The association between containment measures implemented by countries across the world and excess mortality linked to Covid-19: a study based on functional data analysis

Chiara Micheletti[1], Marco Stefanucci[2], Stefano Mazzuco[1]

[1]Università degli Studi di Padova
[2]Università degli Studi di Trieste

26 March 2021

Abstract

In 2020, the world had to face Covid-19 and its tragic consequences. To limit the spread of the new virus, governments imposed restrictions, aiming at reducing risks for disease transmission. This work investigates the association between these policy interventions, measured by the Stringency Index, and the increase in the number of deaths recorded by most countries, measured by excess mortality, between January and August 2020. A functional data approach that envisages a smoothing procedure (B-splines) combined with function-on-function regression has been adopted. The fitted models suggest that restrictions need from four to five weeks to show their effects and they indicate the stay-at-home requirement as the most effective policy.

1 Introduction

Starting from January 2020, the virus known as SARS-CoV-2, causing Covid-19, spread rapidly across the globe, affecting people in different ways: a person infected could develop mild to severe symptoms, including, in the worst case scenario, death. To reduce the diffusion of the virus, governments implemented non-pharmaceutical interventions (NPIs) such as school and workplace closings, restrictions on internal and external movements and so on.

Our work aims at studying if and how these policies limited the surge in death recorded by most countries, from January to August 2020.

Many studies (see for example Li et al. (2021), Flaxman et al. (2020) or Liu et al. (2021)) focus their analysis on the effects of NPIs on the time-varying reproduction number (R_t). We prefer excess mortality, introduced as response variable in our models, since data on reported Covid-19 infections, hospitalisations and deaths are often inaccurate. Excess mortality measures the difference between the observed number of deaths *from all causes* in a specific time period and the numbers of deaths in a comparison period. Considering all causes of death, it includes those directly and indirectly attributed to Covid-19 (for example due to pressures on the health care system), giving a non-biased picture of the impact the pandemic had on countries, also making them more comparable. Kontis et al. (2020) show that indirect effects of the pandemic heavily contributed to the increase of mortality levels and need to be taken into consideration.

To track governments' responses to the health emergency, we use one of the indicators of the the Oxford Coronavirus Government Response Tracker (`OxCGRT`), constructed by the Blavatnik School of Government (see Hale et al. (2021)): the Stringency Index. It is a composite measure based on nine response sub-indices including school closures, workplace closures, gatherings and travel bans, rescaled to a value from 0 to 100 (100 = strictest).

We adopt a functional data approach, smoothing the data with B-splines. Functional analysis allows us to see how government measures influenced the raise of mortality levels, and, more importantly, how this influence changed over time. Thanks to the flexible tools it offers, we are also able to pinpoint the time-lag that policies require to be effective. In the framework of penalized function-on-function regression proposed by Ivanescu et al. (2015), we use a historical functional linear model (Malfait and Ramsay (2003)), evaluating different time windows in which to consider the effect of covariates on the response. We also include new daily Covid-19 cases, as a confounding variable.

2 Data

Excess mortality is measured by a weekly P-score: data are sourced from Our World In Data (2020) and consist of combined information from the Human Mortality Database (2020) Short-term Mortality Fluctuations project and the World Mortality Dataset (Karlinsky and Kobak (2021)).

The P-score calculates excess mortality as the percentage difference between the number of deaths in a given week in 2020 and the average number of deaths in the same week over the years 2015–2019.

$$\text{P-score}_{\text{week}} = \frac{\text{Deaths}_{\text{week2020}} - \text{AverageDeaths}_{\text{week2015-2019}}}{\text{AverageDeaths}_{\text{week2015-2019}}} \cdot 100$$

Our World in Data also supplies the number of new daily Covid cases. Data for England & Wales, Scotland and Northern Ireland were gathered from The official UK government website for data and insights on coronavirus (2020).

The Blavatnik School of Government provides daily observations for the Stringency Index and the sub-indices that compose it: school closings, workplace closings, cancellation of public events, restrictions on gatherings, public transport closings, stay-at-home orders, restrictions on internal movements, restrictions on international travel and presence of public information campaigns. Full descriptions of the policy indicators, their meaning and how they are aggregated can be found in the OxCGRT's codebook (2020a) and in the index methodology (2020b).

Data are collected for a total of 36 countries. Each country has 35 weekly observations for excess mortality (from the first week of January 2020 to the last week of August 2020) and 236 daily observations for the Stringency Index, its sub-indices and the number of new Covid cases (from 1 January 2020 to 23 August 2020).

3 Methods

To assess the association between excess mortality and containment measures, we use a historical functional linear model

$$Y(t) = \beta_0(t) + \int_{s_1}^{s_2} X(s)\beta_1(s,t)ds + \epsilon_i(t) \tag{1}$$

where $X(s)$ is a functional covariate, $s \in [s_1, s_2]$, and $Y(t)$ is the functional response, $t \in [t_1, t_2]$. The regression coefficient β_1 is a surface and the value $\beta_1(s,t)$ determines the impact of X at time s on Y at time t. Model (1) describes the simplest situation, but multiple functional predictors can be considered (see Ivanescu et al. (2015)).

We fit two historical models, with excess mortality as response variable, and two different sets of covariates: the first with the Stringency Index in its aggregated form, the natural logarithm of new daily cases and their interaction; the second with the sub-indices and the natural logarithm of new daily cases. In both models, we also examine different integration ranges $[s_1, s_2]$.

Functional data analysis involves a smoothing procedure to construct the functional observations from the discrete data and filter out measurement errors and noise. Before fitting the models, we use the `R` package `fda` (Ramsay et al. (2020)) to smooth excess deaths with 8 cubic B-splines with equally spaced knots, the Stringency Index with 25, and the natural logarithm of the number of new cases with 20. In order to evaluate different time lags over which to use the covariates information, we assume the sub-indices (which can be seen as categorical variables) to vary continuously over time. To limit the inaccuracy of this assumption we do not smooth the sub-indices data.

4 Preliminary Results

We fitted our models in `R` (2020) with the `pffr` function of the `refund` package (Goldsmith et al. (2020)). As mentioned in the previous section, we regressed excess mortality evaluating several intervals of integration and using two sets of mean centered explanatory variables: the first one with the Stringency Index, the second with the sub-indices considered as single predictors.

As in the general model (1), let t index time in weeks for excess deaths and s index time in days for the covariates. As a measure of accuracy of the fit, we used the functional R^2 (fR^2) defined by Ivanescu et al. (2015): with both sets of

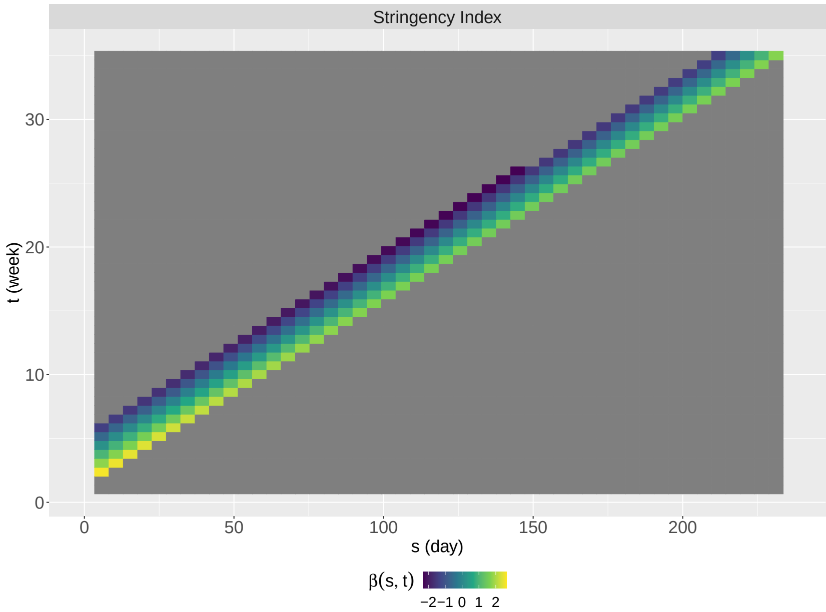


Figure 1. Estimated surface for the Stringency Index's coefficient

regressors, we found that the time window $[s_1, s_2] = [t-5, t-2]$ gave the best goodness of fit. A comparison of the predictions of excess mortality produced by the two models with this integration range is presented at the end of the paper, in Figure 3.

The estimate of the Stringency Index's coefficient is shown in Figure 1. From January to mid-February 2020 ($1 < s < 49$) and from June to the end of August 2020 ($s \geq 151$), the surface gradually shifts from yellow-green to blue: the influence of policies on the number of deaths decreases as the temporal lag increases, until it becomes zero. This is because in these months the numbers of infections and deaths were limited, and few ore no restrictions were implemented. From mid-February to the end of May 2020 ($50 < s < 150$), the coefficient goes from being slightly positive in the week following the day on which the index is taken into account, to being strongly negative between four and five weeks later. The change of sign, from positive to negative, suggests that in this period the tightening of restrictions gradually curbed the increase in excess mortality, with a greater effect after a month.

The model with the sub-indices identified the stay-at-home requirement as the most effective measure (Figure 2, bottom-left). Its coefficient, among all, reaches the greatest absolute value with negative sign. The model also showed that less drastic measures, in the most critical months (March-April), contributed to limiting excess mortality. Figure 2 exhibits the surfaces we obtained for school closings, restrictions on gatherings and international travel controls.

5 Discussion and further extensions

The functional approach allowed us to study how the impact of non-pharmaceutical interventions on the surge of mortality levels evolved and changed over time. In

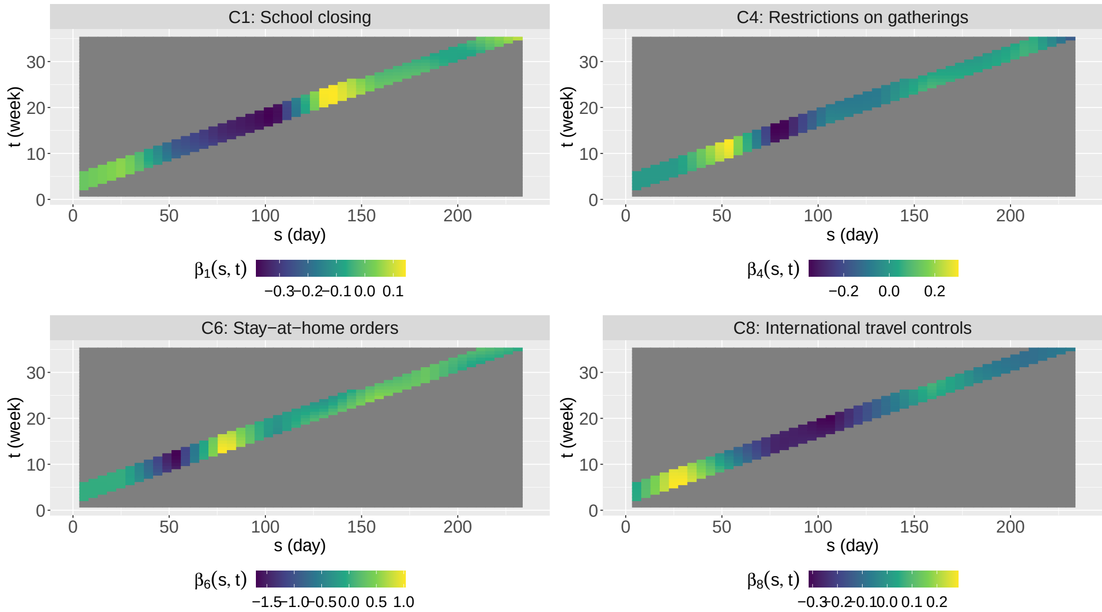


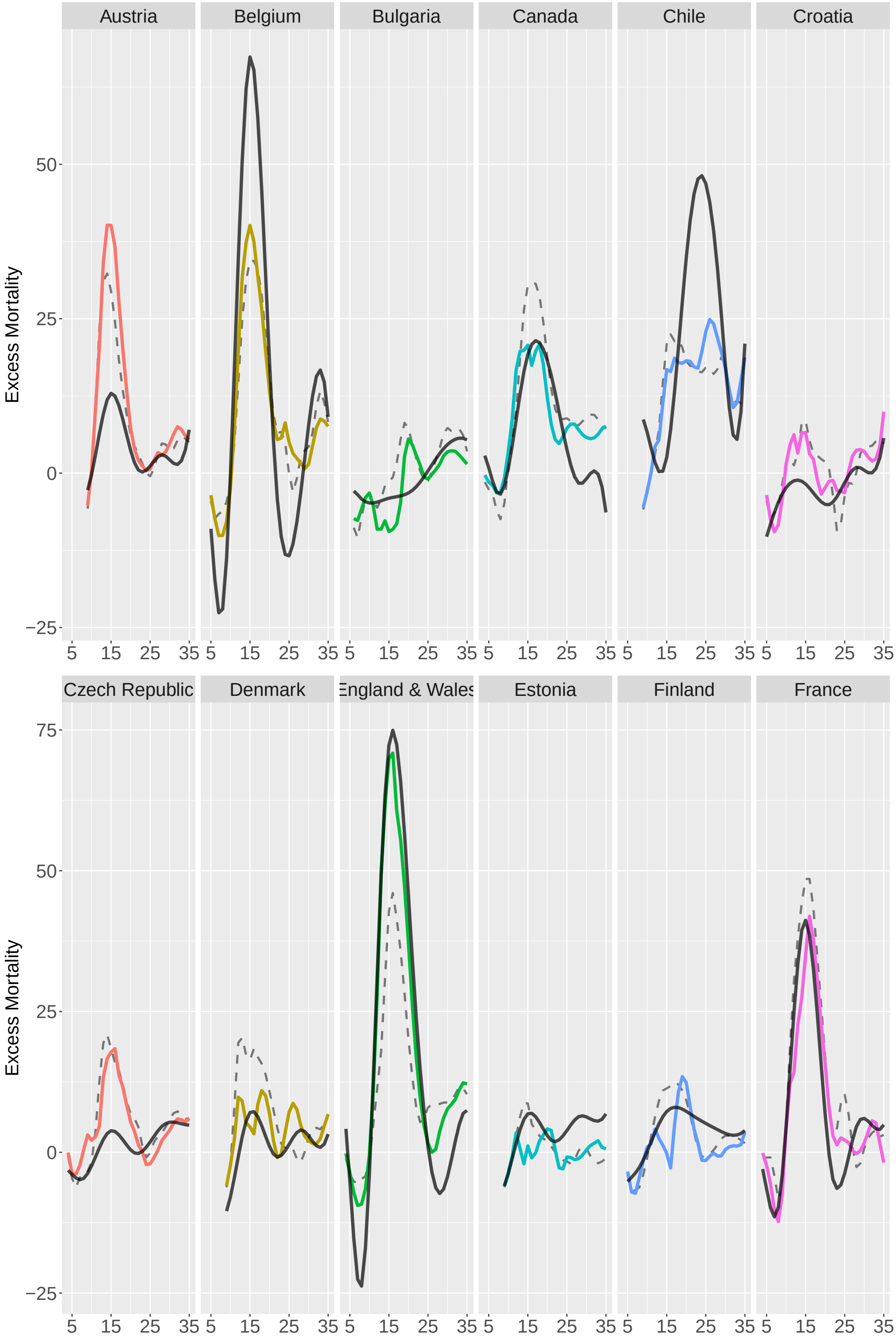
Figure 2. Estimated surfaces for some of the Stringency Index's sub-indices. Darkest areas highlight the period in which every restriction was most effective.

particular, we were able to show that policies' effect increases progressively in the weeks that follow their introduction, becoming more intense one month later.

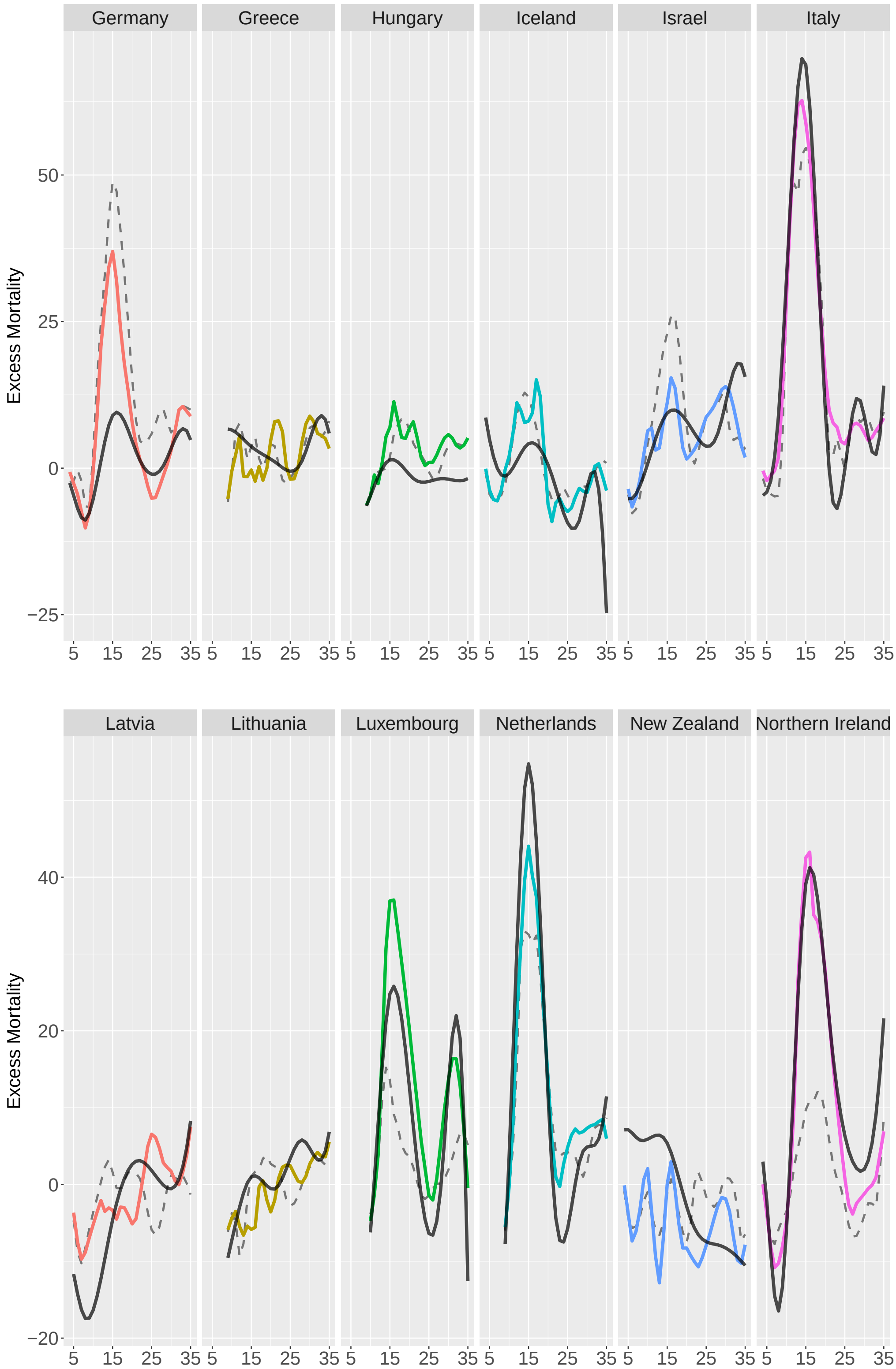
Our sub-indices model confirmed that confining people into their homes has been the most effective solution, as it prevented almost all interactions, making it nearly impossible for the virus to spread, reducing the number of infected people and, consequently, the number of deaths. It also indicated that less strict measures played an important role in curbing the pandemic. It is an important result, especially because imposing self-isolation is only possible for a limited period of time, becoming unsustainable in the long term at an individual, social and economic level.

We are currently working to extend our analysis to Covid-19 second wave, between September and December 2020. We expect the situation to be slightly different, thanks to a deeper knowledge of the disease and governments more prepared to face it.

We also intend to improve our models and how we dealt with the sub-indices variables, looking for example at the work of Preda et al. (2020) on categorical functional data.



Austria
Belgium
Bulgaria
Canada
Chile
Croatia
Excess Mortality
50
25
0
−25
5
15
25
35
Czech Republic
Denmark
England & Wales
Estonia
Finland
France
75


Germany
Greece
Hungary
Iceland
Israel
Italy
Excess Mortality
50
25
0
−25
5
15
25
35
Latvia
Lithuania
Luxembourg
Netherlands
New Zealand
Northern Ireland
Excess Mortality
40
20
0
−20

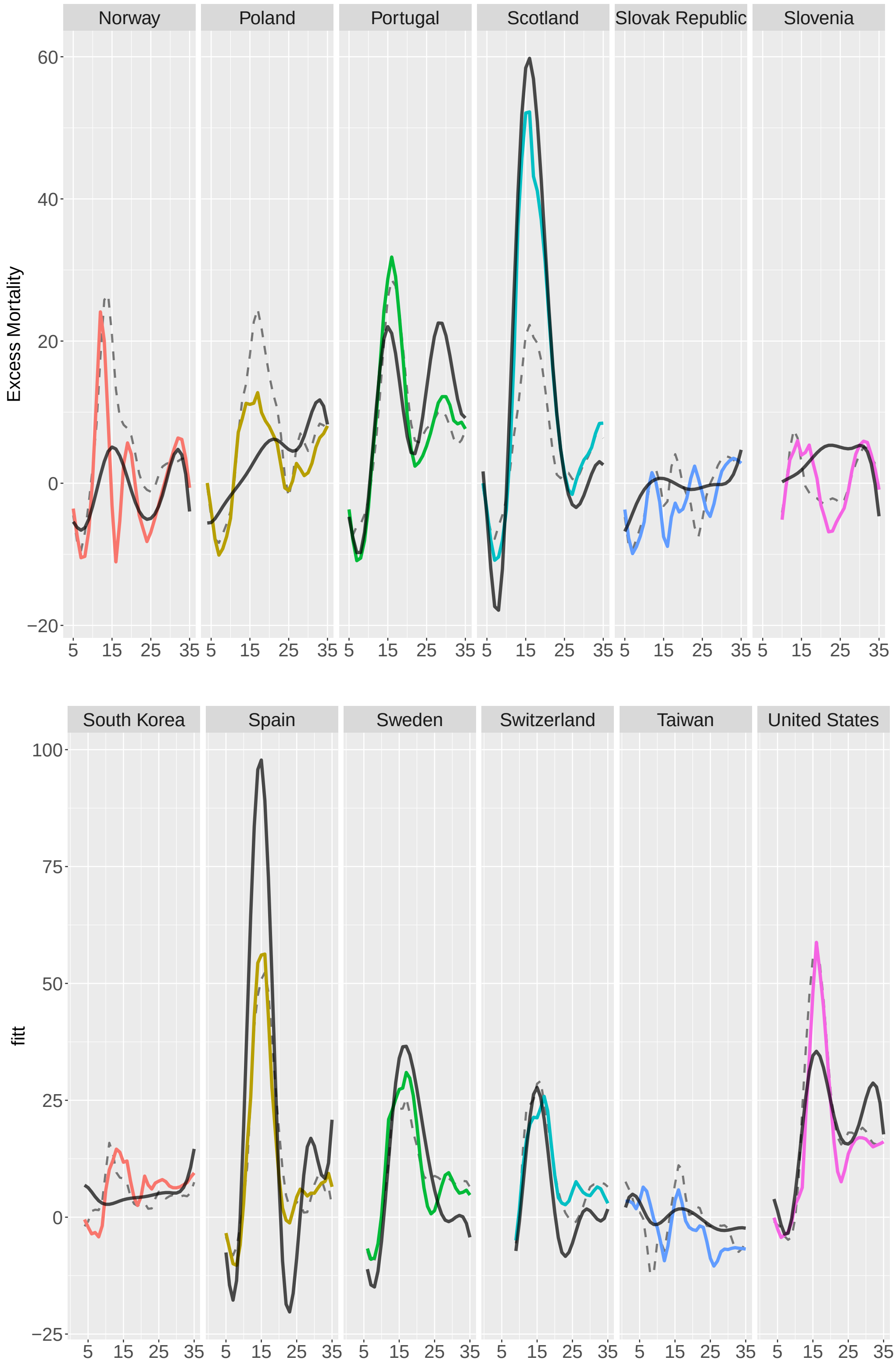


Figure 3. Original data (solid black) and predictions of excess mortality carried out with the Stringency Index model (dashed black) and the sub-indices model (solid coloured).

References

Blavatnik School of Government. University of Oxford (2020a). Codebook for the Oxford Covid-19 Government Response Tracker. Available at `www.github.com/OxCGRT/covid-policy-tracker/blob/master/documentation/codebook.md`.

Blavatnik School of Government. University of Oxford (2020b). Methodology for calculating indices. Available at `www.github.com/OxCGRT/covid-policy-tracker/blob/master/documentation/index_methodology.md`.

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J. W., et al. (2020). Estimating the effects of non-pharmaceutical interventions on covid-19 in europe. *Nature*, 584(7820):257–261.

Goldsmith, J., Scheipl, F., Huang, L., Wrobel, J., Di, C., Gellar, J., Harezlak, J., McLean, M. W., Swihart, B., Xiao, L., Crainiceanu, C., and Reiss, P. T. (2020). *refund: Regression with Functional Data*. R package version 0.1-23.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., et al. (2021). A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature Human Behaviour*, 5(4):529–538.

Human Mortality Database (2020). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at `www.mortality.org`.

Ivanescu, A. E., Staicu, A.-M., Scheipl, F., and Greven, S. (2015). Penalized function-on-function regression. *Computational Statistics*, 30(2):539–568.

Karlinsky, A. and Kobak, D. (2021). The world mortality dataset: Tracking excess mortality across countries during the covid-19 pandemic. *medRxiv*.

Kontis, V., Bennett, J. E., Rashid, T., Parks, R. M., Pearson-Stuttard, J., Guillot, M., Asaria, P., Zhou, B., Battaglini, M., Corsetti, G., et al. (2020). Magnitude, demographics and dynamics of the effect of the first wave of the covid-19 pandemic on all-cause mortality in 21 industrialized countries. *Nature medicine*, pages 1–10.

Li, Y., Campbell, H., Kulkarni, D., Harpur, A., Nundy, M., Wang, X., Nair, H., for COVID, U. N., et al. (2021). The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (r) of sars-cov-2: a modelling study across 131 countries. *The Lancet Infectious Diseases*, 21(2):193–202.

Liu, Y., Morgenstern, C., Kelly, J., Lowe, R., and Jit, M. (2021). The impact of non-pharmaceutical interventions on sars-cov-2 transmission across 130 countries and territories. *BMC medicine*, 19(1):1–12.

Malfait, N. and Ramsay, J. O. (2003). The historical functional linear model. *Canadian Journal of Statistics*, 31(2):115–128.

Our World In Data (2020). Available at `www.ourworldindata.org`.

Preda, C., Grimonprez, Q., and Vandewalle, V. (2020). cfda: an r package for categorical functional data analysis.

R Core Team (2020). R: A Language and Environment for Statistical Computing. R Foundation. Vienna, Austria. Available at `www.R-project.org`.

Ramsay, J. O., Graves, S., and Hooker, G. (2020). *fda: Functional Data Analysis*. R package version 5.1.5.1.

The official UK government website for data and insights on coronavirus (2020). Available at `https://coronavirus.data.gov.uk/`.