FGV IIU Flash Notes

Monitoring the Return to Normality in Six European Countries III: Data Problems Continue, Progress Slow-walks

(week ending on May 31, 2020)

FGV IIU International Intelligence Unit

Rio, June 4, 2020.

FGV IIU is a small think tank directly connected to the Presidency of Fundação Getulio Vargas. **FGV IIU** observes the international scene with the purpose of identifying issues, trends and partners relevant for Brazil, as well as contributes to explain Brazilian issues and characteristics to foreign audiences. The exchange of good practices, in both directions, is a subsidiary aim of the Unit.

FGV IIU Flash Notes are short reports addressing pressing or hot issues in the international scene, with a likely strong interaction with Brazil. They state a viewpoint considered relevant for the moment at stake; by their very nature, the situation they address may change, new information, data or events may come or happen and conclusions accordingly become dated. The views in **FGV IIU Flash Notes** do not engage in any way FGV or its executives and can only be attributed to the respective Authors; if not signed, FGV IIU's Director is solely responsible for them.
FGV IIU Flash Notes are identified by the date of issue, at the left bottom of the front page.

FGV IIU
npii@fgv.br
https://iiu.fgv.br/

Director: Renato G. Flôres Jr.
Praia de Botafogo 210, 12th floor; Rio de Janeiro, Brazil.
Voice: +55 21 3799 6220 (Licia)

1. Introduction: plagued by data problems.

It seems that, with a milder behaviour of the epidemic, and a generalised perception that things are getting better, governments in all six countries at stake -Belgium, France, Germany, Italy, Portugal and Spain- have decided to look closer at their statistics.

That 'new daily cases' was greatly underreported is widely known, so the peaks and troughs that suddenly started to appear, though understandable and largely due to the encompassing testing policy that has (finally) been implemented, jeopardise analyses that have been using this statistic. Even so, the filtering process we've adopted succeeds, to some extent, to attenuate the negative outcomes revisions including new entrants -due to the testing results- have introduced in the series.

In the long run, to have more accurate data is certainly positive, but for short run indicators this creates significant doubts about how the process they try to gauge is evolving.

However, what is more surprising and harder to accept is that daily deaths statistics -as already hinted in the previous Monitoring- are also being reviewed, and largely!, to the extent that negative daily values are becoming a repeated feature of this data, even in well-behaved Spain. Explanations must be produced: this is a serious revision that puts under suspicion many exercises and analyses made so far.

Within this disturbing data scenario, we've produced evidence of the continued extremely slow progress of the fall of the epidemic; it does exist, apparently -though the best evidence comes from the disputable death statistics-, but its staggering dynamics confirms that the left side of the daily contagion curve is a fat tail.

In section 2, the analysis of the two indicators is jointly discussed, in a country basis. Section 3 concludes. All data come from the *worldometers.info* public site.

2. The six analyses by country.

As mentioned and explained in previous Notes[1], and following proposals in Flôres (2020)[2], two indicators have been systematically used: the ratio of new daily cases, at a *basic evolution period of the epidemic*, and the fitting of a straight line using the 21 more recent observations about daily deaths.

The former may be assumed to be a function of two parameters deeply related to the dynamics of the epidemic within a given community: the average number of people infected by a person with the virus at the start of the period, and a synthetic measure of the proportion of infected people who, for a variety of reasons, may be considered outside the contagion group, at the end of the period. The "basic evolution period of the epidemic" has been chosen as 14 days.

The ratios should change along the evolution of the epidemic, notably its decline, signalling how positive the policy package that has been implemented and the return to normality are being. They must reach values below 1 and then (ideally) continue to fall.

As for the regressions, they must fulfil minimal goodness-of-fit criteria: a R^2 of at least 0.45, and coefficients significant at least at 5 per cent.

However, various sources of noise disturb both series[3] and, during the period under analysis, major ones have been the more aggressive and encompassing testing policy, correctly applied in many countries to reduce underreporting on contagion, and governmental revisions of the Covid-19 series, particularly contagion and deaths data.

Testing, by inserting the 'newcomers' uncovered by the new statistical reality, different from the one prevailing beforehand, pushes up the daily new cases data and produces peaks in our ratios. The deaths revisions, harder to explain, plague with outliers -negative values being not absent!- the last 21 days observations.

---

[1] *The 19th Week Effect: Prospects for Flexibilization in Six European Countries.* FGV IIU Flash Notes; May 7, 2020. Rio de Janeiro: FGV International Intelligence Unit; *Monitoring the Return to Normality in Six European Countries I: A Standstill with Positive Signs.* FGV IIU Flash Notes; May 20, 2020. Rio de Janeiro: FGV International Intelligence Unit.

[2] *Corona Data Analyses: Looking for Signs of Recovery in Italy and Spain*, FGV IIU Discussion Papers DP 02/20, R. G. Flôres Jr., with the assistance of L. Garnitskiy, 2020; Rio de Janeiro: FGV International Intelligence Unit. Two statistical procedures to be used in the monitoring of the decline side of the contagion curve are described.

[3] It is neither the case nor the object of this Note to dwell on these sources here; some has been said on them in the works previously cited.

Our public data source, the *worldometers.info* site, is then forced to adjust its corresponding series, oftentimes displaying the negative values, either in 'new cases' or 'new deaths'.

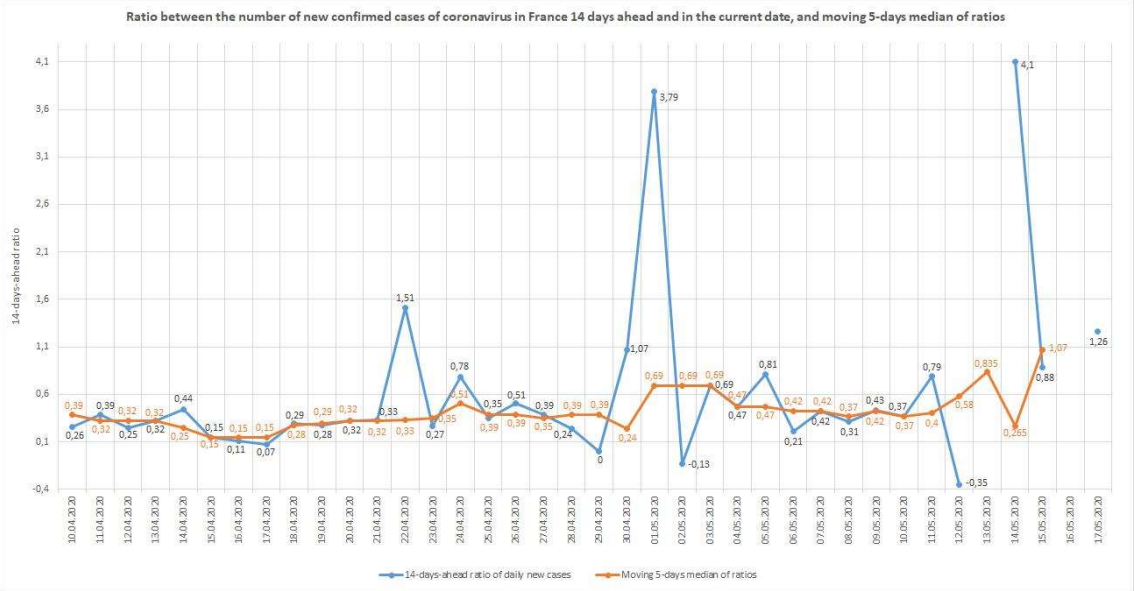
We discuss, by country, the evidences produced by the two approaches.

Exhibit 1 shows graphs for the ratios and regressions for France, the following ones do the same for the other countries.

The *series of daily (two-weeks)* ratios cover the period from April 10 to May 17 (two weeks more, one arrives at May 31, the last observation); in each graph, the five-days moving median, a convenient way to filter the noise in the original series is shown. In the daily deaths graph, the so-called pessimistic line also appears; obtained by subtracting two standard-errors to the intercept and adding two to the angular coefficient[4]. The point this line cuts the horizontal axis produces a notional "zero deaths day"[5], but given the number of outliers sometimes it is senseless to perform this exercise.

Exhibit 1: France.

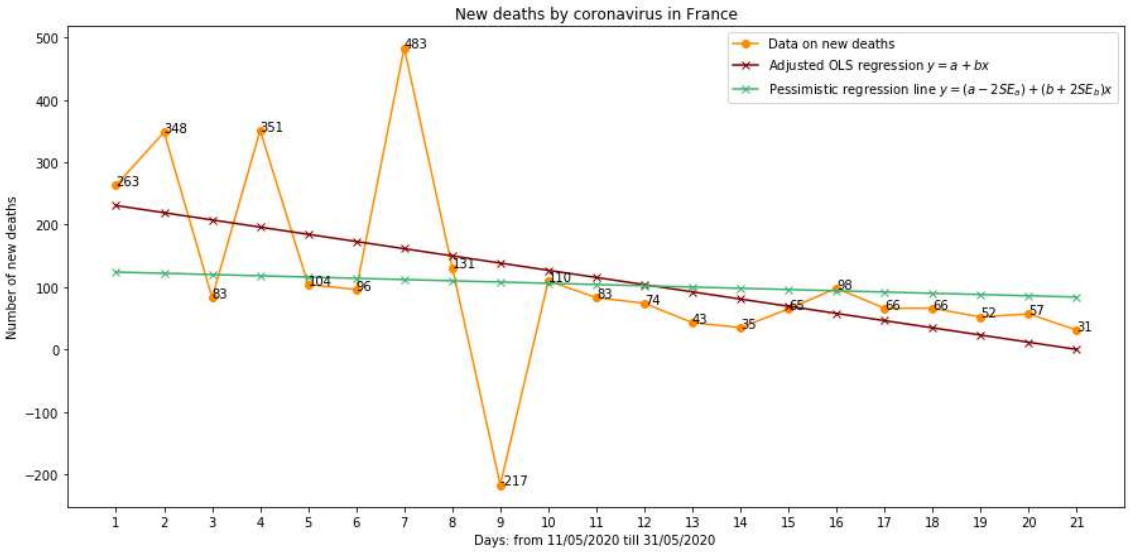
1.a) Ratio between the number of new confirmed cases of coronavirus 14 days ahead and in the shown date, starting on April 10, 2020 (actual values and five-days moving medians).



[4] Basic statistics on the regression results may be demanded at npii@fgv.br .

[5] For more on this idea see the reference in footnote 1. The adjective 'notional' is very important, because this is a mere extrapolation from the curve, a "zero-deaths day" seeming not feasible yet. Nevertheless, as "the period till normality" in section 2, it provides an idea of a date when things will be better.

1.b) The regression line for the last 'new daily deaths' 21 observations.

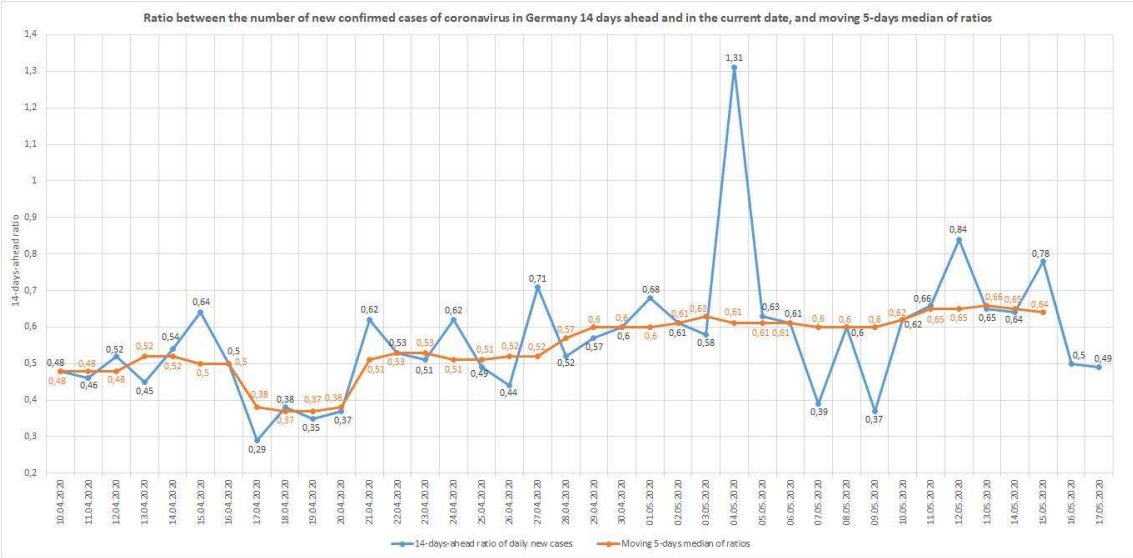


The ratios presented two problems. No value appears for May 13, as zero new cases were registered in this day, making it impossible to obtain a ratio with a zero denominator. For May 16 another odd situation took place: it showed the nonsense value of -141 and then, on May 30, another outlier, 1828 new cases, is present, producing a ratio of -12.96. Low negative values are already present in the series, but this high one has been excluded from the graph, as well as the non-existent May 13 ratio. The "good return" to lower ratios, announced in the previous monitoring and kept until May 11, is prejudiced by the several outliers, and an increasing progression is then observed, *obliging one to wait for a sequence of more reliable observations*.

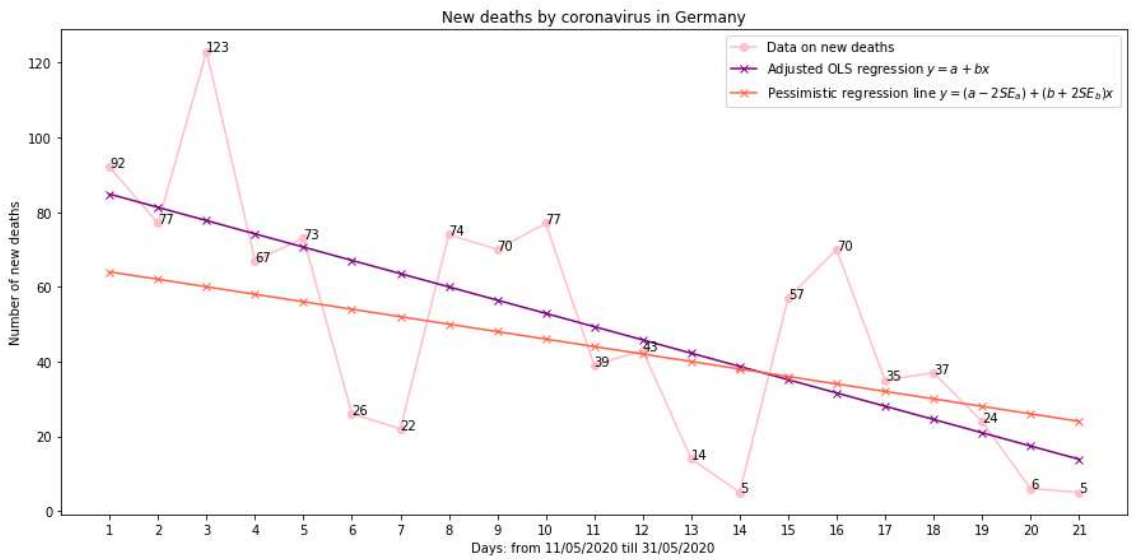
The regression fit showed a borderline significance -a R^2 of 24.6%, thanks to the relative outliers 483 and 217-, and the pessimistic line resulted nearly flat. Even without the outliers, the number of deaths does seem to follow a broadly flat pattern.

Exhibit 2: Germany.

2.a) Ratio between the number of new confirmed cases of coronavirus 14 days ahead and in the shown date, starting on April 10, 2020 (actual values and five-days moving medians).



2.b) The regression line for the last 'new daily deaths' 21 observations.

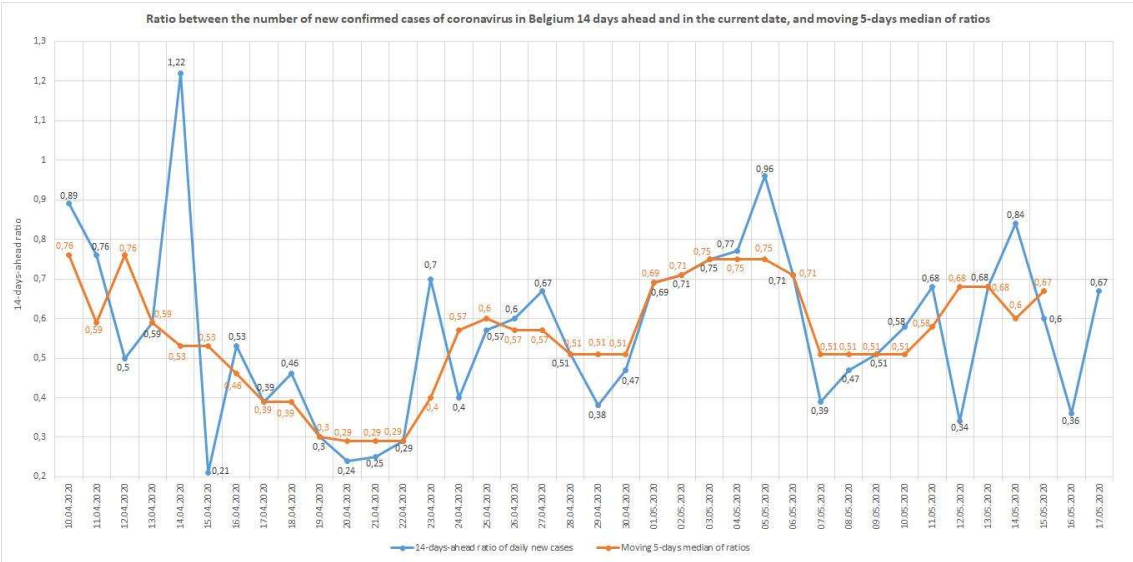


The ratios indicator remains at previous levels, with a slight increase due to peaks likely to be attributed to testing results. Though steady, it still is at somewhat high levels (0.60 or more).

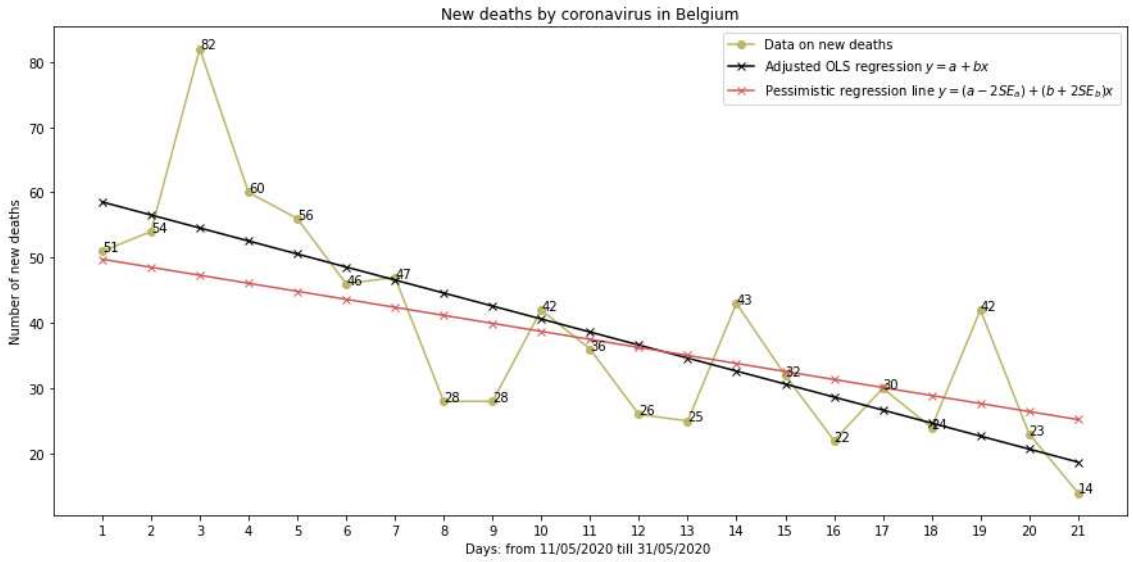
The regressions are significant and show a good result, with the last two observations already equal to 6 and 5 deaths, respectively. Using the pessimistic line, the notional zero deaths day will be June 12-13, signalling an auspicious context for the German case, that must be better matched to the ratios evidence.

Exhibit 3: Belgium.

3.a) Ratio between the number of new confirmed cases of coronavirus 14 days ahead and in the shown date, starting on April 10, 2020 (actual values and five-days moving medians).



2.b) The regression line for the last 'new daily deaths' 21 observations.

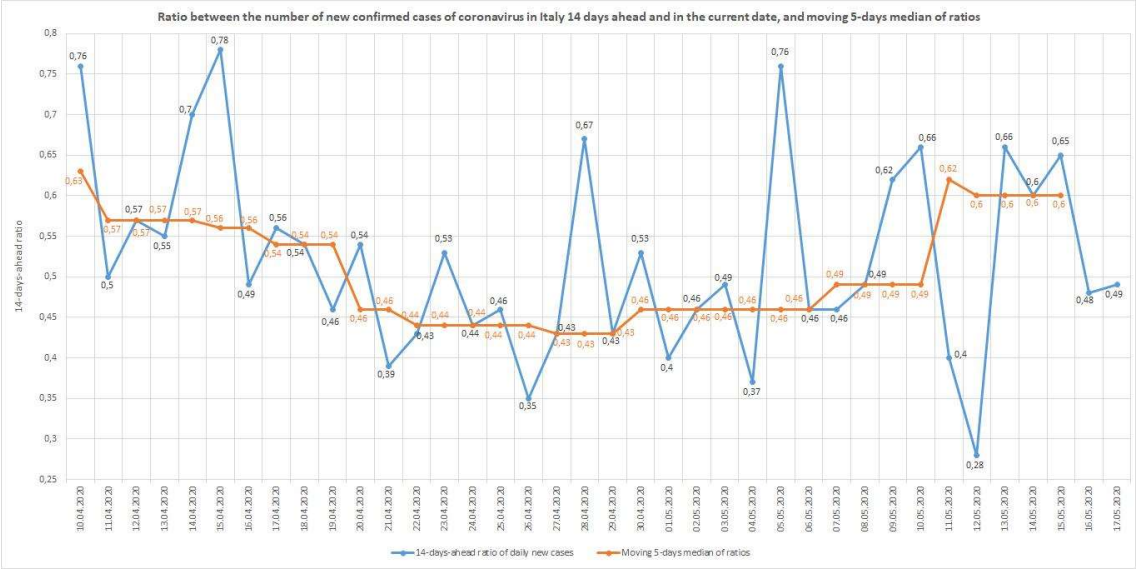


The "normal though still high (0.51) platform", displayed at the end of the last Monitoring, unfortunately increased to values touching 0.68. Again, this seems to be due to the testing effect, and it is hoped that next week results will return to "normality".

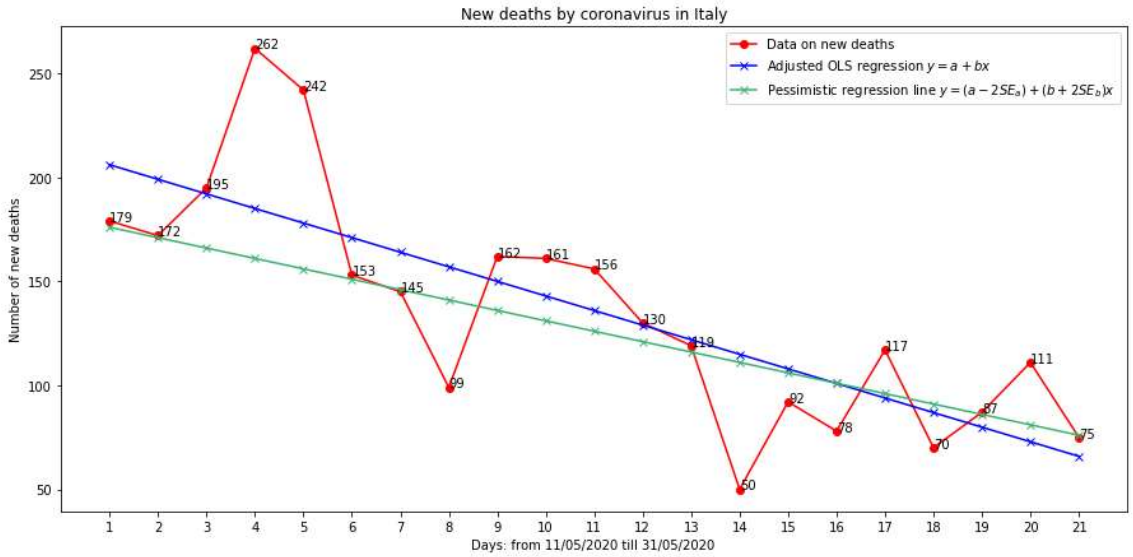
Regression results were significant and auspicious, confirming a decreasing trend. The notional "zero deaths day" is expected at June 20-21.

Exhibit 4: Italy.

4.a) Ratio between the number of new confirmed cases of coronavirus 14 days ahead and in the shown date, starting on April 10, 2020 (actual values and five-days moving medians).



4.b) The regression line for the last 'new daily deaths' 21 observations.

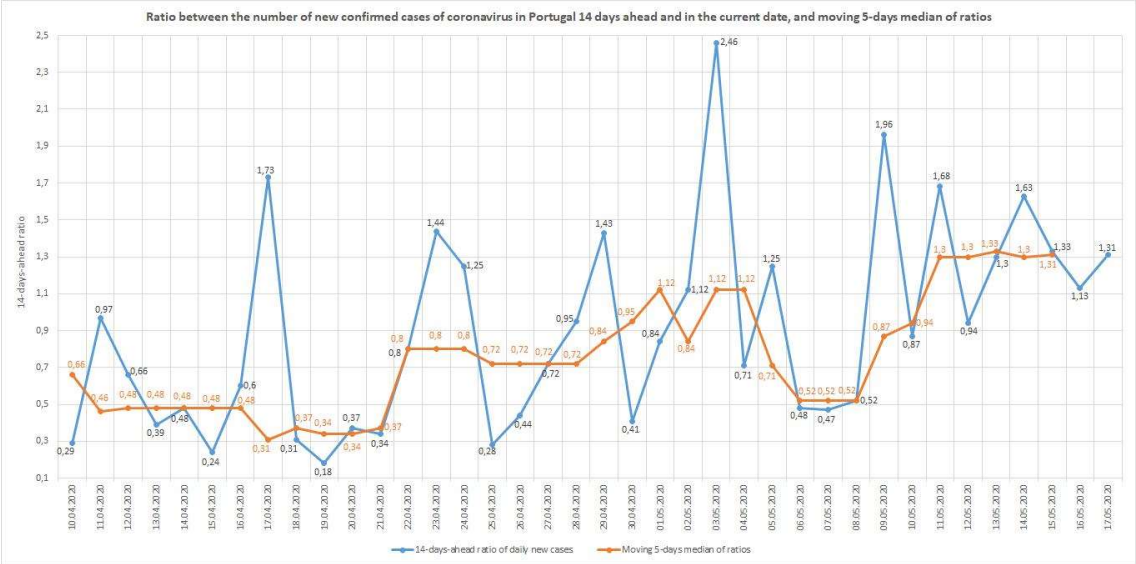


As with Belgium, high oscillation of the new cases statistics moved the steady 0.49 platform (four times repeated) to higher values, reaching 0.60, *but prospects are that the ratios will return to lower values*.

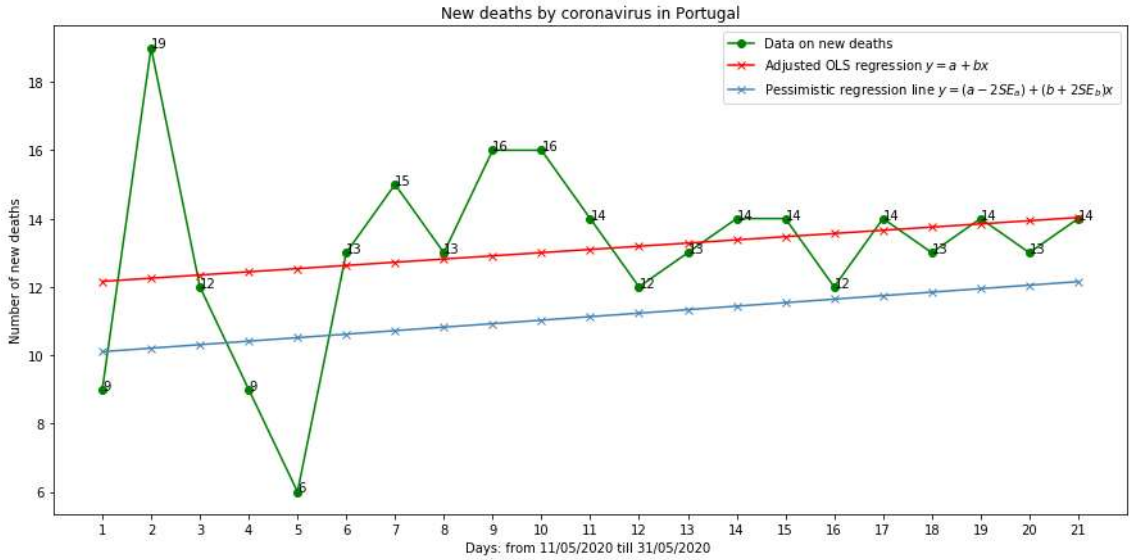
Regressions showed significant and positive results, with the “zero deaths day” expected by June 15-16.

Exhibit 5: Portugal.

5.a) Ratio between the number of new confirmed cases of coronavirus 14 days ahead and in the shown date, starting on April 10, 2020 (actual values and five-days moving medians).



5.b) The regression line for the last ‘new daily deaths’ 21 observations.

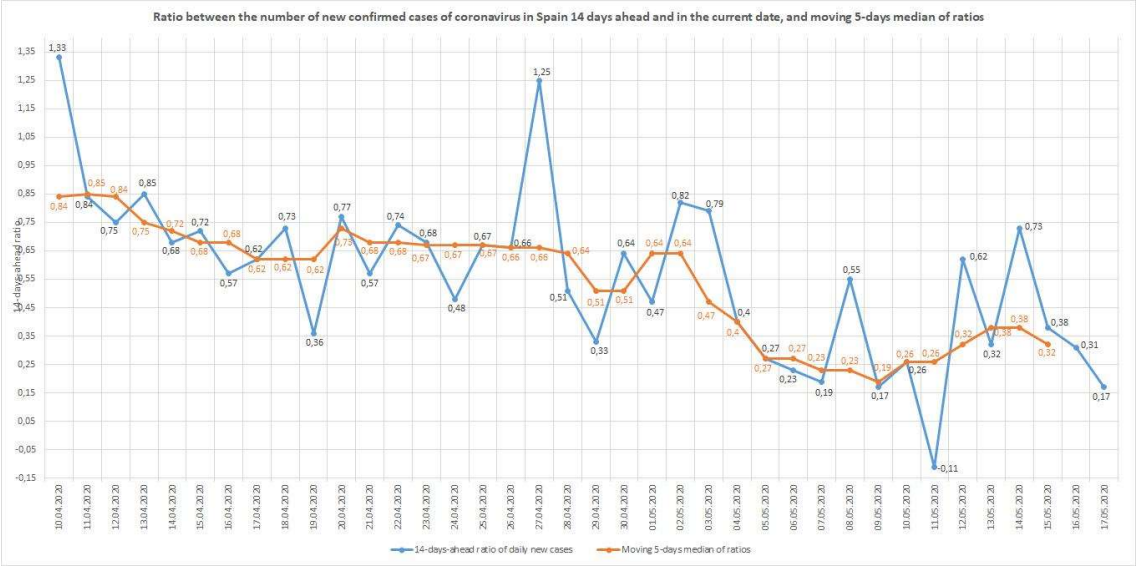


Portugal presented the worst results, rendering conclusions impossible. For the ratios, an oscillating sequence of peaks –certainly due to the continued testing- led them to very high values, greater than one. In the case of the regressions, deaths persisted in the 14 value; thanks to the higher 19 value, on May 12, and the low one (6) on May 15, the original line had a positive inclination. Briefly, as regards daily deaths, the platform around 12 -14 (or even 16) remains. A complete standstill.

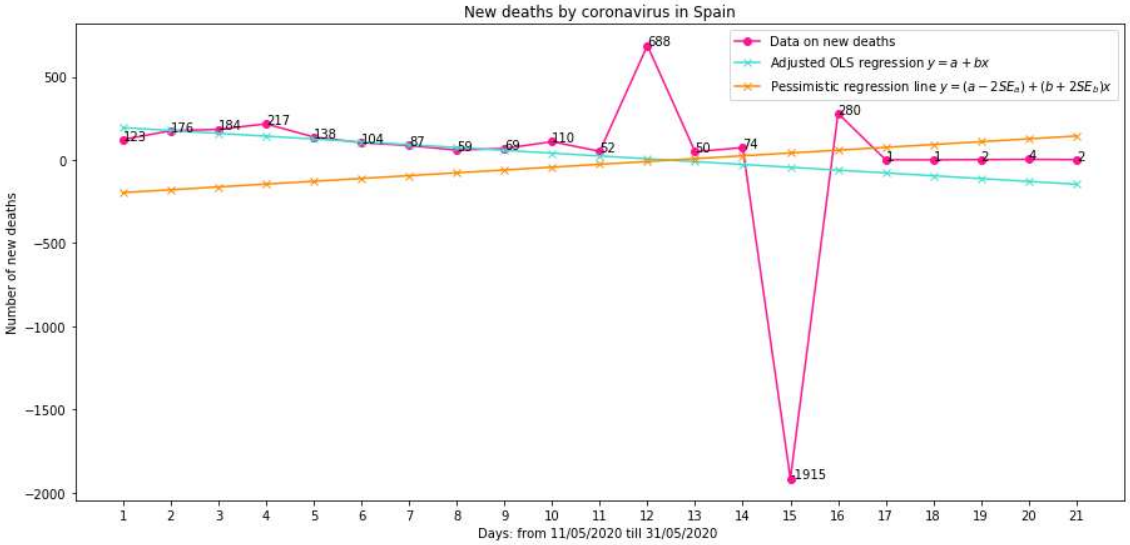
It is hoped that next week, a decreasing trend will be perceived. As for the ratios, it is hard to say something as the three last crude values (in blue) are still higher than one.

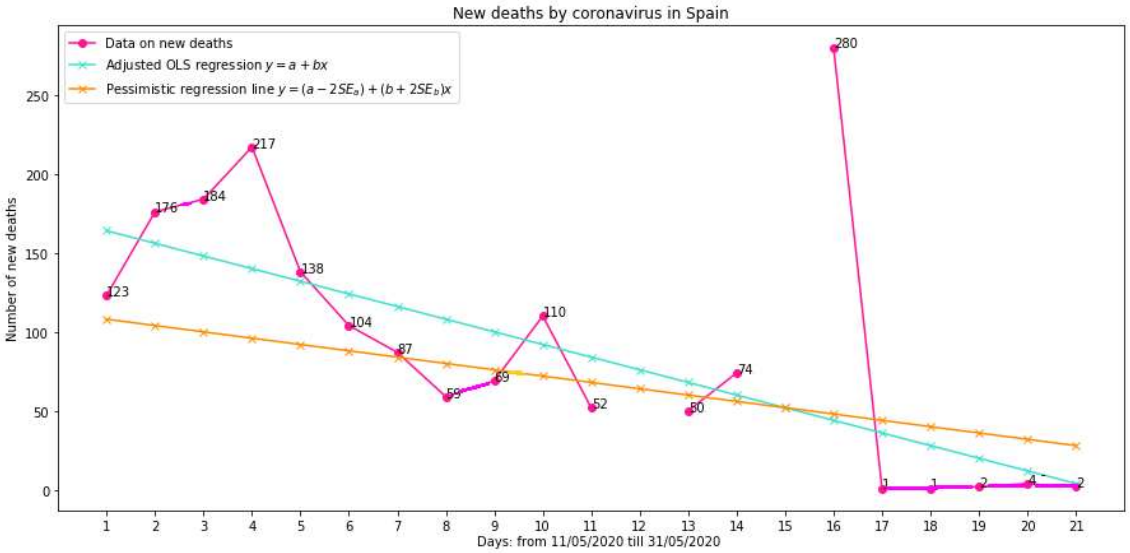
Exhibit 6: Spain.

5.a) Ratio between the number of new confirmed cases of coronavirus 14 days ahead and in the shown date, starting on April 10, 2020 (actual values and five-days moving medians).



5.b) The regression line for the last 'new daily deaths' 21 observations.





Data problems also disturbed the virtuous path Spain had been following. In the case of the ratios, a negative value -due to the revisions- appeared on May 12, and after, a series of increasing, oscillating peaks and troughs took place. The moving median attenuated somewhat this variability, but the textbook decline that was already touching ratios lower than 0.20, changed to a still declining trend but now at 0.32.

As for deaths, the first graph exhibits the disruptive outlier of -1915, on May 25 (the 15th day of the observations), clearly due to data revisions and making any fit senseless. For a better view, we decided to leave aside this ghastly outlier and also the one three days before (688), already signalled in the past Monitoring.

The second graph shows a picture that may be closer to reality, with a significant fit and an expected “zero deaths day” on June 7-8.

3. Summing up.

Evidence on the decline of the epidemic has been blurred this week by continued data problems.

The better results came from the "new daily deaths analysis", leaving aside the persistent plateau in Portugal and no clear progress in France. For the other four countries, daily deaths are indeed decreasing, and though at still somewhat high values in Italy (111 and 75, the two-last data[6]), by the second half of June, very likely no or few deaths will take place in all them, maybe even before in Spain. A nice forecast.

Unfortunately, as regards the speed of contagion, the data are still considerably disturbed. Leaving aside again, the chaotic performance of Portugal, followed by France, in the other four, ratios moved up, lying around or superior to 0.60, with Spain -always better- with a 0.32. A reproduction rate of 0.60 – 0.67 for new cases, though positive, is still high, and does not signal a last-phase of decline, when rates must be even lower than 0.32. The good side, given that deaths in the four countries are clearly decreasing, is that such rates help the four countries to get closer to herd immunity, a point not clear they will be able to reach before autumn.

[6] But decreasing at a faster speed than, for instance, Belgium (according to the fitted lines: 2/day in Belgium, and 7/day in Italy).