Medium-term forecasting of the Coronavirus Pandemic – updated 2020-05-21

Jennifer L. Castle, Jurgen A. Doornik and David F. Hendry*
Department of Economics, Nuffield College, Magdalen College, and Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, UK

Abstract

We have been publishing real-time forecasts of confirmed cases and deaths for COVID-19 online at `www.doornik.com/COVID-19` from mid-March 2020. These forecasts are short-term statistical extrapolations of past and current data, and are described in a companion paper. Here we investigate if the Chinese experience could be used to forecast other countries' progression of cumulative confirmed cases and deaths. We introduce the method of path indicator saturation (PathIS) to obtain the forecasts. We also create estimates of the peaks in daily counts. This is used specify the initial model on which the PathIS forecasts are based. Comparison with other forecasts show that PathIS regularly has the smallest mean absolute percentage error for deaths.

JEL classifications: C51, C22.
KEYWORDS: COVID-19; Model Selection; Robustness; Outliers; Location Shifts; Trend Indicator Saturation; *Autometrics*.

1 Introduction

Our first paper, Castle, Doornik, and Hendry (2020b), aimed to provide short-term forecasts in the early stages of rapid growth of the COVID-19 pandemic, when it was taking hold in Europe. The short-term forecasts are constrained to extrapolating flexible trends from past and current data only, thus limiting forecasts to about one-week ahead.

The methodology proposed here is different: we investigate if the earlier, and more or less completed, experience of China could be used to to provide forecast for other countries. This could allow us to make forecasts beyond the one-week horizon. To incorporate the eventual slowing down in the cumulative counts, we create smoothed scenarios based on the evolution of the pandemic in China. Automatic model selection is then used to select the closest matching scenario or scenario combination. The final forecast is taken as an ensemble of many forecast paths.

We base estimates of the peak in the distribution of daily counts on a smoothed trend, allowing some time to pass for reliability. So these are not forecasts: the dating of the peak is made ex post. When the peak has happened, we have some ability to compare the trajectories of different countries. Graphs

*This is a preliminary draft: please do not quote without permission from the authors. Financial support from the Robertson Foundation (award 9907422), Institute for New Economic Thinking (grant 20029822), and the ERC (grant 694262, DisCont) is gratefully acknowledged. We wish to thank Max Roser and Our World in Data (`ourworldindata.org`) for very helpful information regarding data sources, and their informative COVID-19 page. We also wish to thank Neil Ericsson and Andrew Martinez for helpful comments. Disclaimer: these results are solely the work of the authors, and cannot be considered official forecasts from their respective institutions. Any opinions are those of the authors only. email: jennifer.castle@magd.ox.ac.uk, jurgen.doornik@nuffield.ox.ac.uk and david.hendry@nuffield.ox.ac.uk

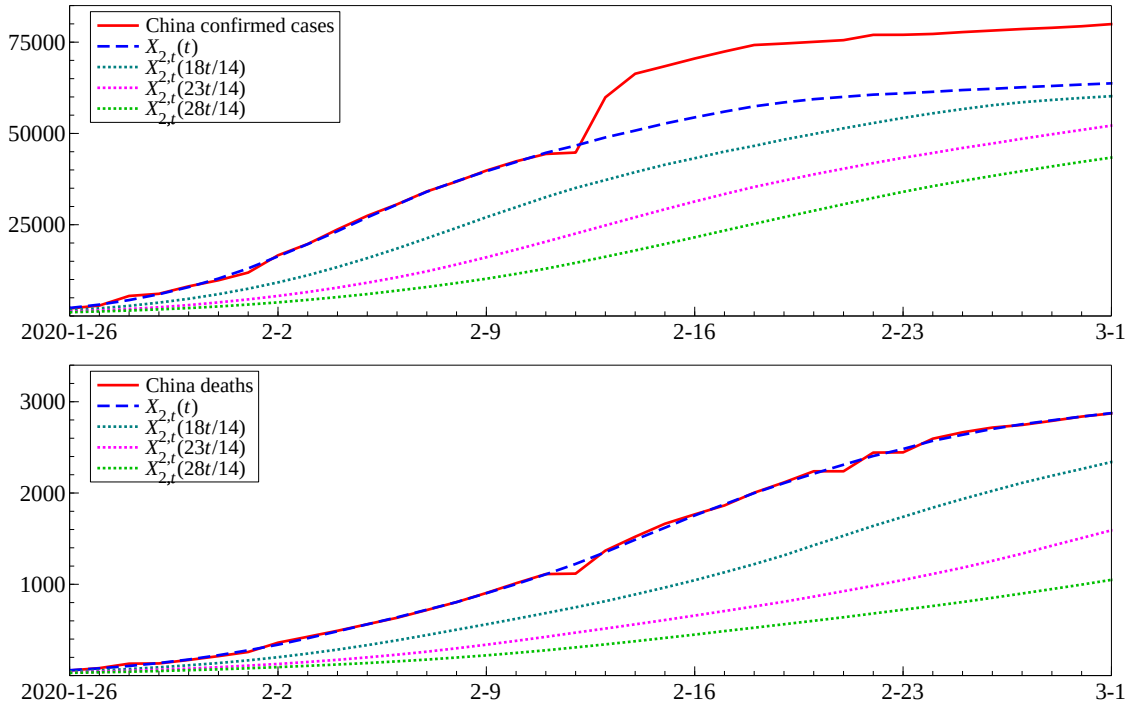

Figure 1: Confirmed cases and deaths (cumulative counts) in China in February 2020, observed and smoothed paths.

show that the trajectories for deaths in European countries were more similar to each other than those of confirmed cases.

The outline of the paper is as follows. We establish the smooth scenarios in §2. Then §3 introduces path indicator saturation (PathIS) as a method that can use the scenarios for forecasting. §4 applies PathIS to create ensemble forecasts for the COVID-19 pandemic. In §5 we define how we estimate the peak and compare the trajectories of several countries. Finally, §6 offers an assessment of the forecast accuracy of PathIS, while §7 concludes.

2 Pandemic scenarios

As a first step, we obtain a smoothed version of the evolution of the COVID-19 epidemic in China, based on the Johns Hopkins/CSSE data, see Appendix A. The substantive part of the epidemic in China took place in February 2020, as shown in Figure 1, with cumulative confirmed cases as the solid line in the top graph, and cumulative deaths in the bottom graph.

Note the jump in confirmed cases on 2020-02-12 (as reported in the WHO report of 2020-02-13). This was caused by the addition of 13332 clinically diagnosed cases. In general, China excludes asymptomatic cases from the counts, effectively classifying it as a false positive. To avoid this definitional change from distorting the results, we replaced the increments for three days from 2020-02-12 by 2000. Next, the path was smoothed using the LATTE specification DDL-SIS, where DDL denotes double differencing after taking logs, SIS is step indicator saturation (see Castle, Doornik, Hendry, and Pretis, 2015), and LATTE is described in Doornik, 2019. This gives the stylized Chinese evolution in $X_{1,t}(t), t =$2020-01-22 to 2020-03-28.

Three stretched versions of the $X_{1,t}(t)$ path were created as

$$X_{1,t}(18t/14), X_{1,t}(23t/14), X_{1,t}(28t/14)$$

using linear interpolation. The last path is the slowest, evolving at half the speed. The four paths were smoothed again with LATTE specification DDL-SIS, resulting in

$$X_{2,t}(t), X_{2,t}(18t/14), X_{2,t}(23t/14), X_{2,t}(28t/14). \tag{1}$$

These are plotted in Figure 1, where $X_{2,t}(t)$ is represented by the dashed line. The aim is to use four time paths to model the evolution in other countries.

3 Path Indicator Saturation

Indicator saturation methods are a general class seeking robust inference in the presence of unknown outliers, shifts, and breaks by incorporating indicators appropriate to the problem. Impulse-indicator saturation for outliers (IIS: Hendry, Johansen, and Santos, 2008, and Johansen and Nielsen, 2009) was the first version; step and trend indicator saturation (SIS and TIS) are both used in the LATTE trends of the short-term COVID-19 forecasts. All saturation methods lead to initial model specifications that have more variables than observations; feasible estimators select a subset so that the resulting model is well specified, see the algorithm in the appendix of Castle, Doornik, and Hendry (2020a) and Hendry and Doornik (2014) for further simulation results.

These indicator saturation methods help forecasting only to the extent that they can fix in-sample model misspecification. Otherwise some judgement may be required: extrapolating linear trends or last-observation adjustments into the future makes a strong assumption that such trends and/or shifts at the end of sample will continue into the future, and therefore these adjustments dominate the forecasts. To address this issue, we introduce the new class of path indicator saturation (PathIS) methods. The main difference is that the paths include future outcomes although the relevant path(s) are selected in sample from a large candidate set. We call these 'scenario forecasts': a selection from many scenarios is used for forecasting. But, because the scenario must match the observed history, the emphasis is on forecasting, not scenario analysis.

The advantage of PathIS is that forecasts can be made from a wide range of possibilities, but limited to those that are empirically relevant. Furthermore, the method is distinct from other forecast approaches, so could be a useful component when averaging over different methods. It is also simple, using only linear methods even when the scenarios may be highly nonlinear, as in the current setting. Of course, the future may not follow the selected paths, and remains subject to unanticipated shocks and structural breaks.

Let y_t denote the variable to forecast and X_t the set of paths, for simplicity we use a single path at first. A model is selected from the following initial specification:

$$y_t = \sum_{x=a}^{b} \beta_x X_{t-x} + u_t,\ t = 1, ..., T, \tag{2}$$

where $u_t \sim D[0, \sigma_u^2]$ after selection (but note that the intial specification is usually overdetermined). Several further decisions need to be made:

1. Whether to include an intercept or not.
2. Estimation in levels, as in (2), in logarithms, or in growth rates.
3. The lag lengths a, b.
4. The significance level for selection.

The estimated model includes a subset of the scenarios in (2):

$$y_t = \sum_{x \in \mathcal{X}} \widehat{\beta}_x X_{t-x} + \widehat{u}_t,\ t = 1, ..., T, \tag{3}$$

and the forecasts are:

$$\widehat{y}_{T+h} = \sum_{x \in \mathcal{X}} \widehat{\beta}_x X_{T-x}. \tag{4}$$

In practice, $\widehat{\sigma}_u^2$ is unlikely to give a good representation of the forecast error variance. This is similar to fitting a trend to a random walk with drift: the trend may fit well in sample, but gives poor forecasts with very narrow forecast confidence intervals. Following this argument, we first compute for the last τ observations:

$$\sigma_\Delta^2 = \frac{1}{\tau} \sum_{t=T-\tau+1}^{T} \left(\Delta y_t - \overline{\Delta y_t}\right)^2,$$

with the mean computed over the same sample. The variance of the forecast $\widehat{y}_{T+h}$ could be taken as:

$$\mathsf{V}[\widehat{y}_{T+h}] = h\sigma_\Delta^2 + \widehat{\sigma}_u^2. \tag{5}$$

This is an ad hoc solution that merits further investigation.

4 Ensemble PathIS forecasts for COVID-19

Returning to the corona virus pandemic, we use the Chinese scenarios of §2 as the candidates for path indicator saturation. Let y_t denote the cumulative counts variable to model. A model is selected from the following initial candidate set of regressors:

$$\{X_{2,t-x}(t), X_{2,t-x}(18t/14), X_{2,t-x}(23t/14), X_{2,t-x}(28t/14);\ x = a, ..., b\}, \tag{6}$$

over the sample $t = T_1, ..., T_2$. We experimented with models in levels, logarithms, and $\Delta \log$, but found that levels provided the best forecasts when tested against a hold back sample. The first version that we created included an intercept in all models, but we found later that it was somewhat better to exclude the intercept. Selection is performed using *Autometrics* at the very tight significance level 0.0001. It is important to protect against selecting an empty model, but that did not happen here.

The selected model $M(y_t, a, b, T_1, T_2)$ is estimated by ordinary least squares (OLS) with forecasts following (4):

$$\widehat{y}_{T_2+h}(a, b, T_1, T_2), h = 1, 2, ...$$

For a relatively short history and when counts change rapidly, the path is still quite uncertain. To improve the robustness, we base an ensemble forecast on the median of many forecasts. Let T again denote the last available observation. The first batch of forecasts is based on a backward expanding sample, with the sample size increasing from 14 to $T - T_1 + 1$:

$$\widehat{y}_{T+h}(a, b, T-13, T), ..., \widehat{y}_{T+h}(a, b, T_1, T). \tag{7}$$

The second batch ends one period earlier, and is recentered on the last known observation y_T:

$$\frac{\widehat{y}_{T+h}(a, b, T-13, T-1)y_T}{\widehat{y}_T(a, b, T-13, T-1)}, ..., \frac{\widehat{y}_{T+h}(a, b, T_1, T-1)y_T}{\widehat{y}_T(a, b, T_1, T-1)}. \tag{8}$$

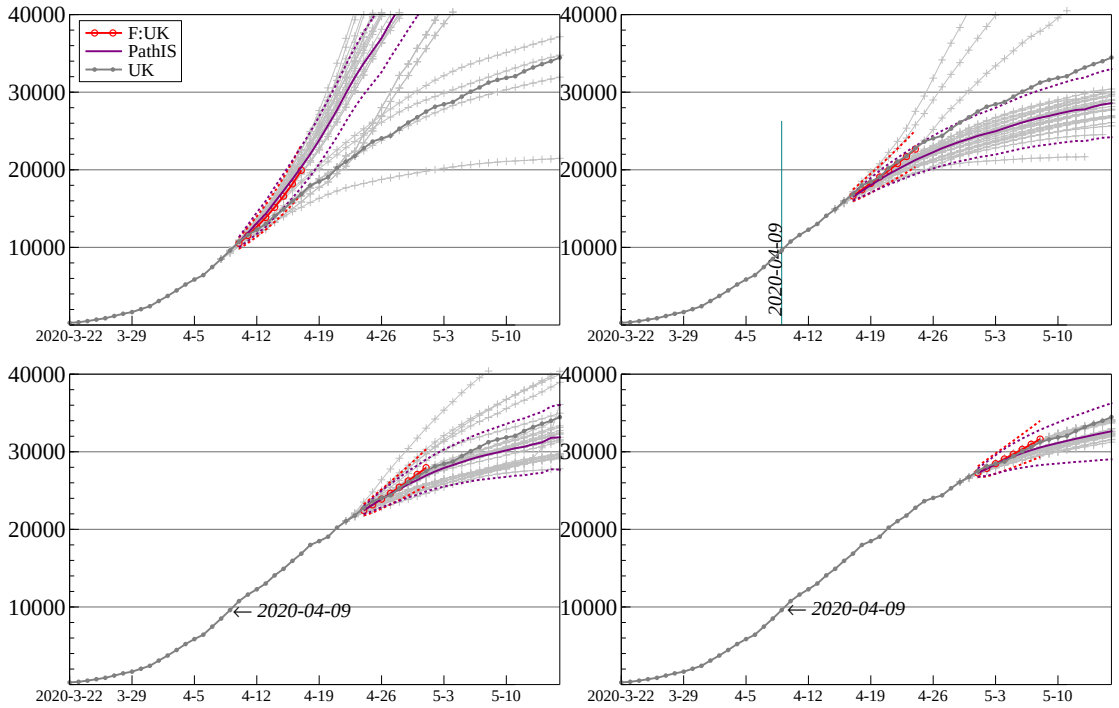

Figure 2: Scenario forecasts (PathIS version 3) of deaths for the UK with 95% forecast intervals. Short-term forecasts F:UK given in red up to eight days ahead, with 80% forecast intervals.

The third and final batch of forecasts drops the last two observations, recentering the forecasts on the penultimate observation:

$$\frac{\widehat{y}_{T+h}(a,b,T-13,T-2)y_{T-1}}{\widehat{y}_{T-1}(a,b,T-13,T-2)},...,\frac{\widehat{y}_{T+h}(a,b,T_1,T-2)y_{T-1}}{\widehat{y}_{T-1}(a,b,T_1,T-2)}. \tag{9}$$

With $T_1 = T-26$ for confirmed cases, the first batch (7) has 14 forecasts, the second 13, and the last 12. For deaths, with $T_1 = T-22$, this is 10,9,8 respectively. The overall forecast is taken as the median at each point in time of all the $M = 39$ forecasts for confirmed ($M = 27$ for deaths). Computation of the forecast standard error deviates from (5):

$$\mathsf{SE}[y_{T+h}] = h^{2/3}\left(\sigma_\Delta + \frac{1}{M}\sum_{i=1}^{M}\widehat{\sigma}_u\right). \tag{10}$$

Figure 2 illustrates the outcomes for cumulative deaths in the UK. The forecasts in panel a start 2020-04-09, so only using data up to 2020-04-08. Then clockwise each panel starts a week later. The solid grey line with dots show the outcomes. The thick solid purple lines are the PathIS 'scenario' forecasts (version 3, see Appendix B), shown with 95% confidence bands. The many thin lines are the individual forecasts. Also shown are the short-term forecasts (see Castle, Doornik, and Hendry, 2020b), labelled 'F:UK' in Figure 2 and plotted with 80% confidence bands. The PathIS forecasts in the first panel are close to the short-term forecasts F:UK. In the next two panels, the short-term forecasts are better initially, while PathIS has better curvature. In the second panel a peak in daily deaths has been detected, as discussed in the next section.

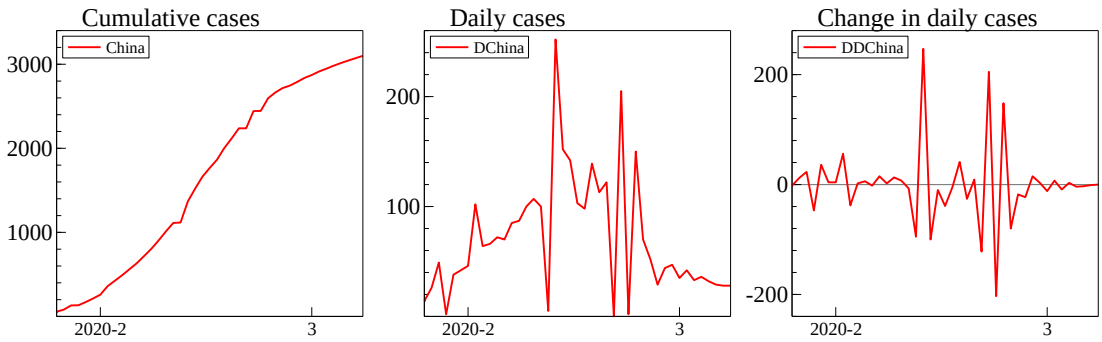

Figure 3: Cumulative daily confirmed cases for China (left), daily count (middle), day-to-day change in daily count (right).

5 Estimating the peak

There has been much media and political interest in the so-called 'peak': the day after which the daily count of confirmed cases or deaths gradually declines. If the confirmed cases have reached a peak, exponential growth is over, and reproduction of infections has slowed. The start of the slowdown could signal potential for a gradual change in lockdown policies. Similarly, for deaths, it could indicate that the demand for intensive care beds may also have peaked. The peak could be inferred from the actual daily increments, but, because these tend to be fairly erratic, basing it on a smoothed trend produces more reliable estimates.

The peak is determined from measured outcomes, so not based on forecasts. Nonetheless a peak can disappear again for several reasons: a resumption of growth after a temporary slowdown, data reporting errors, or data revisions. A flat peak may also make it difficult to pin down the exact date. Our estimate of the peak increment is based on the trend, so not an exact match for the tally of that day.

The first difference of the cumulative cases are the daily increments, and the second difference the change in the daily increments. As the level of differentiation increases, the data becomes more noisy. Figure 3 illustrates for the unadjusted counts for China. A consequence is that it is difficult to estimate the mode of the daily cases, i.e, the point where the cumulative distribution inflects (assuming unimodality). Let $\Delta_s y_t = y_t - y_{t-s}$ denote the sth backward difference and $\nabla_s y_t = y_t - y_{t+s}$ denote the sth forward difference, where we drop subscript s when $s = 1$.

As a more robust alternative we base estimates of the peak on the smoothed trend. Averaged short-term forecasts were constructed from eight forecasts: DDL-TIS and DL-TIS both with damped I(2), each estimated up to $T, T-1, T-2, T-3$ (where TIS denotes trend indicator saturation: details are in Castle, Doornik, and Hendry, 2020b). Then $\mathcal{T}_t$ denotes the averaged smoothed trend from these eight models, and $\mathcal{C}_t = \Delta\mathcal{T}_t$ the daily counts from the averaged smoothed trend. A local peak requires that the observation before and after is lower: $\Delta\mathcal{C}_t > 0, \nabla\mathcal{C}_t > 0$. For more certainty we define a local peak as $\Delta_i\mathcal{C}_t > 0, \nabla_i\mathcal{C}_t > 0, i = 1, ..., p$. For confirmed cases we set $p = 3$, and for deaths, which are more noisy, $p = 4$. A global peak is then defined as a local peak that is also the maximum of $\mathcal{C}_t$; however, it is rejected if it is fewer than 5 days from the end of the sample and at the same time the last three are sloping upwards ($\Delta\mathcal{C}_T > 0, \Delta\mathcal{C}_{T-1} > 0$).

Finally, a peak is called if there is a global peak, or otherwise if there is a global peak in $(\mathcal{C}_{t-1} + 2\mathcal{C}_t + \mathcal{C}_{t+1})/4$ (setting $\mathcal{C}_{T+1} = \mathcal{C}_T$). Having established the presence of a peak, the exponential growth will have ended and we then switch the model from flexible trends to flexible steps which provide more restrained trend forecasts. The actual peak is derived from the averaged trends of the latter model.

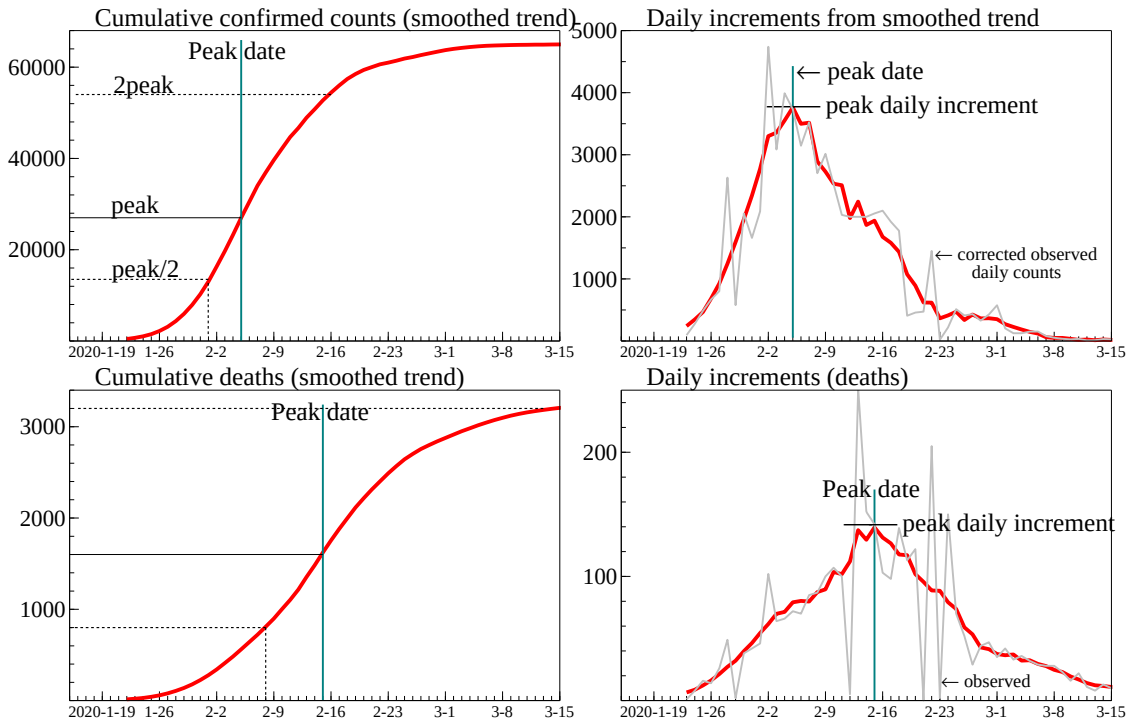

Figure 4: Cumulative daily cases for corrected China data (left), daily count (right). Top row is confirmed cases, bottom row shows deaths.

	UK	*EU*	*AT*	*DE*	*DK*	*ES*	*IT*	*NL*	*SE*
Peak date	04-09	04-07	04-08	04-15	04-04	04-02	03-28	04-07	04-22
Peak daily increment	1014	3241	23	280	18	902	826	163	117
Days from peak/2 to peak	16	24	17	22	14	16	19	18	27
Days since peak in confirmed	−1	10	13	18	0	6	7	−3	−2

Table 1: Dates of peaks in estimated trend of Deaths for selected European countries and EU-27, data up to 2020-05-09.

The right column of Figure 4 shows for the adjusted China data how the date of the peak, with corresponding daily increments, are found from the change in the smoothed cumulative trend, shown on the left. On the date of the peak, cumulative confirmed cases stood at 27 000 and deaths at 1600. Some further observations:

1. It took four days to move from 13 500 to 27 000 confirmed cases, seven days for the equivalent move (from peak/2 to peak) in deaths, reflecting the slower incline in daily deaths.
2. The distribution of confirmed counts is considerably skewed to the left. That of deaths much less so (although it will keep on rising even when the daily confirmed case count falls to zero). The peak in confirmed cases occurred at a cumulative level that was about a third of the final total.
3. Deaths in China are more noisy than confirmed cases. In both cases, the peak in (adjusted) observed counts is before the estimates from the smoothed trend.
4. The peak in deaths was ten days after the peak in confirmed cases.

Table 1 gives the dates of the peak in deaths for some European countries and the EU as a whole. Italy was the only European country to peak in March, the others were all in April. In most cases the

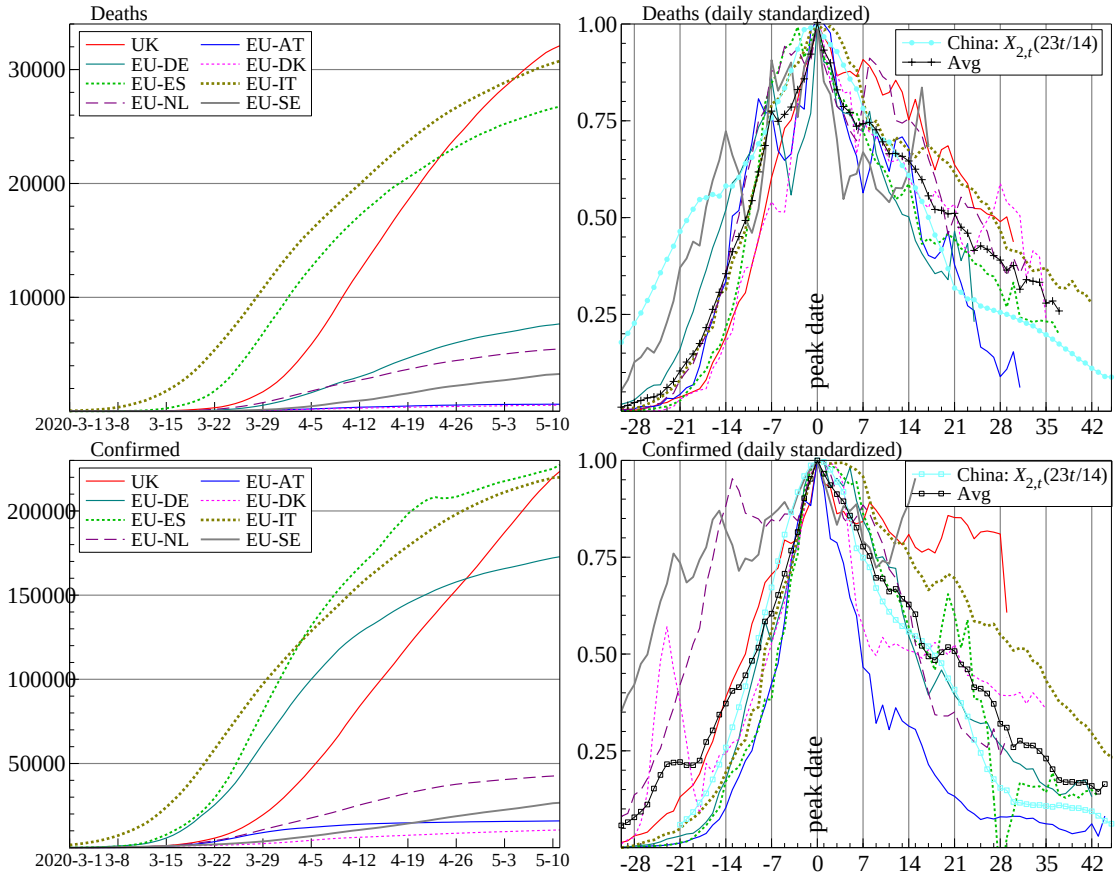

Figure 5: Smoothed cumulated deaths (top left), day-to-day change in smooth deaths (top right), standardized. Smoothed cumulated confirmed cases (bottom left), day-to-day change in smooth cases (bottom right), standardized. Data up to 2020-05-09.

ascent from halfway to the peak to the peak took between two and three weeks, slower than China where it was only a week. It is not necessarily the case that the peak in confirmed cases happens before the peak in deaths, as seen for the UK, Denmark, Netherlands, and Sweden. Germany has the peak in deaths almost three weeks after the peak in confirmed cases. Note that on 27 April the UK deaths data was revised to include those that occurred in care homes. This moved the peak one day forward, and the peak daily increment rose from 883 to 1014.

Figure 5a shows the averaged trends that are used to estimate the peak for the countries in Table 1. These are difficult to compare, although we see that Spain is about four days behind Italy, and the UK started about two weeks later than Italy, but on a steeper slope. To compare these trajectories, we standardize them as follows: the daily change of the lines in panel a is divided by the count at the peak, and the day of the peak is set to day zero. This is shown in Figure 5b, with the average added in. Note that the average of the selected European countries is based on fewer observations as we get nearer the end.

The analogous picture for confirmed cases in the bottom row of Figure 5 is more erratic. This may hamper PathIS estimation, although the averages are surprisingly similar, as shown in Figure 6.

Figure 6 also includes two of the Chinese scenarios: $X_{2,t}(23t/14)$ and $X_{2,t}(28t/14)$ from (1). The closeness of these suggests that PathIS could indeed work with the shifted and scaled Chinese trajectories. For countries that are well behind Europe in their pandemic stage, the European average could also be

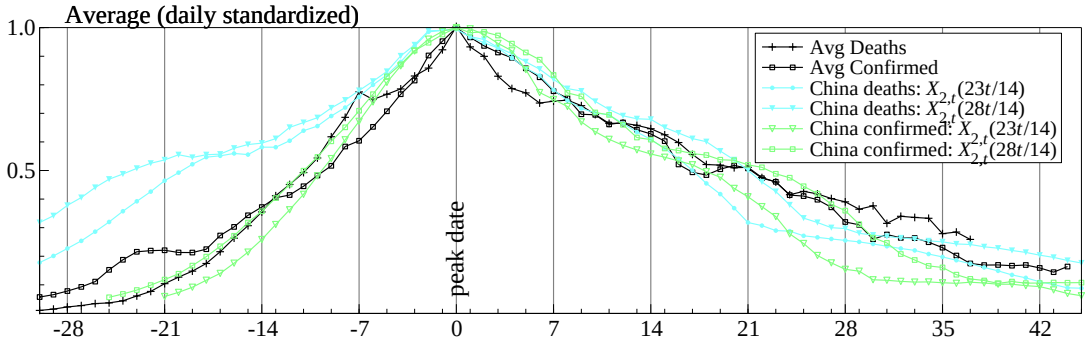

Figure 6: Average standardized day-to-day change in smooth deaths and confirmed.

Horizon	Deaths					Confirmed				
	PathIS	*Avg*	*F:*	*Count*	*Cover*	*PathIS*	*Avg*	*F:*	*Count*	*Cover*
1	2.3	2.2	2.3	974	82	1.2	1.1	1.1	1187	81
2	3.4	3.3	3.7	913	81	1.8	1.8	1.8	1112	82
4	4.6	4.4	5.5	851	80	3.1	3.5	3.4	1038	78
7	6.8	7.4	8.4	787	75	5.4	6.9	6.1	961	67
14	14.2	–	–	562	52	10.8	–	–	684	51
21	24.3	–	–	404	32	16.4	–	–	500	43

Table 2: Forecasts of deaths and confirmed cases: mean absolute percentage errors relative to the outcome for different forecast horizons. PathIS is version 3, F: and Avg are the short-term forecasts. Cover is the percentage of forecasts of PathIS inside the forecast intervals. Count is the number of forecasts.

used for PathIS. We use knowledge of the dates of the peak for a country in relation to that of the scenarios to specify values of a and b in (6), determining the lag lengths used in the initial model, see Appendix B.

6 Assessment of forecast accuracy

The forecast accuracy is measured as the error in percentage of the outcome, computed for almost all the forecasts that we have published. Reported is the mean absolute percentage forecast error (MAPE), which is, for a forecast $\widehat{y}_{j,T+H}$ from group $j = 1, ..., J$:

$$\text{MAPE} = \frac{100}{J}\sum_{j=1}^{J}\frac{|y_{j,T+H} - \widehat{y}_{j,T+H}|}{y_{j,T+H}}. \tag{11}$$

Table 2 gives the average MAPE at selected forecast horizons. Evaluation is over the period 2020-04-11 to 2020-05-16, omitting France and Mexico because of data issues (See the appendix of Castle, Doornik, and Hendry, 2020b). Count is the number of forecast errors J used in (11). The forecast accuracy for deaths is lower than that for confirmed. The PathIS forecast accuracy for confirmed casesis close to that of the short-term forecasts F:, while for deaths there is some gain when forecasting a week ahead.

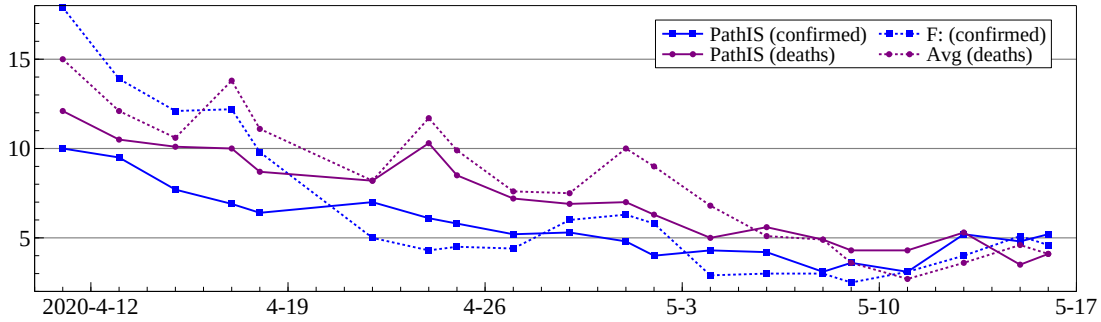

Figure 7: Forecast accuracy over time. MAPE for each target date for 7-day ahead forecasts. Confirmed marked with squares, deaths with circles. Dates on the horizontal axis are the target dates of the forecasts.

Week ending	*Count*	MAPE			MPE		
		PathIS	MCIC	F:	PathIS	MCIC	F:
2020-04-11	23	**14.2**	41.8	16.2	9.1	−39.8	**2.7**
2020-04-18	24	**8.7**	13.2	10.2	**0.2**	−11.6	−1.1
2020-04-25	24	**7.4**	8.7	8.1	3.3	−6.5	**−0.4**
2020-05-02	27	**5.5**	10.9	7.5	**−1.8**	−10.4	−5.5
2020-05-09	28	3.6	2.8	**2.5**	2.7	−0.6	**0.3**
2020-05-16	28	**2.2**	3.2	**2.8**	**0.4**	−2.7	−2.3
2020-05-23	28	3.1	**2.8**	3.7	1.9	−0.6	**−0.4**

Table 3: MAPE and mean percentage error (MPE) of MCIC, PathIS and F: 1-week ahead forecasts. Smallest in bold.

Figure 7 shows one-week ahead accuracies over time, comparing PathIS with short-term F: for confirmed cases, and with Avg (which has a lower MAPE than F here) for deaths. The accuracy increases over time, which is expected because eventually the denominator in (11) settles, and the increments fall to zero. For confirmed cases the two lines (marked with squares) cross over more than three times. But for deaths the PathIS forecasts dominate in the first three weeks.

Forecast standard errors are computed from (10), with a critical value of 1.6, so a confidence interval of 78% assuming normality. The columns labelled *Cover* in Table 2 show the percentage of forecasts that are inside this interval. If we take this pointwise, then the interval is correct one-day ahead, but increasingly too narrow as the horizon grows.

Table 3 compares the PathIS forecasts with those from the MRC Centre for Global Infectious Disease Analysis at Imperial College London (MCIC below, Bhatia, 2020). In each case, the forecasts are compared to their own outcomes (ECDC data for MCIC, Johns Hopkins for ours), using the nearest available outcomes in order to reduce the impact of data revisions. The forecasts are one week ahead for deaths, and PathIS has the smallest MAPE, except for the week ending 2020-05-09. The week ending 2020-04-11 matters, because many European countries (including UK, EU, Belgium, Czechia, France, Netherlands, Portugal) had their peak in that week. In Europe, only Italy, Spain, and Switzerland had their peak earlier.

7 Conclusion

We introduce path indicator saturation as a method that uses available past outcomes (or scenarios) to generate forecasts. This was shown to produce more accurate forecasts for cumulated deaths caused by COVID-19 than some epidemiological models. We use estimates of the peak in daily counts to allow comparison of the distribution of daily change in counts. This shows an unexpected similarity for average deaths and confirmed cases in eight European countries.

As a new method PathIS will need further evaluation. In this application it also needs better estimates of the forecast uncertainty intervals. With the pandemic further advanced, it would be possible to use the European outcomes as scenarios for countries that are further behind in the spread of COVID-19.

A Data source

We use the data repository for the 2019 Novel Coronavirus Visual Dashboard operated by the Johns Hopkins University Center for Systems Science and Engineering (JH/CSSE). This is currently updated daily and located at `github.com/CSSEGISandData/COVID-19`. The data consists of confirmed cases and deaths.

On 2020-03-23 the Johns Hopkins data omitted US states. This gap was filled later by the New York Times, collecting data from state-level health authorities. We used this from the 27th onwards. Their US state data can be downloaded from Github at `github.com/nytimes/covid-19-data`. The New York Times started redefining the US state data on 2020-05-06, but only for the last observation. At that stage, we switched to Johns Hopkins/CSSE for US disaggregate data.

Information on the selection of countries and data revisions is in Castle, Doornik, and Hendry (2020b).

B PathIS implementationm

We introduced PathIS forecasts on 2020-03-26, with a settled specification on 2020-04-02, switching to version 2 on 2020-04-17, and version 3 on 2020-05-13, as summarized in Table 4. Initially, the available sample was short, particularly for deaths, limiting the options for lag lengths in the initial specification. A month later more flexibility was available, and countries (and US states) where divided in three groups, with different values for a and b. We also switched to using the smoothed trend for modelling, rather than the actual data.

In version 3 the values a and b are based on the date of the peak, if one has been detected: a is set to three less than the number of days between a country's peak date and that of $X_{2,t}(28t/14)$ (the 'gap'). If no peak has been detected yet, however, b is limited by the available data; then, if $b - a < 10$, a is set to $b - 10$. The main advantage of version 3 is that the initial model is now different for each country, but automatically choosen. In terms of MAPE there is not much difference between version 2 and 3, while version 1 is a somewhat worse for target dates up to 2020-04-22. All the results in the paper are based on version 3.

References

Bhatia, S. *et al*. (2020). Short-term forecasts of covid-19 deaths in multiple countries. Online reports 2020-04-15 and 2020-04-08, mrc-ide.github.io/covid19-short-term-forecasts/index.html, MRC Centre for Global Infectious Disease Analysis.

	Version 1	*Version 2*	*Version 3*
intercept	always	never	never
y_t	actual data	smoothed trend	smoothed trend
a	30	30 or 40 (LAC)	gap -3
b	43	43 (Set1) or 53 or 63 (LAC)	$a + 20$
T_1	$T - 26$ ($T - 22$ for deaths)	as version 1	as version 1

Table 4: PathIS specifications used for forecasting. LAC refers to Latin American countries; Set1 is EU, AT, BS (Baltic States), CZ, DE, ES, IT, N, CH, Australia, Iran, Philippines.

Castle, J. L., J. A. Doornik, and D. F. Hendry (2020a). Robust discovery of regression models. Economics discussion paper 2020-W04, Nuffield College, University of Oxford.

Castle, J. L., J. A. Doornik, and D. F. Hendry (2020b). Short-term forecasting of the coronavirus pandemic. Economics discussion paper 2020-W06, Nuffield College, University of Oxford.

Castle, J. L., J. A. Doornik, D. F. Hendry, and F. Pretis (2015). Detecting location shifts during model selection by step-indicator saturation. *Econometrics 3(2)*, 240–264.

Doornik, J. A. (2019). Locally averaged time trend estimation and forecasting. mimeo, Nuffield College, Oxford.

Hendry, D. F. and J. A. Doornik (2014). *Empirical Model Discovery and Theory Evaluation: Automatic Selection Methods in Econometrics*. Cambridge, MA: MIT Press.

Hendry, D. F., S. Johansen, and C. Santos (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics 33*, 317–335. Erratum, 337–339.

Johansen, S. and B. Nielsen (2009). An analysis of the indicator saturation estimator as a robust regression estimator. In J. L. Castle and N. Shephard (Eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, pp. 1–36. Oxford: Oxford University Press.