

A positive correlation between mask usage and excess mortality in Europe

Supplementary File 2

1 Diagnostics for the main multivariate regression

We start by presenting typical diagnostic plots for our main regression (PLAA against mask usage, people fully vaccinated/hundred, human development index and CardLife). [Figure 1](#) shows the scatter plot of residuals against fitted values, [Figure 2](#) shows the scatter plot of standardized residuals against fitted values and [Figure 3](#) shows the scatter plot of square root of absolute values of standardized residuals against fitted values (usually called scale location plot). [Figure 1](#) and [Figure 2](#) don't show substantial signs of non linear patterns in the residuals and [Figure 3](#) doesn't show substantial signs of heterocedasticity. For comparison, see the supplementary file [simulations.pdf](#) for the corresponding plots based on 10 simulated datasets in which linear model assumptions are satisfied exactly. We also note that Studentized Breusch–Pagan test for heterocedasticity does not reject the null hypothesis of homocedasticity ($p=0.64$). In [Figure 4](#) we show a QQ plot comparing the distribution of the standardized residuals with a normal distribution. The plot shows no signs of a substantial deviation from normality. For comparison, see the supplementary file [simulations.pdf](#) for the corresponding QQ plots based on 10 simulated datasets in which linear model assumptions are satisfied

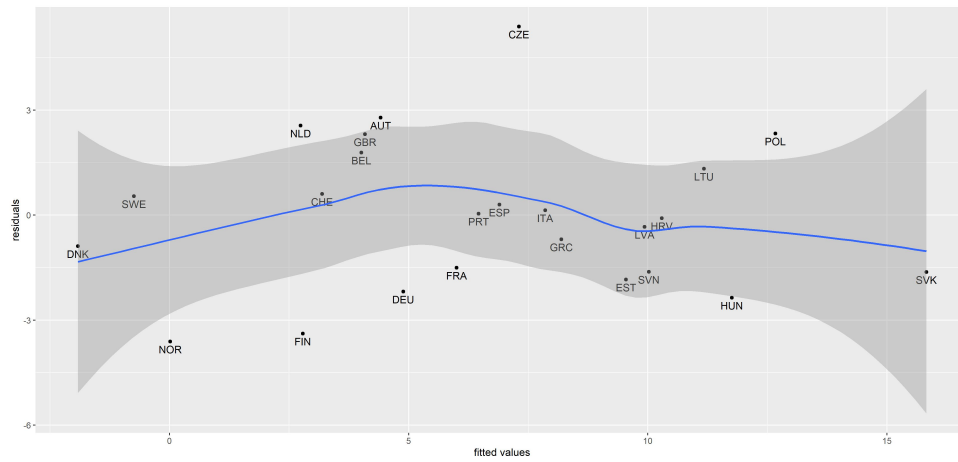


Fig 1. Scatter plot of residuals versus fitted values for main regression

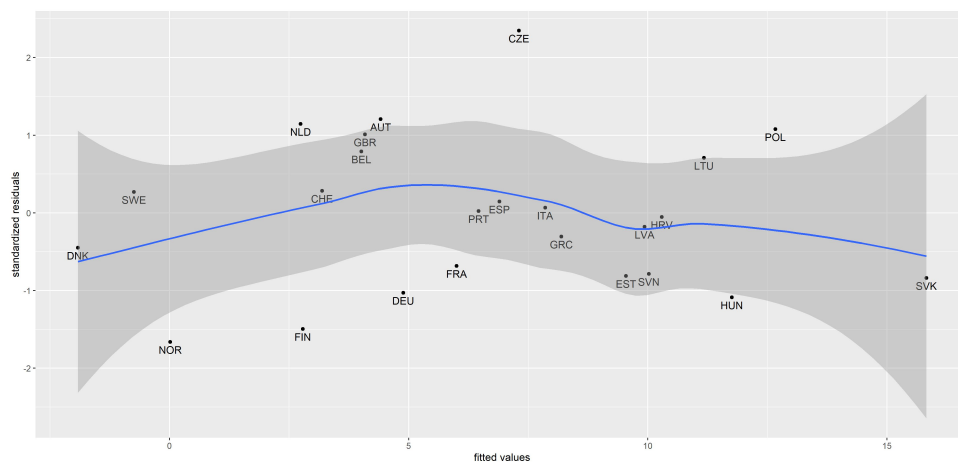


Fig 2. Scatter plot of standardized residuals versus fitted values for main regression

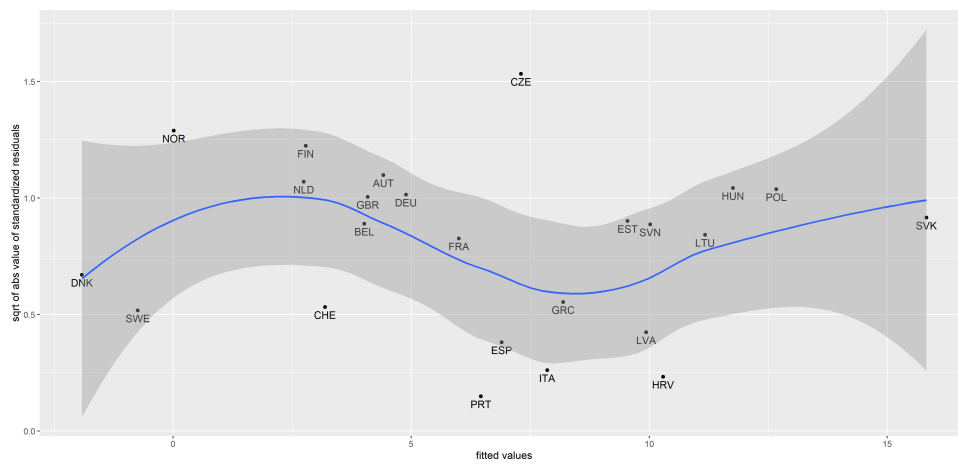


Fig 3. Scale location plot for main regression

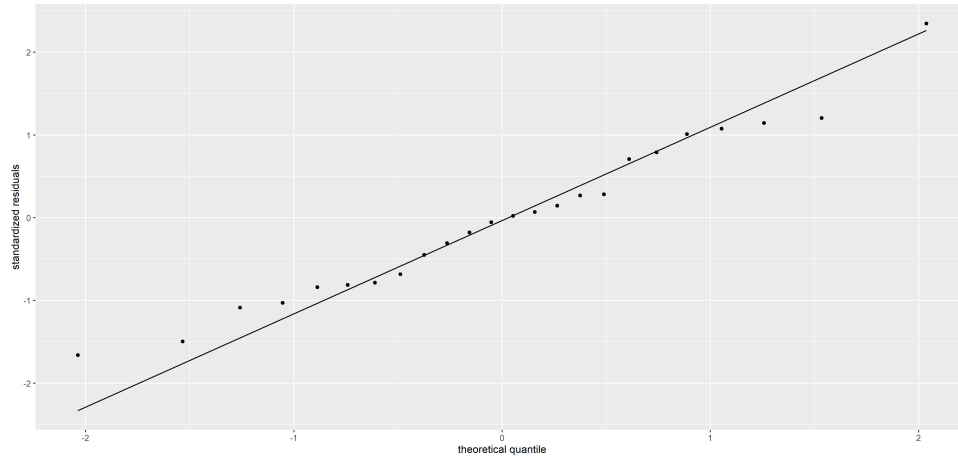


Fig 4. QQ plot of standardized residuals against normal distribution for main regression

exactly. The estimated skewness for the standardized residuals equals 0.40 (95% IC, [-0.31,1.33]) and the estimated excess kurtosis equals -0.009 (95% IC, [-1.077,3.243]). Moreover, Shapiro–Wilk test does not reject the null hypothesis of normality ($p=0.78$). In order to detect the possible presence of multicollinearity, we regressed each independent variable from the main regression against all other independent variables. The values of (unadjusted) R-squareds for each regression are shown in Table 1. All values are substantially below 0.9, indicating absence of relevant multicollinearity. Cook's distance was used to detect possible overly influential data points (see Figure 5). All distances are substantially below 1, indicating that no data points were particularly influential for the regression results.

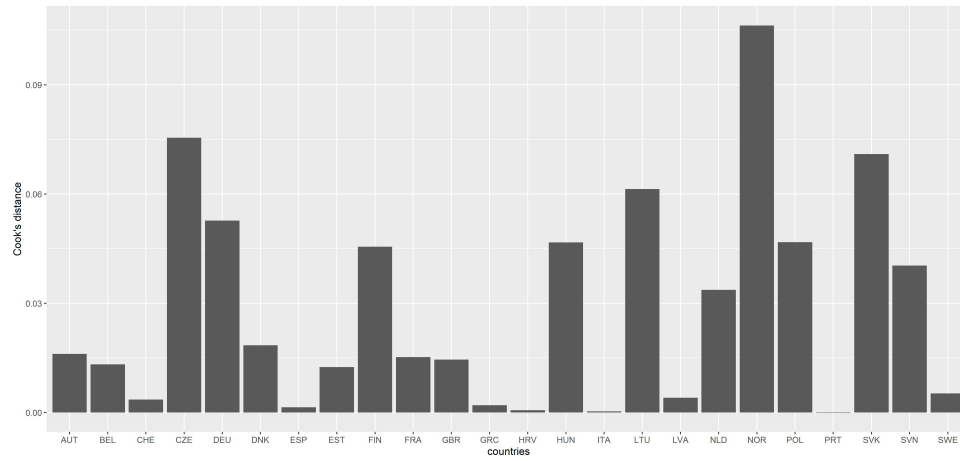


Fig 5. Cook's distances for main regression

Table 1. Multicollinearity check

Independent variable	R^2 against other independent variables
Fully vaccinated/hundred	0.42
Human Development Index	0.69
CardLife	0.60
Mask usage	0.45

2 Sensitivity analyses for the main regression

As sensitivity analyses for our main regression, we used both bootstrap and Huber–White robust sandwich estimators for standard errors. The bootstrap handles violations of normality and heterocedasticity, though it is based on asymptotic theory and works better with large samples. The version of the Huber–White estimators that we used (HC3) works well with violations of heterocedasticity even in small samples (see [Long and Ervin \(2000\)](#)). Table 2 shows the results of the regression with bootstrap and Table 3 shows the results with sandwich estimators.

Table 2. Multivariate regression using bootstrap

Variable	Coefficient	Standardized coefficient	p-value
Intercept	36.7 [-1.3,77.5]		0.063
Fully vaccinated/hundred	-0.2 [-0.3,-0.1]	-0.48 [-0.68,-0.27]	0.00075
Human Development Index	-21.4 [-65.8,18.4]	-0.15 [-0.45,0.13]	0.28
CardLife	0.7 [0.0,1.7]	0.20 [-0.05,0.49]	0.083
Mask usage	13.4 [8.3,18.9]	0.52 [0.30,0.77]	0.00044

Table 3. Multivariate regression using sandwich standard erros

Variable	Coefficient	Standardized coefficient	p-value
Intercept	36.7 [1.0,72.5]		0.045
Fully vaccinated/hundred	-0.2 [-0.4,-0.1]	-0.48 [-0.74,-0.22]	0.00017
Human Development Index	-21.4 [-60.5,17.8]	-0.15 [-0.42,0.12]	0.27
CardLife	0.7 [-0.1,1.5]	0.20 [-0.07,0.48]	0.080
Mask usage	13.4 [8.5,18.3]	0.52 [0.24,0.79]	0.000018

3 Power calculation for alternative multivariate regressions

In a multivariate regression with small sample size the test for nonzero coefficients can become severely underpowered if the number of independent variables included is high. With this in mind, we used power estimates to determine the maximum number of independent variables that could be added to our main regression. The power to detect a nonzero coefficient for a certain variable of interest (in our case, mask usage) depends on sample size, number of independent variables and the values of the following parameters:

- (a) true value of the standardized coefficient of the variable of interest;
- (b) true value of R^2 for the regression of the variable of interest against the other independent variables;
- (c) true value of R^2 for the regression.

We used our main regression as a “pilot study” to obtain an estimate for (a), (b) and (c) (using adjusted R^2 as an estimate of true R^2). The estimates 0.5, 0.4 and 0.75 were used for (a), (b) and (c), respectively, for the power calculations shown in Table 4. Significance level was kept at 0.05.

Table 4. Power estimates

Number of independent variables	Power
4	87.4%
5	85.5%
6	83.5%
7	81.2%
8	78.6%
9	75.7%
10	72.6%

4 Sensitivity analyses using alternative regressions

We performed several sensitivity analyses by adding extra independent variables to our main regression in an attempt to reduce residual confounding. The total number of independent variables was kept not greater than 7 in order to ensure a minimum power of 80% to detect the effect of interest (Table 4). The variables considered were stringency, total tests per thousand inhabitants, population (urban) density, obesity rate, diabetes prevalence, Gini index and percentage of seniors (> 65 years). A total of 64 regressions were computed (one for each subset having at most 3 elements from the set of 7 possible extra independent variables). In all regressions computed vaccination was associated with a decrease in PLAA and mask usage was associated with an increase in PLAA, with both effects being significant at the 0.05 level (see supplemental files [supplementary_file_7-OtherRegressions.pdf](#) and in [this online file](#) for details).

References

Long, J.S., Ervin, L.H., 2000. Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician* 54, 217–224. doi:[10.1080/00031305.2000.10474549](https://doi.org/10.1080/00031305.2000.10474549).