Journal of Population Research (2022) 39:495–511
https://doi.org/10.1007/s12546-021-09257-1

ORIGINAL RESEARCH

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

Ryohei Mogi[1] · **Jeroen Spijker**[1]

Accepted: 4 March 2021 / Published online: 5 April 2021

Abstract

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 propagation. We analyse the association of social, economic, and demographic factors in the initial spread of the coronavirus disease COVID-19 across 23 European countries between March 1 and April 30, 2020. Diagnosed COVID-19 cases from Johns Hopkins University and data from the European Social Survey and other sources were used to estimate bivariate associations between cumulative reported case numbers at ten-day intervals and nine social, demographic, and economic variables. To avoid overfitting, we first reduce these variables to three factors by factor analysis before conducting a multiple regression analysis. We also perform a sensitivity analysis using rates and new cases between two time periods. Results showed that social and economic factors are strongly and positively associated with COVID-19 throughout the studied period, while the association with population density and cultural factors was initially low, but by April, was higher than the earlier mentioned factors. For future influenza-like pandemics, implementing strict movement restrictions from early on will be crucial to curb the spread of such diseases in economically, socially, and culturally vibrant and densely populated countries.

Keywords COVID-19 · Coronavirus · Social ties · Economic development · Nursing homes · Health policy

✉ Jeroen Spijker
jspijker@ced.uab.es

[1] Centre d'Estudis Demogràfics, Carrer de Ca n'Altayó, Edifici E2, Universitat Autònoma de Barcelona, 08193 Bellatarra, Spain

Introduction

As of late April 2020, there were more than three million confirmed cases worldwide of the coronavirus disease COVID-19 (CSSE, 2020) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). While enquiries are still underway regarding the exact origins of the virus, the first case in humans was identified in late 2019 in Wuhan, China. Subsequent person-to-person transmission then mainly took place through respiratory droplets produced through coughs or sneezes of infected persons in close presence of other people (Huang et al. 2020; Peeri et al. 2020).

One well-established cause of the global spread of previous influenza outbreaks is airplane travel (Grais et al. 2003) and without this mode of transportation the coronavirus would not have arrived in Europe so quickly. The first cases were confirmed in Europe in late January-early February and, although international travel restrictions to and from China were quickly imposed, by late March about half of the world's reported cases were in Europe (CSSE, 2020). Once the virus was brought to Europe, human interaction and close contact allowed the virus to spread quickly. Previous research on virus transmission, including on COVID-19, has shown social contact to be very important (Bayer & Kuhn, 2020; Bi et al. 2020; Liu et al. 2020; Mossong et al. 2008; Wallinga et al. 2006). Based on modelled estimates using both empirical (the POLYMOD survey) and synthetic data, Prem, van Zandvoort, et al. (2020) showed how altering (intergenerational) patterns of contact by reducing physical contact or shielding would lead to large reductions in the transmission of the virus. At the same time, cultural differences in Europe are well established. For instance, in Southern European countries intergenerational contact is more frequent than in the less family-oriented Western and Northern European countries, as social norms about providing support to family members and maintaining interpersonal familiar interactions are stronger there (Reher, 1998; Sánchez Rodríguez et al. 2014). In relation to COVID-19 case fatality rates, Arpino et al. (2020) recently showed it to be broadly positively associated with intergenerational co-residence and contacts at the national level across a selection of European countries. However, as conclusive interpretations could not be derived because the association did not hold at the province level in the case of Italy, the authors advocated considering confounding factors when analysing the effect of intergenerational relations.

Notwithstanding, as patterns in social mixing are embedded in socioeconomic and cultural factors and are very different across Europe, this continent remains an excellent geographical area for study of COVID-19 proliferation. The present study therefore analyses statistical associations between different indicators of social and economic ties and the reported number of confirmed cases of COVID-19 in 23 European countries between March 1 and April 30, 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 (CSSE, 2020). Data on the covariates come from different sources and relate to years as close to 2020 as was possible to obtain (see notes under Table 1). Data from the European Social Survey, Eurostat, the World Bank, and the OECD were used to approximate different types of social and cultural ties, which we hypothesize to be positively associated with COVID-19: average number of household members; percentage living in a multi-generational household; proportion of people who have frequent social meetings with friends, relatives, or colleagues; and religious attendance. In addition, we also test the effect of the socioeconomic variables tertiary education and GDP per capita, as we assume that higher educated or more economically developed countries are more likely to pursue activities that require travelling (e.g., international business meetings, skiing), factors which contributed to the initial outbreak of the epidemic. Lastly, we test the effect of demographic variables: the share of the population aged 65 +, population density and per capita number of beds in nursing and residential care facilities (all expected to be positively associated with COVID-19) (Table 1).

Bivariate associations between the covariates and the natural logarithm (ln) of the cumulative number of confirmed COVID-19 cases between March 1 and April 30 are analysed at 10-day intervals to ascertain whether the direction and strength of the associations changed over time. A later date was not analysed due to country-differences in the severity and timing of movement and social contact restrictions that European governments implemented during this period, thus confounding the effect of the tested social and economic tie variables. In the supplementary material file the analyses are repeated for COVID-19 cases per 100,000 population and the number of cases during each 10-day 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 (Harrell Jr et al. 1984; Peduzzi et al. 1996), we first opted to reduce the number of variables by performing a factor analysis. This method not only simplifies the subsequent analysis, it also alerts us to groupings of variables that we would not otherwise have thought of, enabling us to work at a more sophisticated conceptual level (De Vaus, 2002).

The factor analysis yielded three sociodemographic-like latent factors that explained 78% of the country-variation in the selected covariates. Each latent factor is highly associated (> 0.75) with one or more covariates, and this is why Factor 1 has been labelled “socially and economically vibrant”, Factor 2 “relatively young population”, and Factor 3 “densely populated and traditional” (see Tables S1-S4 in the Supplementary file). Correlation coefficients between COVID-19 and the obtained latent factors were first calculated (Table 2) before performing the multivariate regression analyses[1] to obtain the adjusted R^2 of the models (Table 3) and the unstandardized coefficients, i.e., the factors’ slope (Table S5).

[1] As the rotated factors are orthogonal and thus not correlated, each covariate’s standardized coefficient in the multivariate regression analysis is the same as the bivariate Pearson’s correlation coefficient.

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 con-firmed	Cumulative cases of COVID-19 on							Population as at 1/1/2019	Confirmed cases on 31/03/ 100,000 people	Confirmed cases on 30/04/ 100,000 people
				01/03	11/03	21/03	31/03	10/04	20/04	30/04			
Austria	West	AT	25/02	14	246	2814	10,180	13,555	14,795	15,452	8,858,775	114.91	174.43
Belgium	West	BE	04/02	2	314	2815	12,775	26,667	39,983	48,519	11,455,519	111.52	423.54
Bulgaria	East	BG	08/03	0	7	163	399	635	929	1506	7,000,039	5.70	21.51
Czechia	East	CZ	01/03	3	91	995	3308	5732	6900	7682	10,649,800	31.06	72.13
Denmark	North	DK	27/02	4	444	1420	3039	6014	7711	9356	5,806,081	52.34	161.14
Estonia	East	EE	27/02	1	16	306	745	1258	1535	1689	1,324,820	56.23	127.49
Finland	North	FI	29/01	6	59	523	1418	2769	3868	4995	5,517,919	25.70	90.52
France	West	FR	24/01	130	2293	14,463	52,827	91,738	155,393	167,299	67,012,883	78.83	249.65
Germany	West	DE	27/01	130	1908	22,213	71,808	122,171	147,065	163,009	83,019,213	86.50	196.35
Hungary	East	HU	04/03	0	13	103	492	1190	1984	2775	9,772,756	5.03	28.40
Ireland	West	IE	29/02	1	43	785	3235	8089	15,652	20,612	4,904,240	65.96	420.29
Italy	South	IT	31/01	1694	12,462	53,578	105,792	147,577	181,228	205,463	60,359,546	175.27	340.40
Lithuania	East	LT	28/02	1	3	83	537	999	1326	1385	2,794,184	19.22	49.57
Netherlands	West	NL	27/02	10	503	3640	12,667	23,249	33,588	39,512	17,282,163	73.30	228.63
Norway	North	NO	26/02	19	598	2118	4641	6314	7156	7738	5,328,212	87.10	145.23
Poland	East	PO	04/03	0	31	536	2311	5955	9593	12,877	37,972,812	6.09	33.91
Portugal	South	PT	02/03	0	59	1280	7443	15,472	20,863	25,045	10,276,617	72.43	243.71
Serbia	East	RS	06/03	0	12	171	900	3105	6630	9009	6,963,764	12.92	129.37
Slovenia	East	SI	05/03	0	57	383	802	1160	1335	1429	2,080,908	38.54	68.67
Spain	South	ES	01/02	84	2277	25,374	95,923	158,273	200,210	213,435	46,937,060	204.37	454.73
Sweden	North	SE	31/01	14	500	1763	4435	9,685	14,777	21,601	10,230,185	43.35	211.15
Switzerland	West	CH	25/02	27	652	6575	16,605	24,551	27,944	29,586	8,544,527	194.33	346.26
The UK	West	UK	31/01	36	459	5067	25,481	74,605	125,856	178,771	66,647,112	38.23	268.24
EU 23 total			24/01	2176	23,047	147,168	437,763	750,763	1,026,321	1,188,745	490,739,135	89.20	174.43

Table 1 (continued)

Country	Mean number of household members	% living in a multi-generational household	% having a frequent social meeting	% ≥ weekly religious attendance	% tertiary education	Population per km2	% people aged 65 +	GDP per capita	Nursing/ rest home beds / 100,000
Austria	2.2	5	65.5	11.4	31.1	107.21	19.00	51,462	862.4
Belgium	2.3	4.8	68.6	8.6	36.0	377.21	18.79	47,519	1,234.1
Bulgaria	2.4	24.7	51.7	9.3	24.7	64.70	21.02	9273	30.8
Czechia	2.4	7.6	50.8	5.8	21.6	137.60	19.42	23,079	687.5
Denmark	2.0	1.2	71.2	4.3	33.1	138.07	19.81	61,350	816.1
Estonia	2.2	8.9	43.5	3.7	36.5	30.39	19.63	23,266	870.7
Finland	2.0	1.5	65.1	4.6	38.5	18.16	21.72	50,152	1,190.0
France	2.2	2.0	67.9	7.6	33.7	122.34	20.03	41,464	981.5
Germany	2.0	5.3	59.7	7.1	25.9	237.37	21.46	47,603	1,152.2
Hungary	2.3	11.2	20.4	10.6	22.5	107.91	19.16	16,162	853.3
Ireland	2.6	8.1	58.9	33.9	40.7	70.45	13.87	78,806	639.3
Italy	2.3	14.2	60.1	27.4	17.4	205.45	22.75	34,483	415.8
Lithuania	2.2	9.4	31.1	17.1	37.9	44.53	19.71	19,153	726.3
Netherlands	2.2	2.7	74.5	9.8	34.8	511.46	19.20	53,024	1,379.6
Norway	2.0	2.3	78.5	3.9	37.7	14.55	17.05	81,697	765.6
Poland	2.8	16.1	30.9	46.5	28.2	124.04	17.52	15,421	195.3
Portugal	2.5	13.1	77.6	27.8	23.8	112.24	21.95	23,408	555.7
Serbia	2.9	21.3	62.1	11.1	20.6	79.83	18.35	7247	291.5
Slovenia	2.5	16.1	53.1	12.6	29.3	102.64	19.61	26,124	1,012.4

Table 1 (continued)

Country	Mean number of household members	% living in a multi-generational household	% having a frequent social meeting	% ≥ weekly religious attendance	% tertiary education	Population per km2	% people aged 65 +	GDP per capita	Nursing/ rest home beds / 100,000
Spain	2.5	17.9	70.2	15.2	35.1	93.53	19.38	30,371	834.8
Sweden	2.0	1.2	74.3	5.7	37.8	25.00	20.10	54,608	1,388.0
Switzerland	2.2	2.3	71.6	8.3	38.6	215.52	18.62	82,797	1,174.2
The UK	2.3	3.9	62.4	11.7	40.6	274.83	18.40	42,944	821.4

Source: Confirmed COVID-19 cases: Center for Systems Science and Engineering at Johns Hopkins University (CSSE, 2020). Population on 1 January 2019 (Eurostat: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_pjan&lang=en). Mean number of household members 2018 (Eurostat: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_lvph01&lang=en), % 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) and % ≥ weekly religious attendance 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: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=edat_lfse_03&lang=eng), except BGR and SRB (ESS)). Population per km^2 in 2018 (World Bank: https://data.worldbank.org/indicator/EN.POP.DNST), % people aged 65 + in 2018 (World Bank: https://data.worldbank.org/indicator/SP.POP.65UP.TO.ZS) and GDP in 2018 (World Bank: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD). 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 long-term care recipients between Portugal and Spain to the beds available in Spain); Eurostat (https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=hlth_rs_bdsns&lang=en) and OECD (https://stats.oecd.org/Index.aspx?ThemeTreeId=9).

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 between March 1 and April 30, 2020. The highest (and significant) associations are observed for the social meeting and population density variables with the latter association becoming stronger over time. Figure 2 and Table 3 present the association between the three extracted factors and COVID-19. The “socially and economically vibrant” factor has a strong and positive association throughout the study period. The “relatively young population” factor is negatively associated (as expected) with COVID-19 but its p-values are insignificant. On the other hand, the association of the “densely populated and traditional” factor was initially low but increased with time, becoming the most important factor by the end of March. The three factors together explain close to 50% of the cross-country variation in the number of confirmed cases of COVID-19 between March 11 and April 30, compared to just 21% on March 1 (Table 4). The slope of the “socially and economically vibrant” factor was greatest on March 11 and that of the “densely populated and traditional factor” on April 20.

If we analyse the change in COVID-19 cases over 10-day periods rather than the absolute number of cumulative cases, results are virtually the same (Supplementary Table S5a). The “socially and economically vibrant” factor is strongly significant during March and the first 10 days of April, while the “densely populated and traditional factor” was significant throughout the entire study period and became the most important explanatory factor from March 21. The “relatively young population” factor showed little association in any of the models. The proportion of the country differences in change in COVID-19 explained by the three factors equalled 44–48%, dropping down to 33% during the last 10 days of April). Conversely, the factors explain much less of the country differences in the number of cases of COVID-19 per 100,000 population. The correlation with the “socially and economically vibrant” factor is above 0.4 from March 31 onwards and the same applies to the densely populated and traditional factor 10 days later (which again becomes the most important explanatory factor) (Supplementary Table S6). The explanatory power increased steadily over time from 3% (March 1) to 43% (April 30). This is consistent with the fact that towards the end of April, the countries with a 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 (Supplementary Figure S4).

Discussion

Confirmed cases of COVID-19 increased sharply across Europe during March and April of 2020. Throughout most of the studied period, Italy was worst hit by the pandemic in absolute numbers, but Spain surpassed Italy in early April, while Belgium and Ireland did so in terms of cases per 100,000 people. The question we posed is

Table 2 Correlation between the natural log COVID-19 in 23 European countries at six different time periods and the covariates

Variable	March 1	March 11	March 21	March 31	April 10	April 20	April 30
1. Mean number of household members	−0.40	−0.38	−0.28	−0.20	−0.13	−0.08	−0.06
2. % living in a multi-generational household (ln)	−0.31	−0.42*	−0.29	−0.24	−0.23	−0.21	−0.21
3. % having a frequent social meeting (squared)	0.41	0.61**	0.56**	0.53**	0.52*	0.50*	0.50*
4. % ≥ weekly religious attendance (ln)	−0.07	−0.10	0.03	0.12	0.17	0.20	0.21
5. % tertiary education	−0.01	0.06	0.05	0.05	0.05	0.05	0.04
6. Population density (people per km^2) (ln)	0.25	0.37	0.44*	0.50*	0.53**	0.54**	0.54**
7. % people aged 65 + (squared)	0.35	0.22	0.23	0.18	0.14	0.10	0.10
8. GDP per capita	0.35	0.51*	0.44*	0.38	0.35	0.33	0.32
9. Nursing and rest home beds per capita	0.25	0.39	0.31	0.29	0.27	0.24	0.22

**$p < 0.01$, *$p < 0.05$

Table 3 Correlation between the natural log of COVID-19 in 23 European countries 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	April 30
Factor 1: Socially and economically vibrant	0.44*	0.61**	0.52**	0.47**	0.44*	0.41	0.40
Factor 2: Relatively young population	−0.26	−0.16	−0.14	−0.11	−0.07	−0.04	−0.04
Factor 3: Densely populated and traditional	0.25	0.38*	0.48**	0.55**	0.59**	0.60**	0.60**

**$p < 0.01$, *$p < 0.05$

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 ties that may have led to the *initial* spread of the COVID-19 pandemic. While these factors continued to be important throughout the analysed period, population density and cultural factors also contributed to the subsequent diffusion of the virus.

Considering 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. Recent studies have shown that nursing homes may be responsible for 19% to 72% of COVID-19 deaths (Comas-Herrera et al. 2020; Orange, 2020). On the other hand, Italy, which was the initial epicentre of the pandemic in Europe, scored very low on the "relatively young population" factor, but high on the "densely populated and traditional" factor. In other words, its aged population, high population density and traditional values (approximated through the proportion weekly church attendants) is likely to have contributed to their high rate of diagnosed cases. Moreover, the relative position of other traditionally catholic countries in COVID-19, including Spain, Portugal and Belgium worsened markedly between March 21 and April 10 (Fig. 2).

While the country differences in COVID-19 cases per 100,000 people across the 23 European countries could only be weakly explained by the three factors in early March, by mid-April, both the "socially and economically vibrant" and "densely populated and traditional" factors contributed significantly to the explanation of the European country differences, as the number of cases increased markedly in the most socioeconomically developed and densely populated European countries. More specifically, by analysing changes in COVID-19 cases over 10-day periods, we found that during the early stage of the epidemic (early March) social and economic ties appeared to be most important, while population density and church attendance explained more of the growth in diagnosed cases from late March until late April

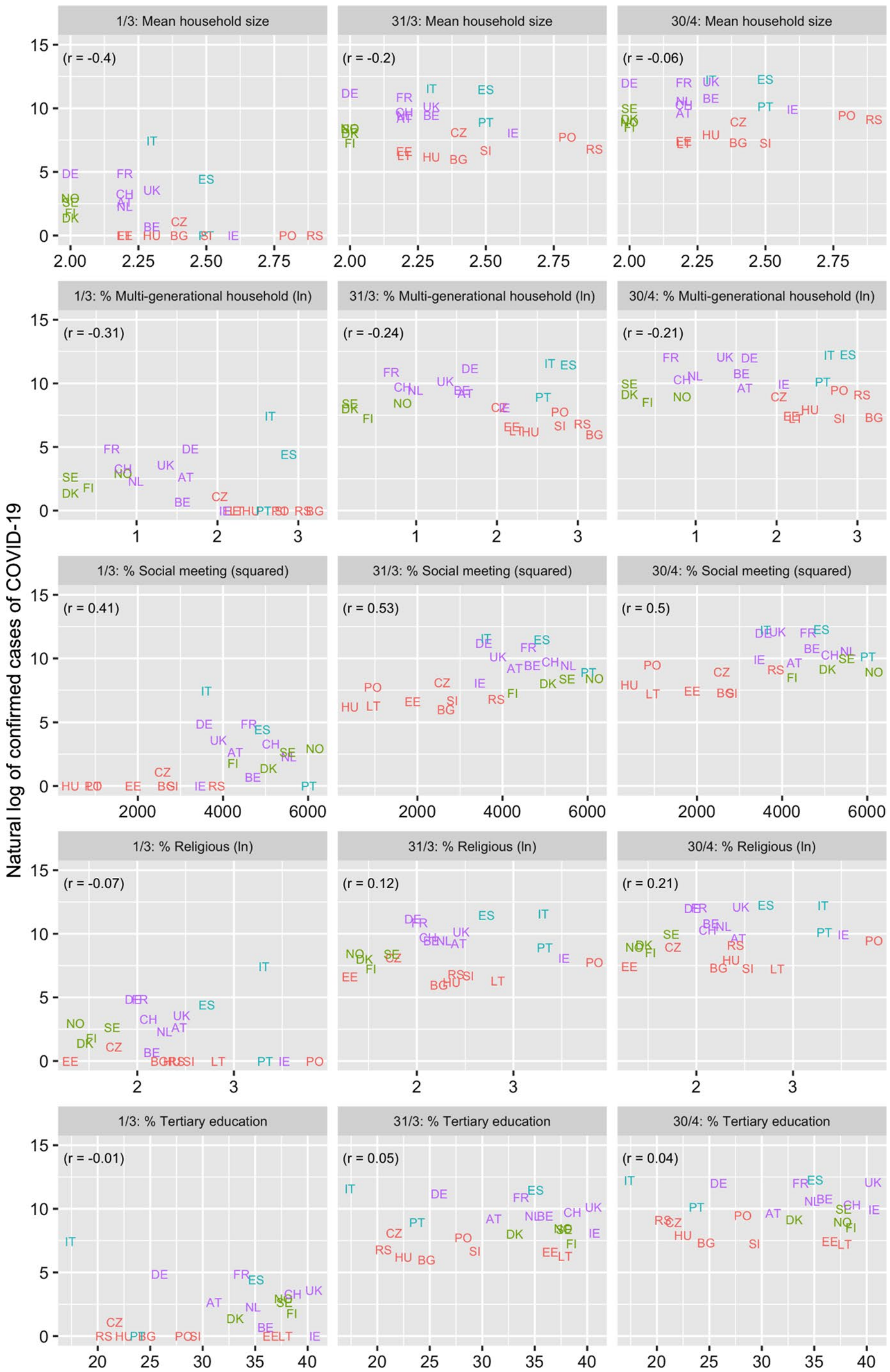

Fig. 1 Association between covariates and natural log of cumulative cases of COVID-19, March 1, 31, April 30, 2020 among 23 European countries

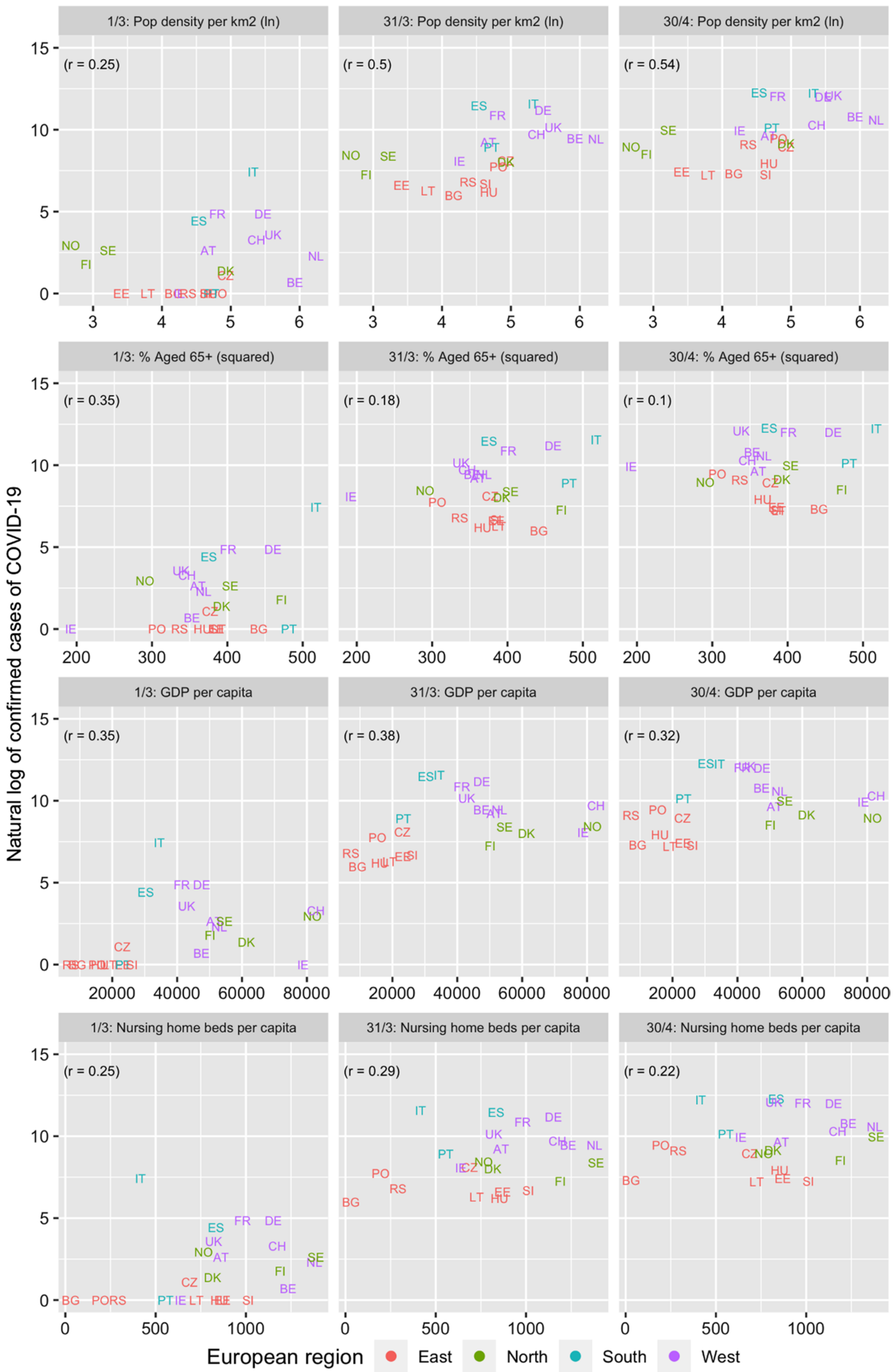

Fig. 1 (continued)

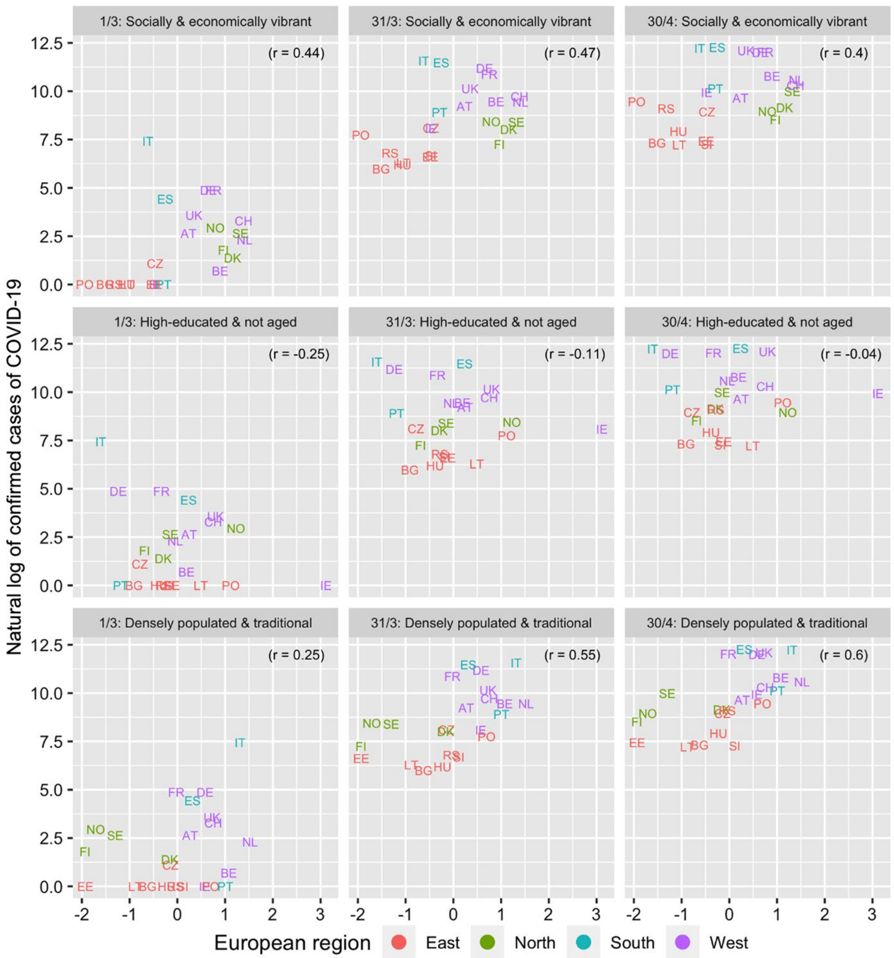

Fig. 2 Association between factor scores and natural log of cumulative cases of COVID-19. March 1, 31, and April 30, 2020 among 23 European countries

(the end of the analysed study period). In light of research from elsewhere, it is noteworthy to mention that the role of religious services in the spread of the coronavirus in South Korea is well-documented through field investigations that established the source of infection for many cases (Prem, Liu, et al., 2020). Likewise, regarding the initial spread in Europe, Bartscher et al. (2020) also found the coronavirus to be initially more prevalent in high social capital areas. On the other hand, the effect of a relatively young, highly educated population (Factor 2) was not associated with the number of registered COVID-19 infections. We know that during the first wave of the pandemic few asymptomatic and/or young people were tested (Kohns Vasconcelos et al. 2021; Surkova et al. 2020), so this could explain why the association was not positive. Conversely, as results also showed that multigenerational households

Table 4 Multivariate regression analysis of social and demographic factors on cumulative cases of the ln of COVID-19 in 23 European countries on March 1, 11, 21 and 31, and April 10, 20, and 30, 2020. Unstandardized coefficients (*p*-value)

	Unstandardized coefficients (p value)						
	March 1	March 11	March 21	March 31	April 10	April 20	April 30
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)**	**0.66 (0.022)**
Factor 2: Relatively young population	−0.53 (0.194)	−0.35 (0.304)	−0.26 (0.377)	−0.18 (0.506)	−0.12 (0.649)	−0.07 (0.795)	−0.07 (0.803)
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)**	**1.00 (0.001)**
R^2 adjusted	0.21	0.48	0.44	0.46	0.47	0.46	0.45

was the most important variable in the “socially and economically vibrant” factor and highly positive with COVID-19, it suggests that it is not the proportion of elderly per se that leads to higher rates of COVID-19, but the level of social interaction (as well as the number people in residential care homes). In this context and supported by evidence from studies that analysed intergenerational co-residence (Esteve et al. 2020) and contact patterns (Prem, van Zandvoort, et al., 2020), tailored public health responses, in particular shielding policy for the elderly, are recommended during the early stage of corona-type of virus epidemics to not only reduce their impact on the health of individuals and public health care systems, but also on the economy (see also Prem, Liu, et al., 2020; Prem, van Zandvoort, et al., 2020; Davies et al. 2020).

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. Governments have differed in the timing of the implementation of measures such as cancelling public events, closing day care centres, schools and universities, social distancing, or partial or total lockdowns (Flaxman et al. 2020). 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 their peak of daily additional cases of COVID-19 (e.g., Italy). Given the estimated average latency period between becoming 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. A French and Austrian ski resort was responsible for initial infections in the UK and other Northern European countries (Flaxman et al. 2020; Hruby, 2020), but tertiary education and GDP variables are likely to capture the influence of winter holidays or international travel on country differences in COVID-19 during the studied period. Smoking is another variable associated with the proliferation of COVID-19 due to its social function (Paul et al. 2010) and because smokers are more likely to touch their face and mouth and have chronic health conditions (GBD, 2015 Tobacco Collaborators 2017; Science Media Centre, 2020). However, we think that smoking is a more important factor to consider in individual or small-area studies, as analysis showed smoking rates to be higher in (mainly Eastern European) countries where the coronavirus was late in getting a hold.

Another issue of concern is country differences in testing for COVID-19. Some countries only test people admitted to hospitals or ramped up the testing program much later during the first 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 as it mainly pertains to symptomatic people (Farge & Revill, 2020; Kohns Vasconcelos et al. 2021; Wikipedia, 2020). Apart from unknown symptomacy, the data did not contain information on the place where the coronavirus was contracted. Such data has only been used in (family) case cluster studies (e.g. Chan et al. 2020; Danis et al. 2020; Fong et al. 2020). In the context of our study, of particular interest would have been being able to distinguish between the proportion of infections that took place outside the home (more likely among children and people employed) and those within the household (more probable among older people and those living in multigenerational or overcrowded households). As we analysed different periods of the first wave, such information could have provided us with more insight into the importance of particular variables for the different settings of infection.

A recurrent problem of any national-level analysis that uses aggregate data is that any association found might not necessarily reflect associations that are observed at the individual level, a shortcoming known as the “ecological fallacy”. That said, we did not obtain results contradictory to what we expected.

Finally, data from the ESS is not available for all European countries, implying that different results may be obtained if data for other countries become available.

To conclude, the main take away message for public health policy is that, while disentangling the effect of a variegated number of social, cultural, economic, and demographic factors on the diffusion of the COVID-19 epidemic is a difficult task, the level of importance of specific determinants in spreading the virus is likely to change over time. In a European setting we found that factors associated with the level of economic development and social ties were particularly important initially, while population density and cultural factors were likely to have facilitated further spread of the virus once it took hold. However, as Chinazzi et al. (2020) showed, the implementation of international travel restrictions would not be enough to curb the initial spread of a virus, as disease transmissibility also needs to be reduced through public health interventions and behavioural changes. Examples of the efficacy of border management policies and stringent public health interventions are Taiwan and New Zealand. Although New Zealand was helped by its isolation, Taiwan managed even better to limit the number of reported infections, despite its close proximity to the Chinese mainland, through its existing disease and outbreak surveillance systems and effective means of face mask distribution and promotion. Both countries also had strict border management policies, quarantining rules, and secure facilities for incoming travellers in place and developed contact tracing (Summers et al. 2020). Conversely, European island nations, particularly Ireland and Iceland, were clearly much less successful in dealing with the first wave of the pandemic as infection rates were one of the highest in the world (CSSE, 2020). Based on our results we therefore recommend for future outbreaks of coronavirus-like epidemics when no vaccine is yet available, quick implementation of travel restrictions and very strict measures of social distancing. This should be especially done in densely populated countries with strong international economic and social ties, in order to minimise the proliferation of cases during the secondary transmission phase that occurs within households and nursing homes. A recommendation for future research is to perform a European analysis at the sub-national level, given the unequal distribution of COVID-19 within countries (e.g., the north vs. the south of Italy).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12546-021-09257-1.

Author contributions RM had the original idea for the paper. JS performed the analysis and wrote the first draft. Both authors contributed to the final draft.

Funding Funding was received from the Spanish Ministry of Science, Innovation and Universities (RYC-2013-14851; CSO2017-89721-R; RTI2018-096730-B-I00), the European Research Council (ERC-2019-CoG-864616) and from the Catalan Government under the CERCA Program.

Data availability and code Details of the sources are provides in the article, but the created data file and codes used to perform the analysis can be obtained from the corresponding author upon request.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Ethical approval We only used publically available secondary data, so no ethics compliance required.

References

Arpino B, Bordone V, & Pasqualini M (2020) Are intergenerational relationships responsible for more COVID-19 cases? A cautionary tale of available empirical evidence. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.2008581117

Bartscher AK, Seitz S, Slotwinski M, Siegloch S, & Wehrhöfer N (2020) Social capital and the spread of Covid-19: Insights from European countries. *CESifo Working Paper, No. 8346*

Bayer C, & Kuhn M (2020) Intergenerational ties and case fatality rates: A cross-country analysis. *IZA Discussion Paper, 13114* https://voxeu.org/article/intergenerational-ties-and-case-fatality-rates.

Bi, Q., Wu, Y., Mei, S., Ye, C., Zou, X., Zhang, Z., et al. (2020). Epidemiology and transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts. *medrxiv*. https://doi.org/10.1101/2020.03.03.20028423

Comas-Herrera A, Zalakain J, Litwin C, Hsu AT, Lane N, & Fernandez-Plotka JL (2020) Mortality associated with COVID-19 outbreaks in care homes: early international evidence. https://ltccovid.org/

CSSE (2020). Novel Coronavirus (COVID-19) Cases. https://github.com/CSSEGISandData/COVID-19.

Chan, J.F.-W., Yuan, S., Kok, K.-H., To, K.K.-W., Chu, H., Yang, J., et al. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet, 395*(10223), 514–523.

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., et al. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science, 368*(6489), 395–400.

Danis, K., Epaulard, O., Bénet, T., Gaymard, A., Campoy, S., Botelho-Nevers, E., et al. (2020). Cluster of coronavirus disease 2019 (COVID-19) in the French Alps, February 2020. *Clinical Infectious Diseases, 71*(15), 825–832.

Davies, N. G., Kucharski, A. J., Eggo, R. M., Gimma, A., Edmunds, W. J., Jombart, T., et al. (2020). Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. The Lancet Public Health. https://doi.org/10.1016/S2468-2667(20)30133-X

De Vaus, D. (2002). *Surveys in social science* (5th ed.). Routledge.

Esteve, A., Permanyer, I., Boertien, D., & Vaupel, J. W. (2020). National age and coresidence patterns shape COVID-19 vulnerability. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.2008764117

Farge E, & Revill J (2020) 'Test, test, test': WHO chief's coronavirus message to world. *Reuters,* March 16. https://www.reuters.com/article/us-healthcare-coronavirus-who/test-test-test-who-chiefs-coronavirus-message-to-world-idUSKBN2132S4.

Flaxman S, Mishra S, Gandy A, Unwin HJT, Coupland H, Mellan TA, et al. (2020) Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. https://arxiv.org/abs/2004.11342.

Fong, M., Cowling, B., Leung, G., & Wu, P. (2020). Letter to the editor: COVID-19 cases among school-aged children and school-based measures in Hong Kong, July 2020. *Eurosurveillance, 25*(37), 2001671.

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. *The Lancet, 389*(10082), 1885–1906.

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*(11), 1065–1072.

Harrell, F. E., Jr., Lee, K. L., Califf, R. M., Pryor, D. B., & Rosati, R. A. (1984). Regression modelling strategies for improved prognostic prediction. *Statistics in medicine, 3*(2), 143–152.

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.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan China. *The Lancet, 395*(10223), 497–506.

Kohns Vasconcelos, M., Renk, H., Popielska, J., Nyirenda Nyang'wa, M., Burokiene, S., Gkentzi, D. et al. (2021). SARS-CoV-2 testing and infection control strategies in European paediatric emergency departments during the first wave of the pandemic. *European Journal of Pediatrics*, *180*, 1299–1305.

Liu, Y., Eggo, R. M., & Kucharski, A. J. (2020). Secondary attack rate and superspreading events for SARS-CoV-2. *The Lancet, 395*(10227), e47.

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., et al. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine, 5*(3), e74.

Orange, R. (2020). 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.

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*(1), 211.

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*(12), 1373–1379.

Peeri, N.C., Shrestha, N., Rahman, M.S., Zaki, R., Tan, Z., Bibi, 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*, *49*(3), 717–726. https://doi.org/10.1093/ije/dyaa033.

Prem, K., Liu, Y., Russell, T.W., Kucharski, A.J., Eggo, R.M., Davies, N., et al. (2020). The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. The Lancet Public Health. *Lancet Public Health*, *5*, e261–70. https://doi.org/10.1016/S2468-2667(20)30073-6.

Prem, K., van Zandvoort, K., Klepac, P., Eggo, R. M., Davies, N. G., Cook, A. R., et al. (2020). Projecting contact matrices in 177 geographical regions: an update and comparison with empirical data for the COVID-19 era. *medrxiv*. https://doi.org/10.1101/2020.07.22.20159772

Reher, D. S. (1998). Family ties in Western Europe: persistent contrasts. *Population and Development Review, 24*(2), 203–234.

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 and Society, 34*(2), 330–354.

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/.

Summers, D. J., Cheng, D. H. Y., Lin, P. H. H., Barnard, D. L. T., Kvalsvig, D. A., Wilson, P. N., et al. (2020). Potential lessons from the Taiwan and New Zealand health responses to the COVID-19 pandemic. The Lancet Regional Health - Western Pacific, 100044.

Surkova, E., Nikolayevskyy, V., & Drobniewski, F. (2020). False-positive COVID-19 results: hidden problems and costs. *The Lancet Respiratory Medicine*, *8*(12), 1167–1168. https://doi.org/10.1016/S2213-2600(20)30453-7.

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*(10), 936–944.

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