Supplementary Material

Differential impact of physical distancing strategies on social contacts relevant for the spread of SARS-CoV-2: evidence from a cross-national online survey, March–April 2020

Contents

Additional methods

Survey design

The data used in the analysis came from the Covid-19 Health Behaviour Survey (CHBS), an online survey carried out in multiple countries (and multiple languages) to collect key information about people's health and behaviour during the early phase of the pandemic (March – April 2020) [1]. Participation was voluntary and open to people who were at least 18 years old. Recruitment occurred through targeted advertisements implemented through the Facebook Ads Manager (FAM), which is a tool that can be used to quickly reach large numbers of survey participants across several countries [2]. The ads were stratified by age group, sex, and region of residence (largely following the NUTS-1 classification in Europe and the census regions in the United States, both aggregated into larger macro-regions) to ensure that a minimum number of respondents was reached in all demographic strata.

The questionnaire was divided into multiple sections covering different areas of interest: socio-demographic indicators (age, gender, country of birth, region of residence, level of education, and household size); health indicators (symptoms experienced in the previous seven days, among others); opinions and behaviours' (adoption of protective behaviours and disruption to daily routines, among others); and social contact data. For the purposes of validation and comparability, the questionnaire included standard questions taken from relevant sources, such as the European Social Survey for socio-demographic questions [3], and an Ipsos poll for questions related to public opinion on the COVID-19 outbreak [4].

Informed consent was obtained from all participants, enabling the collection, storage, and processing of their answers. Participants' data was treated anonymously. Ethical approval for the study was obtained from the Ethics Council of the Max Planck Society.

Recruitment Strategy

Marketers use the FAM to advertise their products or businesses. The FAM enables advertisers to create advertising campaigns that can have various goals, such as creating salience for a given service or product among Facebook users or generating traffic to an external website. Following previous examples of Facebook surveys [5–7], we used the FAM to run campaigns to recruit respondents for our study. The countries included in the study, selected because of the number of COVID-19 cases they were experiencing, were Germany, Italy, Spain, the UK, the US, France, Belgium, and the Netherlands. The FAM allowed us to perform non-probabilistic quota sampling by targeting user groups based on demographic characteristics such as sex, age, and region of residence. In this way, we could not only ensure a minimum number of participants in each combination of strata, but we were also able to control for variables possibly associated the survey outcomes and the propensity to take the survey [8,9].

At the beginning of March 2020, prior to the start of the actual survey, we carried out separate pilots in Italy, the UK, and the US to test the FAM and the performance of our ads, allowing the study to officially begin on March 13. Fig. S1 shows the question that prompted survey participants to report the information on their social contacts in the English-language version of the questionnaire.

Due to technical problems on the Facebook platform for some countries, very few respondents accessed the survey on certain days between March 20 and March 25. We thus excluded from the analysis those days with less than ten respondents. Although this issue might have increased the uncertainty in the model predictions (especially the daily ones), we believe that the analysis by week is robust and less affected by this issue.

Poststratification

Since online surveys are not random samples of the population, we adjusted our survey data using a poststratification weighting approach. In particular, given that our aim was to achieve a nationally representative sample, we stratified the survey by sex, age group, and region of residence, which are all important variables that are related to differences in people's responses to the pandemic, as well as to survey participation. In order to create poststratification weights, we divided the true population proportion in each stratum, obtained from nationally representative data available through Eurostat (2019) [3] and the US census (2018) [10], by the sample proportion from the same stratum in our survey. These poststratification weights were used in all the statistical analyses presented in this paper.

Social contact data

We defined social contacts, which are the focus of this paper, as any interaction involving either physical contact (such as a handshake or a hug) or a conversation of three or more words in the physical presence of another person. This definition is consistent with the definition employed in past social contact surveys, which we used as the baseline for comparison in this work [11,12]. In more detail, we asked respondents to report the number of individuals with whom they interacted on the day before the survey in different settings: i.e., at home, at school or college, at work, and in the general community (such as during commuting or leisure activities); while making it clear that the respondent should not report multiple interactions with the same person in different settings. However, unlike previous contact surveys, we did not ask about the characteristics of the contacted individuals (e.g., age and sex) to avoid overburdening respondents, given the nature of the online survey. The text describing the social contact question is shown in Fig. S1.

Health behaviours and perceived risk

We asked respondents to rate the level of perceived threat they thought the COVID-19 posed to different levels of society (Q13), namely, (i) to themselves, (ii) to their family, (iii) to their local community, (iv) to their country, and (v) to the world, using a 5-point Likert scale. For the sake of our analysis, we rescaled participants' responses to the range 0 – 1, where zero indicated "very low threat" and one "very high threat". To account for a feeling of threat that might have affected social contact patterns of respondents, we combined in one variable the perceived level of threat to themselves and to their families, taking their simple mean. These two indicators were found to have a good internal consistency, having a Cronbach's alpha of 0.84.

Moreover, we asked respondents which health behaviours, in the form of specific actions, they took up to protect themselves from the virus (Question (Q18): "Which of the following actions, if any, have you already taken to protect yourself from the coronavirus?"). Among the listed actions, coded as dummy variables, we asked whether respondents avoided shaking hands, social activities (e.g., meeting friends), crowded places (e.g., restaurants, cinemas, gym, or playground), travelling by public transportation (e.g., by bus, tram, subway, or train) or by taxi, whether they stockpiled medicines or food, whether they used a face mask, and whether they washed their hands or used a sanitizing hand gel more often. For our analysis, we picked three of these actions, which might have an impact on the transmissibility of the virus, namely, avoiding social activities ("Avoided social activities (e.g. meeting friends)"), wearing a face mask ("Worn a face mask"), and washing the hands more often ("Washed hands more often") [13,14].

Epidemiological analysis

The net reproduction number R_t can be estimated as the dominant eigenvalue, denoted as ρ, of the next generation matrix N, i.e., $R_t = \rho(N)$, where N provides information on the numbers of newly infected individuals by age group at a given time [15]. Under the "social contact hypothesis", $N = DqC$, where C is the matrix containing the average number of contacts between age groups, q is the disease transmissibility parameter, and D is the length of the infectiousness period [16]. We use the proportional relationship between N and C, i.e., $N = \frac{R_t}{\rho(C)} C$, to assess changes in R_t due to changes in social contact numbers [17–19].

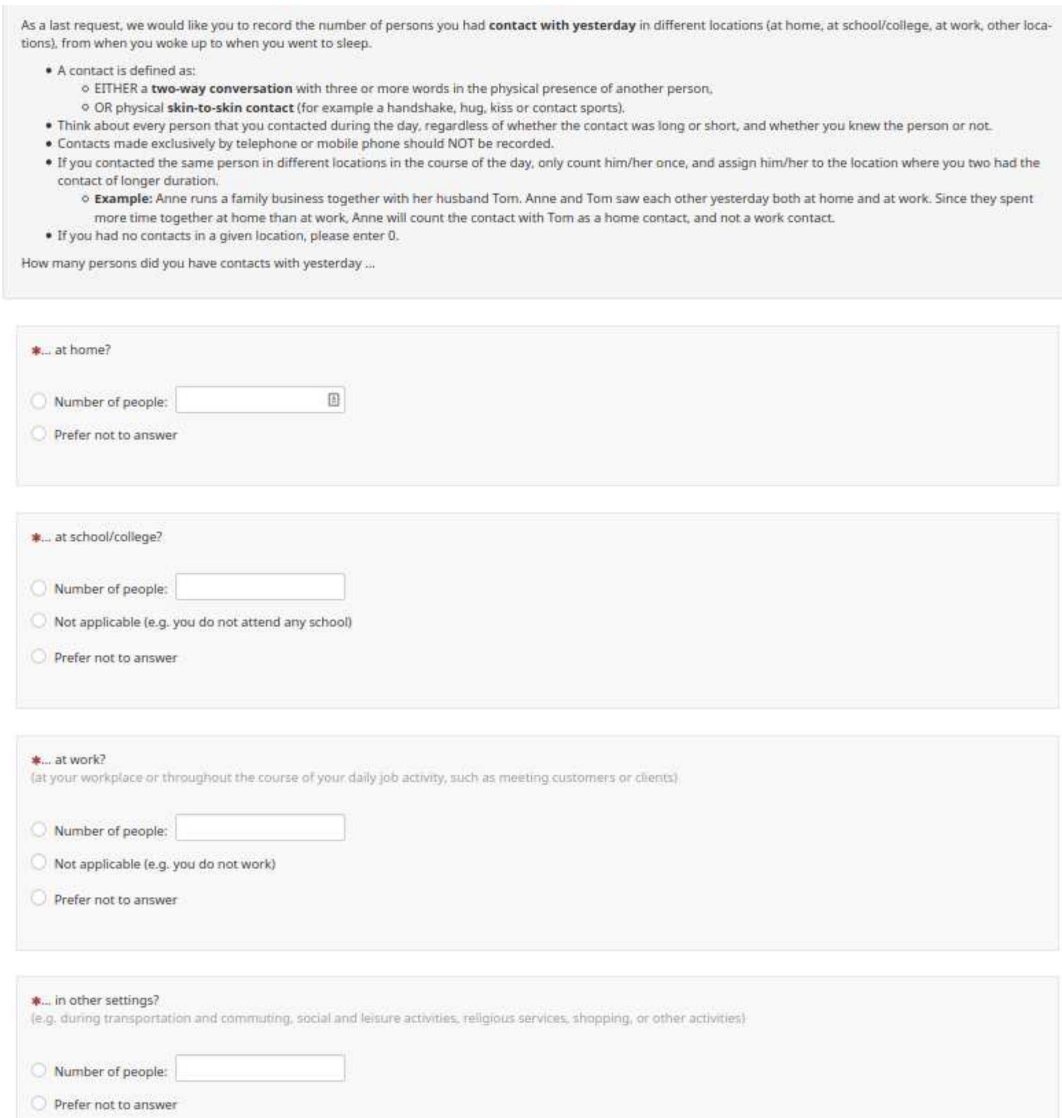
As we did not collect information on the age of the individuals encountered by the study participants, we derived the contact matrices C for each study week using the following procedure. First, for each country, based on the data collected in the POLYMOD study between 2005 and 2006 [11], and in the Comes-F studies in 2012 [12], we constructed pre-COVID age-specific social contact matrices C^{pre}, using the same age groups of the survey (18 – 24 years, 25 – 44 years, 45 – 64 years, and 65 years or more), which contained the average number of contacts between participants in the ith age group and their contactees in the jth age group, using a constraint for the reciprocity of contact when considering the total number of contacts at the population level. This constraint implies that, in a closed population, the total number of contacts from age class i to age class j must be equal to the total number of contacts from age class j to age class i. Hence, if we had N_i individuals in age class i and N_j individuals in age class j, reciprocity would entail that $C_{ij}N_i = C_{ji}N_j$. For this purpose, we started by dividing the total number of contacts, M_{ij}, for the number of participants in each age group, n_i, i.e., $m_{ij} = \frac{M_{ij}}{n_i}$. We then constructed the matrix with the total number of contacts between age classes at the population level, by multiplying the expected mean number of contacts reported by participants in the age group i with contactees in the age group j, m_{ij}, for the total number of individuals in the age group i in the population, i.e., $Z_{ij} = m_{ij}N_i$. At this point, we adjusted Z_{ij} for reciprocity at the population level by averaging the total number of contacts in one direction, Z_{ij}, with those in the other direction, Z_{ji}, weighting for the sample size by age group, namely, $Z_{ij}^{(rec.)} = \frac{(Z_{ij}n_i + Z_{ji}n_j)}{(n_i + n_j)}$ [20]. We finally obtained the expected average number of contacts under reciprocity at the population level dividing again by the population N_i, i.e., $C_{ij}^{pre} = Z_{ij}^{(rec.)}N_i$.

Second, we projected the matrices C^{pre} to the population structure of each country in 2020 [20] and, to remove differences in the contact levels between countries, we normalized each matrix C^{pre} to the average number of contacts at the population level, so that the mean number of contacts of the matrix $C^{pre+norm}$ was equal to one.

Finally, we multiplied the normalized matrix $C^{pre+norm}$ by the average number of daily contacts that we predicted for each group and study week, thus obtaining the weekly COVID matrices C_w^{cov}. This final passage was motivated by the fact that we did not collect in our survey the information on the age of the individuals contacted by survey participants, but only their reported number of contacts. Had we collected the age of the contacted individuals, we could have directly built the age-specific matrices. In this case, we had to assume that, due to the physical distancing, people reduced their number of contacts without fundamentally changing the structure of their mixing patterns, which is represented by $C^{pre+norm}$.

As we did not collect data for people younger than 18 years old, we computed their average number of daily contacts per week in the following way. To reflect school closure during the survey period, we excluded school contacts from the contact data prior to COVID-19, considering only the overlapping age groups between the two types of matrices. Then, for each country, we computed a weekly overall scaling factor, and we obtained the average number of contacts for individuals aged less than 18 by multiplying their average number of pre-COVID-19 contacts in the population for such factor.

To account for the uncertainty in the pre-COVID contact data, we applied a non-parametric bootstrap procedure to the original data, resampling with replacement the participants, proportionally to the population age distribution, and assigning them to their original contact data.



As a last request, we would like you to record the number of persons you had **contact with yesterday** in different locations (at home, at school/college, at work, other locations), from when you woke up to when you went to sleep.

- A contact is defined as:
  - EITHER a **two-way conversation** with three or more words in the physical presence of another person,
  - OR physical **skin-to-skin contact** (for example a handshake, hug, kiss or contact sports).
- Think about every person that you contacted during the day, regardless of whether the contact was long or short, and whether you knew the person or not.
- Contacts made exclusively by telephone or mobile phone should NOT be recorded.
- If you contacted the same person in different locations in the course of the day, only count him/her once, and assign him/her to the location where you two had the contact of longer duration.
  - **Example:** Anne runs a family business together with her husband Tom. Anne and Tom saw each other yesterday both at home and at work. Since they spent more time together at home than at work, Anne will count the contact with Tom as a home contact, and not a work contact.
- If you had no contacts in a given location, please enter 0.

How many persons did you have contacts with yesterday ...

*... at home?

Number of people:

Prefer not to answer

*... at school/college?

Number of people:

Not applicable (e.g. you do not attend any school)

Prefer not to answer

*... at work?
(at your workplace or throughout the course of your daily job activity, such as meeting customers or clients)

Number of people:

Not applicable (e.g. you do not work)

Prefer not to answer

*... in other settings?
(e.g. during transportation and commuting, social and leisure activities, religious services, shopping, or other activities)

Number of people:

Prefer not to answer

Fig. S1. CHBS questions on daily social contacts.

Additional figures

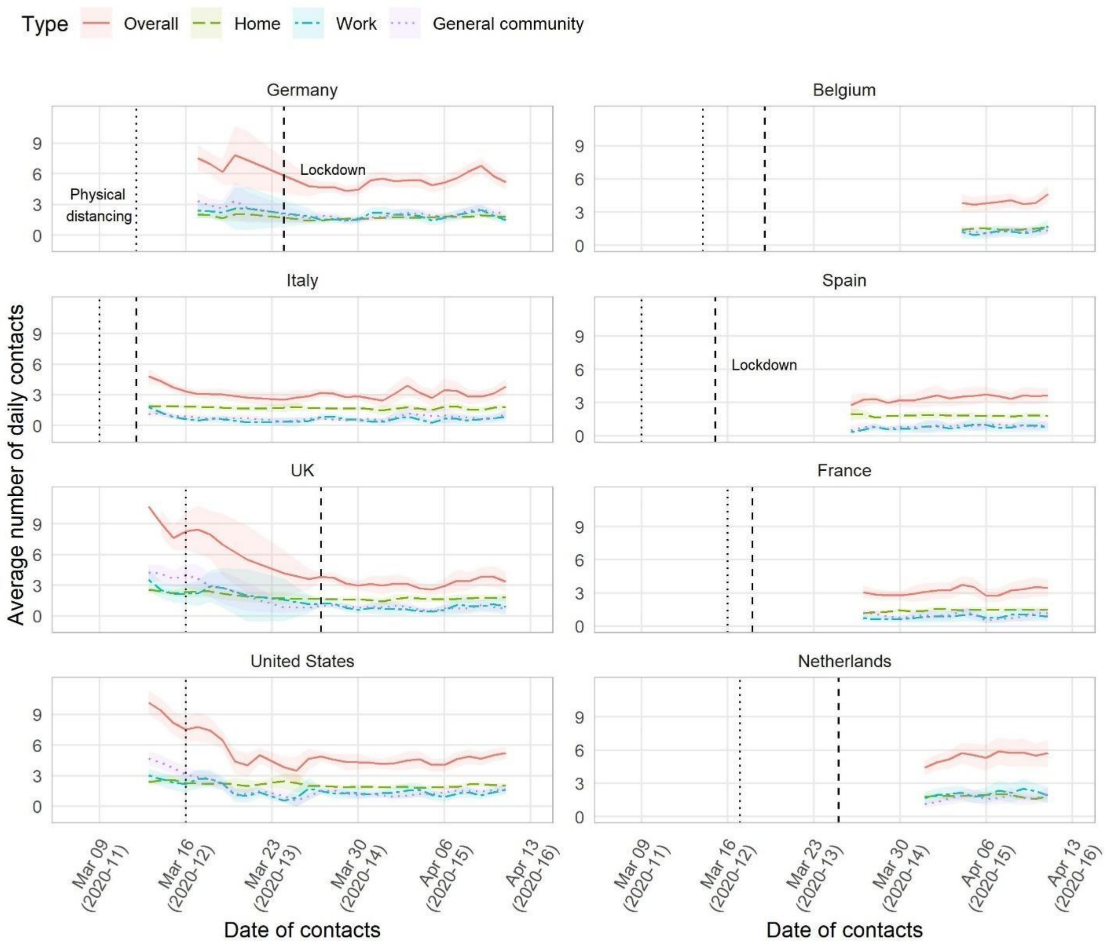


Fig. S2. Mean overall number of daily social contacts (with 95% CI), smoothed by a simple two-day moving average, by country and study day. The dotted line corresponds to the date in which the physical distancing guidelines were introduced at the national level; the dashed line corresponds to the date in which the lockdown, regardless of being full or partial, was ordered. Respondents with contacts above the $\leq$29 threshold were removed from the analysis.

We performed a sensitivity analysis of the threshold chosen to curtail the right tail of the distribution of social contacts, in response to the extremely high values reported by some of the participants. In the following figures and tables, we present the results of the analysis using a fixed cut-off point at 29 contacts per day, removing respondents reporting 30 contacts per day or more in each of the settings.

Two features of the highly skewed distribution of the overall social contacts (Fig. S2) are worth noting. First, we found that the percentages of participants who reported having fewer than one contact per day were relatively high, at more than 35% in all countries except for Germany, the Netherlands, and the US. Second, the distribution was characterized by a long right tail, the length of which depended on the threshold used to remove the outliers at the top of the distribution (Tables S9-S16).

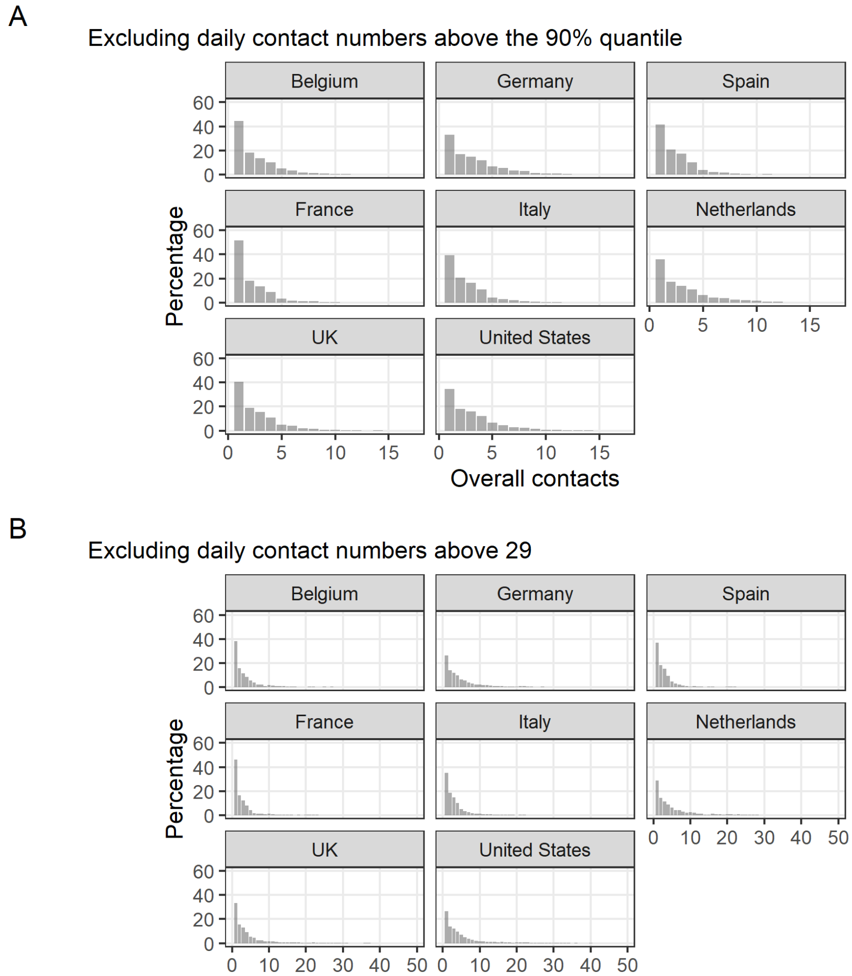


Fig. S3. Histogram with the weighted distribution of the overall number of contacts by country. All contacts exclude the top 10% (A) and those equal to 30 or more (B).

In Figures S4-S5, we showed the evolution of the weekly average number of contacts, overall and by setting, after applying the post-stratification weights, under the two threshold scenarios.
Home contacts (which show little variability over the weeks) are the main component of the overall number of contacts, followed by contacts in the general community and at work. Eventually, the contacts at school, which are reduced to zero due to the closure of secondary schools and university/colleges in every surveyed country.

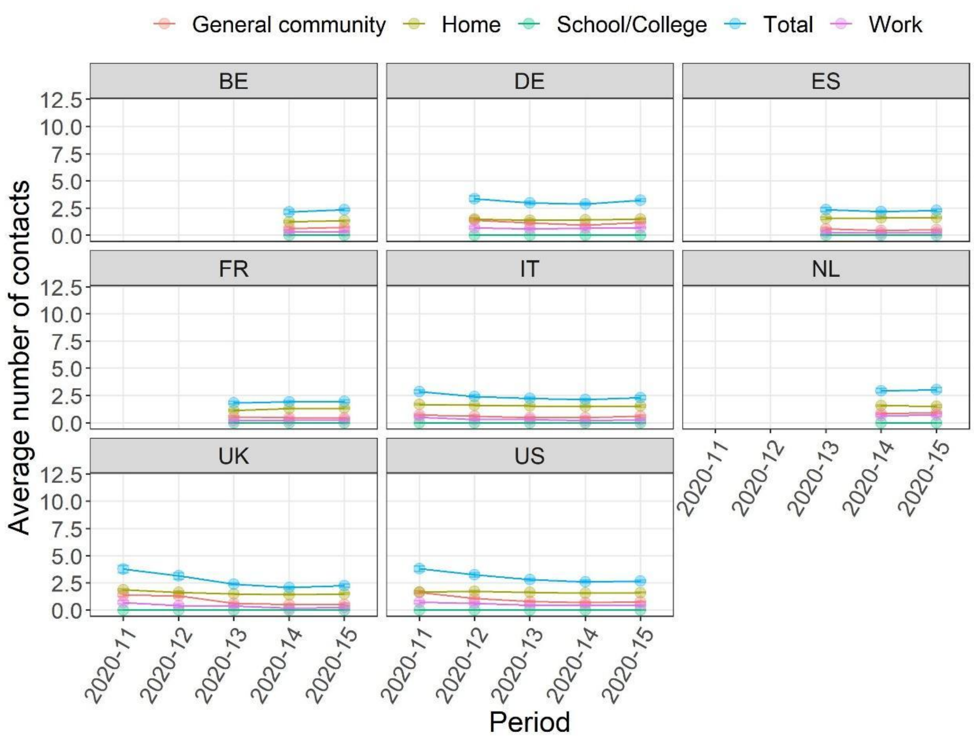


Fig. S4. Average number of daily social contacts, overall and by setting, per week and country. Respondents with contacts above the 90% quantile threshold were removed from the analysis.

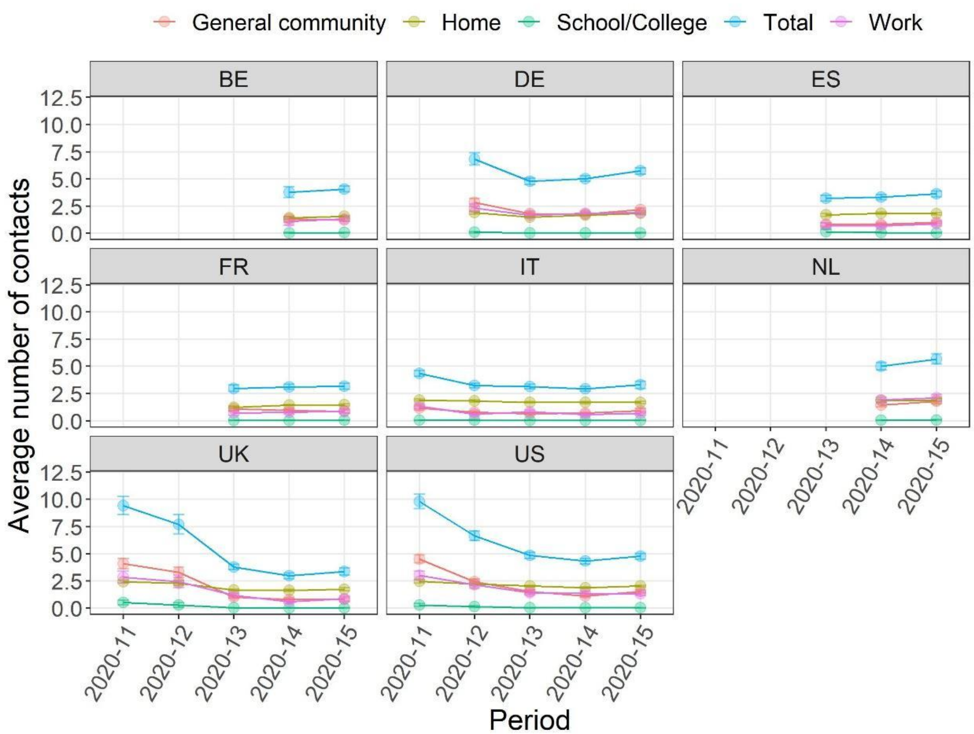


Fig. S5. Average number of daily social contacts, overall and by setting, per week and country. Respondents with contacts above the ≤29 threshold were removed from the analysis.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

Fig. S6 shows the consistency of the household size and the number of home contacts under the two thresholds. More than half of respondents lived either alone or with another person. Respondents living in large households, with five members or more, were more frequent in the US (above 10%). Home contacts ranged between around half a contact per day, reported by people living alone, to three contacts per day or more, reported by people living in households of four or more components, regardless of the threshold.

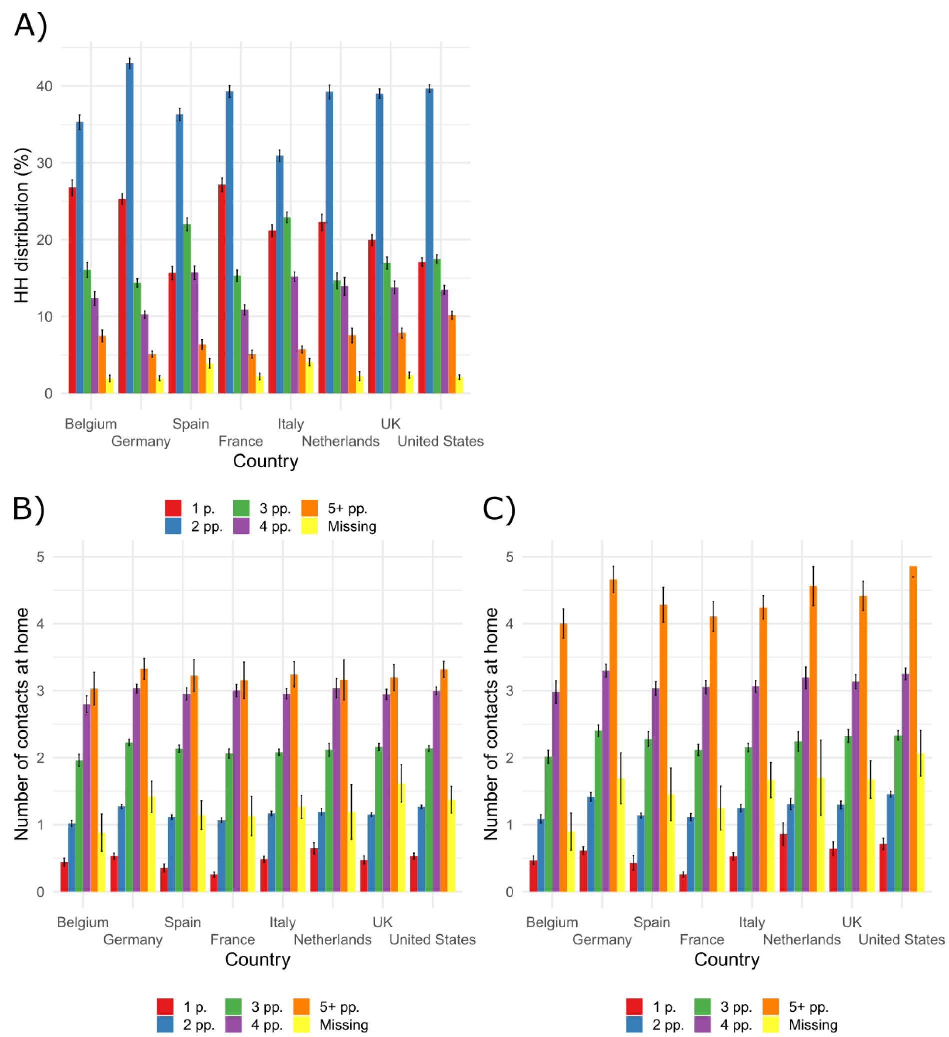


Fig. S6. Household size distribution by country (A), compared to the average number of contacts at home by household size and country, under the 90% quantile threshold (B), and under the ≤29 threshold.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

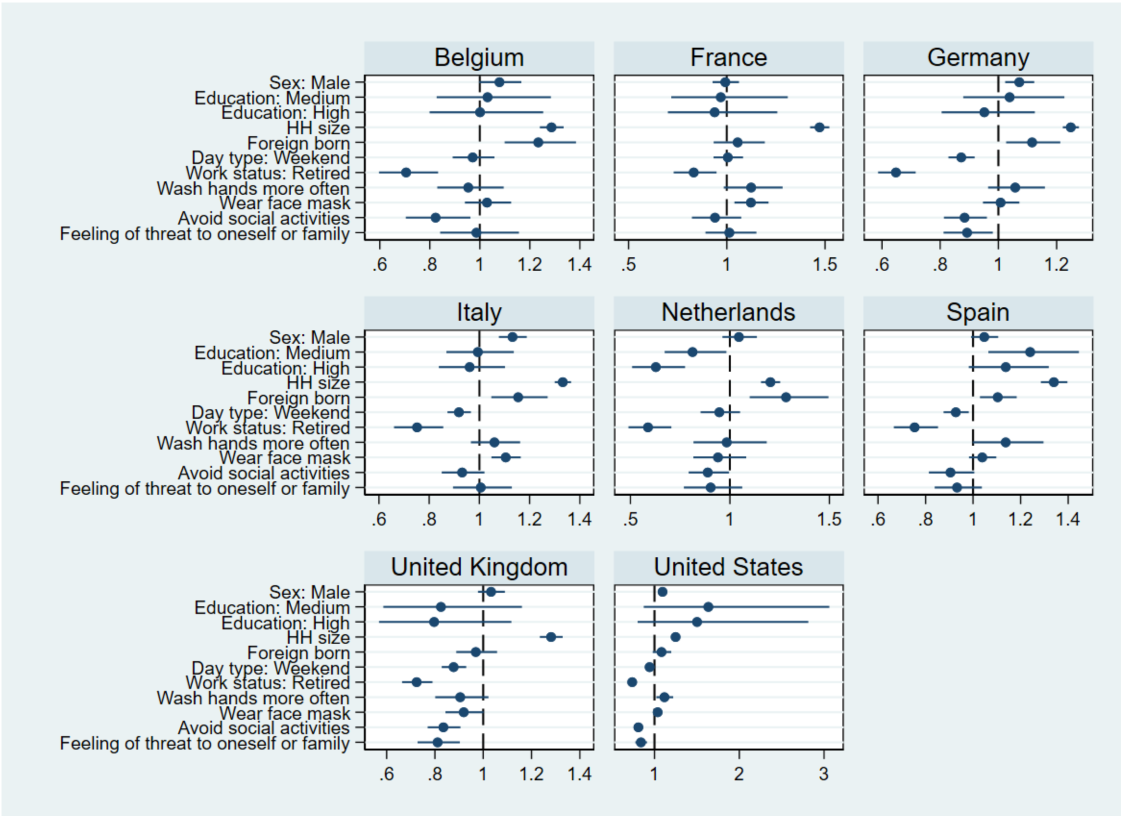


Fig. S7. Exponentiated coefficients (with 95% confidence interval) for the socio-demographic and behavioural variables included in the negative binomial regression model for the number of overall daily contacts reported by participants in each country. Respondents with contacts above the 90% quantile threshold were removed from the analysis.

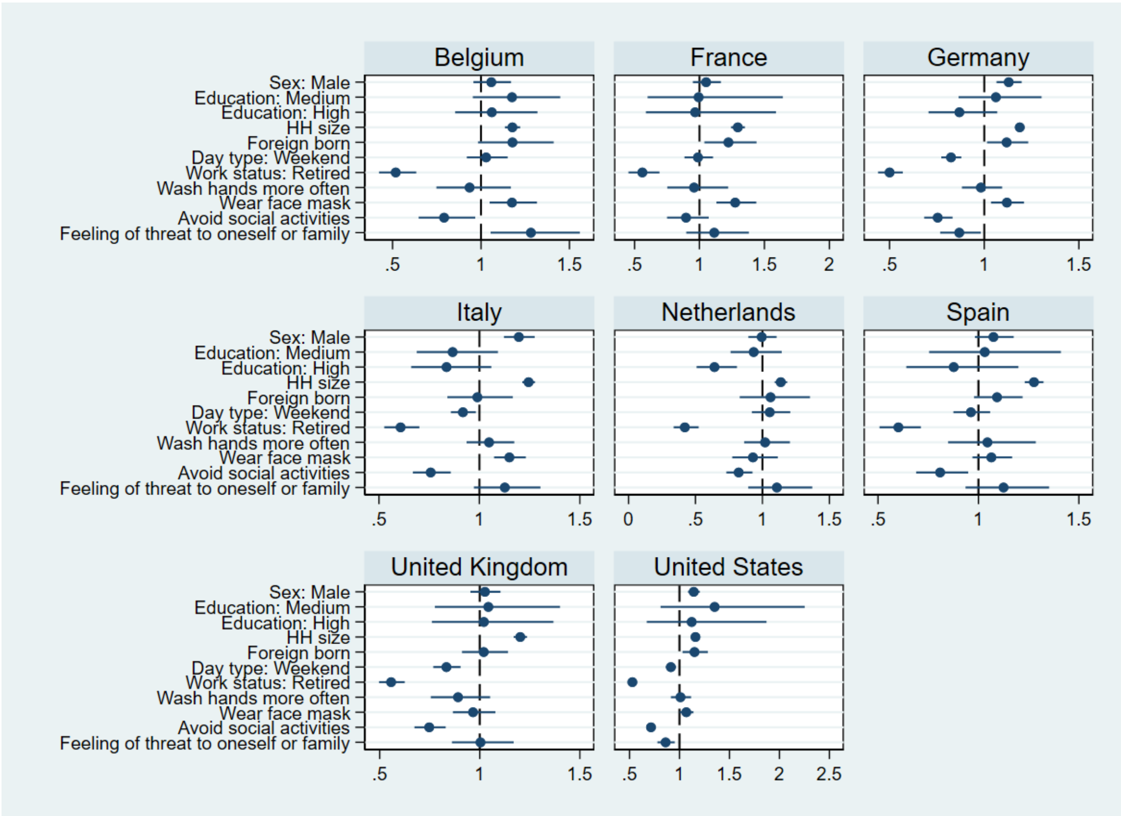
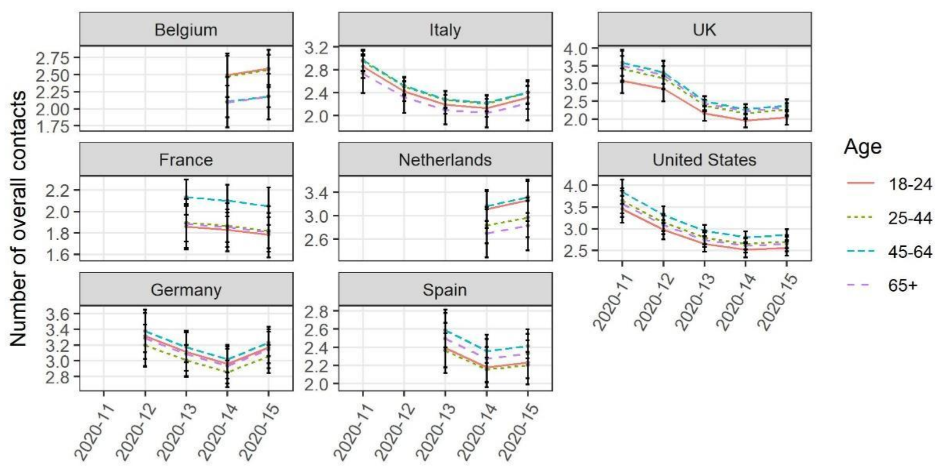


Fig. S8. Exponentiated coefficients (with 95% confidence interval) for the socio-demographic and behavioural variables included in the negative binomial regression model for the number of overall daily contacts reported by participants in each country. Respondents with contacts above the ≤29 threshold were removed from the analysis.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

(A)



(B)

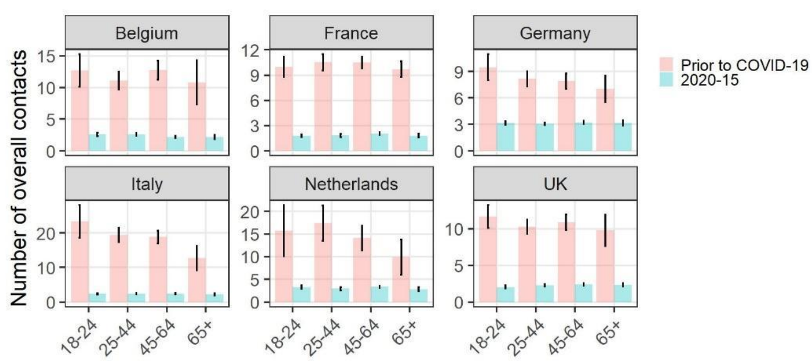


Fig. S9. Model-predicted daily number of overall social contacts, by age group, country, and week, March-April 2020. (A). Comparison of model-predicted overall contact numbers between the pre-COVID period and calendar week 2020-15, by country (B). Respondents with contacts above the 90% quantile threshold were removed from the analysis.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

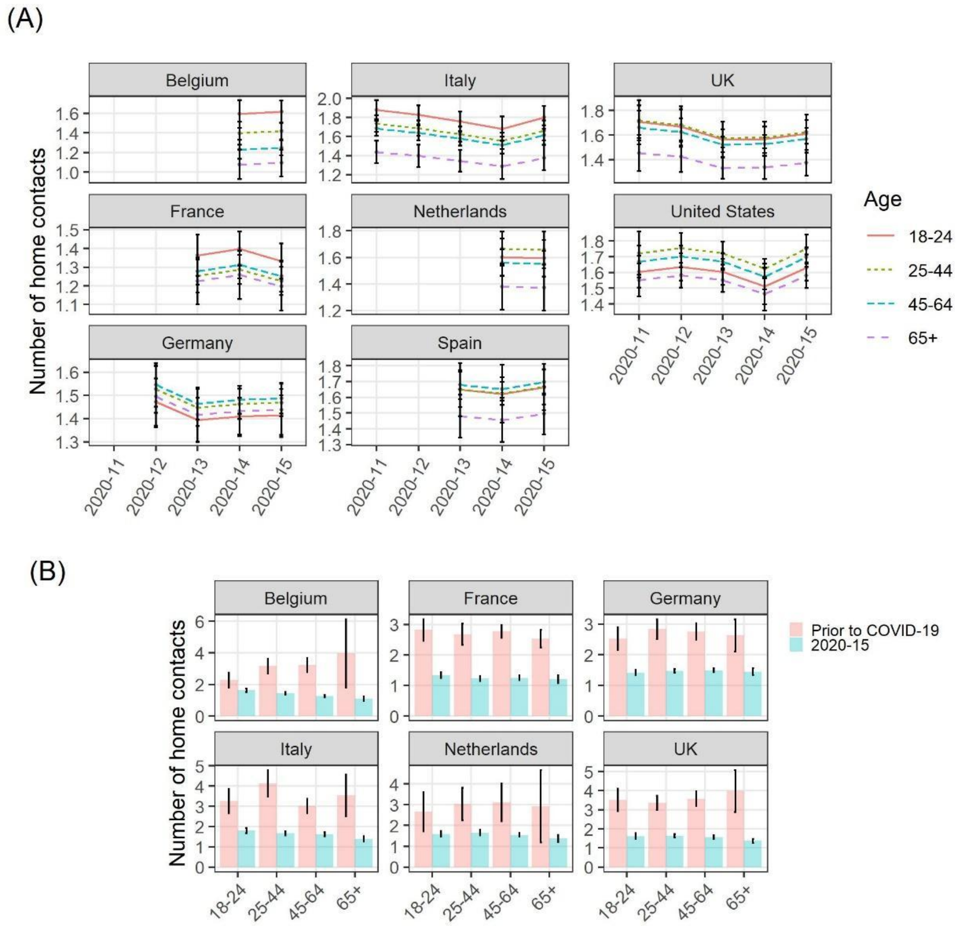


Fig. S10. Model-predicted daily number of social contacts at home, by age group, country, and week, March-April 2020. (A). Comparison of model-predicted overall contact numbers between the pre-COVID period and calendar week 2020-15, by country (B). Respondents with contacts above the 90% quantile threshold were removed from the analysis.

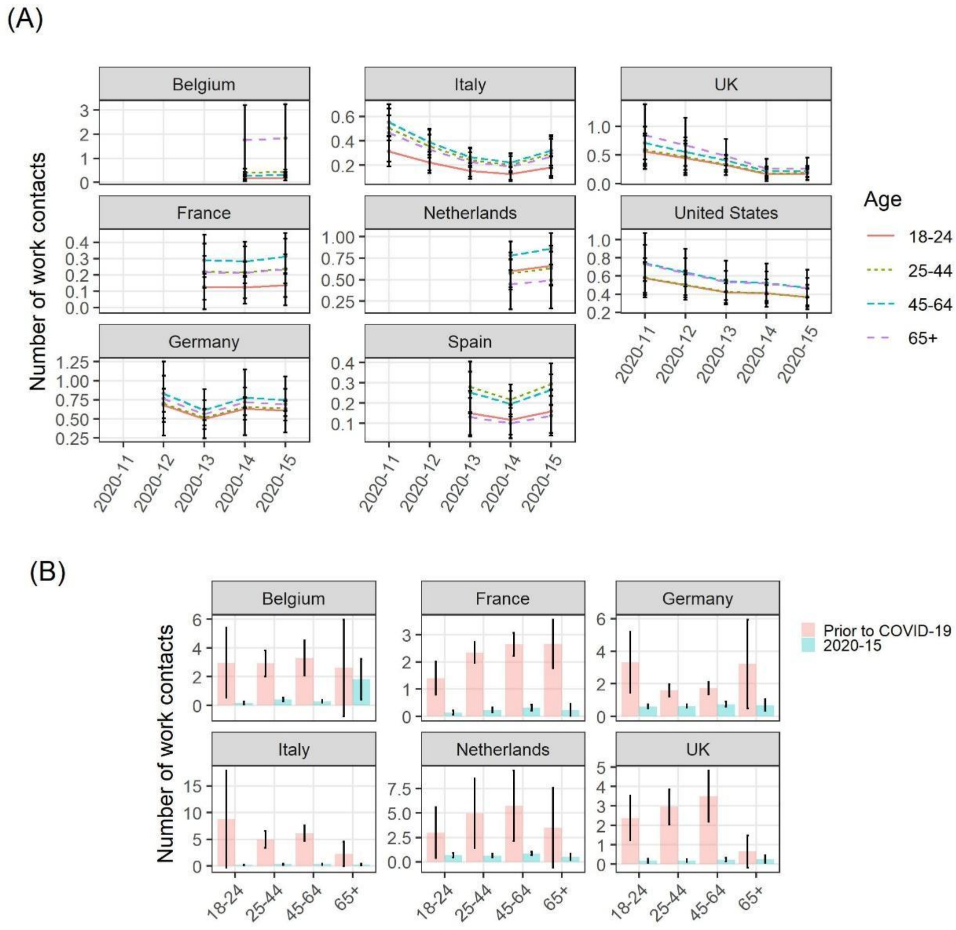


Fig. S11. Model-predicted daily number of social contacts at work, by age group, country, and week, March-April 2020. (A). Comparison of model-predicted contact numbers at work between the pre-COVID period and calendar week 2020-15, by country (B). Respondents with contacts above the 90% quantile threshold were removed from the analysis.

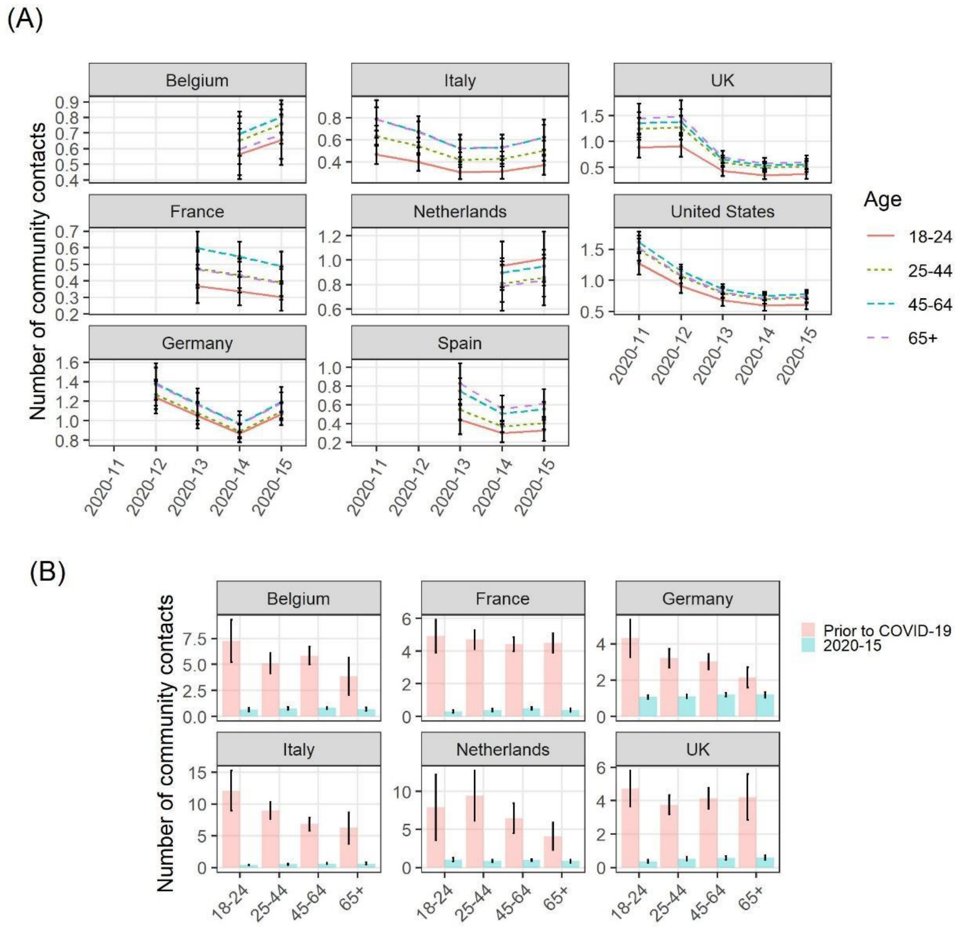


Fig. S12. Model-predicted daily number of social contacts in the general community, by age group, country, and week, March-April 2020. (A). Comparison of model-predicted contact numbers in the general community between the pre-COVID period and calendar week 2020-15, by country (B). Respondents with contacts above the 90% quantile threshold were removed from the analysis.

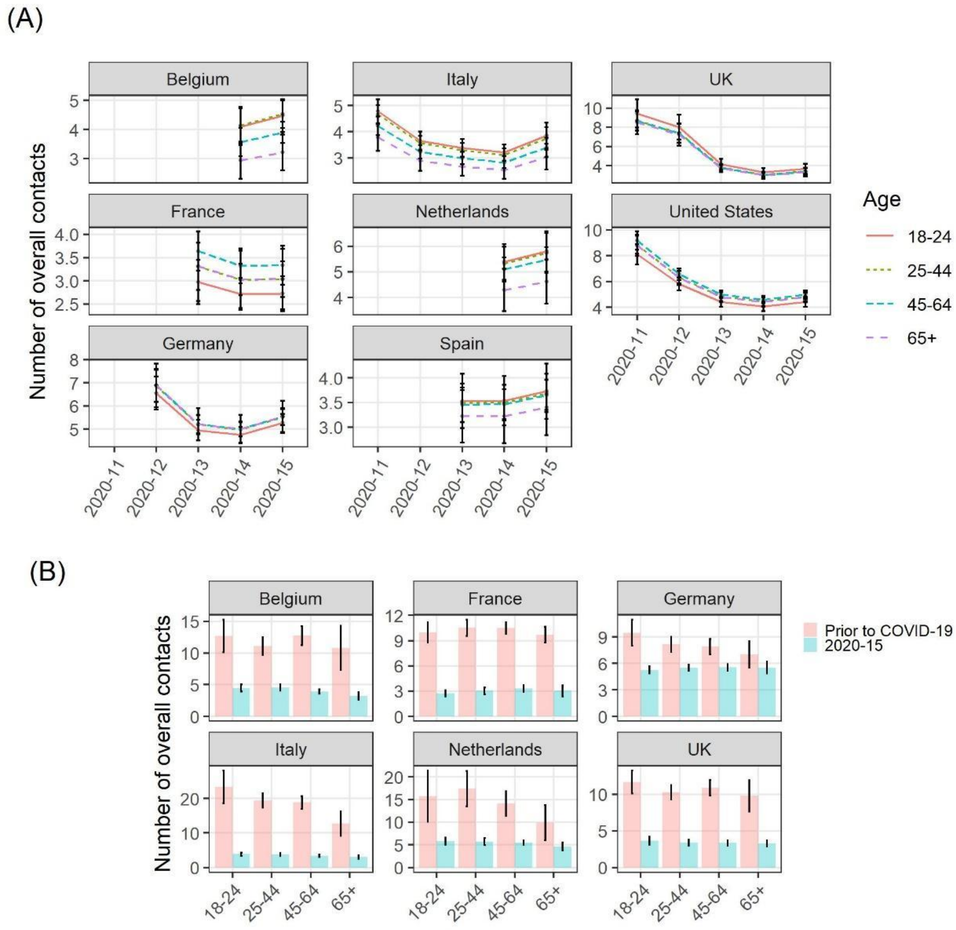


Fig. S13. Model-predicted daily number of overall social contacts, by age group, country, and week, March-April 2020. (A). Comparison of model-predicted overall contact numbers between the pre-COVID period and calendar week 2020-15, by country (B). Respondents with contacts above the ≤29 threshold were removed from the analysis.

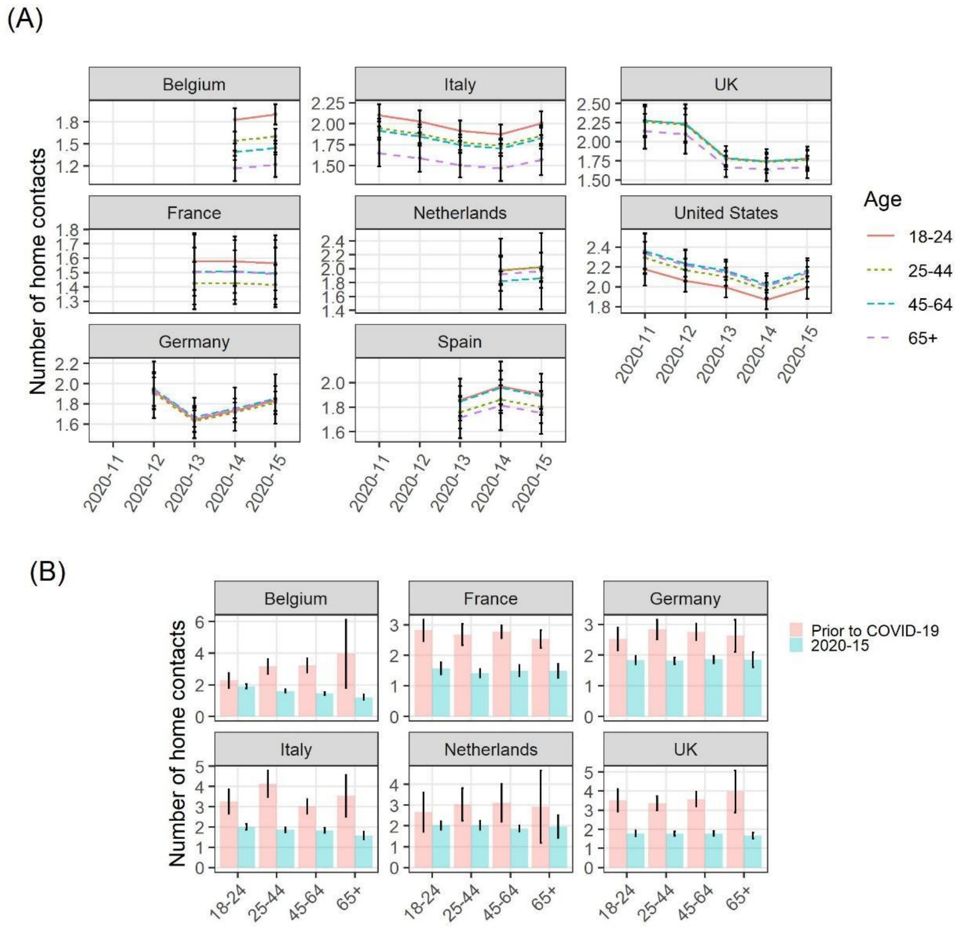


Fig. S14. Model-predicted daily number of social contacts at home, by age group, country, and week, March-April 2020. (A). Comparison of model-predicted contact numbers at work between the pre-COVID period and calendar week 2020-15, by country (B). Respondents with contacts above the ≤29 threshold were removed from the analysis.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

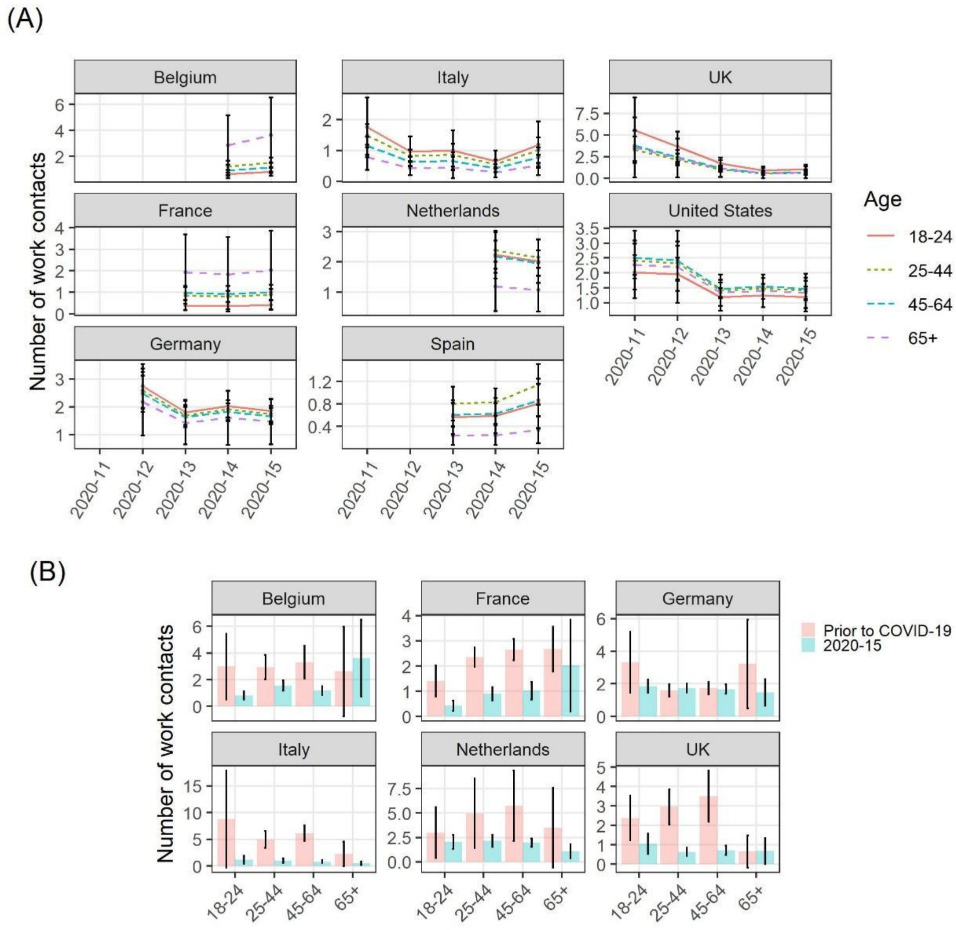


Fig. S15. Model-predicted daily number of social contacts at work, by age group, country, and week, March-April 2020. (A). Comparison of model-predicted contact numbers at work between the pre-COVID period and calendar week 2020-15, by country (B). Respondents with contacts above the ≤29 threshold were removed from the analysis.

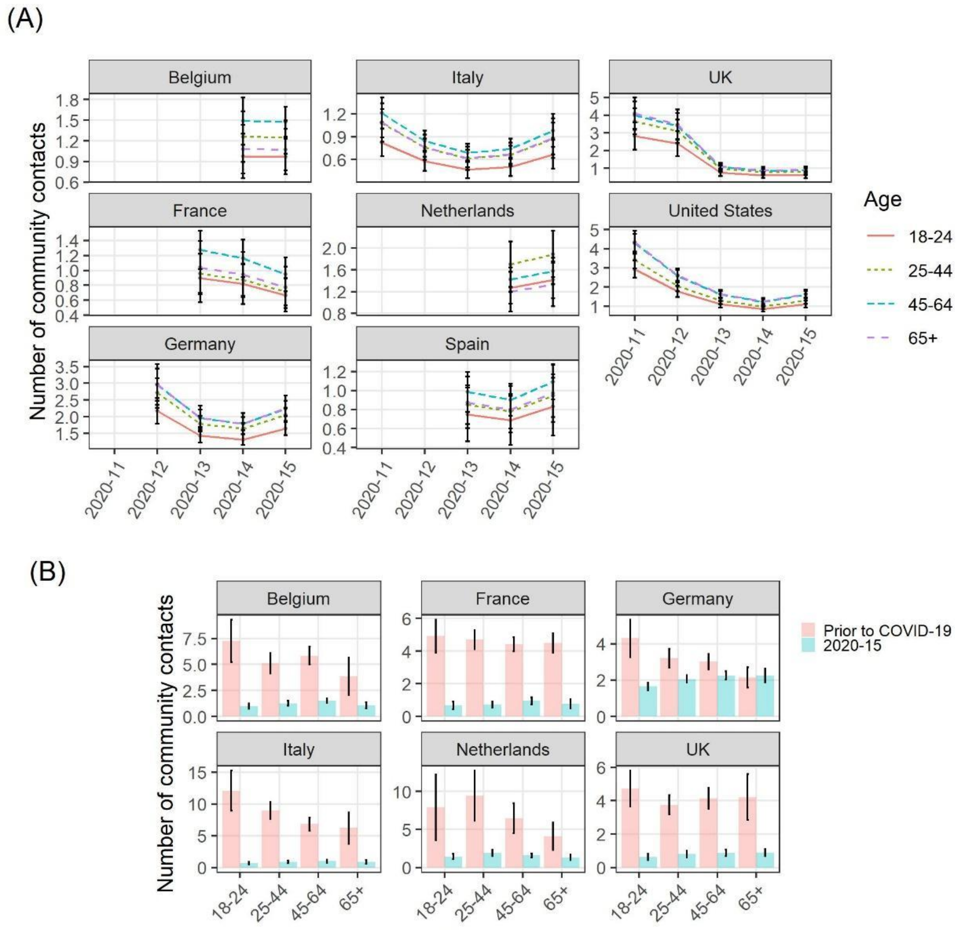


Fig. S16. Model-predicted daily number of social contacts in the general community, by age group, country, and week, March-April 2020. (A). Comparison of model-predicted contact numbers in the general community between the pre-COVID period and calendar week 2020-15, by country (B). Respondents with contacts above the ≤29 threshold were removed from the analysis.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

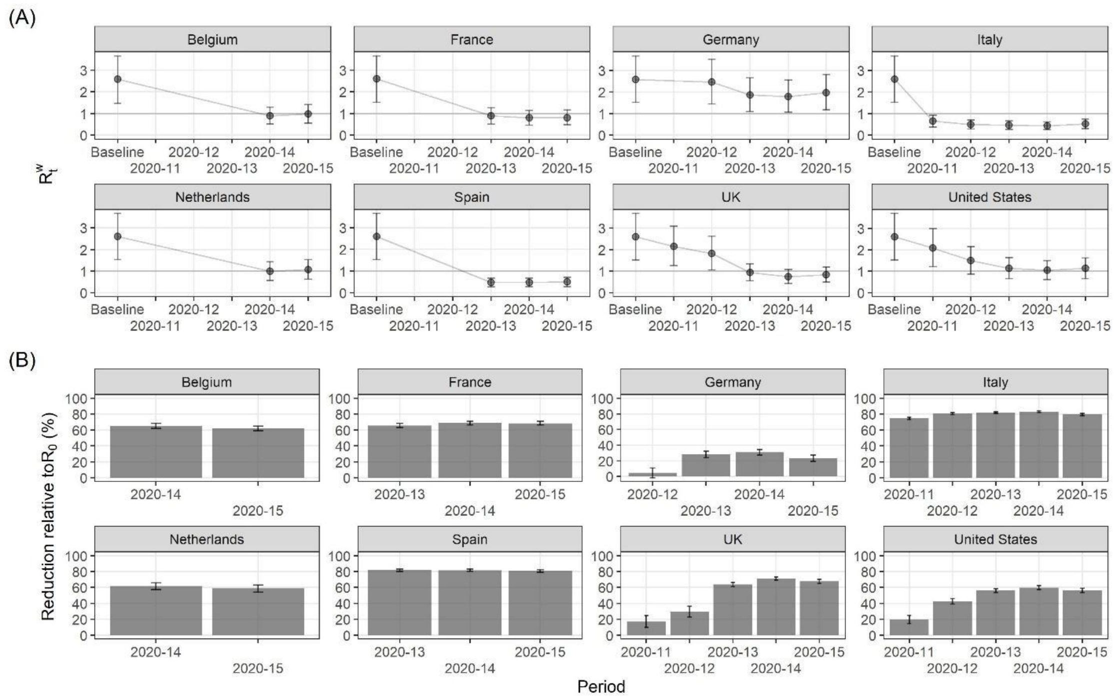


Fig. S17. Absolute change in the weekly net reproduction number R_t^w with respect to the basic reproduction number R_0 at baseline (A) and percent reduction with respect to R_0 (B), by country and calendar week. The 95% CIs are based on 5,000 replicates. Respondents with contacts above the ≤29 threshold were removed from the analysis.

Additional tables

Comparison between the raw and the weighted samples

In Tables S1-S8, for each country, we describe participants in terms of the variables used for the stratification and the construction of the Facebook advertising campaigns, comparing the unweighted sample to the weighted sample, and to the overall population.

We found that, in all countries, women were overrepresented in the sample Moreover, respondents were more uniformly spread across regions than in the real population. Finally, no trend in terms of age was found when comparing the raw and the adjusted sample distributions. We also noted that the unweighted sample size for the 65+ age group (who are at higher risk of death from COVID-19 [21], and who usually have lower social media participation rates [22]) was fairly large, with the Netherlands, UK, and the US showing higher percentages than in the population.

Table S1. Characteristics of survey participants in Belgium (N=3504), by region of residence, age group, and sex. We compare the raw sample ("unweighted") to the sample adjusted with the post-stratification weights ("weighted"), and to the distribution of each stratum in the population.

Variable	**Category**	**Unweighted**		**Weighted**		**Population**
		N	**%**	**N**	**%**	**%**
Region	**Brussels**	605	*17.3%*	359	*10.3%*	*10.3%*
	Flanders	1329	*37.9%*	2033	*58.0%*	*58.0%*
	Wallonia	1570	*44.8%*	1112	*31.7%*	*31.7%*
Age group	**18-24**	623	*17.8%*	479	*13.7%*	*13.7%*
	25-44	1075	*30.7%*	1092	*31.2%*	*31.2%*
	45-64	1172	*33.5%*	1135	*32.4%*	*32.4%*
	65+	634	*18.1%*	797	*22.8%*	*22.8%*
Sex	**Female**	2395	*68.4%*	1791	*51.1%*	*51.1%*
	Male	1109	*31.7%*	1713	*48.9%*	*48.9%*

Table S2. Characteristics of survey participants in Germany (N=8135), by region of residence, age group, and sex. We compare the raw sample ("unweighted") to the sample adjusted with the post-stratification weights ("weighted"), and to the distribution of each stratum in the population.

Variable	Category	Unweighted		Weighted		Population
		N	%	N	%	%
Region*	**East**	1996	*24.5%*	1436	*17.7%*	*17.7%*
	North	1871	*23.0%*	1472	*18.1%*	*18.1%*
	South	36	*25.0%*	2362	*29.0%*	
	West	2232	*27.4%*	2865	*35.2%*	*35.2%*
Age group	**18-24**	1764	*21.7%*	977	*12.0%*	*12.0%*
	25-44	3142	*38.6%*	2346	*28.8%*	*28.8%*
	45-64	2144	*26.4%*	2785	*34.2%*	*34.2%*
	65+	1085	*13.3%*	2028	*24.9%*	*24.9%*
Sex	**Female**	5149	*63.3%*	4147	*51.0%*	*51.0%*
	Male	2986	*36.7%*	3988	*49.0%*	*49.0%*

* The "East" region includes Berlin, Brandenburg, Sachsen, Sachsen-Anhalt, and Thüringen; the "North" regions includes Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen, and Schleswig-Holstein; the "South" region includes Baden-Württemberg and Bayern; the "West" region includes Hessen, Nordrhein-Westfalen, Rheinland-Pfalz, and Saarland.

Table S3. Characteristics of survey participants in Spain (N=4965), by region of residence, age group, and sex. We compare the raw sample ("unweighted") to the sample adjusted with the post-stratification weights ("weighted"), and to the distribution of each stratum in the population.

Variable	Category	Unweighted		Weighted		Population
		N	%	N	%	%
Region*	**Central**	608	*12.3%*	592	*11.9%*	*11.9%*
	East	1166	*23.5%*	1444	*29.1%*	*29.1%*
	Insular	387	*7.8%*	238	*4.8%*	*4.8%*
	Madrid	774	*15.6%*	697	*14.0%*	*14.0%*
	Northeast	569	*11.5%*	474	*9.6%*	*9.6%*
	Northwest	690	*13.9%*	471	*9.5%*	*9.5%*
	South	771	*15.5%*	1049	*21.1%*	*21.1%*
Age group	**18-24**	392	*7.9%*	573	*11.6%*	*11.6%*
	25-44	1858	*37.4%*	1582	*31.9%*	*31.9%*
	45-64	2016	*40.6%*	1680	*33.8%*	*33.8%*
	65+	699	*14.1%*	1130	*22.8%*	*22.8%*
Sex	**Female**	3490	*70.3%*	2553	*51.4%*	*51.4%*
	Male	1475	*29.7%*	2412	*48.6%*	*48.6%*

* The region "Central" includes Castilla y León, Castilla-la Mancha, and Extremadura; the region "East" includes Cataluña, Comunidad Valenciana, and Illes Balears; the region "Insular" includes Canarias; the region "Northeast" includes País Vasco, Comunidad Foral de Navarra, La Rioja, and Aragón; the region "Northwest" includes Galicia, Principado de Asturias, and Cantabria; the region "South" includes Andalucía, Región de Murcia, Ciudad Autónoma de Ceuta, and Ciudad Autónoma de Melilla.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

Table S4. Characteristics of survey participants in France (N=4637), by region of residence, age group, and sex. We compare the raw sample ("unweighted") to the sample adjusted with the post-stratification weights ("weighted"), and to the distribution of each stratum in the population.

Variable	Category	Unweighted		Weighted		Population
		N	%	N	%	%
Region	**Île de France**	678	*14.6%*	862	*18.6%*	*18.5%*
	Northeast	971	*20.9%*	1027	*22.2%*	*22.0%*
	Southeast	978	*21.1%*	941	*20.3%*	*20.2%*
	Southwest	994	*21.4%*	871	*18.8%*	*18.7%*
	West	1016	*21.9%*	936	*20.2%*	*20.1%*
Age group	**18-24**	818	*17.6%*	660	*14.2%*	*14.2%*
	25-44	1448	*31.2%*	1365	*29.4%*	*29.4%*
	45-64	1576	*34.0%*	1467	*31.6%*	*31.6%*
	65	795	*17.1%*	1146	*24.7%*	*24.7%*
Sex	**Female**	3300	*71.2%*	2422	*52.2%*	*52.2%*
	Male	1337	*28.8%*	2215	*47.8%*	*47.8%*

* The region "Northeast" includes Bourgogne - Franche-Comté, Nord-Pas-de-Calais – Picardie, and Alsace - Champagne-Ardenne – Lorraine; the region "Southeast" includes Auvergne - Rhône-Alpes, Provence-Alpes-Côte d'Azur, and Corse; the region "Southwest" includes Aquitaine - Limousin - Poitou-Charentes, and Languedoc-Roussillon - Midi-Pyrénées; the region "West" includes Centre - Val de Loire, Normandie, Pays-de-la-Loire, and Bretagne.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

Table S5. Characteristics of survey participants in Italy (N=7532), by region of residence, age group, and sex. We compare the raw sample ("unweighted") to the sample adjusted with the post-stratification weights ("weighted"), and to the distribution of each stratum in the population.

Variable	Category	Unweighted		Weighted		Population
		N	%	N	%	%
Region*	**Central**	1605	*21.3%*	1504	*20.0%*	*20.0%*
	Insular	795	*10.6%*	829	*11.0%*	*11.0%*
	Northeast	1885	*25.0%*	1453	*19.3%*	*19.3%*
	Northwest	2257	*30.0%*	2011	*26.7%*	*26.7%*
	South	990	*13.1%*	1735	*23.0%*	*23.0%*
Age group	**18-24**	1550	*20.6%*	846	*11.2%*	*11.2%*
	25-44	3083	*40.9%*	2114	*28.1%*	*28.1%*
	45-64	2034	*27.0%*	2591	*34.4%*	*34.4%*
	65+	865	*11.5%*	1981	*26.3%*	*26.3%*
Sex	**Female**	4905	*65.1%*	3897	*51.7%*	*51.7%*
	Male	2627	*34.9%*	3635	*48.3%*	*48.3%*

* The region "Central" includes Toscana, Umbria, Marche, and Lazio; the region "Insular" includes Sicilia and Sardegna; the region "Northeast" includes Provincia Autonoma di Bolzano/Bozen, Provincia Autonoma di Trento, Veneto, Friuli-Venezia Giulia, and Emilia-Romagna; the region "Northwest" includes Piemonte, Valle d'Aosta/Vallée d'Aoste, Liguria, and Lombardia; the region "South" includes Abruzzo, Molise, Campania, Puglia, Basilicata, and Calabria.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

Table S6. Characteristics of survey participants in the Netherlands (N=3172), by region of residence, age group, and sex. We compare the raw sample ("unweighted") to the sample adjusted with the post-stratification weights ("weighted"), and to the distribution of each stratum in the population.

Variable	Category	Unweighted		Weighted		Population
		N	%	N	%	%
Region*	**East**	793	*25.0%*	664	*20.9%*	*20.9%*
	North	654	*20.6%*	319	*10.1%*	*10.1%*
	South	802	*25.3%*	682	*21.5%*	*21.5%*
	West	923	*29.1%*	1506	*47.5%*	*47.5%*
Age group	**18-24**	447	*14.1%*	465	*14.7%*	*14.7%*
	25-44	793	*25.0%*	928	*29.3%*	*29.3%*
	45-64	1191	*37.6%*	1056	*33.3%*	*33.3%*
	65+	741	*23.4%*	723	*22.8%*	*22.8%*
Sex	**Female**	2134	*67.3%*	1606	*50.6%*	*50.6%*
	Male	1038	*32.7%*	1566	*49.4%*	*49.4%*

* The region "East" includes Overijssel, Gelderland, and Flevoland; the region "North" includes Groningen, Friesland, and Drenthe; the region "South" includes Noord-Brabant and Limburg; the region "West" includes Utrecht, Noord-Holland, Zuid-Holland, and Zeeland.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

Table S7. Characteristics of survey participants in the United Kingdom (N=7371), by region of residence, age group, and sex. We compare the raw sample ("unweighted") to the sample adjusted with the post-stratification weights ("weighted"), and to the distribution of each stratum in the population.

Variable	Category	Unweighted		Weighted		Population
		N	%	N	%	%
Region	**England**	3381	*45.9%*	5225	*70.9%*	*70.9%*
	London	763	*10.4%*	974	*13.2%*	*13.2%*
	Northern Ireland	547	*7.4%*	203	*2.8%*	*2.8%*
	Scotland	1548	*21.0%*	618	*8.4%*	*8.4%*
	Wales	1132	*15.4%*	351	*4.8%*	*4.8%*
Age group	**18-24**	577	*7.8%*	1056	*14.3%*	*14.3%*
	25-44	1831	*24.8%*	2346	*31.8%*	*31.8%*
	45-64	3044	*41.3%*	2316	*31.4%*	*31.4%*
	65+	1919	*26.0%*	1653	*22.4%*	*22.4%*
Sex	**Female**	4789	*65.0%*	3762	*51.0%*	*51.0%*
	Male	2582	*35.0%*	3609	*49.0%*	*49.0%*

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

Table S8. Characteristics of survey participants in the United States (N=11917), by region of residence, age group, and sex. We compare the raw sample ("unweighted") to the sample adjusted with the post-stratification weights ("weighted"), and to the distribution of each stratum in the population.

Variable	Category	Unweighted		Weighted		Population
		N	%	N	%	%
Region*	**Midwest**	3206	*26.9%*	2490	*20.9%*	*20.9%*
	Northeast	2673	*22.4%*	2088	*17.5%*	*17.5%*
	South	2950	*24.8%*	4507	*37.8%*	*37.8%*
	West	3088	*25.9%*	2832	*23.8%*	*23.8%*
Age group	**18-24**	1308	*11.0%*	1923	*16.1%*	*16.1%*
	25-44	3227	*27.1%*	3889	*32.6%*	*32.6%*
	45-64	4075	*34.2%*	3757	*31.5%*	*31.5%*
	65+	3307	*27.8%*	2348	*19.7%*	*19.7%*
Sex	**Female**	7900	*66.3%*	6098	*51.2%*	*51.2%*
	Male	4017	*33.7%*	5819	*48.8%*	*48.8%*

* The region "Midwest" includes Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; the region "Northeast" includes Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont; the region "South" includes Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia; the region "West" includes Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

Effect for the contact threshold values

Tables S9-S16 compare the contact threshold scenario presented in the manuscript, where individuals who reported more contacts than the 90% quantile of the contact distribution per setting were removed, with the scenario considered in the sensitivity analysis, where individuals who reported 30 or more contacts per setting were removed. For each scenario, we show the minimum, the median, and the first and the third quartile, and the maximum number of contacts for each setting and overall, as well as the number of missing values (either because the participant did not report it or because he or she reported a number of contacts above the considered threshold).

Under the 90% quantile scenario used in the main analysis, the number of missing values (due either to respondents who did not provide any contact number or who reported a number of contacts above the threshold) was much higher than under the ≤ 29 scenario. Even though these thresholds might lead to different average contacts numbers, which is a metric not robust to outliers, the effect on the median and the quartiles was much smaller, with differences between the two outliers' scenarios mainly found in Germany, the Netherlands, and the United States for the home contacts (and hence for the overall contacts).

Table S9. Contact distribution by threshold in each setting and overall, for Belgium (N=3704).

Threshold	Statistic	Overall	Home	School	Work	General community
90% quantile	**Minimum**	0	0	0	0	0
	First quartile	1	0	0	0	0
	Median	2	1	0	0	0
	Third quartile	3	2	0	0	1
	Maximum	14	4	0	8	5
	No. missing values	625	162	52	304	195
≤29	**Minimum**	0	0	0	0	0
	First quartile	1	0	0	0	0
	Median	2	1	0	0	0
	Third quartile	4	2	0	0	1
	Maximum	45	13	17	26	29
	No. missing values	144	1	2	127	23

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

Table S10. Contact distribution by threshold in each setting and overall, for Germany (N=8135).

Threshold	Statistic	Overall	Home	School	Work	General community
90% quantile	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	3	1	0	0	0
	Third quartile	5	2	0	0	2
	Maximum	16	4	0	8	5
	No. Missing values	2124	490	92	1154	867
≤29	**Minimum**	0	0	0	0	0
	First quartile	2	1	0	0	0
	Median	4	1	0	0	1
	Third quartile	7	3	0	1	2
	Maximum	61	29	24	28	27
	No. Missing values	527	14	8	402	167

Table S11. Contact distribution by threshold in each setting and overall, for Spain (N=4965).

Threshold	Statistic	Overall	Home	School	Work	General community
90% quantile	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	2	1	0	0	0
	Third quartile	3	3	0	0	0
	Maximum	14	4	0	8	5
	No. Missing values	717	193	71	333	229
≤29	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	2	2	0	0	0
	Third quartile	4	3	0	0	1
	Maximum	76	28	25	28	25
	No. Missing values	188	12	6	139	52

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

Table S12. Contact distribution by threshold in each setting and overall, for France (N=4637).

Threshold	Statistic	Overall	Home	School	Work	General community
90% quantile	**Minimum**	0	0	0	0	0
	First quartile	1	0	0	0	0
	Median	2	1	0	0	0
	Third quartile	3	2	0	0	0
	Maximum	15	4	0	8	5
	No. Missing values	706	159	63	350	246
≤29	**Minimum**	0	0	0	0	0
	First quartile	1	0	0	0	0
	Median	2	1	0	0	0
	Third quartile	4	2	0	0	1
	Maximum	54	18	25	28	25
	No. Missing values	214	4	7	169	60

Table S13. Contact distribution by threshold in each setting and overall, for Italy (N=7532).

Threshold	Statistic	Overall	Home	School	Work	General community
90% quantile	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	2	2	0	0	0
	Third quartile	3	3	0	0	0
	Maximum	16	4	0	8	5
	No. Missing values	1034	380	82	474	213
≤29	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	2	2	0	0	0
	Third quartile	4	3	0	0	1
	Maximum	42	22	22	25	23
	No. Missing values	194	12	5	167	25

Table S14. Contact distribution by threshold in each setting and overall, for the Netherlands (N=3172).

Threshold	Statistic	Overall	Home	School	Work	General community
90% quantile	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	2	1	0	0	0
	Third quartile	4	2	0	0	2
	Maximum	17	4	0	8	5
	No. Missing values	718	204	43	398	226
≤29	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	3	1	0	0	0
	Third quartile	7	3	0	0	2
	Maximum	46	20	18	29	25
	No. Missing values	135	8	2	106	28

Table S15. Contact distribution by threshold in each setting and overall, for the UK (N=7371).

Threshold	Statistic	Overall	Home	School	Work	General community
90% quantile	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	2	1	0	0	0
	Third quartile	3	2	0	0	1
	Maximum	15	4	0	8	5
	No. Missing values	1633	426	175	773	713
≤29	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	2	1	0	0	0
	Third quartile	5	3	0	0	2
	Maximum	79	25	25	29	29
	No. Missing values	478	16	34	353	170

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

Table S16. Contact distribution by threshold in each setting and overall, for the US (N=11,917).

Threshold	Statistic	Overall	Home	School	Work	General community
90% quantile	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	2	1	0	0	0
	Third quartile	4	2	0	0	1
	Maximum	17	4	0	8	5
	No. Missing values	3131	1036	197	1330	1340
≤29	**Minimum**	0	0	0	0	0
	First quartile	1	1	0	0	0
	Median	3	1	0	0	0
	Third quartile	6	3	0	0	2
	Maximum	60	28	22	29	28
	No. Missing values	819	27	47	561	306

Table S17. Model-predicted mean number (with standard error (SE)) of daily contacts per person compared with pre-pandemic model predictions, by country, setting, and week, March-April 2020. The "Overall" category encompasses contacts reported in all four surveyed settings, e.g., home, school, work, and general community. Respondents with contacts above the ≤29 threshold were removed from the analysis.

Country	Setting	Prior to COVID-19		Week 2020-11 (Mar 9 – 15)		Week 2020-12 (Mar 16 – 22)		Week 2020-13 (Mar 23 – 29)		Week 2020-14 (Mar 30 – Apr 5)		Week 2020-15 (Apr 6 – 12)	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Belgium	Home	3.15	0.15							1.49	0.05	1.55	0.03
	Community	5.34	0.3							1.25	0.14	1.24	0.06
	Work	3.05	0.36							1.07	0.19	1.36	0.13
	Total	11.78	0.48							3.77	0.23	4.11	0.13
France	Home	2.7	0.07					1.49	0.08	1.49	0.06	1.48	0.08
	Community	4.59	0.13					1.07	0.1	0.97	0.08	0.8	0.08
	Work	2.37	0.12					0.87	0.15	0.82	0.13	0.9	0.12
	Total	10.3	0.21					3.37	0.19	3.08	0.12	3.08	0.15
Germany	Home	2.73	0.1			1.93	0.08	1.65	0.04	1.74	0.04	1.83	0.05
	Community	2.97	0.14			2.78	0.21	1.84	0.1	1.68	0.07	2.11	0.08
	Work	1.82	0.17			2.54	0.31	1.67	0.16	1.88	0.14	1.71	0.12
	Total	8.03	0.27			6.83	0.32	5.16	0.18	4.96	0.12	5.49	0.14
Italy	Home	3.56	0.19	1.9	0.05	1.84	0.05	1.74	0.04	1.7	0.04	1.82	0.06
	Community	7.91	0.36	1.1	0.08	0.77	0.05	0.63	0.05	0.67	0.05	0.89	0.1
	Work	5.28	0.53	1.29	0.15	0.72	0.08	0.74	0.16	0.48	0.06	0.88	0.17
	Total	18.19	0.61	4.35	0.15	3.32	0.11	3.08	0.1	2.91	0.1	3.5	0.19
Netherlands	Home	3	0.23							1.92	0.06	1.97	0.07
	Community	7.13	0.76							1.43	0.1	1.59	0.1
	Work	4.97	1.08							2.23	0.2	2.02	0.17
	Total	15.03	1.2							5.11	0.18	5.5	0.21
Spain	Home							1.8	0.05	1.91	0.06	1.84	0.05
	Community							0.89	0.08	0.82	0.07	0.99	0.08
	Work							0.68	0.11	0.7	0.09	0.97	0.14
	Total							3.44	0.15	3.45	0.13	3.64	0.14
United Kingdom	Home	3.54	0.14	2.25	0.09	2.21	0.11	1.76	0.04	1.72	0.05	1.76	0.05
	Community	4.07	0.21	3.73	0.31	3.16	0.3	1	0.06	0.8	0.06	0.81	0.09

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

	Work	2.89	0.39	3.76	0.9	2.47	0.37	1.17	0.11	0.63	0.1	0.71	0.12
	Total	10.58	0.35	8.77	0.47	7.46	0.46	3.85	0.13	3.07	0.11	3.41	0.16
United States	Home			2.29	0.08	2.17	0.05	2.1	0.04	1.97	0.03	2.1	0.04
	Community			3.77	0.19	2.27	0.12	1.42	0.07	1.09	0.05	1.42	0.08
	Work			2.36	0.26	2.29	0.27	1.39	0.11	1.45	0.16	1.39	0.15
	Total			8.81	0.31	6.3	0.2	4.81	0.13	4.42	0.13	4.79	0.14

Del Fava E, *et al. BMJ Open* 2022; 11:e050651. doi: 10.1136/bmjopen-2021-050651

References

1 Perrotta D, Grow A, Rampazzo F, *et al.* Behaviours and attitudes in response to the COVID-19 pandemic: insights from a cross-national Facebook survey. *EPJ Data Sci* 2021;**10**:1–13. doi:10.1140/epjds/s13688-021-00270-1

2 Grow A, Perrotta D, Del Fava E, *et al.* Addressing Public Health Emergencies via Facebook Surveys: Advantages, Challenges, and Practical Considerations. *J Med Internet Res* 2020;**22**:e20653. doi:10.2196/20653

3 Eurostat. *Eurostat regional yearbook: 2019 edition.* 2019. http://publications.europa.eu/publication/manifestation_identifier/PUB_KSHA19001ENN (accessed 15 May 2020).

4 Ipsos. Public Opinion on the Coronavirus outbreak: a multi-country poll from Ipsos. 2020.https://www.ipsos.com/en/coronavirus-public-opinion (accessed 15 May 2020).

5 Pötzschke S, Braun M. Migrant Sampling Using Facebook Advertisements: A Case Study of Polish Migrants in Four European Countries. *Soc Sci Comput Rev* 2016;**35**:633–53. doi:10.1177/0894439316666262

6 Hoffman Pham K, Rampazzo F, Rosenzweig LR. Online Surveys and Digital Demography in the Developing World: Facebook Users in Kenya. *ArXiv191003448 Cs* Published Online First: 8 October 2019.http://arxiv.org/abs/1910.03448 (accessed 1 Dec 2019).

7 Zhang B, Mildenberger M, Howe PD, *et al.* Quota sampling using Facebook advertisements. *Polit Sci Res Methods* 2018;:1–7. doi:10.1017/psrm.2018.49

8 Keiding N, Louis TA. Perils and potentials of self-selected entry to epidemiological studies and surveys. *J R Stat Soc Ser A Stat Soc* 2016;**179**:319–76. doi:10.1111/rssa.12136

9 Mercer AW, Kreuter F, Keeter S, *et al.* Theory and Practice in Nonprobability Surveys Parallels between Causal Inference and Survey Inference. *Public Opin Q* 2017;**81**:250–71. doi:10.1093/poq/nfw060

10 US Census Bureau. 2018 Population Estimates by Age, Sex, Race and Hispanic Origin. U. S. Census Bur. https://www.census.gov/newsroom/press-kits/2019/detailed-estimates.html (accessed 15 May 2020).

11 Mossong J, Hens N, Jit M, *et al.* Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLoS Med* 2008;**5**:e74. doi:10.1371/journal.pmed.0050074

12 Béraud G, Kazmercziak S, Beutels P, *et al.* The French Connection: The First Large Population-Based Contact Survey in France Relevant for the Spread of Infectious Diseases. *PLoS ONE* 2015;**10**:e0133203. doi:10.1371/journal.pone.0133203

13 Teslya A, Pham TM, Godijk NG, *et al.* Impact of self-imposed prevention measures and short-term government-imposed social distancing on mitigating and delaying a COVID-19 epidemic: A modelling study. *PLoS Med* 2020;**17**:e1003166. doi:10.1371/journal.pmed.1003166

14 Chu DK, Akl EA, Duda S, *et al.* Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The Lancet* 2020;**395**:1973–87. doi:10.1016/S0140-6736(20)31142-9

15 Diekmann O, Heesterbeek JAP, Metz JAJ. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *J Math Biol* 1990;**28**:365–82. doi:10.1007/BF00178324

16 Wallinga J, Teunis P, Kretzschmar M. Using Data on Social Contacts to Estimate Age-specific Transmission Parameters for Respiratory-spread Infectious Agents. *Am J Epidemiol* 2006;**164**:936–44. doi:10.1093/aje/kwj317

17 Coletti P, Wambua J, Gimma A, *et al.* CoMix: comparing mixing patterns in the Belgian population during and after lockdown. *Sci Rep* 2020;**10**:21885. doi:10.1038/s41598-020-78540-7

18 Feehan DM, Mahmud AS. Quantifying population contact patterns in the United States during the COVID-19 pandemic. *Nat Commun* 2021;**12**:893. doi:10.1038/s41467-021-20990-2

19 Jarvis CI, Van Zandvoort K, Gimma A, *et al.* Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Med* 2020;**18**:124. doi:10.1186/s12916-020-01597-8

20 Arregui S, Aleta A, Sanz J, *et al.* Projecting social contact matrices to different demographic structures. *PLoS Comput Biol* 2018;**14**:e1006638. doi:10.1371/journal.pcbi.1006638

21 WHO. Transmission of SARS-CoV-2: implications for infection prevention precautions. 2020.https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions (accessed 24 Aug 2020).

22 Quinn K. Social media and social wellbeing in later life. *Ageing Soc* 2019;:1–22. doi:10.1017/S0144686X19001570