

Cadamuro, Leonardo; Marcus, Scott; Papadia, Francesco

Working Paper

COVID-19 in the European Union: Health impacts and effects on economic activity

Bruegel Working Paper, No. 13/2022

Provided in Cooperation with:

Bruegel, Brussels

Suggested Citation: Cadamuro, Leonardo; Marcus, Scott; Papadia, Francesco (2022) : COVID-19 in the European Union: Health impacts and effects on economic activity, Bruegel Working Paper, No. 13/2022, Bruegel, Brussels

This Version is available at:

<https://hdl.handle.net/10419/270467>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

COVID-19 IN THE EUROPEAN UNION: HEALTH IMPACTS AND EFFECTS ON ECONOMIC ACTIVITY

LEONARDO CADAMURO, J. SCOTT MARCUS AND FRANCESCO PAPADIA

This paper quantitatively describes different aspects of the COVID-19 pandemics: new cases, hospitalisations, intensive-care admissions and deaths, while illustrating their changing relationships over time. It then assesses how the different variables have affected relevant sectoral and macroeconomic indicators. Finally, it concludes that, from an economic perspective, what matters when it comes to managing the pandemic is to prevent intensive-care admissions and deaths arising from COVID-19. The success of vaccination should be measured in terms of its ability to prevent the most serious consequences, rather than its ability to prevent infections and hospitalisations.

Leonardo Cadamuro is a PhD Student at Hitotsubashi University
J. Scott Marcus (scott.marcus@bruegel.org) is a Senior Fellow at Bruegel
Francesco Papadia (francesco.papadia@bruegel.org) is a Senior Fellow at Bruegel



Recommended citation:

Cadamuro, L., J.S. Marcus and F. Papadia (2022) 'COVID-19 in the European Union: health impacts and effects on economic activity', *Working Paper 13/2022*, Bruegel

1 Introduction

Many countries are easing, or have eliminated altogether, COVID-19 related restrictions. Has the situation with the pandemic, in late June/early July 2022, really improved to the point where this is warranted?

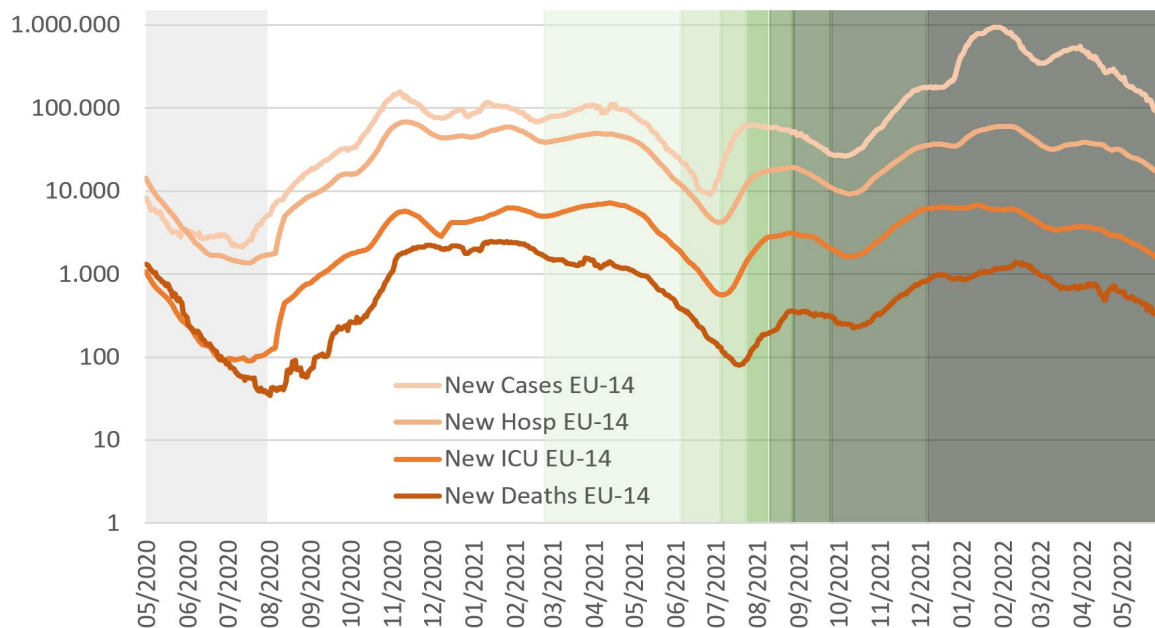
This paper first provides a synthetic and quantitative assessment of the epidemiological parameters that underlie the general feeling that the situation has improved. It does so through empirical examination of developments in the number of new cases, hospitalisations, intensive-care admissions and deaths, and of their changing relationships over time. Specifically, we assess the ways in which the relationships between these variables have evolved, and how they impact on sectoral and macroeconomic indicators.

Section 2 briefly examines raw data about the COVID-19 pandemic, while Section 3 focuses on ratio analysis to explain the evolution of key trends. Section 4 adds economic data to the study and applies regression analysis, while Section 5 concludes. Annex 1 describes how we dealt with a specific aspect of poor data quality, while Annex 2 clarifies the nature and the sources of the data.

2 Synthetic infection developments

Figure 1 depicts the 7-day average of new COVID-19 cases, hospitalisations, intensive-care admissions and deaths in selected European Union countries with daily frequency. The figure shows the time series for the 14 EU countries for which we have sufficiently comprehensive data. We use a log scale to clarify the trend of these figures and to make it easier to compare them despite the different scales. The figure starts only in May 2020 because the figures for infections (new cases) are not credible for the first months of the pandemic. Indeed, even the data between May and August, shaded in grey in the left part of Figure 1, is of poor quality, as discussed in Annex 1. Therefore, our analysis concentrates on the data starting in August 2020. The green shading on the right of Figure 1 reflects increasing degrees of vaccination, with the colour becoming darker for every additional 10 percent of the population fully vaccinated.

Figure 1: New cases, hospitalisation, intensive-care admissions and deaths per million inhabitants (7-day average), selected EU countries, May 2020-May 2022, logarithmic scale



Source: Bruegel based on Our World in Data in Covid-19 dataset (<https://github.com/owid/covid-19-data/tree/master/public/data>). The 14 countries covered are Cyprus, Czechia, Estonia, France, Germany, Greece, Ireland, Italy, Latvia, Malta, Netherland, Slovenia, Spain and Sweden. These 14 countries represent 71.25 percent of the total EU population.

The number of reported new cases grew exponentially over the second half of 2020, plateaued in the following six months, and exhibited irregular behaviour before reaching a much higher peak at the beginning of 2022 than in 2020, followed by a sustained decrease, bringing it back to the level reached in the spring of 2021. Thus, the progressive increase in the vaccination rate has not been accompanied by a visible and sustained reduction in the number of new cases reported. This development may in part reflect better testing and reporting, leading to an improvement in the proportion of documented cases. It may also reflect in part the fact that for the Delta and Omicron variants of COVID-19, vaccines are more effective in preventing severe outcomes than in preventing infection.

The number of hospitalisations followed a similar pattern to that of new cases in 2020 and 2021. However, around the middle of 2022, the level of hospitalisations was significantly lower, by more than 60 percent, than it was prior to vaccination rollouts. Intensive-care admissions have followed the same waves as new cases and hospitalisations, but their level in the summer of 2022 is nearly 80 percent lower than the previous peak in the spring of 2021. The strongest effect of vaccination seems to apply to deaths: while the waves in this variable follow a similar pattern to the previous three variables, at the

end of the period, the number of deaths had decreased by about 90 percent relative to the peak that was reached between the end of 2020 and March 2021. In conclusion, the strongest effect of vaccination seems to be on deaths, followed by hospitalisations and intensive-care admissions. But the number of reported new cases has not declined consistently. This is further explored in Section 3.

3 Ratio analysis

We use ratio analysis to study how the relationships between pandemic variables changed over time. Our primary rationale for exploring these ratios is to better understand the degree to which vaccination has had differential effects on infection versus hospitalisation, intensive-care admissions and deaths, and likewise to understand the degree to which the Delta and Omicron variants have differed from one another and from previous variants.

We calculate four ratios among pandemic variables with daily frequency. The five ratios are:

1. $r_1 = h_t/i_{t-k}$
2. $r_2 = I_t/h_{t-k}$
3. $r_3 = d_t/I_{t-k}$
4. $r_{3.5} = d_t/h_{t-k}$
5. $r_4 = d_t/i_{t-k}$

Where:

i_t is the number of new cases at time t

h_t is the number of hospitalisations at time t

I_t is an indicator of ICU admissions at time t

d_t is the number of deaths at time t

t is time; k is number of lags, different for different variables¹.

¹ The denominator variable is lagged when useful. The precise definition of the variables is in Annex 2.

The ratios respect the following relationship:

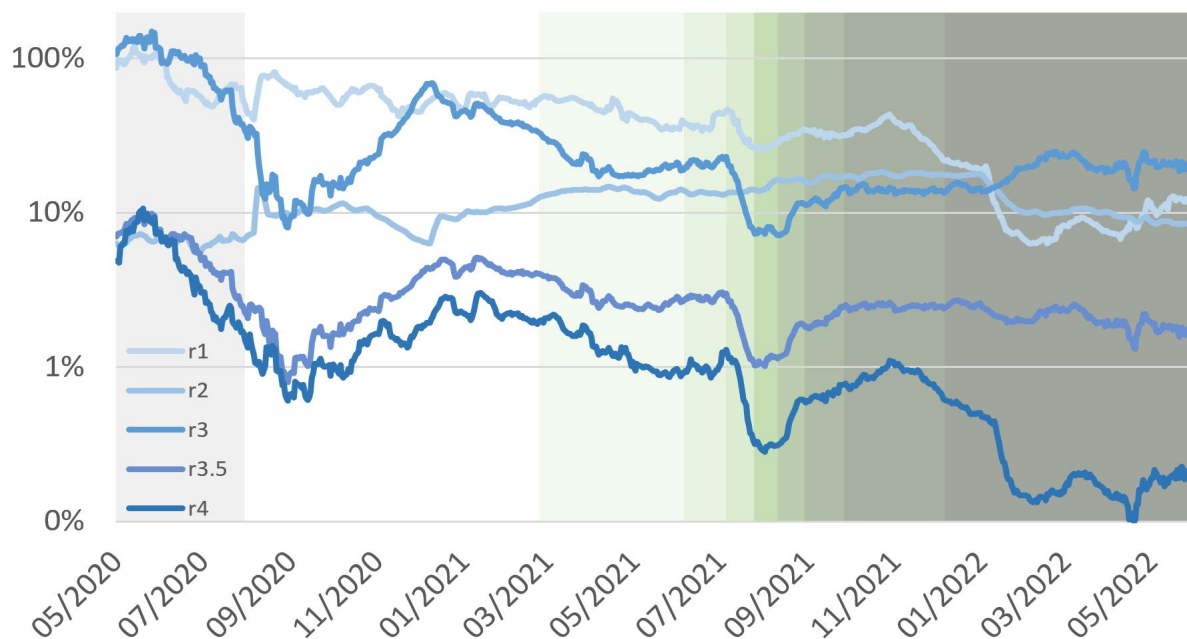
$$r_4 = r_1 \times r_2 \times r_3$$

We use ratio analysis to test whether r_4 , an indicator of mortality per infection, has decreased and, if so, by how much and what has been the contribution to this decline from the lower number of hospitalisations relative to the number of infections (r_1), the lower recourse to intensive care for hospitalised patients (r_2), and of the number of deaths relative to patients admitted to intensive care (r_3). Furthermore, we can document the behaviour of the ratios over time, connecting them to vaccination campaigns.

The evidence in Figure 2 can be summarised as follows:

- The strongest effect of vaccination seems to be on deaths relative to documented infections (r_4), which decreased, albeit irregularly, from around 2 percent of documented infections in August 2020 to less than 0.2 percent in the spring of 2022. One cannot exclude, however, that part of the decline was because the reporting of infections became more complete over time.
- The reduction in hospitalisations relative to documented new infections (r_1), from about 56 percent to 8 percent over the same time period as above, is also dramatic, meaning that far fewer people known to be infected required hospitalisation.
- By contrast, the ratio between intensive-care admissions and hospitalisations (r_2) does not show any sustained change.
- The behaviour of the ratio between deaths and intensive-care admissions (r_3) shows the greatest variation, with no obvious or consistent trend until the beginning of vaccinations; there was then a sustained decline until the autumn of 2021, followed by a creeping up until the end of the period.
- There is also a noteworthy reduction in the ratio of deaths to hospitalisations ($r_{3.5}$), but only by a factor of 2 instead of a factor of 10, as in the ratio of deaths relative to documented infections (r_4). Particularly interesting is the fact that this ratio stayed stable from October 2021 to May 2022, while the ratio of deaths relative to documented infections (r_4) decreased, and the ratio of deaths relative to intensive-care admissions (r_3) increased. Taking these trends together, it seems likely that hospitals changed their practices as to the fraction of admitted patients that wound up in intensive care.

Figure 2: Ratios among documented infections, hospitalisations, ICU admissions, and deaths, selected EU countries (May 2020 – May 2022)



Source: Bruegel. Note: the ratios are calculated by lagging hospitalisations, ICUs, and deaths by 6, 8 and 14 days relative to infections. See note to Figure 1 for countries covered.

Overall, the deadliness of COVID-19 appears to have declined dramatically between August 2020 and the summer of 2022, and fewer of those with documented infections required hospitalisation. By contrast, there was little change in the number of infected people going from normal hospital beds into intensive care and the number of people dying while in intensive care.

Several distinct factors have likely contributed to the marked decline of deaths relative to the number of reported cases in the EU, as is clearly demonstrated by our ratio analysis (notably for r_4).

First and foremost, the widespread take-up of effective vaccines against the SARS-CoV-2 virus (even if perhaps less than ideal) has made a huge difference in reducing the frequency of severe illness and death. This finding is unsurprising and is in line with multiple epidemiological studies undertaken in the United Kingdom and the United States.

Second, and relatedly, while the most recent Omicron virus variant (including both the BA.1 and BA.2 sub-variants) is more contagious than Delta, it is less deadly than prior variants, including Delta. For the UK, Nyberg *et al* (2022) found that the overall likelihood of death from Omicron was 69 percent lower than from Delta (with significant differences by age), which is fully in line with the 67 percent decrease

found by Ward *et al* (2022). Differences in inherent virulence have been offset partly by the fact that the vaccines have been somewhat less effective against Omicron than against Delta.

Nyberg *et al* (2022) also found that the likelihood of hospital admission from Omicron was 59 percent lower than from Delta.

Third, European health systems have become less prone to overload than in the early days of the pandemic. Overloaded health systems, in particular overloaded intensive-care units, have often been linked to increases in the number of infections leading to deaths.

The evolution over time among these variables likely reflects a mix of factors, including:

- The percentage of the population immune due to vaccination, prior infection or both.
- Restrictions (non-pharmaceutical interventions, or NPIs) imposed by health authorities.
- Changes in behaviour voluntarily implemented by individuals based on their own perception of the epidemiological risk.
- Differences in the virulence of the Delta and Omicron variants relative to one another, and to prior COVID-19 variants.
- Improved care overall.
- Changes in hospital practices in relation to the fraction of admitted COVID-19 patients who are admitted to intensive care.
- Better reporting over time, especially on the number of individuals infected.

4 Estimating economic effects

This section explores the relationships between the different COVID-19 variables and economic activity. We split the economic variables into two groups: the first includes broader indicators – GDP and industrial production – which thus cover a range of sectors that have suffered differently in the pandemic; the second group of variables comprises airline flights and tourism, which are examples of specific sectors that might be expected to be most exposed to the pandemic and to resultant restrictions. We selected these variables because they seemed likely to shed light on the underlying relationships between infections and the economy.

In Table 1, we report regression results in which tourism, GDP, industrial production and airline passenger numbers are regressed against the different COVID-19 variables (infections, hospitalisations, intensive-care admissions and deaths) and against a stringency index² summarising different degrees of 'lockdown measures' for the period from August 2020 to May 2022. To regress economic variables (typically available only at monthly frequency) against pandemic variables, we switched to monthly frequency and calculated the percentage of change compared with the same month in 2019 for both pandemic variables and economic variables. Statistically significant results are highlighted in green (99 percent significance), orange (95 percent significance) and yellow (90 percent significance). In the regressions for GDP and industrial production, the effective exchange rate of the euro and the global manufacturing Purchasing Managers' Index (PMI) are included as control variables.

Table 1: Regression results

		Alpha	Beta Variable	Beta Stringency	Gamma REER	Gamma Global PMI	R ²	Adj. R ²	Durbin Watson	Obs
Tourism	New Cases	-0,31 ***	1,72E-07 ***	-7,33E-03 ***			0,27	0,26	1,88	184
		0,05	6,10E-08	8,91E-04						
	Hospitalization	-0,40 ***	-2,12E-07	-4,92E-03 ***			0,17	0,16	1,88	145
		0,07	1,73E-07	1,22E-03						
	ICU	-0,41 ***	6,40E-08	-5,01E-03 ***			0,12	0,10	1,80	119
0,08		1,47E-06	1,36E-03							
New Deaths	-0,35 ***	-5,64E-06	-5,97E-03 ***			0,25	0,24	1,88	184	
	0,06	3,71E-06	1,04E-03							
GDP	New Cases	-0,06 **	-2,65E-09	7,46E-06	4,28E-05	5,93E-03 ***	0,39	0,38	1,82	241
		0,03	3,45E-09	4,60E-06	3,32E-04	4,94E-04				
	Hospitalization	-0,10 ***	-1,60E-08 **	1,14E-05 *	4,08E-04	6,07E-03 ***	0,39	0,38	1,90	204
		0,04	7,87E-09	6,08E-06	3,65E-04	5,59E-04				
	ICU	-0,03	-9,87E-07 ***	7,14E-06	-1,52E-04	5,10E-03 ***	0,40	0,38	2,10	147
0,05		1,59E-07	2,08E-04	4,52E-04	6,68E-04					
New Deaths	-0,06 *	-6,66E-07 ***	1,44E-05 ***	1,70E-05	5,79E-03 ***	0,40	0,39	1,89	241	
	0,03	2,51E-07	5,04E-06	3,27E-04	4,76E-04					
Industrial Production	New Cases	0,87 ***	1,53E-09	5,36E-05	-9,29E-03 ***	1,44E-02 ***	0,07	0,05	2,42	219
		0,31	1,40E-08	3,80E-05	3,26E-03	4,64E-03				
	Hospitalization	1,30 ***	-1,92E-08	8,59E-05	-1,37E-02 ***	1,85E-02 ***	0,11	0,09	2,42	175
		0,37	5,94E-08	5,48E-05	3,82E-03	6,16E-03				
	ICU	1,67 ***	-4,50E-06 ***	-1,83E-04	-1,66E-02 ***	1,66E-02 **	0,14	0,12	2,57	136
0,55		1,70E-06	2,17E-03	4,94E-03	7,66E-03					
New Deaths	0,86 ***	-1,71E-06	7,85E-05	-9,13E-03 ***	1,45E-02 ***	0,07	0,06	2,42	219	
	0,31	2,22E-06	4,76E-05	3,23E-03	4,61E-03					
Flight Passengers	New Cases	-0,25 ***	6,60E-08 ***	-5,85E-03 ***			0,25	0,24	1,87	261
		0,04	1,37E-08	7,20E-04						
	Hospitalization	-0,23 ***	8,11E-08	-5,66E-03 ***			0,16	0,16	1,60	218
		0,05	1,10E-07	8,96E-04						
	ICU	-0,26 ***	1,16E-06	-4,85E-03 ***			0,12	0,11	1,57	175
0,06		9,50E-07	1,02E-03							
New Deaths	-0,25 ***	-2,46E-06	-5,43E-03 ***			0,19	0,18	1,62	261	
	0,05	2,90E-06	8,41E-04							
		0,1 >>	0,05 >>>	0,01 >>>>	0					

Source: Bruegel.

² This is the overall index developed by Oxford University that summarises the restriction level for COVID-19 in each country. See <https://covidtracker.bsg.ox.ac.uk/>.

The results are interesting as they show that:

- Broad variables (GDP and Industrial production) dropped in periods when intensive-care admissions were high, and the relationship is statistically significant. GDP is affected also by new deaths. Overall, severe outcomes from the pandemic are important for both broad variables.
- We do not find, by contrast, a significant relationship between broad economic variables and milder occurrences, such as the number of cases and hospitalisations. This may be in part a function of lower reliability of data on new infections, which depend heavily on the level of testing (see Annex 1) and many other potentially confounding factors; however, the most important message to draw from this lack of relationship is that the numbers of reported cases and hospitalisations, in and of themselves, have little impact on these key macroeconomic indicators. The *severity* of cases is a better indicator.
- Our analysis does not specifically reveal the mechanism through which severe pandemic outcomes such as intensive-care admissions impact GDP and industrial production. It seems clear that the reduction in the number of people in the labour force was not large enough to account for the observed changes in GDP and industrial production. A more likely explanation is that voluntary distancing on the part of the public increased during the most acute phases of the crisis, when intensive care units were saturated, and the risk of death was significant (see Toxvaerd, 2020; and Goolsbee and Syverson, 2021). Media attention likely acted as a signal to the public to withdraw from both production and consumption, with strong effects on these two variables.
- Both control variables, the effective exchange rate and global industrial PMI, are highly significant for industrial production. Only the global PMI is significant for GDP and has strong explicatory power.
- New cases and hospitalisations are, by contrast, very significant for tourism and air passenger numbers.
- The stringency index is highly significant for tourism and airline passengers beyond any effect from COVID-19 variables. This was to be expected insofar as restrictions impact directly on the ability to travel.

- From August 2020 to May 2022, the stringency index does not seem to play an important role in GDP or industrial production. For the initial phase of the pandemic, from February 2020 to July 2020, however, Sapir (2020) found that “countries with stricter lockdown measures witnessed a greater economic collapse, with the causality running from the strictness of the lockdown measures to the economic collapse.” In Sapir (2020), the stringency index explained about 50 percent of the GDP loss recorded in the initial phase of the pandemic. When we apply our methodology to a regression analysis between the stringency index and GDP (taking COVID deaths into account), we likewise find a significant impact for the same period analysed by Sapir.

The results of our analysis, together with those in Sapir (2020), naturally raise two questions. First, why are the results dramatically different between the early phase of pandemic (February 2020 to July 2020) and the later phase?

Our prior expectation was that the stringency index would demonstrate an inverse relationship with GDP, ie that more stringent restrictions would lead to a loss of GDP. This expectation was fulfilled for the period from February to July 2020, but not for the period after August 2020. How should one interpret these results?

The most likely explanation is that the lockdown measures put in place from February 2020 to July 2020 were extreme by today’s standards, limiting the ability to work in nonessential businesses or to open stores other than those that sell food, medicines and other necessities of daily life. Over the past two years, governments and public health authorities have become progressively more sophisticated in implementing restrictions (NPIs), so that they damage the broader economy as little as possible. It is likely that the Oxford stringency index alone does not fully capture these nuances. If this is so, then the difference between the results from February to July 2020 (a significant impact on GDP) versus those from August 2020 to May 2022 (no impact) is likely a measure of the growing success of health authorities in imposing restrictions when needed, in such a way as to do minimal harm to the broader economy.

Under this interpretation, factors other than restrictions that were formally imposed likely played the dominant role during the later stages of the pandemic. The relationships between the number of cases, the number of deaths, the restrictions imposed and their impact on individual mobility are complex (see, for instance Marcus and Guetta-Jeanrenaud, 2021). One explanation for the lack of statistically significant relationship between the stringency index and GDP during the later stages of the pandemic

could be that voluntarily-adopted measures had a greater impact than those imposed by health authorities. This explanation is in line with Goolsbee and Syverson (2020): *“While overall consumer traffic [in the US] fell by 60 percentage points, legal restrictions explain only 7 percentage points of this. Individual choices were far more important and seemed tied to fears of infection”* (see also Toxvaerd, 2020).

5 Conclusions

Our results suggest that, as long as COVID-19, or any other infection, only leads to new cases but not to the more serious consequences of intensive-care admissions or deaths, the macroeconomic effects need not be so large. What matters is preventing intensive-care admissions and deaths arising from COVID-19. From a macroeconomic perspective, the success of vaccination should be measured in terms of its ability to prevent the most serious consequences, rather than its ability to prevent infections and hospitalisations.

The tendency in Europe over roughly the period February to May 2022 has been to ease up on restrictions, even if the number of new Omicron cases per day continues to be extremely high. Our findings support this course: where health facilities are not overloaded and where serious outcomes (intensive-care admissions and deaths) are at moderate and manageable levels, easing up on restrictions may be a sensible policy course in order to reduce negative economic effects.

Our results also have a bearing on possible new pandemics. First, the lessons learned during the current pandemic of how to impose restrictions when necessary, without totally crippling the economy, must not be forgotten. Second, overall health policies should primarily seek to limit severe occurrences such as intensive-care admissions and deaths. Limiting less grave consequences, such as infections and hospitalisations, is less critical.

Our analysis has numerous limitations, most of which stem in one way or another from the huge uncertainties in the future evolution of COVID-19. First of all, our conclusions about how to deal with future pandemics, or with new virus variants as part of the current pandemic, reflect experience with the variants of the virus that have circulated to date in Europe and worldwide. If a future variant were to have substantially different characteristics in terms of contagiousness, virulence, or both, our conclusions might well need to be revisited.

Second, we do not deal with the possible long-term health consequences of COVID-19, which could have persistent negative effects on economic activity. Relatedly, the effect of repeated infections is not known.

Third, our analysis is economic, and does not seek to consider many possible dynamic developments in the evolution of COVID-19. On the one hand, greater circulation of the virus encourages natural immunity; on the other hand, it carries the risk of more dangerous mutations.

Any of these developments might justify a different approach to testing and to restrictions (NPIs) than what we have sketched out in this paper.

References

Goolsbee, A. and C. Syverson (2020) 'Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline 2020', University of Chicago, Becker Friedman Institute for Economics *Working Paper* No. 2020-80, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3631180

Marcus, J.S. and L. Guetta-Jeanrenaud (2021) 'What will be the impact of Europe's next round of COVID-19 restrictions?' *Bruegel Blog*, 23 December, available at <https://www.bruegel.org/2021/12/what-will-be-the-impact-of-europes-next-round-of-covid-19-restrictions/>

Nyberg, T., N.M. Ferguson, S.G. Nash, H.H. Webster, S. Flaxman, N. Andrews ... S. Thelwall (2022) 'Comparative analysis of the risks of hospitalisation and death associated with SARS CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study', *The Lancet* vol. 399, available at [https://doi.org/10.1016/S0140-6736\(22\)00462-7](https://doi.org/10.1016/S0140-6736(22)00462-7)

Sapir, A. (2020) 'Why has COVID-19 hit different European Union economies so differently?', *Policy Contribution* 2020/18, Bruegel, available at <https://www.bruegel.org/policy-brief/why-has-covid-19-hit-different-european-union-economies-so-differently>

Toxvaerd, F. (2020) 'Equilibrium Social Distancing', *Cambridge-INET Working Paper Series* No: 2020/08, Cambridge-INET Institute, available at <https://www.inet.econ.cam.ac.uk/working-paper-pdfs/wp2008.pdf>

Vilches, T.N., S.M. Moghadas, P. Sah, M.C. Fitzpatrick, A. Shoukat, A. Pandey and A.P. Galvani (2022) 'Estimating COVID-19 Infections, Hospitalizations, and Deaths Following the US Vaccination Campaigns

During the Pandemic', *JAMA Netw Open* 2022; 5(1): e2142725, available at <https://doi.org/10.1001/jamanetworkopen.2021.42725>

Ward, I.L., C. Bermingham, D. Ayoubkhani, O.J. Gethings, K.B. Pouwels, T. Yates, K. Khunti, J. Hippisley-Cox, A. Banerjee, A.S Walker and V. Nafilyan (2022) 'Risk of COVID-19 related deaths for SARS-CoV-2 Omicron (B.1.1.529) compared with Delta (B.1.617.2)', medRxiv, 25 February, available at <https://doi.org/10.1101/2022.02.24.22271466> [not yet peer reviewed].

+++

Annex 1: data quality

As in all empirical analyses, we were confronted with an issue of data quality in examining the developments of different indicators of the COVID-19 pandemic. We focus in this annex on the relationship between the number of tests carried out and the number of reported new cases. In principle, the true number of infections should be independent of the number of tests: the virulence of the virus does not depend on how many tests are carried out. However, the number of *known, documented* cases is obviously dependent on the number of tests administered. Even the earliest results based on the outbreak in Wuhan, China made clear that the actual number of infections was four or five times greater than the number of documented cases.

A first indication that the data on new cases is downward biased because of the low number of tests carried out in the first phase of the pandemic is that there were, until June 2020, systematically more COVID-19 hospitalisations than infections, which cannot be. To further explore the data-quality issue affecting new infections and to decide how to deal with it, we document the empirical relationship between reported infections and tests.

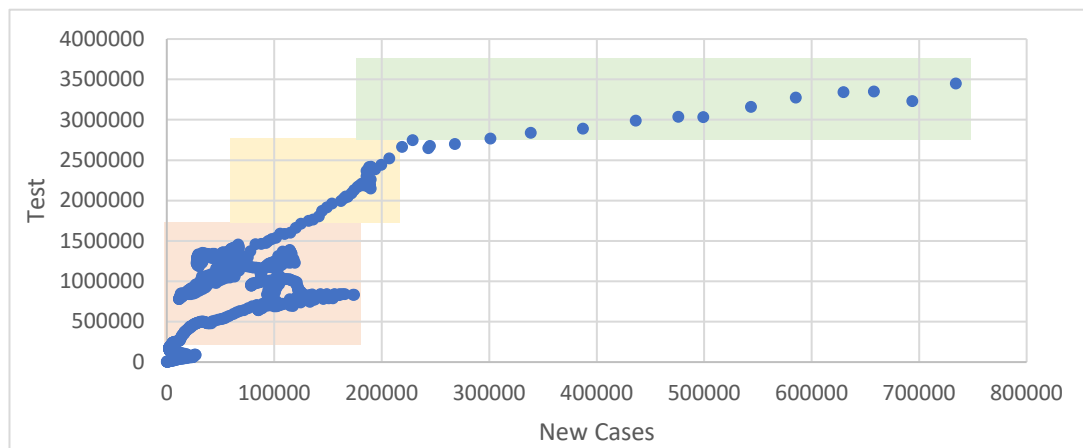
Table 2 reports the correlation between the number of tests and the number of reported new cases, while changing the starting date. Shortening the period by starting later, there is a slight decrease in the correlation between the two variables. In other words, the two variables tend to move less together when the number of tests increases. However, the correlation remains high, confirming the problem mentioned above.

Table 2: Correlation between the number of tests and new cases: shifting initial dates

2020:03	2020:04	2020:05	2020:06	2020:07	2020:08	2020:09	2020:10
0.8210	0.8190	0.8158	0.8128	0.8102	0.8081	0.8035	0.8004

Figure 3 reports the relationship between new cases and the number of tests. We can see three different boxes: the first is when the number of tests is less than 1,750,000 (in orange), the second is between 1,750,000 and 2,700,000 (in yellow), and the last is over 2,700,000 (in green) reached around the end of 2021.

Figure 3: Scatter diagram between tests and new reported cases



In the first box, the relationship between the tests and new cases is unclear. This was the case in the first months of a pandemic, when there wasn't a clear organisation and tests were carried out randomly among the population. In the second box, we see a clear and steep positive relationship. When the number of test increases so does the number of reported cases. Finally, in the third box, the connection becomes less intense, and the number of reported cases is nearly independent from the number of tests.

Table 3 reports a test of stationarity of the series of reported new cases, using samples starting from different dates, to test when the time series become stationary.

Table 3: Adjusted Dickey-Fuller test.

Adjusted Dickey-Fuller with Trend			
Data	Test Value	P-Value	Stationary
01/05/2020	-3.2072	0.0866	NO
15/05/2020	-3.2506	0.0792	NO
01/06/2020	-3.2851	0.0732	NO
15/06/2020	-3.3326	0.0662	NO
01/07/2020	-3.3595	0.0604	NO
15/07/2020	-3.3989	0.0536	NO
01/08/2020	-3.4298	0.0491	YES

The series of new reported cases becomes stationary in August 2020. So, we feel that we can trust data on new infections since that date. Data on hospitalisations, intensive care and deaths do not suffer from the same problem as new infections since they relate to more visible events than infections. Therefore, when useful, we use data going further back in the past, namely starting in May 2020.

Annex 2: data description

We organise this section by dividing data into three types: COVID-19 data, economic data and control variables. The COVID-19 data indicates all data relative to the pandemic situation: new cases, hospitalisations, intensive-care admissions, deaths, stringency index and vaccination rate. Economic data which could be affected by the COVID-19 pandemic includes the overall economy or some specific sectors. They are GDP, industrial production, tourism and airline passengers. Finally, control variables represent the possible exogenous variables (different from the COVID-19 pandemic) that could affect economic variables. They are the real effective exchange rate of the euro (REER) and the global Purchasing Managers' Index (PMI).

COVID-19 data is provided by Our World in Data, with daily frequency from March 2020 to May 2022. Only the stringency index is downloaded from a different dataset, specifically the Oxford University Systematic dataset of Covid-19 policy (OxCGRT) for the same time period and daily frequency. All data is provided per country and we computed the European aggregate for the 14 countries that have data for all variables: Cyprus, Czech Republic, Estonia, France, Germany, Greece, Ireland, Italy, Latvia, Malta, Netherland, Slovenia, Spain and Sweden. These 14 countries represent around 71.25 percent of the total EU population. We define variables in this dataset as:

- *New cases*: indicate the number of reported new infections from COVID-19;
- *Hospitalisations*: indicate the number of people that start to be hospitalised for COVID-19;
- *ICUs*: indicate the number of people that are admitted to intensive-care units for COVID-19;
- *New Deaths*: indicate the number of people who died from COVID-19;
- *Vaccination rate*: indicates the percentage of the population fully vaccinated with an approved COVID-19 vaccine; we do not take into account the type of vaccine or third doses;
- *Stringency index*: an overall index developed by Oxford University that summarises the restriction level for COVID-19 in each country.

We used daily data to have an overall view of the COVID-19 pandemic and to compute the ratio analysis. Since the data is heavily affected by daily seasonality, we calculated the average of the previous week (7-days average), to avoid spikes due to different collection times among weekdays and to maintain the daily frequency. Where we analyse the impact of the COVID-19 pandemic on economic data, we moved to monthly frequency due to the economic data frequency. In that case, we used the monthly average calculated from daily data.

In the case of economic data, we have monthly data representing macro data as GDP and industrial production, and sectoral data as the number of nights spent in tourist accommodation (called ‘tourism’) and the number of airline passengers. We use *Leading Indicators OECD: Reference series: Gross Domestic Product (GDP)* from the FRED dataset as a proxy for monthly GDP while using data from International Monetary Fund datasets for industrial production. Tourism and airline passengers were provided by Eurostat. For each variable, we used the percentage difference compared to the same month in 2019. At time of writing, not all economic variables were available for May 2022 so we used data up to April 2022.

Finally, control variables are the real effective exchange rate (or REER) and global Purchasing Managers’ Index (or PMI). REER for the euro and the main European currencies are available at the European Central Bank Statistical Data Warehouse, while we use the JP Morgan Global Manufacturing PMI as a proxy for global PMI. Both datasets have monthly frequency. To check for multicollinearity, we calculated the variance inflation factor test as reported in Table 4.

Table 4: Variance inflation factor test for each regression

		Variable	Stringency Index	REER	Global PMI
Tourism	New Cases	1.05	1.05		
	Hospitalisation	1.33	1.33		
	ICUs	1.17	1.17		
	New Deaths	1.39	1.39		
GDP	New Cases	1.92	2.66	1.79	1.07
	Hospitalisation	2.72	3.35	1.60	1.01
	ICUs	1.17	1.76	1.49	1.11
	New Deaths	2.44	3.28	1.79	1.02
Industrial Prod	New Cases	1.35	2.27	2.12	1.03
	Hospitalisation	2.42	3.07	1.93	1.01
	ICUs	1.16	1.61	1.35	1.10
	New Deaths	2.60	3.57	2.09	1.02
Flight Pass	New Cases	1.00	1.00		
	Hospitalisation	1.14	1.14		
	ICUs	1.10	1.10		
	New Deaths	1.26	1.26		

Only few variables have a variance inflation factor value higher than 2.5, which would indicate some concern about multicollinearity, but none have a value higher than 5, which would indicate a serious multicollinearity problem. Thus, we proceeded using REER and global PMI as control variables.



© Bruegel 2022. All rights reserved. Short sections, not to exceed two paragraphs, may be quoted in the original language without explicit permission provided that the source is acknowledged. Opinions expressed in this publication are those of the author(s) alone.

Bruegel, Rue de la Charité 33, B-1210 Brussels
(+32) 2 227 4210
info@bruegel.org
www.bruegel.org