A peer-reviewed version of this preprint was recently published in the *Journal of Population Research*:

Mogi, R., & Spijker, J. (2021). The influence of social and economic ties to the spread of COVID-19 in Europe. *Journal of Population Research (online first)*. https://doi.org/10.1007/s12546-021-09257-1

For those citing our paper, please now use the above citation as reference.

The preprint version can still be consulted for those who do not have access to the journal version.

# The influence of social and economic ties to the spread of COVID-19 in Europe

Ryohei Mogi*, Jeroen Spijker*

Centre d'Estudis Demogràfics, Universitat Autònoma de Barcelona

rmogi@ced.uab.es @rmogimogi @popageing

*Equal contribution

**Background**: By late January 2020, the 2019 novel coronavirus (SARS-CoV-2) had reached Europe and most European countries had registered cases by March 1. However, the spread of the virus has been uneven in both prevalence and speed of its propagation. Given the role of social factors in the spread of contagious diseases and the known demographic, economic and cultural differences in Europe, we analyse the contribution of such factors in the coronavirus disease COVID-19 caused by SARS-CoV-2.

**Methods**: Diagnosed COVID-19 cases from the Center for Systems Science and Engineering at Johns Hopkins University and the European Social Survey for 20 European countries were used to estimate bivariate associations between cumulative reported case numbers ten-day intervals between March 1 and April 20, 2020 and ten social, demographic and economic variables. To avoid overfitting, we first reduce these variables to three factors (labelled "socially and economically vibrant", "high-educated and not aged", and "densely populated and traditional") by means of factor analysis before conducting a multiple regression analysis. We also perform a sensitivity analysis using rates and new cases between two time periods.

**Findings**: The socially and economically vibrant factor has a strong and positive association with COVID-19 throughout the study period. The association with the

densely populated and traditional factor was initially low but by mid-March similar and from April higher than the socially and economically vibrant factor.

**Conclusions**: For future influenza-like pandemics where a vaccine is not yet available, early movement restrictions will be especially crucial to curb the spread of such diseases in economically, socially, and culturally vibrant and densely populated countries.

**Introduction**

As of late April 2020, there were more than 3 million confirmed cases worldwide of the coronavirus disease COVID-19 (*1*) that is caused by the 2019 novel coronavirus (SARS-CoV-2). It is thought that the virus was initially transmitted late 2019 to humans via live animals contaminated by bats that were sold at the Huanan seafood market in Wuhan, China, after which it began to be spread through person-to-person transmission. Person-to-person transmission occurs primarily through respiratory droplets produced when an infected person coughs or sneezes in the close presence of other people (*2, 3*). Although the epidemic initially only appeared to be affecting China and other countries quickly placed international travel restrictions to and from China, the first cases were being confirmed in Europe by early February and by late March about half of the world's reported cases were in Europe (*4*).

One well-established cause of global spread of previous influenza outbreaks is airplane travel (*5*) and without this mode of transportation the coronavirus would not have arrived in Europe so quickly. However, once someone brought the virus to Europe, human interaction and close contact was also required for the virus to spread quickly. Previous research on virus transmission, including on COVID-19, has shown that social contact is very important (*6-10*). At the same time, cultural differences in Europe are well established. In terms of social contact, in southern European countries intergenerational

3

contact is more frequent than in other parts of Europe, as social norms about providing support to family members and maintaining interpersonal familiar interactions are stronger than in the less family-oriented Western and Northern European countries (*11-12*).

Patterns in social mixing are therefore embedded in socioeconomic and cultural factors and are different across Europe, making this continent an excellent geographical area to the study of COVID-19 proliferation. This study aims therefore to ascertain whether countries where social interaction is considered to be more common have more confirmed cases of COVID-19. We examine this by analyzing statistical associations between different indicators of social ties and the reported number of confirmed cases of COVID-19 in 23 European countries between March 1 and April 20, 2020.

**Data and method**

Data on the number of confirmed cases of COVID-19 come from the Center for Systems Science and Engineering at Johns Hopkins University (*1*) (https://coronavirus.jhu.edu/map.html). Cross-sectional data for 18 countries from Round 9 of the European Social Survey (ESS), conducted in 2018, 4 from Round 8 (2016) and 1 from Round 7 (2014), was used to create variables to proximate different types of social ties, which we hypothesize to be positively associated with COVID-19. i) the average number of household members; ii) percentage living in a multi-generational household; iii) proportion of people who have a frequent social meetings with friends, relatives, or colleagues; iv) religious attendance. In addition, we also test the effect of the socioeconomic variables: v) tertiary education and vi) GDP per capita (current US$) as we assume that countries with higher proportions of tertiary educated or that are more economically developed are more likely to have people employed and pursue non-

employment activities (e.g. international business meetings; skiing) that require travelling, which increases person-to-person contact. Lastly, we test the effect of the following demographic variables: vii) the share of the population who live in (semi-) rural areas (expected to be negatively associated with COVID-19); viii) the share of the population that is 65 years or older and ix) population density and x) per capita number of beds in nursing and residential care facilities (expected to be positively associated with COVID-19. More detail on the data sources (*13-19*) and the descriptive statistics of the covariates can be found in Table 1.

We first analyze the bivariate associations between each variable and the ln of the cumulative number of confirmed COVID-10 cases. We do this for six different dates (March 1, 11, 21, 31, April 10 and 20) to check whether the type of predictor responsible for the diffusion changed over time (Table 2). March 1 was chosen as the first date as most of the 20 European countries had reported their first cases of COVID-19 (Table 1). April 20 is the last date analysed as by this time all studied countries had restrictions of movement and social interaction in place for over a month. This implies that country-differences in cumulative or new cases at later dates will be more and more influenced by the different severities and timing of these restrictions rather than due to differences in social and demographic factors. The supplementary material section provides robustness checks by repeating the analyses by using as the dependent variable the number of COVID-19 cases per 100.000 population and the ln of the change in cases between each period.

Ordinary linear regression analysis was used to assess the unique association between confirmed cases of COVID-19 and the covariates. However, as covariate data could only be obtained for 23 countries, i.e. too few to test all covariates simultaneously without overfitting (*21-22*), we first opted to reduce the number of variables by

performing a factor analysis. This yielded three sociodemographic-like factors that explained 78% of the country-variation in the selected covariates. These have been labelled "socially and economically vibrant", "high-educated and not aged", and "densely populated and traditional" (see also Tables S1, S2, S3 and S4 and the explanation under Table S1 in the supplementary material). Correlation coefficients between COVID-19 and the obtained factors were subsequently calculated. As the rotated factors are orthogonal and thus not correlated, each covariate's standardized coefficient is the same as the correlation coefficient in a multivariate regression analysis. However, multivariate regression analyses were still performed to obtain the adjusted $R^2$ of the models as well as the unstandardized coefficients as they represent the slope of the factors (Table S5).

**Results**

Figure 1 and Table 2 present the association between the number of confirmed cases of COVID-19 and the different covariates for the six dates. The highest (and significant) associations are observed for the social meeting and population density variables. Particularly regarding the latter, this association gets stronger over time. As the relatively small sample of countries does not permit us to use all covariates in a multivariate analysis, we used the three factors that were extracted though factor analysis. Figure 2 and Table 3 present the association of each factor with the number of confirmed cases of COVID-19 and are consistent with the results of the specific variables. The socially and economically vibrant factor has a strong and positive association throughout the study period and the highly-educated and not aged factor has a statistically insignificant negative and declining association with the number of confirmed cases of COVID-19. The association of densely populated and traditional factor was initially low but increased with time, becoming significant on March 11 and the most important factor by the end of

6

March. The three factors together explain 51% of the cross-country variation in the number of confirmed cases of COVID-19 on both March 31 and April 20, compared to just 21% on March 1 (Table 4). The slope of the socially & economically vibrant factor was greatest on March 11 and that of the densely populated and traditional factor on April 20.

Our sensitivity analysis shows virtually identical statistical associations between the change in COVID-19 cases in a space of 10 days and the three sociodemographic factors (see Supplementary Table S5). The socially and economically vibrant factor is strongly significant across all time periods, except for the last period between April 10 and 21, while the densely populated and traditional factor was significant throughout the entire study period and became the most important explanatory factor at the end of March. The high-educated and not aged factor did not prove to be significant in any of the models. The proportion of the country differences in change in COVID-19 explained by the three factors was about the same during all but the last period (44-48%, dropping down to 38%). The factors explain much less of the country differences in the number of cases of COVID-19 per 100,000 population (see Supplementary table S6): the socially and economically vibrant factor from March 31 and the densely populated and traditional factor from April 10 but, again, becoming the most important explanatory factor. The explanatory power did increase steadily over time from 3% (March 1) to 40% (April 20). This is consistent with the fact that towards the end of April, the countries with high number of cases per 100,000 population included not only Italy, but also the densely populated Belgium and the Netherlands, while COVID-19 rates were (still) quite low in the sparsely populated Scandinavian and Baltic countries.

**Discussion**

Confirmed cases of COVID-19 increased sharply across Europe during the course of March and April, 2020. Throughout most of the studied period, Italy was worst hit by the pandemic in absolute number of cases, but Spain surpassed Italy early April, while others also did so in terms of cases per 100,000 people (see Table 1). The question we posed is whether social, economic and demographic factors could explain the observed differences in Europe.

Our results suggest that it is not so much how aged countries are but their (historical) level of economic development and (associated) social contact that determined the initial spread of the COVID-19 pandemic and continued to be important for its subsequent spread during late March and early April. However, population density and religious practice were particularly important demographic and cultural country-specific factors associated with frequency of human-to-human contact necessary for its subsequent spread in the most affected countries. Considering some specific examples, the Netherlands, Switzerland and Sweden all scored high on the socially and economically vibrant factor and saw their number of coronavirus infections quickly increase during March despite households being almost exclusively single-person or nuclear. An important component of this factor, however, is also the number of available beds in nursing and residential care facilities, in which all three countries score high. One recent study showed that nursery homes may be responsible for anything between 19% and 72% of COVID-19 deaths in, respectively, Australia and Canada, with the seven European countries included in the study observing proportions within this range; *23*, see also *24*). On the other hand, Italy having been the initial epicenter of the epidemic in Europe scored very low on "high-educated and not aged" factor, but high on "densely populated and traditional" factor, i.e. its high population density and traditional values

(approximated through the proportion attending church at least once a week) is likely to have contributed to their high levels. Interestingly, the relative position of other traditionally catholic countries in COVID-19, including Spain, Portugal and Belgium worsened markedly between March 21 and April 10 (Figure 2). Our sensitivity analysis shows almost identical results if change in COVID-19 cases are analyzed over 10-day periods, while the number of cases of COVID-19 per 100,000 population in the 23 European countries could only be weakly explained by these three factors from March 31 onwards. Again, on the last analysed date, 20 April, the population density factor proved to be the most discriminant one, although having a socially and economically vibrant population remained to contribute significantly to the explanation of the European country differences as the number of cases had increased markedly in the most socioeconomically developed European countries.

Some limitations of our study should be mentioned. First, we did not consider country differences in (the timing of) government (and individual) responses to the COVID-19 pandemic. In particular, governments have differed in the timing of the implementation of measures such as public event cancellation, the closing of daycare centers, schools and universities, social distancing, or partial or total lockdowns (see *25* for details). This implies that the effect of social and demographic factors on COVID-19 cases may be confounded by these measures in those countries that were quickest at adopting them and had already past its peak of daily additional cases of COVID-19 (e.g. Italy). Given the estimated average latency period between the moment that someone becomes infected by the coronavirus and reported COVID-19, we think that only the last two data points may be affected by this.

Other factors are also likely to be responsible for the spread of the coronavirus in Europe that we did not consider. We know that a French and Austrian ski resort were

responsible for initial infections in the UK and other Northern European countries (*26-27*). However, the influence that winter holidays or international travel has on the country difference in COVID-19 during the studied period is likely to be captured by the tertiary education and GDP variables, two well-known proxies for average disposable income. Smoking is another variable associated with the proliferation of COVID-19. For instance, there is evidence that it is considered a social function (*28*), smokers are more likely to touch their face and mouth (*29*) and have chronic health conditions (*29-30*). However, we think that smoking is a more important factor to consider in individual or small-area studies and when studying COVID-19 mortality, as analysis showed smoking rates to be highest in Eastern European countries, which are least "socially and economically" vibrant and observed the lowest levels of COVID-19.

Another issue of concern is country differences in the approach to testing for COVID-19. Some countries only test people admitted to hospitals or ramped up the testing program much later during the outbreak than other countries. This implies that particularly some of the earlier data points will be an underestimate of the real prevalence of COVID-19 (*31-32*).

Finally, data from the ESS is not available for all European countries. This implies that different results may be obtained if data for more countries becomes available. Despite its availability, we did not use data from Round 9 (2018) as it contains fewer countries that the previous round. However, this is unlikely to have biased the results as cross-country demographic, social and cultural differences are unlikely to change in a few years.

To conclude, our policy recommendation based on our results for similar future outbreaks of coronavirus-like epidemics is that very strict measures of social distancing would be required to be implemented quickly in countries with strong social ties in order

to minimize the proliferation of cases during the secondary transmission phase that occurs within households. These social ties are, however, not only related to population density, household characteristics and religious and cultural practices, but also to economic development. This is probably why it is particularly the north of Italy, Switzerland and some of the main urban centers of Spain (particularly Madrid) that have been hit the hardest by COVID-19. A recommendation for future research would therefore be to perform analysis at the sub-national level for the countries that have been studied here.

Table 1. Descriptive statistics of the number of reported confirmed COVID-19 cases and the covariates used in the analysis.

| Country | Area | Code | Date of the first case confirmed | Cumulative cases of COVID-19 on | | | | | | Population | Confirmed cases on 31/03/ 100,000 people | Confirmed cases on 20/04/ 100,000 people |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | 01/03 | 11/03 | 21/03 | 31/03 | 10/04 | 20/04 | | | |
| Austria | West | AUT | 25/02 | 14 | 246 | 2814 | 10180 | 13555 | 14795 | 8858775 | 114.91 | 167.01 |
| Belgium | West | BEL | 04/02 | 2 | 314 | 2815 | 12775 | 26667 | 39983 | 11455519 | 111.52 | 349.03 |
| Bulgaria | East | BGR | 08/03 | 0 | 7 | 163 | 399 | 635 | 929 | 7000039 | 5.70 | 13.27 |
| Czechia | East | CZE | 01/03 | 3 | 91 | 995 | 3308 | 5732 | 6900 | 10649800 | 31.06 | 64.79 |
| Denmark | North | DNK | 27/02 | 4 | 444 | 1420 | 3039 | 6014 | 7711 | 5806081 | 52.34 | 132.81 |
| Estonia | East | EST | 27/02 | 1 | 16 | 306 | 745 | 1258 | 1535 | 1324820 | 56.23 | 115.86 |
| Finland | North | FIN | 29/01 | 6 | 59 | 523 | 1418 | 2769 | 3868 | 5517919 | 25.70 | 70.10 |
| France | West | FRA | 24/01 | 130 | 2293 | 14463 | 52827 | 91738 | 155393 | 67012883 | 78.83 | 231.89 |
| Germany | West | DEU | 27/01 | 130 | 1908 | 22213 | 71808 | 122171 | 147065 | 83019213 | 86.50 | 177.15 |
| Hungary | East | HUN | 04/03 | 0 | 13 | 103 | 492 | 1190 | 1984 | 9772756 | 5.03 | 20.30 |
| Ireland | North | IRL | 29/02 | 1 | 43 | 785 | 3235 | 8089 | 15652 | 4904240 | 65.96 | 319.15 |
| Italy | South | ITA | 31/01 | 1694 | 12462 | 53578 | 105792 | 147577 | 181228 | 60359546 | 175.27 | 300.25 |
| Lithuania | East | LTU | 28/02 | 1 | 3 | 83 | 537 | 999 | 1326 | 2794184 | 19.22 | 47.46 |
| Netherlands | West | NLD | 27/02 | 10 | 503 | 3640 | 12667 | 23249 | 33588 | 17282163 | 73.30 | 194.35 |
| Norway | North | NOR | 26/02 | 19 | 598 | 2118 | 4641 | 6314 | 7156 | 5328212 | 87.10 | 134.30 |
| Poland | East | POL | 04/03 | 0 | 31 | 536 | 2311 | 5955 | 9593 | 37972812 | 6.09 | 25.26 |
| Portugal | South | PRT | 02/03 | 0 | 59 | 1280 | 7443 | 15472 | 20863 | 10276617 | 72.43 | 203.01 |
| Serbia | East | SRB | 06/03 | 0 | 12 | 171 | 900 | 3105 | 6630 | 6963764 | 12.92 | 95.21 |
| Slovenia | East | SVN | 05/03 | 0 | 57 | 383 | 802 | 1160 | 1335 | 2080908 | 38.54 | 64.15 |
| Spain | South | ESP | 01/02 | 84 | 2277 | 25374 | 95923 | 158273 | 200210 | 46937060 | 204.37 | 426.55 |
| Sweden | North | SWE | 31/01 | 14 | 500 | 1763 | 4435 | 9685 | 14777 | 10230185 | 43.35 | 144.45 |
| Switzerland | West | CHE | 25/02 | 27 | 652 | 6575 | 16605 | 24551 | 27944 | 8544527 | 194.33 | 327.04 |
| The UK | North | GBR | 31/01 | 36 | 459 | 5067 | 25481 | 74605 | 125856 | 66647112 | 38.23 | 188.84 |
| EU 23 total | | | 24/01 | 2176 | 23047 | 147168 | 437763 | 750763 | 1026321 | 490739135 | 89.20 | 209.14 |

Table 1(Cont).

| Country | Mean number of household members | % living in a multi-generational household | % having a frequent social meeting | % ≥ weekly religious attendance | % tertiary education | Popula-tion per km² | % pop (semi) rural | % people aged 65+ | GDP per capita | Nursing/ rest home beds / 100,000 |
|---|---|---|---|---|---|---|---|---|---|---|
| Austria | 2.2 | 5 | 65.5 | 11.4 | 31.1 | 107.21 | 43.90 | 19.00 | 51462 | 862.4 |
| Belgium | 2.3 | 4.8 | 68.6 | 8.6 | 36.0 | 377.21 | 50.10 | 18.79 | 47519 | 1234.1 |
| Bulgaria | 2.4 | 24.7 | 51.7 | 9.3 | 24.7 | 64.70 | 32.20 | 21.02 | 9273 | 30.8 |
| Czechia | 2.4 | 7.6 | 50.8 | 5.8 | 21.6 | 137.60 | 29.30 | 19.42 | 23079 | 687.5 |
| Denmark | 2.0 | 1.2 | 71.2 | 4.3 | 33.1 | 138.07 | 26.00 | 19.81 | 61350 | 816.1 |
| Estonia | 2.2 | 8.9 | 43.5 | 3.7 | 36.5 | 30.39 | 28.00 | 19.63 | 23266 | 870.7 |
| Finland | 2.0 | 1.5 | 65.1 | 4.6 | 38.5 | 18.16 | 33.20 | 21.72 | 50152 | 1190.0 |
| France | 2.2 | 2.0 | 67.9 | 7.6 | 33.7 | 122.34 | 34.20 | 20.03 | 41464 | 981.5 |
| Germany | 2.0 | 5.3 | 59.7 | 7.1 | 25.9 | 237.37 | 32.80 | 21.46 | 47603 | 1152.2 |
| Hungary | 2.3 | 11.2 | 20.4 | 10.6 | 22.5 | 107.91 | 36.70 | 19.16 | 16162 | 853.3 |
| Ireland | 2.6 | 8.1 | 58.9 | 33.9 | 40.7 | 70.45 | 41.10 | 13.87 | 78806 | 639.3 |
| Italy | 2.3 | 14.2 | 60.1 | 27.4 | 17.4 | 205.45 | 46.80 | 22.75 | 34483 | 415.8 |
| Lithuania | 2.2 | 9.4 | 31.1 | 17.1 | 37.9 | 44.53 | 23.00 | 19.71 | 19153 | 726.3 |
| Netherlands | 2.2 | 2.7 | 74.5 | 9.8 | 34.8 | 511.46 | 45.40 | 19.20 | 53024 | 1379.6 |
| Norway | 2.0 | 2.3 | 78.5 | 3.9 | 37.7 | 14.55 | 35.30 | 17.05 | 81697 | 765.6 |
| Poland | 2.8 | 16.1 | 30.9 | 46.5 | 28.2 | 124.04 | 44.30 | 17.52 | 15421 | 195.3 |
| Portugal | 2.5 | 13.1 | 77.6 | 27.8 | 23.8 | 112.24 | 30.50 | 21.95 | 23408 | 555.7 |
| Serbia | 2.9 | 21.3 | 62.1 | 11.1 | 20.6 | 79.83 | 35.60 | 18.35 | 7247 | 291.5 |
| Slovenia | 2.5 | 16.1 | 53.1 | 12.6 | 29.3 | 102.64 | 53.00 | 19.61 | 26124 | 1012.4 |
| Spain | 2.5 | 17.9 | 70.2 | 15.2 | 35.1 | 93.53 | 47.30 | 19.38 | 30371 | 834.8 |
| Sweden | 2.0 | 1.2 | 74.3 | 5.7 | 37.8 | 25.00 | 28.00 | 20.10 | 54608 | 1388.0 |
| Switzerland | 2.2 | 2.3 | 71.6 | 8.3 | 38.6 | 215.52 | 55.00 | 18.62 | 82797 | 1174.2 |
| The UK | 2.3 | 3.9 | 62.4 | 11.7 | 40.6 | 274.83 | 25.20 | 18.40 | 42944 | 821.4 |

Source: Confirmed COVID-19 cases: the Center for Systems Science and Engineering at Johns Hopkins University: https://coronavirus.jhu.edu/map.html. Population on 1 January 2019 (13). Mean number of household members 2018 (Eurostat; 14), % living in a multi-generational household (people who live with i) parents, ii) parents and children, iii) parents and relatives, iv) parents, children, and relatives, v) children and relatives, or vi) relatives), % having a frequent social meeting (people who socially meet with their friends, relatives, or colleagues at least once a week), % ≥ weekly religious attendance, and % pop (semi) rural are created from European Social Survey (ESS) Rounds 7 (2014) (Denmark), 8 (2016) (Lithuania, Portugal, Spain and Sweden) and 9 (2018) (all other countries). % tertiary education of 25-64 year-olds in 2018; Eurostat (15), except BGR and SRB (ESS)). Population per km² in 2018, % people aged 65+ in 2018 and GDP in 2018; World Bank (16-18). Available beds in nursing and residential care facilities per 100.000 inhabitants in 2017, except for Belgium (2012) and Denmark (2011) and Portugal (applied 2017 ratio of number long-term care recipients between Portugal and Spain to the beds available in Spain); Eurostat (19-20).

Table 2. Correlation between the natural log COVID-19 at six different time periods and the covariates.

| Variable | March 1 | March 11 | March 21 | March 31 | April 10 | April 20 |
|---|---|---|---|---|---|---|
| 1. Mean number of household members | -0.40 | -0.38 | -0.28 | -0.20 | -0.13 | -0.08 |
| 2. % living in a multi-generational household (ln) | -0.31 | -0.42* | -0.29 | -0.24 | -0.23 | -0.21 |
| 3. % having a frequent social meeting (squared) | 0.41 | 0.61** | 0.56** | 0.53** | 0.52* | 0.50* |
| 4. % ≥ weekly religious attendance (ln) | -0.07 | -0.10 | 0.03 | 0.12 | 0.17 | 0.20 |
| 5. % tertiary education | -0.01 | 0.06 | 0.05 | 0.05 | 0.05 | 0.05 |
| 6. Population density (people per km$^2$) (ln) | 0.25 | 0.37 | 0.44* | 0.50* | 0.53** | 0.54** |
| 7. % population in (semi-)rural areas | 0.18 | 0.32 | 0.33 | 0.30 | 0.27 | 0.25 |
| 8. % people aged 65+ (squared) | 0.35 | 0.22 | 0.23 | 0.18 | 0.14 | 0.10 |
| 9. GDP per capita | 0.35 | 0.51* | 0.44* | 0.38 | 0.35 | 0.33 |
| 10. Nursing and rest home beds per capita | 0.25 | 0.39 | 0.31 | 0.29 | 0.27 | 0.24 |

**p<0.01, *p<0.05

Table 3. Correlation between the natural log of COVID-19 at six different time periods and the extracted factors.

| | Date of cumulative cases of COVID-19 | | | | | |
|---|---|---|---|---|---|---|
| | March 1 | March 11 | March 21 | March 31 | April 10 | April 20 |
| Factor 1: Socially and economically vibrant | 0.44* | 0.61** | 0.52** | 0.47** | 0.44* | 0.41* |
| Factor 2: High-educated and not aged | -0.26 | -0.16 | -0.14 | -0.11 | -0.07 | -0.04 |
| Factor 3: Densely populated and traditional | 0.25 | 0.38* | 0.48** | 0.55** | 0.59** | 0.60** |

**p<0.01, *p<0.05

Table 4. Multivariate regression analysis of social and demographic factors on cumulative cases of ln of COVID-19 on March 1, 11, 21 and 31, and April 10 and 20, 2020, for 23 European countries. Unstandardized coefficients.

| | Unstandardized coefficients (p value) | | | | | |
|---|---|---|---|---|---|---|
| | March 1 | March 11 | March 21 | March 31 | April 10 | April 20 |
| Factor 1: Socially & economically vibrant | **0.91 (0.033)** | **1.30 (0.001)** | **0.93 (0.004)** | **0.81 (0.007)** | **0.75 (0.010)** | **0.70 (0.017)** |
| Factor 2: High educated & not aged | -0.53 (0.194) | -0.35 (0.304) | -0.26 (0.377) | -0.18 (0.506) | -0.12 (0.649) | -0.07 (0.795) |
| Factor 3: Densely populated & traditional | 0.53 (0.197) | **0.82 (0.022)** | **0.86 (0.007)** | **0.96 (0.002)** | **0.99 (0.001)** | **1.01 (0.001)** |
| $R^2$ adjusted | 0.21 | 0.48 | 0.44 | 0.46 | 0.47 | 0.46 |

**Fig. 1. Association between the natural log of cumulative cases of COVID-19 on March 1, 11, 21 and 31 2020 and different covariates 20 European countries.**
Sources: the Center for Systems Science and Engineering at Johns Hopkins University, European Social Survey Wave 8, World Bank Database, OECD, and WHO Database.
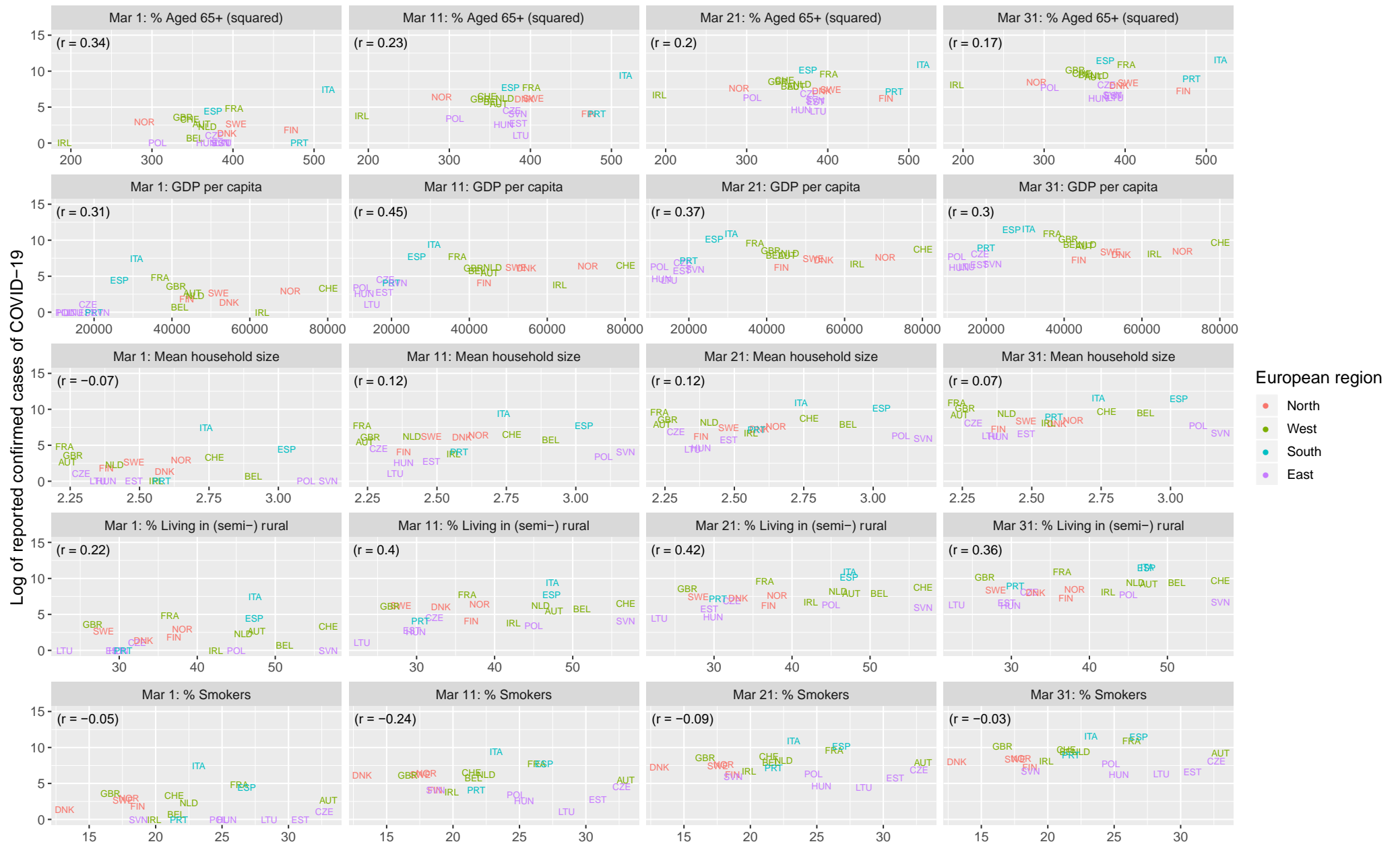
Fig. 1. (Cont.)

**Fig. 2. Factor scores vs. natural log of COVID-19. March 1 and 31, 2020, for 20 European countries.** Sources: the Center for Systems Science and Engineering at Johns Hopkins University, European Social Survey Wave 8, World Bank Database, OECD, and WHO Database. Note: * p<0.05; ** p<0.01.

**References**

1.  Johns Hopkins University Center for Systems Science and Engineering (2020). Novel Coronavirus (COVID-19) Cases. https://github.com/CSSEGISandData/COVID-19. Accessed: April 30, 2020.

2. Huang, C., Wang, Y., Li, W., & et al. (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*, 395, 497-506.

3. Peeri, N.C., Shestha, N., Rahman, M.S., & et al. (2020) The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? *International Journal of Epidemiology*. In press.

4.  WHO (2020) Coronavirus disease (COVID-19) situation reports. www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.

5. Grais, R.F., Ellis, J.H., & Glass, G.E. (2003). Assessing the impact of airline travel on the geographic spread of pandemic influenza. *European journal of epidemiology*, 18, 1065-1072.

6. Bayer, C. & Kuhn, M. (2020). Intergenerational ties and case fatality rates: A cross-country analysis. Available from: https://voxeu.org/article/intergenerational-ties-and-case-fatality-rates.

7. Mossong J., Hens N., Jit M., Beutels P., Auranen K., Mikolajczyk R., . . . Wallinga J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine*, *5*, e74.

8. Liu, Y., Eggo, R.M. & Kucharski, A.J. (2020). Secondary attack rate and superspreading events for SARS-CoV-2. *Lancet*, *395*, e47.

9. Bi, Q., Wu, Y., Mei, S., & et al. (2020), Epidemiology and transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts. *medRxiv*.

10. Wallinga, J., Teunis, P., & Kretzschmar. M. (2006). Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *American Journal of Epidemiology*, 164, 936-944. Doi:10.1093/aje/kwj317.

11. Reher, D.S. (1998). Family ties in Western Europe: Persistent contrasts. *Population and Development Review*, 24, 203-234.

12. Sánchez Rodríguez, M.M., Gierveld, J.D.J., & Buz, J. (2014). Loneliness and the exchange of social support among older adults in Spain and the Netherlands. *Ageing & Society*, 34, 330-354.

13.  Eurostat (2020). Population on 1 January by age and sex. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_pjan&lang=en. Accessed: April 30, 2020.

14. Eurostat (2020). Average household size - EU-SILC survey. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_lvph01&lang=en. Accessed: April 30, 2020.

15. Eurostat (2020). Population by educational attainment level, sex and age (%). http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=edat_lfse_03&lang=eng. Accessed: April 30, 2020.

16. World Bank (2020). GDP per capita (current US$) https://data.worldbank.org/indicator/NY.GDP.PCAP.CD? Accessed April 30, 2020.

17. World Bank (2020). Population ages 65 and above (% of total population), 2019 https://data.worldbank.org/indicator/SP.POP.65UP.TO.ZS? Accessed April 30, 2020.

18. World Bank (2020). Population density (people per sq. km of land area), 2019 https://data.worldbank.org/indicator/EN.POP.DNST Accessed April 30, 2020.

19. Eurostat (2019). Available beds in nursing and residential care facilities. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=hlth_rs_bdsns&lang=en. Accessed April 30, 2020.

20. Eurostat (2019). Long-Term Care Resources and Utilisation: Long-term care recipients. https://stats.oecd.org/Index.aspx?ThemeTreeId=9. Accessed April 30, 2020.

21. Harrell Jr, F.E., Lee, K.L., Califf, R.M., Pryor, D.B., & Rosati, R.A. (1984). Regression modelling strategies for improved prognostic prediction. *Statistics in Medicine*, 3, 143-152.

22. Peduzzi, P., Concato, J., Kemper, E., Holford, T.R., & Feinstein, A.R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, 49, 1373-1379.

23. Comas-Herrera A, Zalakain J, Litwin C, Hsu AT, Lane N, Fernandez-Plotka J-L (2020) Mortality associated with COVID-19 outbreaks in care homes: early international evidence. International Long-Term Care Policy Network, CPEC-LSE. https://LTCcovid.org

24. Orange, R. (2020, Sunday 19). Anger in Sweden as elderly pay price for coronavirus strategy. The Guardian. 19 April. https://www.theguardian.com/world/2020/apr/19/anger-in-sweden-as-elderly-pay-price-for-coronavirus-strategy

25. Flaxman, S., Mishra, S., Gandy, A., & et al. (2020). Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. London: Imperial College London.

26. Hodcroft, E.B. (2020). Preliminary case report on the SARS-CoV-2 cluster in the UK, France, and Spain. *Swiss Medical Weekly*, 150. doi:10.4414/smw.2020.20212.

27. Hruby, D. (2020). How an Austrian ski resort helped coronavirus spread across Europe. https://edition.cnn.com/2020/03/24/europe/austria-ski-resort-ischgl-coronavirus-intl/index.html.

28. Paul, C.L., Ross, S., Bryant, J., Hill, W., Bonevski, B., & Keevy, N. (2010). The social context of smoking: A qualitative study comparing smokers of high versus low socioeconomic position. *BMC public health*, 10, 211.

29. Science Media Centre (2020). Expert reaction to questions about smoking and COVID-19. https://www.sciencemediacentre.org/expert-reaction-to-questions-about-smoking-and-covid-19/

30. GBD 2015 Tobacco Collaborators (2017). Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: a systematic analysis from the Global Burden of Disease Study 2015. *Lancet*, 389, 1885-1906. Doi: 10.1016/S0140-6736(17)30819-X.

31. Farge, E. & Revill, J. (2020). 'Test, test, test': WHO chief's coronavirus message to world. Reuters. https://www.reuters.com/article/us-healthcare-coronavirus-who/test-test-test-who-chiefs-coronavirus-message-to-world-idUSKBN2132S4.

32. Wikipedia (2020). COVID-19 testing. https://en.wikipedia.org/wiki/COVID-19_testing.

## Supplementary material

Table S1. Correlation matrix between the covariates

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| 1. Mean number of household members | | | | | | | | | |
| 2. % living in a multi-generational household (ln) | 0.77** | | | | | | | | |
| 3. % having a frequent social meeting (squared) | -0.32 | -0.54** | | | | | | | |
| 4. % ≥ weekly religious attendance (ln) | 0.69** | 0.60** | -0.27 | | | | | | |
| 5. % tertiary education | -0.40 | -0.59** | 0.29 | -0.26 | | | | | |
| 6. Population density (pop per km$^2$) (ln) | 0.19 | 0.10 | 0.05 | 0.32 | -0.25 | | | | |
| 7. % population in (semi-)rural areas | 0.29 | 0.19 | 0.14 | 0.34 | -0.05 | 0.40 | | | |
| 8. % people aged 65+ (squared) | -0.34 | 0.02 | 0.08 | -0.14 | -0.45* | 0.05 | -0.18 | | |
| 9. GDP per capita | -0.53** | -0.73** | 0.62** | -0.27 | 0.62** | -0.02 | 0.22 | -0.35 | |
| 10. Beds in nursing/residential care homes per 100,000 | -0.64** | -0.70** | 0.40 | -0.49* | 0.54** | 0.11 | 0.14 | 0.04 | 0.54** |

\*\*$p<0.01$, \*$p<0.05$ The Shapiro-Wilk test for normality was used to check the cross-country distribution of the variables. As a result, variables 2, 3, 4, 6 and 8 were transformed prior to analysis. Two variables were not included in the final Factors Analysis: the mean number of household members because of its high correlation with multigenerational household but a slightly lower communality; and % population in (semi-) rural areas because the factor loadings were almost identical to the population density variable but its association with COVID-19 was the opposite than expected. The remaining variables were used to obtain the final set of factors. The KMO test score after excluding variables 1 and 7 equals 0.65.

Table S2. Communalities, factors, eigenvalues and percentage of explained variance in unrotated solution

| Variable | Communality | Factor | Eigenvalue | % of variable | Cum % |
|---|---|---|---|---|---|
| 2. % living in a multi-generational household | 0.858 | 1 | **3.59** | 44.82 | 44.82 |
| 3. % having a frequent social meeting (squared) | 0.582 | 2 | **1.42** | 17.78 | 62.61 |
| 4. % ≥ weekly religious attendance (ln) | 0.766 | 3 | **1.23** | 15.40 | 78.01 |
| 5. % tertiary education | 0.794 | 4 | 0.72 | 9.03 | 87.03 |
| 6. Population density (pop per km$^2$) (ln) | 0.854 | 5 | 0.44 | 5.53 | 92.56 |
| 8. % people aged 65+ (squared) | 0.860 | 6 | 0.27 | 3.39 | 95.95 |
| 9. GDP per capita | 0.829 | 7 | 0.20 | 2.50 | 98.46 |
| 10. Beds in nursing/residential care homes per 100,000 | 0.697 | 8 | 0.12 | 1.54 | 100.00 |

Table S3. Rotated factor matrix

| Variable | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| 2. % living in a multigenerational household | **-0.888** | -0.105 | 0.241 |
| 3. % having a frequent social meeting (squared) | **0.749** | -0.020 | 0.146 |
| 4. % ≥ weekly religious attendance (ln) | *-0.565* | 0.249 | **0.620** |
| 5. % tertiary education | *0.540* | **0.652** | -0.279 |
| 6. Population density (pop per km$^2$) (ln) | 0.120 | -0.142 | **0.905** |
| 8. % people aged 65+ (squared) | 0.072 | **-0.924** | -0.031 |
| 9. GDP per capita | **0.772** | *0.482* | 0.024 |
| 10. Beds in nursing/residential care homes per 100,000 | **0.833** | 0.005 | -0.050 |

Note: The principal component matrix was rotated using the Varimax method. Given the factor loadings (highest values in bold), Factor 1 could be labelled as "socially and economically vibrant", Factor 2 as "high-educated and not aged", and Factor 3 as "densely populated and traditional". Factor 1 is also typified by a high number of nursing care facilities per population. The countries that score highest/lowest in Factor 1 are the Netherlands, Sweden and Switzerland/Eastern European countries; Factor 2: Ireland/Italy, Germany and Portugal; Factor 3 Netherlands, Italy/Estonia, Nordic countries except Denmark. See also Table S4.

Table S4. Factor scores for each country

| Country | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| Austria | 0.2338 | 0.2490 | 0.2696 |
| Belgium | 0.8971 | 0.1933 | 1.0795 |
| Bulgaria | -1.5110 | -0.9075 | -0.6210 |
| Czechia | -0.4680 | -0.7845 | -0.1445 |
| Denmark | 1.1588 | -0.2939 | -0.1596 |
| Estonia | -0.4807 | -0.1151 | -1.9281 |
| Finland | 0.9651 | -0.6853 | -1.9351 |
| France | 0.7554 | -0.3312 | -0.0198 |
| Germany | 0.6553 | -1.2374 | 0.5822 |
| Hungary | -1.0652 | -0.3820 | -0.2253 |
| Ireland | -0.4701 | 3.1165 | 0.5764 |
| Italy | -0.6188 | -1.6056 | 1.3142 |
| Lithuania | -1.0465 | 0.4921 | -0.8800 |
| Netherlands | 1.4106 | -0.0443 | 1.5242 |
| Norway | 0.7975 | 1.2269 | -1.7080 |
| Poland | -1.9381 | 1.1215 | 0.6966 |
| Portugal | -0.2882 | -1.1869 | 1.0072 |
| Serbia | -1.3220 | -0.2815 | -0.0398 |
| Slovenia | -0.4741 | -0.1887 | 0.1143 |
| Spain | -0.2511 | 0.2381 | 0.3171 |
| Sweden | 1.3230 | -0.1462 | -1.3008 |
| Switzerland | 1.3879 | 0.7528 | 0.7532 |
| The UK | 0.3494 | 0.7999 | 0.7274 |

Note: See Table S3

Table S5a. Correlation between the ln of change in cases of COVID-19 between March 1-11, 11-21, 21-31, March 31-April 10, and April 10-20, 2020 and the extracted factors.

| | March 1-11 | March 11-21 | March 21-31 | Mar 31-April10 | April 10-20 |
|---|---|---|---|---|---|
| Factor 1: Socially & economically vibrant | 0.61** | 0.50** | 0.44* | 0.41* | 0.32 |
| Factor 2: High educated & not aged | -0.16 | -0.14 | -0.08 | -0.04 | 0.00 |
| Factor 3: Densely populated & traditional | 0.39* | 0.50** | 0.58** | 0.60** | 0.60** |

**p<0.01, *p<0.05

Table S5b. Multivariate regression analysis of social and demographic factors on the ln of change in cases of COVID-19 between March 1-11, 11-21, 21-31, March 31-April 10, and April 10-20, 2020, for 23 European countries. Unstandardized coefficients.

| | Unstandardized coefficients (p value) | | | | |
|---|---|---|---|---|---|
| | March 1-11 | March 11-21 | March 21-31 | Mar 31-April10 | April 10-20 |
| Factor 1: Socially & economically vibrant | **1.31 (0.001)** | **0.88 (0.006)** | **0.77 (0.011)** | **0.69 (0.017)** | **0.57 (0.010)** |
| Factor 2: High educated & not aged | -0.35 (0.194) | -0.25 (0.392) | -0.15 (0.599) | -0.07 (0.799) | 0.01 (0.983) |
| Factor 3: Densely populated & traditional | **0.84 (0.020)** | **0.88 (0.006)** | **1.01 (0.001)** | **1.02 (0.001)** | **1.07 (0.002)** |
| $R^2$ adjusted | 0.48 | 0.44 | 0.47 | 0.46 | 0.38 |

Table S6a. Correlation between the natural log of cumulative cases of COVID-19 per 100,000 people at six different time periods and the extracted factors.

| | March 1 | March 11 | March 21 | March 31 | April 10 | April 20 |
|---|---|---|---|---|---|---|
| Factor 1: Socially & economically vibrant | -0.03 | 0.29 | 0.35 | 0.40* | 0.44* | 0.42* |
| Factor 2: High educated & not aged | -0.32 | -0.17 | -0.08 | 0.01 | 0.10 | 0.22 |
| Factor 3: Densely populated & traditional | 0.24 | 0.14 | 0.28 | 0.39 | 0.46* | 0.50** |

**p<0.01, *p<0.05

Table S6b. Multivariate regression analysis of social and demographic factors on cumulative cases of COVID-19 per 100,000 people on March 1, 11, 21 and 31, 2020, and April 10 and 20, for 23 European countries. Unstandardized coefficients.

| | Unstandardized coefficients (p value) | | | | | |
|---|---|---|---|---|---|---|
| | March 1 | March 11 | March 21 | March 31 | April 10 | April 20 |
| Factor 1: Socially & economically vibrant | -0.02 (0.886) | 1.36 (0.197) | 8.04 (0.103) | **23.06 (0.050)** | **38.76 (0.021)** | **48.78 (0.020)** |
| Factor 2: High educated & not aged | -0.18 (0.146) | -0.81 (0.434) | -1.82 (0.702) | 0.57 (0.959) | 8.7 (0.580) | 25.47 (0.199) |
| Factor 3: Densely populated & traditional | 0.14 (0.261) | 0.65 (0.528) | 6.44 (0.186) | 22.84 (0.052) | **40.86 (0.016)** | **58.10 (0.007)** |
| $R^2$ adjusted | 0.03 | -0.01 | 0.08 | 0.21 | 0.33 | 0.40 |

**Fig. S1. Association between cumulative cases of COVID-19 per 100,000 population on March 1, 11, 21 and 31 2020 and different covariates 20 European countries.** Sources: the Center for Systems Science and Engineering at Johns Hopkins University, European Social Survey Wave 8, World Bank Database, OECD, and WHO Database. Note: Population data in 2019 is used.
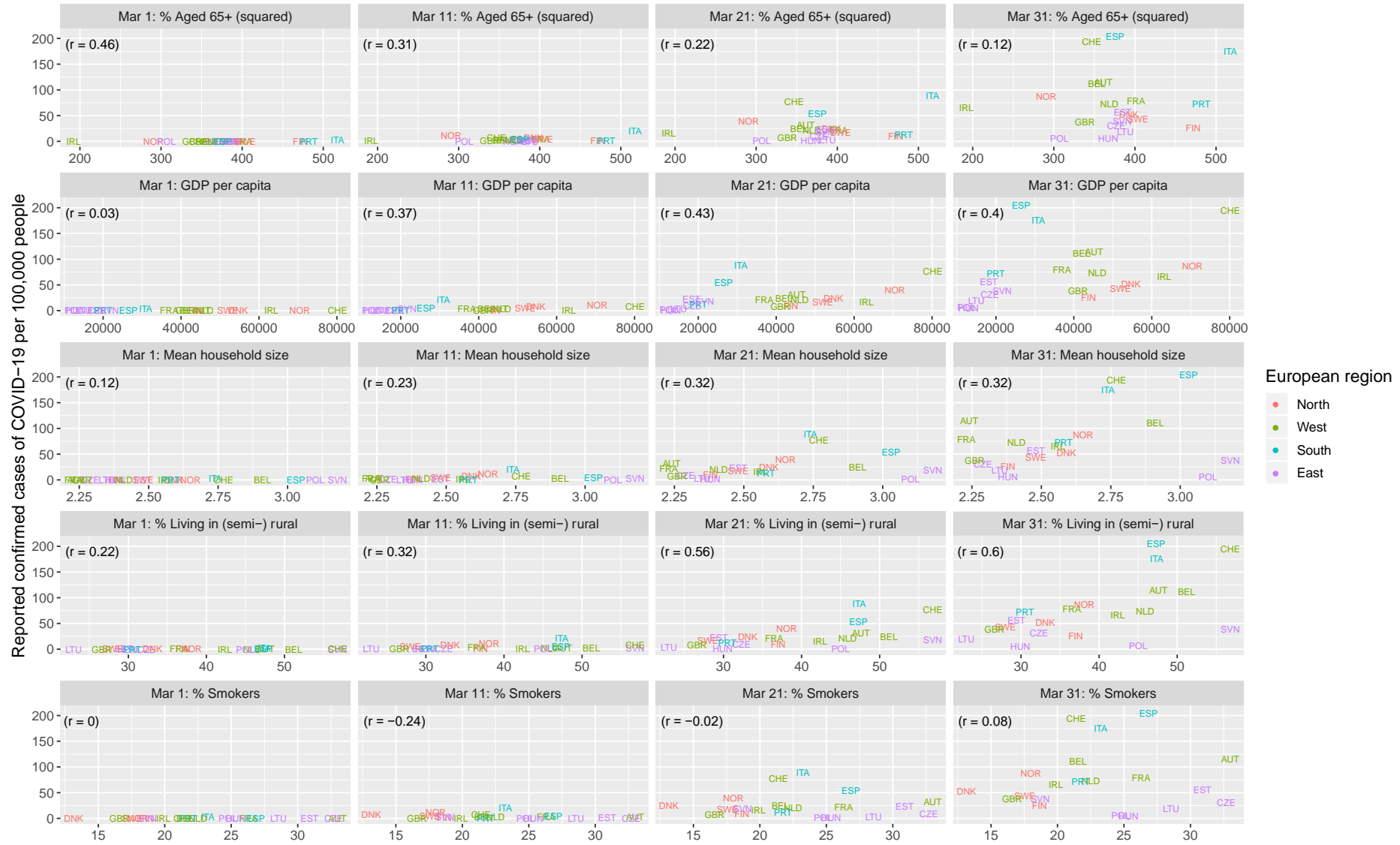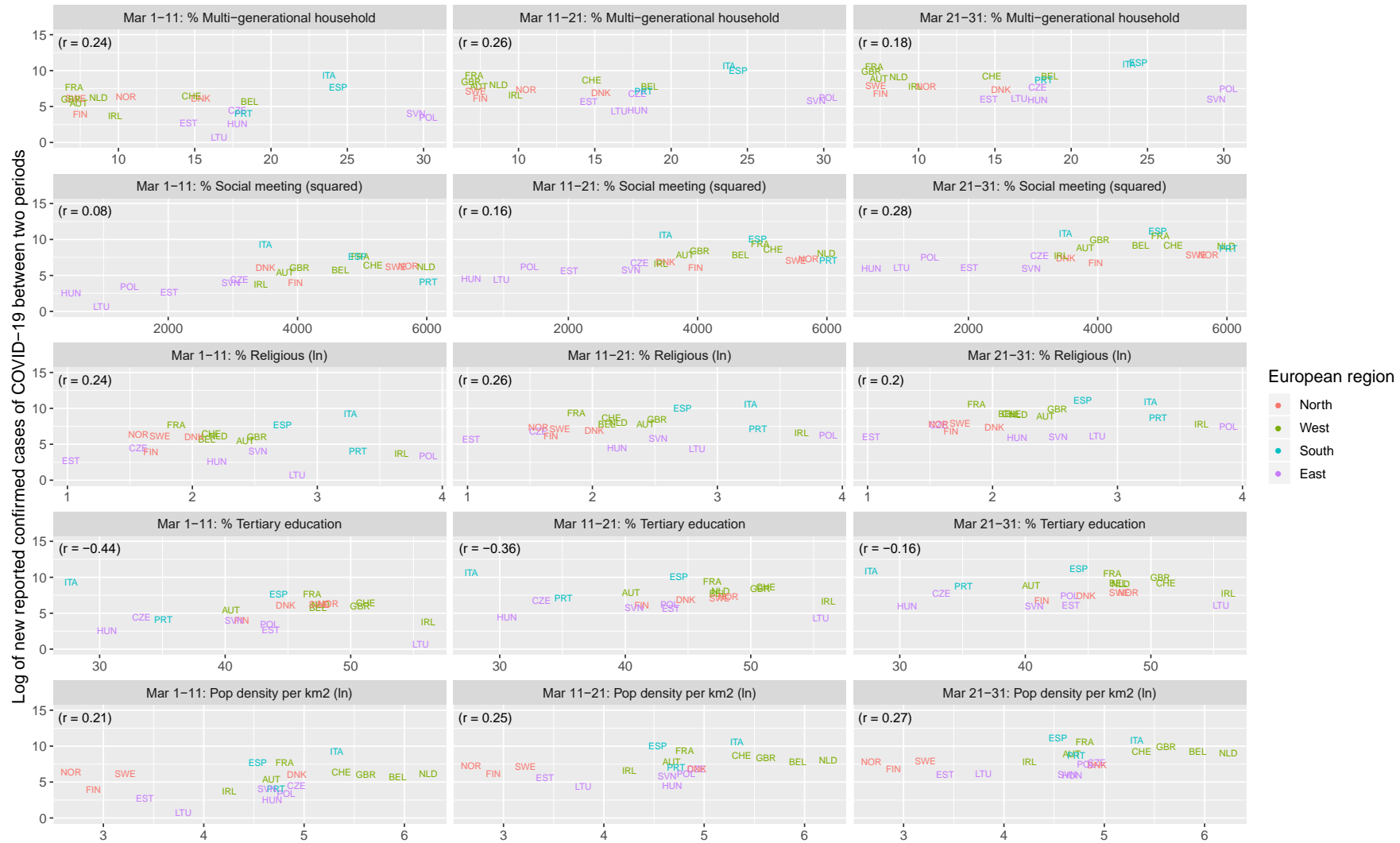
Fig. S1. (Cont.)

**Fig. S2. Association between the ln of new cases of COVID-19 during the periods March 1-11, 11-21 and 21-31 2020 and different covariates for 20 European countries.** Sources: the Center for Systems Science and Engineering at Johns Hopkins University, Eurostat, European Social Survey Wave 8, World Bank Database, OECD, and WHO Database.
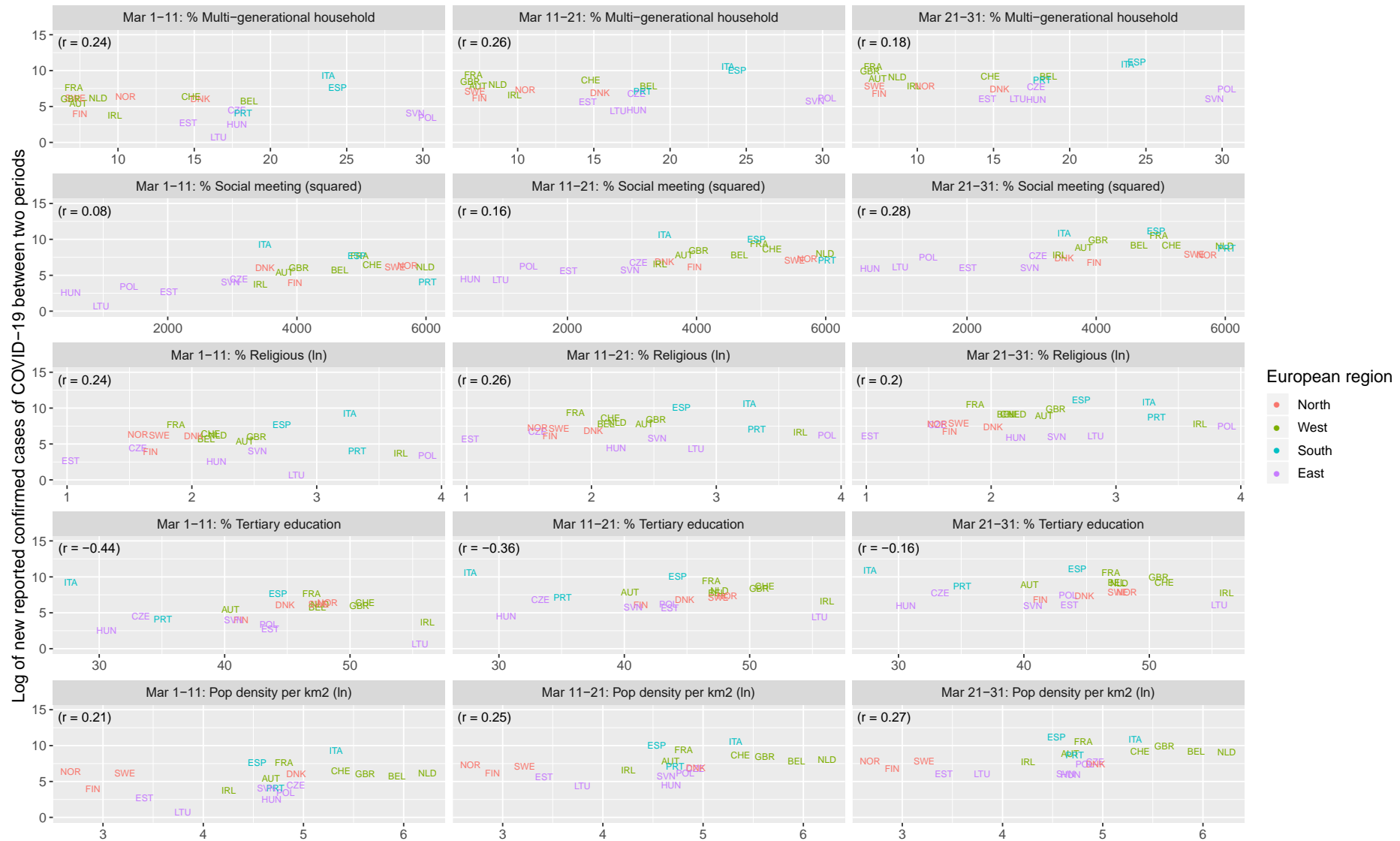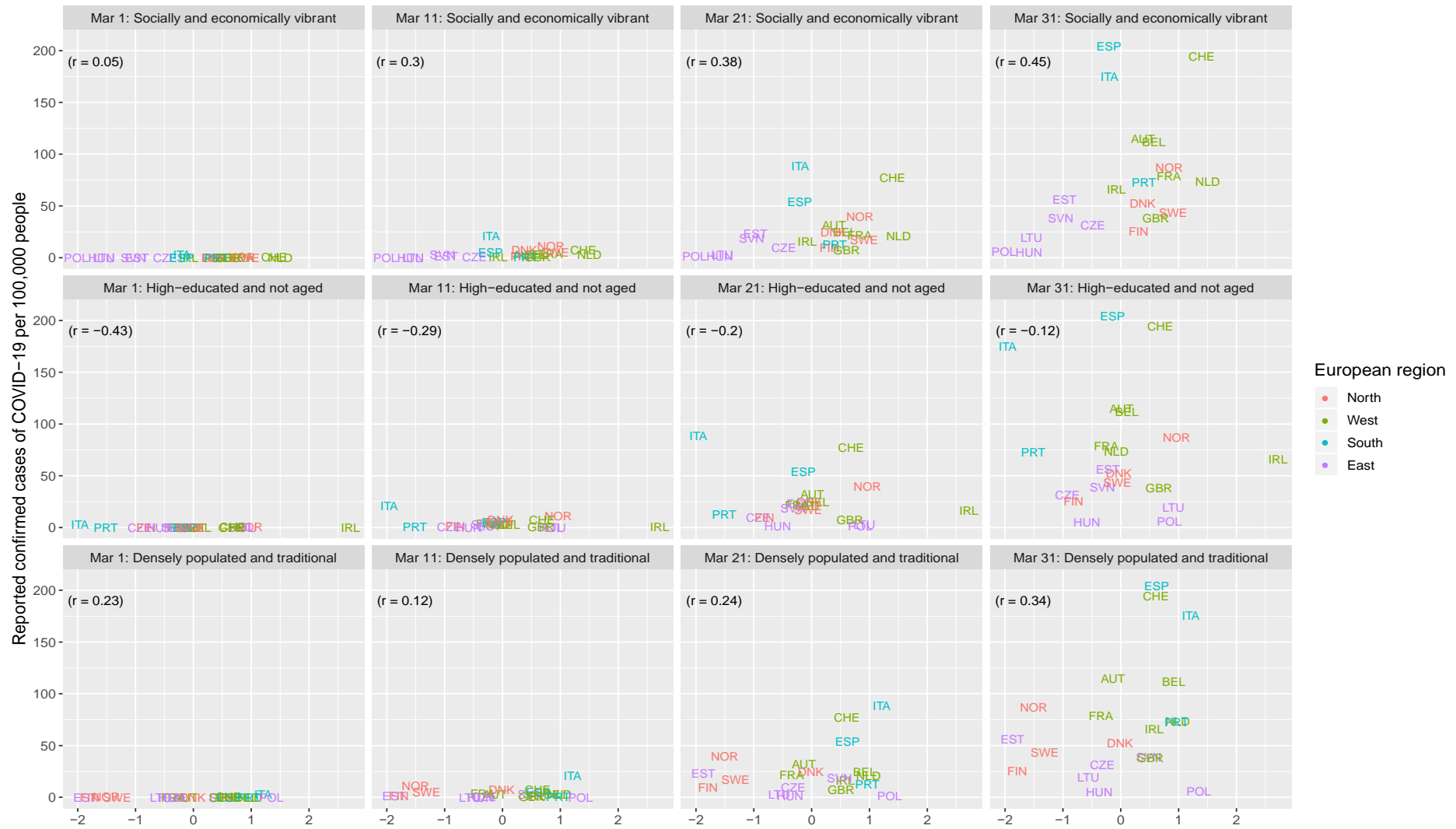
Fig. S2. (Cont.)

**Fig. S3. Association between cumulative cases of COVID-19 per 100,000 population on March 1, 11, 21 and 31 2020 and factor scores for 20 European countries.** Sources: the Center for Systems Science and Engineering at Johns Hopkins University, Eurostat, European Social Survey Wave 8, World Bank Database, OECD, WHO Database, and UNDP. Note: Population data in 2019 is used.
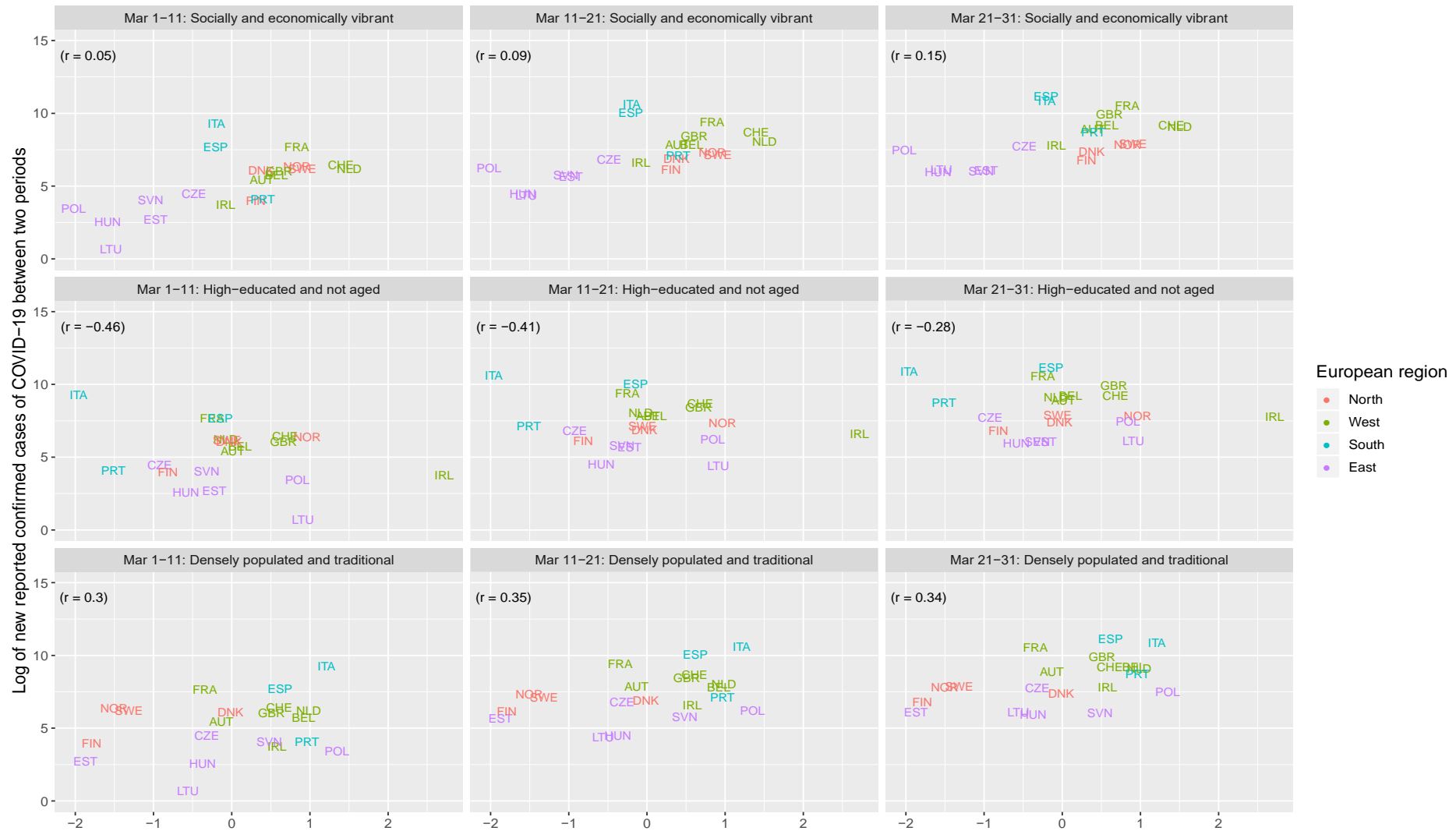
**Fig. S4. Association between the ln of new cases of COVID-19 during the periods March 1-11, 11-21 and 21-31 2020 and factor scores for 20 European countries.** Sources: the Center for Systems Science and Engineering at Johns Hopkins University, European Social Survey Wave 8, World Bank Database, OECD, and WHO Database.