Leipzig, Institute of Sociology

Lockdown benefit varies among countries and sub-national units: a reanalysis of the data by Bendavid et al. (2021)

Abstract

Are the lockdown measures limiting the propagation of COVID-19? Recent analyses on the effectiveness of non-pharmaceutical interventions in reducing COVID-19 growth rates delivered conflicting conclusions. While Haug et al. (2020) did find strong empirical support for reductions in COVID-19 growth rates, Bendavid et al. (2021) did not. Here, I present the results of a reanalysis of the data by Bendavid et al. (2021). Instead of relying on pairwise comparisons between 10 countries with fixed-effects regression models to isolate the effect of lockdown measures, I modelled the development of the pandemic with and without lockdown measures for the entire period and all countries included in the data with one mixed-effects regression model. My results reconciled the conflicting conclusions of Haug et al. (2020) and Bendavid et al. (2021): while no general decrease in COVID-19 growth rates was attributable to the implementation of lockdown measures, mandatory stay-at-home orders did lead to substantial decreases in COVID-19 growth rates in some countries and sub-national units. The heterogeneity in the effect of mandatory stay-at-home orders on the spread of COVID-19 is challenging from a scientific and political point of view.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.17.21251898; this version posted February 23, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 4.0 International license .

Introduction

Are non-pharmaceutical interventions (NPIs) helping to limit the spread of COVID-19 infections? Most importantly, are lockdown measures, the more restrictive NPIs (mrNPIs) based on mandatory stay-at-home and business closure orders, efficient in reducing COVID-19 growth rates? Two recently published analysis resulted in contradicting evidence. On the one hand, Haug et al. (2020) concluded that mrNPIs were the most effective NPIs but showed a considerable variation in effectiveness across countries. On the other hand, Bendavid et al. (2021) denied the possibility of large declines in COVID-19 growth rates due to mrNPIs: "While modest decrease in daily growth (under 30%) cannot be excluded, the possibility of large decreases in daily growth due to mrNPIs is incompatible with the accumulated data." (2021, p. 8).

In this paper, I present a reanalysis of the data used by Bendavid et al. (2021). I run one mixed-effects regression that compared the daily changes in COVID-19 growth rates across all countries and sub-national units included in the data to model the development of the pandemic with and without the implementation of mrNPIs. I believe that this approach provides a more accurate assessment of the effectiveness of mrNPIs in decreasing COVID-19 growth rates than the one taken in Bendavid et al. (2021), which relied on the unrealistic assumption that Sweden and South Korea are counterfactuals of England, France, Germany, Iran, Italy, the Netherlands, Spain, and the United States.

Methods

Bendavid et al. (2021) used an analytical framework that was based on the assumption that countries that did not implement mrNPIs were counterfactuals to countries that did implement mrNPIs: „Here, we use Sweden and South Korea as the counterfactuals to isolate the effects of mrNPIs in countries that implemented mrNPIs“ (Bendavid et al. 2021, p. 5). Based on this assumption, the authors run pairwise fixed-effects regression models to compare the combined effect size of all NPIs on the daily COVID-19 growth rates of countries that did implement mrNPIs (the treatment countries England, France, Germany, Iran, Italy, the Netherlands, Spain, and the United States) with the combined effect size on the daily COVID-19 growth rate of all NPIs of countries that did not implement mrNPIs (the control countries South Korea and Sweden). The analysis was based on a time-series of COVID-19 case counts at the sub-national unit level, matched with data on the implementation of a variety of NPIs during winter/spring 2020. The dependent variable was the daily difference in the natural log of the number of cumulated COVID-19 cases. Only cumulated daily differences equal or larger ten were considered.

The authors claimed that their approach allowed to isolate the effect of mrNPIs. A closer look, however, revealed that this was a misleading view. First, the classification of countries and their sub-national units either as control or treatment ignored the huge differences in the implementation of the two mrNPIs. For example, only three treatment countries implemented mandatory business closure, and this with varying intensity. Furthermore, while seven sub-national units in the USA did not implement mandatory home isolation at all, only the minority of French and Italian sub-national units did so with maximal intensity. Second, this approach did not allow to tease apart the effects of business closure and home isolation, with the risk of drawing the wrong conclusion: if one is positive and the other negative, the combined effect is zero. Finally, the treatment and control countries are not counterfactuals. Even if the additional benefit of implementing mrNPIs was neglectable in a direct comparison of countries, we still not know how daily COVID-19 growth rates would have evolved if the countries had/had not implemented mrNPIs.

I solved these issues by modelling the counterfactual outcomes for each country. Based on the data provided by Bendavid et al. 2021, I built one dataset that included all observations of all sub-national units of all ten countries of interest for the entire period available. In order to model a dynamic effect of the mrNPIs on the daily growth rate in COVID-19 cases, I generated the variable home, which was the cumulated sum of days after a sub-national unit implemented home isolation, weighted with the intensity of implementation, and the variable busi, which was the cumulated sum of days after a sub-national unit implemented business closure, weighted with the intensity of

medRxiv preprint doi: https://doi.org/10.1101/2021.02.17.21251898; this version posted February 23, 2021. 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 4.0 International license .

implementation. To control for the unobserved heterogeneity and model the time trend, I computed the variable days, which simply was a count of the epidemic age, in days, after a sub-national unit had reached cumulative confirmed cases of 10. I log transformed all three variables (home and busi after the addition of one). Finally, I narrowed down the extent of the time series such as to include the same time period as in Bendavid et al. (2021). In total, the data set used for the reanalysis included 5324 observations, clustered within 10 countries and 109 sub-national units. The dates of the time series spanned from 18.02.2020 to 06.04.2020. The number of observations at the sub-national units level ranged from three to 49 (mean=25.47; sd=8.33). Epidemic age ranged from one to 67 days (mean=15.71; sd=10.38) Weighted home isolation days from zero to 29 (mean=4.84; sd=6.83) and weighted business closure days from zero to 32.75 (mean=2.87; sd=5.56).
I modelled the daily growth rate of COVID-19 cases, g_{ci}, as a function of a deviation from an overall mean B_0. The overall mean varied across countries (country random intercept RI_{0c}) and sub-national units (units random intercept RI_{0i}). Furthermore, g_{ci} was also determined by an overall time trend, B_{0days}, which also varied across countries (country random slope RS_{cdays}) and sub-national units (units random slope RS_{idays}). These 6 terms captured the entire variation in g_{ci} due to the countries, the sub-national units and their specific time trends that included the combined effects of all implemented NPIs (excluding home isolation and business closure). This was the baseline model. By adding a fixed effect for business closure (B_{0busi}) and a fixed effect as well as random slopes for home isolation (B_{0home}; RS_{chome}; RS_{ihome}), the effect of home isolation and business closure on the change in daily COVID-19 growth rates was estimated, while controlling the country and sub-national unit specific time trend. This was the full model (equation 1). I did not include random slopes for B_{0busi} as only the minority of countries and sub-national units implemented business closure. I then used the model estimates to predict the evolution of daily COVID-19 growth rates for each country in both counterfactual conditions (no implementation of mrNPIs vs. maximal implementation of mrNPIs).

$$g_{ci} = B_0 + RI_{0c} + RI_{0i} + B_{0days} + RS_{cdays} + RS_{idays} + B_{0home} + RS_{chome} + RS_{ihome} + B_{0busi} + e_{ci}$$

equation 1. The mixed-effects regression model used to estimate the effect of mrNPIs.

The model in equation 1 was estimated with the glmer function of the lme4 package (Bates et al. 2015) in R (R Core Team 2019). I used a gamma error distribution with a log link to keep estimated values in the range of defined values (exclusion of negative values) and to avoid violating distributional assumptions made for a model with Gaussian error distribution. Prior to estimation, growth rates equal to zero were replaced with the smallest growth rate observed, divided by ten (6.78196e-05). P-values for the fixed effects were obtained by comparing the fit of the full model with the fit of reduced models excluding single fixed effects. The multicollinearity among predictors was assessed by applying the vif function of the car package (Fox & Weisberg 2019) to a multiple linear regression model that only included the fixed effects from the full model (appendix 1). I ran a model stability analysis, based on the countries, to check for robustness of estimates (appendices 2). I checked for violations of distributional assumptions of the random terms with histograms (appendices 3) and calculated the dispersion parameter of the model (appendix 4). The significance of the combined contribution of both mrNPIs on the daily change in COVID-19 growth rate was assessed by comparing the fit of the full model with the fit of a model excluding the fixed effects of the predictors for the mrNPIs. P-values for the random slopes were computed by comparing the fit of the full model with the fit of models excluding single random slopes.

Results

The estimates for home isolation and business closure were both negative and non-significant. On the other hand, their combined inclusion into the model significantly increased the fit and the predictive power of the model (Chisq.: 7.742; Df: 2; Pval.: 0.021). All random slopes, as well as the fixed effect for the epidemic age, were highly significant (table 1). While the model predictions for the counterfactual control outcome did not differ from the model predictions for the counterfactual

medRxiv preprint doi: https://doi.org/10.1101/2021.02.17.21251898; this version posted February 23, 2021. 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 4.0 International license .

treatment outcome in eight countries (above South Korea and Sweden this was also the case for England, France, Iran, the Netherlands, Spain and the USA), the model predictions suggested that Germany and Italy clearly decreased daily COVID-19 growth rates by implementing mandatory home isolation. For both countries, predicted COVID-19 growth rates were higher at the start of the implementation, compared to the same period without implementation of home isolation. In the end, predicted COVID-19 growth rates were lower after the implementation of home isolation than otherwise (figure 1).

However, even if the effect of home isolation was zero on average for a specific country, this did not mean that home isolation did not affect the COVID-19 growth rates within this country, as the effect for the sub-national units differed significantly from the country average (table 1). Of course, this also holds for an average negative effect. For example, in Italy, the overall effect of home isolation on COVID-19 growth rates was negative. Nevertheless, not all Italian regions were able to decrease daily COVID-19 growth rates through the implementation of home isolation to the same extent. In the regions of Emilia Romagna and Marche, the predicted effect of home isolation on COVID-19 growth rates was zero. In Sardegna, Lombardia and Veneto, it was slightly below average. In Toscana, Lazio, Campania and Liguria it was slightly above average and in Piemonte, the predicted effect of home isolation on the decrease in daily COVID-19 growth rates was much stronger than on average (figure 2).

Table 1.Result of the estimation of the full model (equation 1)

Random effects			
Group	Name	Std. Dev.	P-val.
Sub-national unit	RI	0.113	Not applicable
	RS: days	0.163	<0.001
	RS: home	0.125	<0.001
Country	RI	0.393	Not applicable
	RS: days	0.150	<0.001
	RS: home	0.179	<0.001
Number obs: 5324, Number countries: 10, Number sub-national units: 209			
Fixed effects			
Name	Estimate	Std. Err.	P-val.
Intercept	-1.895	0.002	Not applicable
days*	-0.252	0.002	<0.001
home#	-0.166	0.002	0.194
busi#	-0.118	0.002	0.052

Random effects correlations were estimated but are not shown for purpose of clarity; * log transformed; # log transformed after addition of one.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.17.21251898; this version posted February 23, 2021. 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 4.0 International license .

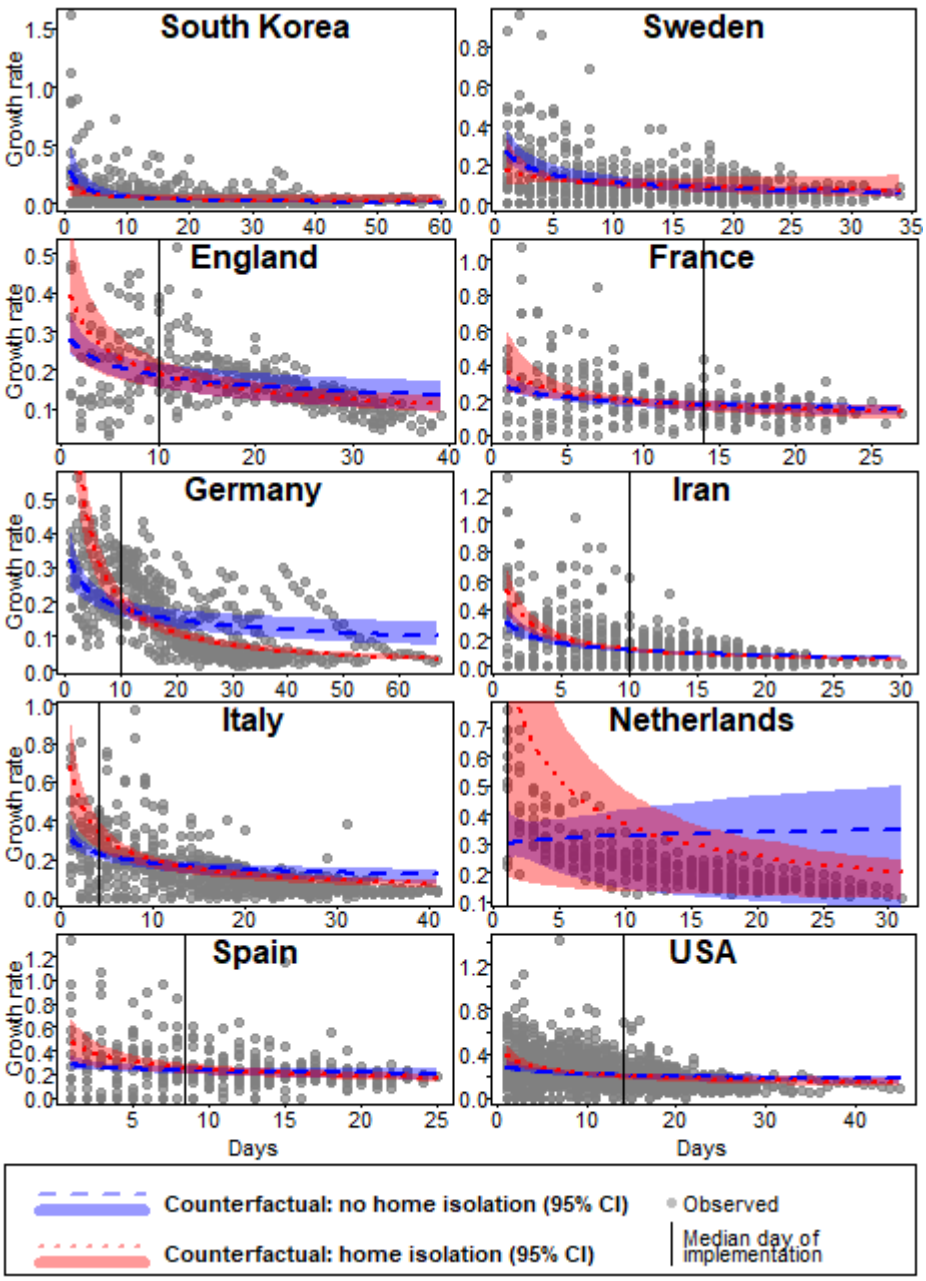


Figure 1. Modelled counterfactual daily COVID-19 growth rates for control (no implementation of home isolation) and treatment (implementation of home-isolation from the beginning and with full intensity), by country and for the entire period of the time series. All model predictions are for average days with business closure. The predictions for the counterfactual control outcome were calculated by building a sequence of days covering the entire period of the time-series of the country and multiplying each value of this sequence with the fixed effect of epidemic age and the country-specific effect of epidemic age. The population intercept and the country-specific intercept were added to the resulting values. The predictions for the counterfactual treatment outcome were calculated by adding the product between the sequence of days with the fixed effect of home isolation and the country-specific effect of home isolation to the control outcome. 95% confidence intervals were computed based on the predictions of 1000 bootstrapped models. For South Korea and Sweden, the model predictions for the counterfactual treatment outcome entirely lacked data support. For all other countries, the model predictions for the counterfactual treatment outcome prior to the implementation of the measure (left of the vertical line) were an extrapolation of the predicted trend after implementation.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.17.21251898; this version posted February 23, 2021. 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 4.0 International license .

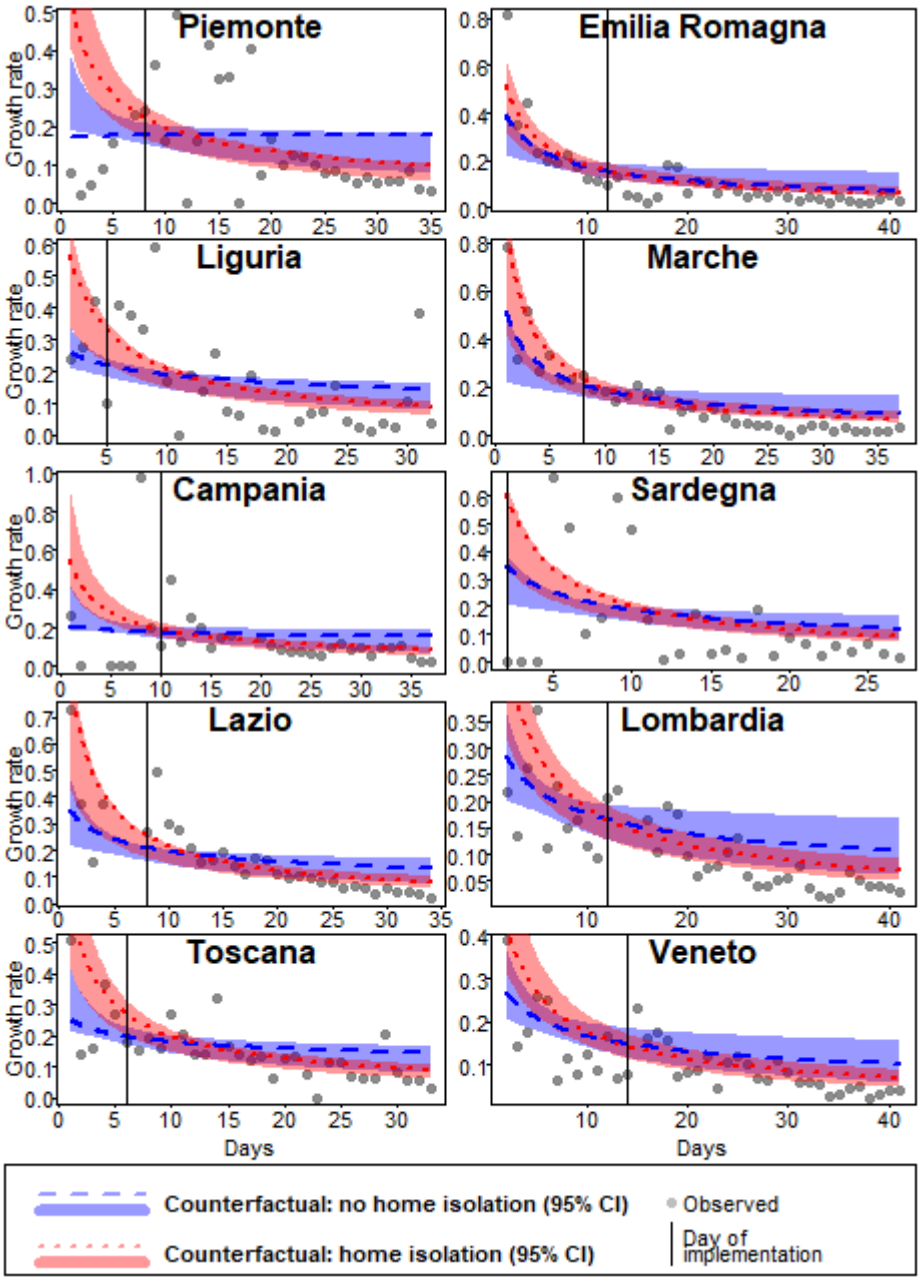


Figure 2. Modelled counterfactual daily COVID-19 growth rates for control (no implementation of home isolation) and treatment (implementation of home-isolation from the beginning and with full intensity), for the entire time series of the ten Italian regions with the strongest deviations from the country average. All model predictions are for average days with business closure. The predictions for the counterfactual control outcome were calculated by building a sequence of days covering the entire time-series of the region and multiplying each value in this sequence with the fixed effect of epidemic age, the country-specific effect of epidemic age and the region-specific effect of epidemic age. The overall intercept, the country-specific intercept and the region-specific intercept were added to the resulting values. The predictions for the counterfactual treatment outcome were calculated by adding the product between the sequence of days with the fixed effect of home isolation, the country-specific effect of home isolation and the region-specific effect of home isolation to the values of the control outcome. 95% confidence intervals were computed based on the predictions of 1000 bootstrapped models. For all regions, the model predictions for the counterfactual treatment outcome prior to the implementation of the measure (left of the vertical line) were an extrapolation of the predicted trend after implementation.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.17.21251898; this version posted February 23, 2021. 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 4.0 International license .

Conclusion

In this paper, I presented a reanalysis of the data used by Bendavid et al. (2021). On the one hand, I replicated the main finding of the authors: the data do not support the hypothesis that the implementation of mrNPIs leads to a general and strong decline in daily COVID-19 growth rates. Although the estimates for both mrNPIs (home isolation and business closure) were negative, suggesting that COVID-19 growth rates did decline with the implementation of the measures, the statistical ground for rejecting the null-hypothesis was too weak. Considering the fundamental difference in analytical approaches (16 pairwise fixed-effects regression models on 16 different data subsets (Bendavid et al. 2021) vs. one mixed-effects regression model on one data set), this zero-finding can be considered as robust.

On the other hand, the average effect of home isolation showed a statistically significant variation across levels of countries and sub-national units. The inclusion of home isolation random slopes at both levels substantially increased the predictive power of the model. While for few countries and sub-national units, the predicted effect of home isolation was strongly negative, in most countries and sub-national units the predicted effects were moderately negative or zero and in few sub-national units, predicted COVID-19 growth rates even increased after the implementation of home isolation. The heterogeneity of the effectiveness of NPIs, their interaction with the countries of implementation, was a central finding of the study by Haug et al. (2020). Here, I showed that this heterogeneity was also present in the data by Bendavid et al. (2021) and also applies to the sub-national unit levels.

To some extent, my findings are recomforting, as they allow to reconcile the opposing conclusions by Haug et al. (2020) and Bendavid et al. (2021). However, this reconciliation results from an artificial methodological separation between the average treatment effect, on the one side, and the heterogeneity of the treatment effect, on the other side. From a theoretical point of view, as long as there is a systematic variation of the average treatment effect, which is based on some population clusters, the treatment does have an effect on the outcome, even when zero on average. Furthermore, as long as the mechanisms driving the interaction between population clusters and the treatment effect have not been uncovered, it remains unknown how the treatment truly works.

This is the position we are in now. We know that there is no substantial average effect of home isolation on the COVID-19 growth rates. However, we also know that home isolation can be effective in decreasing daily COVID-19 growth rates in some countries and sub-national units. As long as the mechanisms responsible for the interaction between home isolation and the social context have not been uncovered, it will not be possible to predict where and when home isolation will lead to decreased COVID-19 growth rates. Moreover, our knowledge is based on a deprecated database. Most countries in the world experience the 2nd or even 3rd wave of the COVID-19 pandemic, with records of implemented NPIs that span periods of up to one year. The first wave is a selective and unrepresentative sample of the epidemic. The assessment of the effectiveness of NPIs must be conducted as soon as possible with an updated database that includes all phases of the pandemic and not only the beginning.

Acknowledgement

I thank Roger Mundry for his valuable advices and support, Christophe Boesch for reviewing my manuscript and Roger Berger for his general support.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.17.21251898; this version posted February 23, 2021. 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 4.0 International license .

Literature

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Bendavid, E., Oh, C., Bhattacharya, J., & Ioannidis, J. P. A. (2021) Assessing mandatory stay at home and business closure effects on the spread of COVID-19. *European Journal of Clinical Investigation.* https://doi.org/10.1111/eci.13484

Fox, J., & Weisberg, S. (2019). An {R} Companion to Applied Regression, Third Edition. Thousand Oaks CA: Sage. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Haug, N., Geyrhofer, L., Londei, A., Dervic, E., Desvars-Larrive, A., Loreto, V., Pinior, B., Thurner, S., & Klimek, P. (2020) Ranking the effectiveness of worldwide COVID-19 government interventions. *Nature Human Behavior.* https://doi.org/10.1038/s41562-020-01009-0

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.17.21251898; this version posted February 23, 2021. 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 4.0 International license .

Appendix

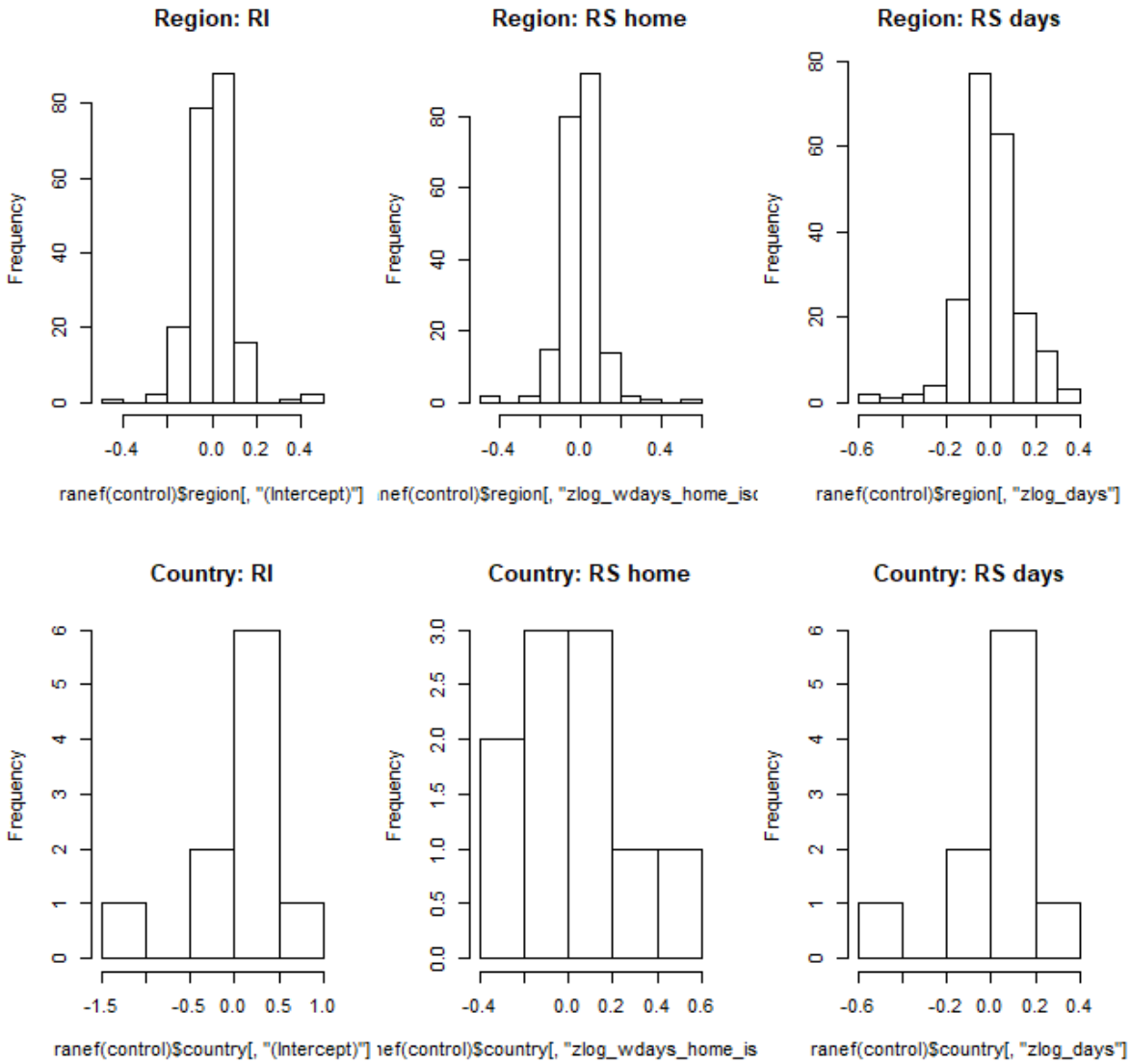
1) VIF values among predictors in the model from table 1.

days	home	busi
1.381	1.436	1.134

2) Range of estimates from the model in table 1 when excluding single countries.

	Min	original	max
(Intercept)	-2.100	-1.895	-1.607
days	-0.339	-0.252	-0.180
home	-0.241	-0.166	0.459
busi	-0.315	-0.118	-0.090

3) Distributions of random terms from the model in table 1.



4) Dispersion parameter of the model in table 1.
dispersion parameter: 0.521