European Economic Review 141 (2022) 103992

Contents lists available at ScienceDirect

European Economic Review

journal homepage: www.elsevier.com/locate/eer

# Religion, social interactions, and COVID-19 incidence in Western Germany

Ioannis Laliotis [a,b], Dimitrios Minos [c,*]

[a] University of Peloponnese, Department of Economics, Greece
[b] City University of London, Department of Economics, United Kingdom
[c] King's College London, King's Business School, United Kingdom

ARTICLE INFO

*JEL codes:*
H12
I10
Z10

*Keywords:*
COVID-19
Religion
Mortality

ABSTRACT

This paper investigates how social interactions, as shaped by religious denomination, are related to COVID-19 incidence and associated mortality in Western Germany. We observe that the number of infections and deaths during the early pandemic phase were much higher in predominantly Catholic counties with arguably stronger family and social ties. The relationship was confirmed at the county level through numerous robustness checks, and after controlling for a series of characteristics and county fixed effects. At the individual level, we confirmed that Catholics, relative to non-Catholics, have tighter and more frequent interactions with their family and friends. Moreover, the intensity of social interaction was able to partially explain the relationship between COVID-19 incidence and share of Catholics at the county level. Our results highlight the number of dimensions that have to be taken into account when designing and implementing mitigation measures in the early stages of disease outbreaks.

## 1. Introduction

Since the onset of the COVID-19 pandemic, millions of cases and related fatalities have been recorded across the globe. In Europe, Italy and Spain were the first countries to be hit the hardest in early 2020, with France following suit. Initial containment measures were relatively homogeneous across countries, e.g. school closures, local and nation-wide lockdowns, social distancing etc. However, the numbers of cases and related deaths varied substantially by country. The reasons are difficult to uncover. The timing of both the pandemic onset and the various policies to constrain it clearly played a role in March 2020; for example, in Sweden, Belgium or the UK. However, evidence suggests that the virus was present in Europe long before it was initially believed, e.g. as early as December 2019 in France (Deslandes et al., 2020). Yet, in the first phase some countries were more severely affected compared to others, resulting in vastly overburdened health systems and excess mortality rates, before most governments were able to enact their mitigation policies.

Variation in social preferences could partly explain differences in population health outcomes between countries. Such preferences are crucial in determining the degree of social distancing and social interactions and they are largely culturally inherent (Remland et al., 1995; Sorokowska et al., 2017). Especially when vulnerable groups such as the elderly are concerned, lower levels of social interactions have been linked to a lower incidence of the virus in some countries during the early pandemic phase (Bayer and Kuhn, 2020). The early COVID-19 literature promptly highlighted that, for as long as a medical treatment was not widely available, social capital and human behaviour are important in keeping the number of cases low (Bartscher et al., 2021; Borgonovi and Andrieu, 2020; Ding et al., 2020; Durante et al., 2021; Van Bavel et al., 2020). For example, Bargain and Aminjonov (2020) argue that compliance to

* Corresponding author.

https://doi.org/10.1016/j.euroecorev.2021.103992
Received 27 February 2021; Received in revised form 22 October 2021; Accepted 12 November 2021
Available online 30 November 2021

containment policies was higher in European regions where people trust their politicians more.

Recent empirical evidence demonstrated that variations in the number of COVID-19 cases and related fatalities are crucially driven by the type and frequency of interactions among people, especially in early pandemic stages (Alfaro et al., 2020; Liu et al., 2020; Platteau and Verardi, 2020). Dowd et al. (2020) argued that in countries like Italy and Spain, co-residence and intergenerational social interaction can partially explain variation in COVID-19 incidence. However, substantial differences in socioeconomics, demographics, health systems, testing and healthcare capacity, social distancing policies and various types of policy response at different time periods, and event data collection processes, render cross-country comparisons problematic (Georganas et al., 2021).[1] There is also evidence on how genetic differences affect the cross-country variation of the number of cases (Delanghe et al., 2020). Moreover, such cross-country differences did not allow for a coordinated virus mitigation and lockdown exit strategy at the EU level in the first phase (Platteau and Verardi, 2020). Therefore, empirical research shifted towards subnational analyses in order to identify how social structures, and differences in societal and individual behaviour and attitudes are linked to within-country variations in the number of COVID-19 infections and deaths. Bayer and Kuhn (2020) observed a country-level positive correlation between the number of cases and the share of people 30–49 years old living with their parents, arguing that social integration might be the factor behind the death toll in Italy. However, subsequent research by Belloc et al. (2020) using Italian regions as their units of observation revealed that the correlation has the opposite sign, suggesting policy recommendations based on cross-country results should be read with caution. In Europe, policy responses were rather similar across several countries at the onset of the pandemic, yet they resulted in varying degrees of efficiency. Even though European governments implemented nation-wide measures during the early phases of the pandemic, the number of infections and the associated death toll exhibited a great degree of variation at the local level within countries.

For these reasons, we focused on Western Germany where concerns around heterogeneity are mitigated (Spenkuch, 2017). Our main purpose is to examine whether the strength of social and family ties as determined by religious denomination played a role. More specifically, we focused on a religious divide within Western German counties, i.e. Catholics and non-Catholics. This setting has been widely used to explain how behaviour is affected by cultural differences related to religion (Iannaccone, 1998; Ekelund et al., 2002; Arruñada, 2009; Spenkuch, 2017; Becker and Woessmann, 2018; Spenkuch and Tillmann, 2017; Becker and Pascali, 2019).[2] Rather than discussing religion or Catholicism *per se*, we explore whether differences in specific behaviours within an otherwise very similar population could have impacted groups differently in terms of infections and related deaths. Our paper is the first one to rely on regional variation of religious group composition as a measure of cultural norms and preferences in this context. Hence, it contributes to the literature on how social preferences are linked to the virus incidence within a country. Some earlier attempts examined the implications of local variation in mitigation measures compliance. Specifically, two studies for Switzerland focused on the cultural and behavioural gradient behind the regional variation of COVID-19 cases. Proxying culture through language and using the twenty-six Swiss cantons as their units of analysis, Mazzonna, (2020) argued that in the early pandemic phase, elderly people living in Latin speaking cantons were more severely affected by the virus spread, relative to those living in German-speaking cantons. This was attributed to more frequent social contacts at that time, although people from all ages in Latin, relative to German, cantons strictly complied after the lockdown. This was in line with Brodeur et al. (2021), Durante et al. (2021), and Deopa and Fortunato (2020) who also reported greater mobility declines in areas with higher levels of civic capital and trust.[3]

In this paper we use religious denomination to proxy culture. Guiso et al. (2006) defined culture as those customary values and beliefs being unchangeably transmitted from generation to generation within ethnic, religious, and social groups. The geographic dispersion of Catholics in West Germany has deep historic origins and has been covered extensively (Becker and Woessmann, 2009; Spenkuch and Tillmann, 2018). Cultural and behavioural dimensions shaped by religion at the local level are inherited from previous generations, they can be considered lifetime invariant, and they may have manifested themselves as a local way of life that is resilient to changes (Becker, 1996; Guisoet al., 2006). Our paper is the first to contribute to this literature by shedding light on how cultural variations, that can be considered as exogenous and largely unclouded by unobserved heterogeneity, affect the within-country variation in the number of COVID-19 cases and associated deaths.

Using data from the initial stage of the pandemic, we observe that COVID-19 infection and related death rates were strikingly correlated with the local share of Catholics at the county level. The relationship remained strong after partialing out the effects of socioeconomic characteristics, county fixed effects, time trends, and after ruling out explanations related to geographical proximity to Northern Italy, differences in mobility patterns, and differences to pre-pandemic total mortality trends between groups using a difference-in-differences approach. Our OLS, instrumental variables and difference-in-differences estimates confirmed that Catholic counties were more affected by the virus, in terms of cases and associated fatalities during the early stages of the pandemic. Our argument behind this finding is that Catholics, compared to non-Catholics exhibit stronger family and social ties, triggering a possible

[1] Differences in the data collection process might arise from the fact that countries lie on different points of their epidemic curve, hence not applying the same COVID-19 testing and reporting standards (Belloc et al., 2020).

[2] We focused on Western Germany, as the Eastern part of the country historically exhibits systematic differences (Becker et al., 2020). Moreover, including it in our estimation samples could favourably distort the results, due to the very low number of COVID-19 cases at the pandemic onset and the low prevalence of Catholics in Eastern Germany.

[3] Although Deopa and Fortunato (2020) argue that mobility becomes less relevant for the virus spread due to increasing compliance to infection prevention and control norms.

mechanism behind the county-level relationship that we document.[4] Previous research in economics has established that Catholics are more bound to close social circles and networks, i.e. family and friends, and they exhibit different patterns of social interactions compared to non-Catholics (Arruñada, 2009; Glaeser and Glendon, 1998; Ekelund et al., 2002; Satyanath et al., 2017). To formally test our argument, we relied on individual-level information. First, we used data from the European Social Survey (ESS) and the European Values Study (EVS) to confirm that, relative to non-Catholics, Catholics exhibit systematic differences in social preferences and behaviours that are crucial determinants for the number of COVID-19 cases. Next, we showed that the correlation between the share of Catholics and the COVID-19 incidence at the county level, can be partially explained by the intensity of social interactions, as the latter was proxied by cohabitation with parents, closeness to family and friends, frequent religious-service attendance etc.

Our results could support epidemiologists and public health policy makers to better understand how cultural factors can influence the initial spread of pathogens in a society and incorporate such differences as well when designing their optimal response policies (Platteau and Verardi, 2020). The remainder of the paper is organised as follows: Section 2 offers a background regarding the early pandemic phase in Western German counties. Section 3 describes the data sources used in the empirical analysis. Section 4 presents some descriptive evidence and outlines the adopted empirical strategy. Section 5 discusses the results, and Section 6 concludes.

## 2. The early pandemic in Western Germany

According to data released from the Robert Koch Institute, the first reported COVID-19 infections in Western Germany were recorded on January 27, 2020, in Starnberg, Bavaria. By the end of March 2020, there were confirmed infections in all Western German states. The first deaths linked to the virus were reported on March 09, 2020, in North Rhine-Westphalia and by the end of the month hospitals in all Western German states confirmed virus-related fatalities. State and federal government response was swift (Stafford, 2020). In early March, the federal government and the Robert Koch Institute issued the National Pandemic Plan to be carried out across the country. The states (*Länder*) themselves were given some autonomy in handling the pandemic, although the response was coordinated and crucial to avoid inter-state travel. All states in Western Germany enacted strict social distancing measures and closures of schools, shops, and workplaces. Based on the Coronavirus Government Response Tracker developed by the University of Oxford, these social distancing measures were implemented on February 29 (public events cancellation), March 16 (school closings), March 22 (shops and workplaces closings) and April 09 (public transport closing).[5] According to the Robert Koch Institute, the number of daily cases peaked in early April and started receding from that point on. Some of the nationwide restrictions such as the closure of small shops were lifted three weeks later on April 22, 2020.

In terms of testing, widespread Polymerase Chain Reaction (PCR) testing was made available on March 25, 2020, and guidelines restricting testing only to severe cases were lifted, so that more than 2 million people received a test by late April. As of April 23, 2020, more than 148,000 confirmed cases and more than 5000 related deaths were reported, based on data from the Robert Koch Institute. Fig. A.1 displays how the cumulative numbers of reported cases and deaths per 100,000 population were scattered across the country at the end of April.

The early pandemic phase, spanning from January 28, 2020, to May 01, 2020, is the period we will focus on, as the initial stages of the pandemic largely determine the incidence of the virus across the country (Zhao et al., 2020). This is important, if we consider the relatively long incubation period, i.e. up to two weeks, and the large number of asymptomatic carriers, due to which cases may rise during the early stage before mitigating strategies become effective (Gudbjartsson et al., 2020; Lauer et al., 2020).

## 3. Data sources

Information on COVID-19 cases and related fatalities were drawn from the Robert Koch Institute (RKI) which is the federal government research institute responsible for disease control and prevention. They have been publishing data on confirmed COVID-19 cases and associated deaths since January 28, 2020. The data are broken down by day, gender, age group and county, for a total of 401 counties (NUTS-3 regions; *Landkreise*). Information on cases and deaths is collected by counties, reported to the Federal Ministry of Health and published by the RKI after validation. Using these data, we constructed a balanced panel of Western German counties running from January 28, 2020, to May 01, 2020, i.e. more than 32,000 observations.

Demographic and socioeconomic information at the county level, validated until 2017, was taken from the Federal Statistical Office of Germany (*Statistisches Bundesamt)*. These include population, population density, share of people over 65 years old, share of foreigners, share of males and females, share of those who completed secondary education, GDP per capita, number of hospital beds, and number of nights spend per person as an indicator for tourism-related activity. Moreover, we used Google Maps data in order to calculate the distance between the capital of each county and Wittenberg, Germany, and the fastest highway driving distance between the capital of each county and Milan, Italy. We also used the 2011 German Census in order to calculate the number of Catholics, Protestants and other/no denominations in each county and state. Information on overall mortality at the state and at the county level

[4] The comparison group mainly consists of Protestants and non-religious. In terms of social interactions Spenkuch (2017) argues that non-religious exhibit more individualistic behaviours and seem closer to Protestants. Moreover, they argue that individuals that opt out of religion are twice as likely to stem from Protestant families.

[5] All 10 Western German states agreed to close schools in March. Despite their autonomy, only a couple of states acted slightly differently. The district of Heinsberg closed for a week in late February and Bavaria announced a curfew six days before all other states in Western Germany. We take those actions into consideration in our estimations.

was taken from the Federal Statistical Office, and monthly weather data from the Climate Data Centre (CDC) of the German Meteorological Service (*Deutscher Wetterdienst*).[6] The share of Catholics relative to the total population in the county will be our main variable of interest. A county is classified as a predominantly Catholic one if the share of Catholics is higher than the respective state (NUTS-1 region; *Länder*) average.[7] Appendix Table A.1 presents some basic descriptive statistics on county-level COVID-19 incidence and various demographic and socioeconomic variables. The share of Catholics in the population of a predominantly Catholic county is more than twice their share in a non-Catholic county, on average. The number of COVID-19 cases and related fatalities per 100,000 county population was considerably higher in Catholic counties right before the lockdown date. With respect to the other county-level characteristics considered, both groups look quite similar.[8]

Data at the individual level were taken from the 2018 European Social Survey (ESS) and the 2017 European Values Study (EVS). These are two representative datasets of approximately 1900 and 4300 adults in Western Germany, respectively. Their questionnaires cover a wide range of issues related to values, preferences, behaviour and social norms. We used both surveys to construct variables indicating family and social ties at the individual level. These include the frequency of social interactions, the perceived frequency compared to individuals of the same age, the number of people someone is comfortable confiding in, the frequency of attending religious services, the importance of family, the importance of friends, the level of trust towards family members and whether the respondents reside with their parents and/or parents-in-law. Further, we used information on individual and household characteristics. These include the age of the respondent, gender, employment status, education level, subjective health status, age of the youngest household member (EVS only), household size, household income, size of the city or town (EVS only) and state fixed effects. Appendix Table A.2 shows some basic descriptive statistics on ESS and EVS variables, in Panels A and B, respectively. In both datasets, Catholic and non-Catholic individuals seem quite similar in terms of standard demographics. With respect to the social interaction variables in the ESS data, Catholics seem more likely to meet other people more than several times per week relative to non-Catholics, and they are also more likely to meet other people more frequently compared to individuals of the same age. Based on the EVS data, Catholics consider family and friends to be more important in their lives, they attend religious services more frequently, they consider religion to be more important, and they are more likely to believe that children should provide care for their parents, relative to non-Catholics. Also, youngest individuals within Catholic households are older compared to those within non-Catholic households, on average.

Finally, to investigate for differences in mobility by different means of transportation, we used Apple mobility data. They are obtained through GPS tracking and they are available after January 13, 2020, and relative to that date, at the city level and for eighteen major cities in total. They provide information regarding requests for directions by transport type, i.e. walking, driving and transit and they are useful in comparing mobility patterns between cities.

## 4. Empirical strategy

### *4.1. Descriptive evidence, overall mortality trends and mobility*

Appendix Fig. A.1 depicts our empirical observation. Western German counties with a higher share of Catholics in their total population appear more severely affected by the virus. There is a striking similarity between the prevalence of the disease in the early pandemic and the dispersion of Catholics across the country. On the sample of Western German counties, the correlation coefficient between the Catholic share over the total county population and the cumulative COVID-19 cases and deaths per 100,000 inhabitants was 0.42 and 0.32, respectively. This difference in COVID-19 incidence between Catholic counties and non-Catholic counties is also shown in Appendix Table A.1. Apart from this pattern, demographic and economic characteristics are balanced between counties. A comparison of some county-level characteristics, suggests that Western German counties are rather homogeneous with respect to a series of characteristics, other than religion.[9]

Despite the observed similarity in terms of demographic and socioeconomic characteristics, we inspect for differences in overall monthly mortality rates before the pandemic (Appendix Fig. A.2). This is to ensure that the observed pandemic-related mortality patterns in Western German counties are not picking up pre-existing mortality differences between Catholic and non-Catholic counties. The time series run from January 2011 to December 2017. They begin considerably after the last major swine flu outbreak in 2009, with no other major public health threats that could be affected from social contacts occurring since then. Both series are quite close

[6] The data are available for over 400 weather stations in Germany. We excluded those in Eastern Germany, as well as those located on mountain ranges in counties with more than one weather stations. This resulted in weather data available for 136 counties in Western Germany.

[7] This is the main definition of Catholic counties we adopted throughout this paper. However, results are robust across alternative definitions of a Catholic county e.g. Catholics are the majority, Catholics are more than Protestants, the share of Catholics in a county is higher than the Western Germany average.

[8] There are some differences in a couple of observable characteristics, however, we account for those when estimating our models. Specifically, compared to Catholic counties, non-Catholic ones have a higher share of people completed secondary education and they are more densely populated.

[9] Some differences are statistically significant, but small. For example, there are no notable differences in characteristics like GDP per capita, share of foreigners, hospital beds per person, and average number of nights spent by tourists. Non-Catholic counties have a higher share of people over 65, more people completed secondary education and they are more densely populated. With respect to mortality, only education would work against us, but the difference is rather small.

**Table 1**
All-cause mortality rate in Western Germany before and after the pandemic onset: Difference-in-differences estimates.

| | [1] | [2] | [3] | [4] |
|---|---|---|---|---|
| Catholic state × Period after first COVID-19-related death | .056*** | .057** | .057** | .055** |
| | (0.014) | (0.019) | (0.019) | (0.020) |
| Daily time trend | No | No | Yes | Yes |
| State fixed effects | No | No | No | Yes |
| Sample until April 22, 2020 | No | Yes | Yes | Yes |
| *R*-squared | .028 | .056 | .061 | .411 |
| Observations | 1380 | 1090 | 1090 | 1090 |

Source: Federal Statistical Office of Germany.
Notes: OLS estimates. Treated (Catholic) states are those where Catholics are the majority, i.e. Bayern, Baden-Württemberg, Saarland, Rheinland-Pfalz, and North Rhine-Westphalia. Control (non-Catholic) states are Schleswig-Holstein, Hamburg, Bremen, Niedersachsen, and Hessen. Standard errors in parentheses are clustered at the state level. Logged all-cause mortality rate since January 01, 2020, is the outcome variable. Days after first COVID-19-related death are state-specific. Regressions are weighted by the size of the state population. Clustered standard errors at the state level in parentheses. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

and they move in parallel implying that before the COVID-19 pandemic nothing noticeable caused Catholics to pass at a higher rate relative to non-Catholics. Any systematic differences in genetic predisposition, pre-existing conditions, overall health status, socio-economic characteristics, local infrastructure, health infrastructure or behaviours such as risky attitudes would be likely reflected in systematically different past mortality rates between groups. This would also capture most of the unobserved heterogeneity that could drive mortality rates. Therefore, any divergence in mortality patterns between counties observed during the first half 2020 should be due to the incidence of COVID-19 cases and related deaths.[10] The non-existence of pre-pandemic systematic mortality differences was also validated in a regression-based framework. More specifically, we regressed (logged) annual mortality rates from all causes (2011–2017) on a set of controls to formally test the parallel trends assumption at the county level. As seen in Appendix Table A.3, there were no systematic differences in overall monthly mortality between counties before the pandemic, conditional on county observables and fixed effects.

Ideally, the existence of such differences should be formally tested under a difference-in-differences (DiD) framework using daily mortality data at the county level. However, such daily information is only available at the state level. Given this limitation, we used data from January 01, 2020, onwards to test whether mortality rates were higher in predominantly Catholic states relative to non-Catholic ones after the pandemic onset. Predominantly Catholic states were defined as those where Catholics are the majority in their population.[11] Having shown that the pre-pandemic mortality trends in Catholic and non-Catholic counties were similar (Appendix Fig. A.2), we estimated the following DiD specification at the state level:

$$m_{st} = \beta_0 + \beta_1 Catholic_s + \beta_2 Onset_t + \beta_3 Catholic_s {*} Onset_t + u_{st} \quad (1)$$

where $m_{st}$ is the logged mortality rate (total deaths per 100,000 population) in the $s$-th state on day $t$ of 2020, *Catholic* is a binary indicator switched on for predominantly Catholic states, and *Onset* is a dummy variable switched on after the first pandemic-related death in that state was reported. The DiD parameter of interest in Eq. (1) is $\beta_3$ and indicates whether differences in mortality rates between Catholic and non-Catholic states have changed significantly after the onset of the pandemic. Based on our hypothesis about a higher COVID-19 incidence in Catholic counties and hence, states, we expect that $\beta_3 > 0$.

Table 1 displays the DiD parameters obtained after estimating Eq. (1). Column 1 indicates an increased mortality rate in Catholic states during the exposure period. This is confirmed when the estimation sample is limited until April, 22, i.e. when some restrictions were lifted, covering roughly one month since the nation-wide lockdown. This would allow to capture fatalities resulting from the initial stage of the pandemic.[12] Column 3 adds a linear time trend running from the day when the first related death was reported to capture the already higher spread in some states compared to others. Finally, we add state fixed effects in column 4. In every case, the DiD parameter is positive and significant. Given that pre-pandemic mortality trends were similar across all states and counties, these results provide further evidence that mortality increased in predominantly Catholic states after the pandemic onset, following a higher number of coronavirus infections in those areas.

Moreover, in Fig. 1, there is a clear discrepancy on how the number of COVID-19 cases per 100,000 population during the early pandemic phase has evolved over time in Catholic counties. At the onset of the pandemic, COVID-19 incidence is higher in Catholic counties, and this is the reason behind higher levels of related fatalities per 100,000 population in those counties as well. Moreover, apart from aggregating across Western German counties, Fig. 1 displays the results for individual states with a sufficient number of counties. The similar pattern emerges across states with a sufficient number of counties, with the difference being more pronounced in some states, e.g. in Bayern, Rheinland-Pfalz, and Niedersachsen.

[10] Pre-pandemic mortality rates are also parallel for people over 65 years old, who face a higher risk of dying.
[11] Therefore, Bayern, Baden-Württemberg, Saarland, Rheinland-Pfalz, and North Rhine-Westphalia were defined as Catholic states. Schleswig-Holstein, Hamburg, Bremen, Niedersachsen, and Hessen were defined as non-Catholic states.
[12] The results are robust to limiting the timeframe until 31 March which would still capture infection prior to the lockdown and would exclude potential infection at around Easter (12 April).

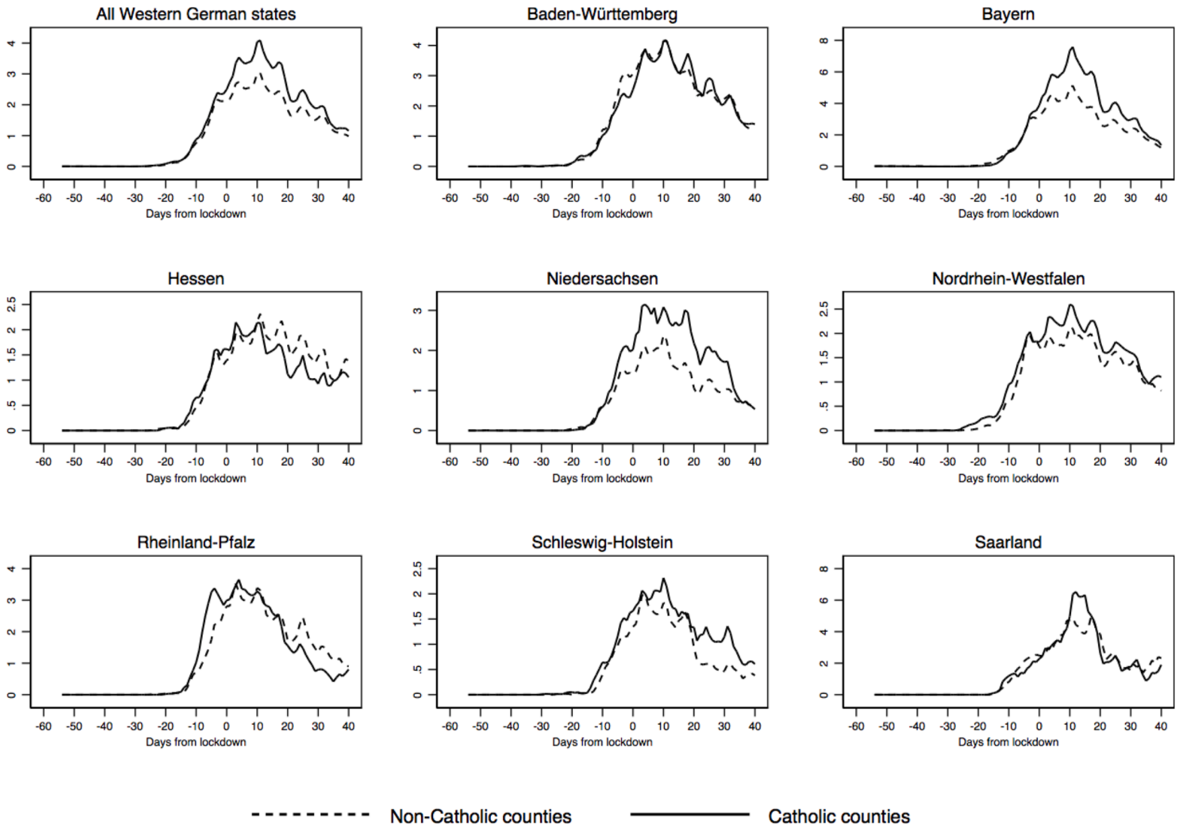


**Fig. 1. Reported COVID-19 cases per 100,000 population in Western German counties.** Source: Robert Koch Institute (RKI). Notes: Horizontal axes are centred at the national lockdown date (March 22, 2020). Catholic (non-Catholic) counties are defined as those where their local share of Catholics is higher (lower) from their respective state average. From the 10 Western German states, only those with a considerable number of counties (given in parentheses) are shown here: Baden-Württemberg (44), Bayern (96), Hessen (26), Niedersachsen (45), Nordrhein-Westfalen (53), Rheinland-Pfalz (36), Saarland (6), and Schleswig-Holstein (15). Figures for Bremen and Hamburg are not shown here, due to the small number of counties (2 and 1, respectively).

Another issue could be that the higher number of cases could be due to the carnival festivities that took place in North Rhine-Westphalia in February (Pluemper and Neumayer, 2020). We used Apple mobility data to examine if this could induce problems of any sort. We classified the 18 Western German cities available in those data into Catholic and non-Catholic ones and we calculated the mean mobility indicators for each group, weighted by the size of the local population. Fig. 2 displays the trends by transportation type before and after the lockdown (March 22, 2020). The use of public transport spikes about 30 days before the lockdown, which coincides with the culmination of carnival festivities in late February. However, this spike appeared at the same period for both Catholic and for non-Catholic cities, and mobility patterns were almost indistinguishable from each other, implying that there were no noticeable mobility differences between Catholic and non-Catholic cities. If the number of COVID-19 cases during the early pandemic was due to carnival festivities in Catholic cities, we would observe substantially higher mobility in those cities compared to non-Catholic ones. However, despite carnival being a Catholic tradition, it has evolved into a nationwide celebration for the youth, so that the spike in mobility occurred throughout Western Germany, and thus transmission of an infectious disease is just as likely anywhere in the country.

### *4.2. COVID-19 incidence and share of catholics at the county level*

Having established that observable characteristics, mobility patterns and pre-pandemic mortality rates were similar in Catholic and non-Catholic counties and states, the main goal of our analysis is to demonstrate that the reported COVID-19 incidence varies with

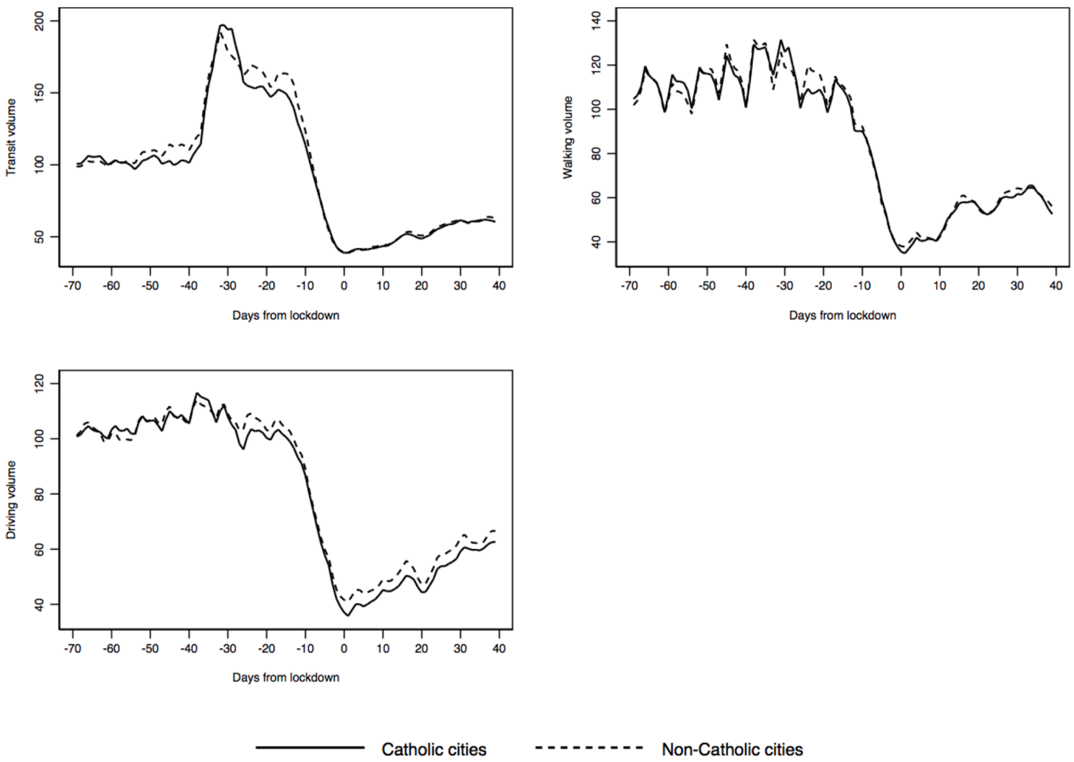


**Fig. 2. Mobility patters in major Western German cities.** Source: Apple mobility data (www.apple.com/covid19/mobility); Federal Statistical Office of Germany. Notes: Horizontal axes are centred at the national lockdown date (March 22, 2020). Catholic (non-Catholic) cities are defined as those where their local share of Catholics is higher (lower) from their respective state average. Relative mobility volumes are weighted by the size of the respective local population.

religion at the county level. This is challenging because demographic information at the county level does not arrive as frequently as data on infectious diseases do. Hence, we start our empirical application by estimating variants of the following specification:

$$Y_{rd} = \alpha Y_{rd-t} S_{rd-t} + \beta Catholic_r + \gamma X_r + t_d + \lambda_R + \varepsilon_{rd} \tag{2}$$

where $Y$ is the logged count of COVID-19 cases (or deaths) per 100,000 local population in county $r \in \{1,\ldots,N\}$ in day $d$ of the early pandemic phase, *Catholic* is the logged share of Catholics in county $r$ (based on the most recent data available before the virus outbreak), $t_d$ is a linear time trend starting from the day when the first COVID-19 case was reported in each county, $\lambda_R$ is a set of state fixed effects so they are not perfectly collinear with demographic and economic predictors varying at the county level, and $\varepsilon_{rd}$ is the disturbance term. For robustness, Eq. (2) is also estimated with *Catholic* being expressed as a binary variable that is switched on for counties where the share of Catholics over the total population is higher than their respective state average.

The variable $S$ represents the proportion of susceptible individuals in each county, where their stock has been approximated as the number of people after removing those reported deceased from the virus in each day (Adda, 2016).[13] The $t$ parameter represents the incubation period and has been set equal to 14 days.[14] We do not consider any spatial variation in the model so the incidence rate in each county is solely determined by its own past realisations. Therefore, parameter $\alpha$ could be interpreted as an estimate of the within-county spread. All models control for the size of the local population and for a series for economic and demographic controls.[15] Models also include state-specific time trends in order to account for confounders or policies that may vary at the state level, given the country's federal system. Estimating parameter $\beta$ through these empirical models will provide a first indication about how the count of COVID-19 cases and related fatalities per 100,000 population varied with the local share of Catholics in each county during the early pandemic, conditional on past incidence, observable characteristics, state fixed effects and time trends. We test the robustness of

[13] We deviate from Adda (2016) by focusing on parameter $\beta$ that measures level differences in the incidence of COVID-19 between Catholic and non-Catholic, rather than focusing on differences in transmission rates between counties through a Standard Inflammatory Response (SIR) model. However, we are also able to adapt their methodology as an additional robustness check in Appendix section A.1.

[14] We have also considered alternative values for the incubation period, i.e. from 3 to 20 days (Lauer et al., 2020).

[15] In robustness checks, models control for other variables. For example, driving distance from Milan which was the first major pandemic centre in continental Europe, as well as factors that have been argued to affect the number of COVID-19 incidence, e.g. temperature, precipitation etc.

**Table 2**
COVID-19 incidence and share of Catholics in Western German counties: Baseline estimates.

| | All Western German states | | | | North Rhine-Westphalia | |
|---|---|---|---|---|---|---|
| | Cases[1] | Cases[2] | Deaths[3] | Deaths[4] | Cases[5] | Deaths[6] |
| | **Panel A: Variable of interest is continuous** | | | | | |
| Local Catholic share | .050** (0.023) | .050** (0.022) | .012** (0.005) | .012** (0.005) | .092** (0.037) | .019** (0.008) |
| Within county ($\alpha$) | .364*** (0.013) | .367*** (0.014) | .053*** (0.004) | .055*** (0.004) | .356*** (0.031) | .058*** (0.008) |
| $R$-squared | .178 | .188 | .046 | .047 | .118 | .049 |
| | **Panel B: Variable of interest is dichotomous** | | | | | |
| Catholic county (0/1) | .058*** (0.021) | .058*** (0.021) | .010** (0.005) | .010** (0.005) | .099*** (0.032) | .015** (0.007) |
| Within county ($\alpha$) | .362*** (0.013) | .367*** (0.014) | .053*** (0.004) | .055*** (0.004) | .355*** (0.031) | .058*** (0.008) |
| $R$-squared | .183 | .193 | .047 | .048 | .120 | .049 |
| Daily time trend | Yes | Yes | Yes | Yes | Yes | Yes |
| County controls | Yes | Yes | Yes | Yes | Yes | Yes |
| State fixed effects | Yes | Yes | Yes | Yes | No | No |
| State linear trends | No | Yes | No | Yes | No | No |
| Observations | 16,952 | 16,952 | 16,952 | 16,952 | 2938 | 2938 |
| Number of counties | 312 | 312 | 312 | 312 | 53 | 53 |

Source: Robert Koch Institute (RKI); Federal Statistical Office of Germany.
Notes: OLS estimates. Outcomes are expressed as the logged count of cases (or deaths) per 100,000 local population. In Panel A, the local (logged) Catholic share is calculated relative to the county's total population. In Panel B, Catholic (non-Catholic) counties are defined as those where the local share of Catholics is higher (lower) than their respective state average. County controls include the share of people over than 65 years old, the share of foreign-born, GDP per capita, the share of those completed secondary education, the average number of nights spent by tourists, the number of hospital beds per capita. Regressions are weighted by the size of the county population. Standard errors in parentheses are clustered at the county level. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

baseline estimates by considering factors that could arguably affect the incidence of COVID-19, i.e. including more control variables, controlling for distance from places that suffered a lot during the early pandemic, excluding extreme counties and counties close to any border, and controlling for weather conditions at the county level.

The problem with Eq. (2) is that county-specific time-invariant unobserved heterogeneity is not partialled out, because including a county fixed-effect would not allow $\beta$ to be estimated. Hence, the results based on Eq. (2) are conditional only to state-level fixed effects.[16] To address this concern within a context where $N$ is large and $T$ is small and fixed, we applied the 2-stage method of Pesaran and Zhou (2018). In the first stage, the predicted residual cases (or deaths) per population from a fixed-effects estimation were averaged for each county over the entire period under consideration. In the second stage, they were used as the dependant variable in regressions over the cross-sectional sample of Western German counties, where models controlled for the local share of Catholics and other county-level characteristics. More specifically, in the first stage we run the following daily panel data model using a fixed effects estimator:

$$Y_{rd} = a_r + \alpha Y_{rd-t} S_{rd-t} + t_d + u_{rd} \quad (3)$$

In the second stage, the obtained residuals, $\widehat{u}_{rd}$, were averaged over time and they were regressed on time-invariant county characteristics, using the sample of Western German counties:

$$\widehat{u}_r = \beta_0 + \beta_1 Catholic_r + \gamma X_r + e_r \quad (4)$$

where $r \in \{1,\ldots,N\}$, and $\widehat{\beta}_1$ and $\widehat{\gamma}$ are consistent fixed effect filtered (FEF) estimators of the local share of Catholics and other time-invariant county-level controls. As shown by Pesaran and Zhou (2018), this two-stage approach can be modified in the case where instruments are available for the explanatory variables in Eq. (2). To address issues related to the endogeneity, the share of Catholics in the county was instrumented using the distance between Wittenberg and each county's capital. Because Wittenberg was the centre of the Protestant Reformation in the early 16th century, this instrument has been used and extensively discussed in the literature to predict the geographic dispersion of Catholics, e.g. in Becker and Woessmann (2009). Counties located further away from Wittenberg remained predominantly Catholic and the historical composition closely resembles the contemporary composition (Spenkuch, 2017).

[16] In fact, in the COVID-19 literature, most studies focusing on one country report estimates that either condition on some higher-level fixed effects or they are unconditional to fixed spatial differences, e.g. Sa (2020), and Pluemper and Neumayer (2020).

**Table 3**
COVID-19 incidence and share of Catholics in Western German counties: Fixed effects estimates.

| | Outcome variable: COVID-19 cases per 100,000 population | | | Outcome variable: COVID-19 deaths per 100,000 population | | |
|---|---|---|---|---|---|---|
| | OLS [1] | OLS [2] | 2SLS [3] | OLS [4] | OLS [5] | 2SLS [6] |
| | **Panel A: Total period** | | | | | |
| Local Catholic share | .135*** (0.014) | .114*** (0.017) | .107*** (0.037) | .016*** (0.003) | .014*** (0.003) | .010* (0.006) |
| *R*-squared | .174 | .330 | .245 | .057 | .145 | .113 |
| First-stage result | – | – | 1.364*** (0.117) | – | – | 1.364*** (0.117) |
| | **Panel B: Before lockdown** | | | | | |
| Local Catholic share | .123*** (0.017) | .105*** (0.020) | .201*** (0.038) | .012*** (0.003) | .011*** (0.003) | .019** (0.008) |
| *R*-squared | .122 | .189 | .175 | .037 | .052 | .045 |
| First-stage result | – | – | 1.364*** (0.117) | – | – | 1.364*** (0.117) |
| | **Panel C: After lockdown** | | | | | |
| Local Catholic share | .168*** (0.018) | .138*** (0.021) | .133*** (0.044) | .023*** (0.004) | .020*** (0.004) | .016* (0.009) |
| *R*-squared | .170 | .307 | .229 | .072 | .178 | .140 |
| First-stage result | – | – | 1.364*** (0.117) | – | – | 1.364*** (0.117) |
| | **Panel D: Persons below 60 years old** | | | | | |
| Local Catholic share | .082*** (0.011) | .069*** (0.012) | .074*** (0.029) | .002 (0.003) | .002 (0.003) | .003 (0.006) |
| *R*-squared | .126 | .324 | .268 | .002 | .025 | .026 |
| First-stage result | – | – | 1.364*** (0.117) | – | – | 1.364*** (0.117) |
| | **Panel E: Persons above 60 years old** | | | | | |
| Local Catholic share | .097*** (0.010) | 085*** (0.011) | .085*** (0.024) | .020*** (0.003) | .017*** (0.003) | .015** (0.007) |
| *R*-squared | .158 | .322 | .241 | .077 | .174 | .133 |
| First-stage result | – | – | 1.364*** (0.117) | – | – | 1.364*** (0.117) |
| Observations | 312 | 312 | 312 | 312 | 312 | 312 |
| County controls | No | Yes | Yes | No | Yes | Yes |

Source: Robert Koch Institute (RKI); Federal Statistical Office of Germany; Google Maps.
Notes: Results from Pesaran and Zhou (2018) two-stage approach. In the first stage, the logged count of COVID-19 cases (or deaths) per 100,000 local population is regressed on a full set of county fixed effects, lagged number of cases and time trends. In the second stage, the dependant variable is the mean residual obtained from the fixed effects estimation in the first stage. The (logged) local Catholic share is calculated relative to the county's total population. County controls include the share of people over than 65 years old, the share of foreign-born, GDP per capita, population density, the share of those completed secondary education, the average number of nights spent by tourists, the number of hospital beds per capita. Regressions are weighted by the size of the county population. In columns [3] and [6], the local Catholic share is instrumented with the logged distance between the county's capital and Wittenberg, Germany. Robust standard errors in parentheses. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

## 5. Results

### 5.1. County-level baseline results

Table 2 displays the results from Eq. (2) using the logged daily count of new COVID-19 cases and associated deaths per 100,000 population at the county level as outcomes. In column 1, Panel A, models control for a series of observable county-level characteristics and state fixed effects. The estimated coefficient of the county-level share of Catholics is positive and statistically significant, confirming the empirical observation from Appendix Fig. A.1. According to the estimated $\alpha$ parameter, past infections strongly predict the current within-county COVID-19 incidence, as expected.[17] Despite focusing on regions within a single country, there might be some room for confounders pertaining to a state-level policy making; especially in a decentralised federal state like Germany. Therefore, we included linear daily state trends, in column 2, but the results remained unaffected. Further to the graphical evidence and the results in columns 1–2, it follows that COVID-related mortality also varies with the county-level share of Catholics conditional on county-level observables, state fixed effects and state time trends, in columns 3 and 4. For robustness, Panel B replaces the logged share of Catholics with a binary variable which is equal to one for predominantly Catholic counties and zero otherwise. However, this leaves the Panel A estimates unaffected, ensuring that both variables capture the dispersion of Catholics in Western Germany.

By the end of April 2020, the three states with most cases and deaths were Bavaria, Baden-Württemberg and North Rhine-Westphalia. The share of Catholics relative to the total population is 56 percent, 38 percent, and 42 percent, respectively. However, Bavaria and Baden-Württemberg border on Austria and are geographically closer to Northern Italy, the alleged epicentre of the pandemic in continental Europe. As this geographical proximity might raise concerns about our baseline estimates, we restricted the

[17] The 14-day lagged COVID-19 incidence was multiplied by a factor representing the fraction of susceptible individuals in the county population. This factor is an approximation, and the county population is calculated as the population minus the cumulative daily number of COVID-19-related deaths. Therefore, it does not consider population changes due to county-specific fertility and mortality from other causes, and does not address any endogeneity concerns. The results are robust when alternative lag structures are considered, i.e. ranging from 3 to 20 days.

**Table 4**
Differences in behavioural indicators between Catholics and non-Catholics in Western Germany.

| | Catholic parameter estimate [1] | Unbiased Catholic parameter estimate [2] | Individual controls [3] | State fixed effects [4] | *R*-squared [5] | Observations [6] |
|---|---|---|---|---|---|---|
| | **Panel A: Social interaction outcomes (ESS data)** | | | | | |
| Meet other people more than several times a week | .044 (0.027) | .050 | Yes | Yes | .047 | 1660 |
| Meet other people more frequently relative to peers | .056* (0.026) | .056 | Yes | Yes | .048 | 1660 |
| Discuss intimate and private matters with other people | .156*** (0.016) | .161 | Yes | Yes | .072 | 1660 |
| | **Panel B: Social and family ties outcomes (EVS data)** | | | | | |
| Family important in life | .040*** (0.011) | .034 | Yes | Yes | .079 | 3247 |
| Friends important in life | .092*** (0.022) | .099 | Yes | Yes | .055 | 3237 |
| Trust towards family members | .056*** (0.017) | .014 | Yes | Yes | .050 | 2338 |
| Reside with parents or parents-in-law | .017*** (0.007) | .018 | Yes | Yes | .226 | 3247 |
| | **Panel C: Social and family ties outcomes for major groups (EVS data)** | | | | | |
| Family important in life | .047*** (0.012) | .042 | Yes | Yes | .082 | 2970 |
| Friends important in life | .075*** (0.024) | .085 | Yes | Yes | .055 | 2963 |
| Trust towards family members | .062*** (0.016) | .051 | Yes | Yes | .051 | 2133 |
| Reside with parents or parents-in-law | .015* (0.008) | .014 | Yes | Yes | .226 | 3008 |
| | **Panel D: Adherence to rules outcomes (EVS data)** | | | | | |
| Justified to avoid taxes | −0.023 (0.058) | −0.010 | Yes | Yes | .024 | 2322 |
| Justified to accept bribe | −0.062 (0.064) | −0.051 | Yes | Yes | .021 | 2333 |
| Justified to avoid fare in public transport | −0.159** (0.064) | −0.107 | Yes | Yes | .085 | 2333 |
| Confidence in government | .064* (0.034) | .062 | Yes | Yes | .081 | 2269 |

Source: European Social Survey 2018 (ESS); European Values Study 2017 (EVS).
Notes: OLS estimates. All outcomes are binary. Each row refers to a separate regression with a different outcome variable. Individual-level controls include perceived health status, gender, age, household income, education level, employment status, age of youngest household member, as well as city size. The unbiased Catholic parameter estimate has been obtained using the methodology proposed by Oster (2019). Standard errors in parentheses are clustered at the state level. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

**Table 5**
COVID-19 incidence, local Catholic share and social interactions in Western Germany: Fixed effects estimates.

| | OLS [1] | OLS [2] | 2SLS [3] | 2SLS [4] | 2SLS [5] | 2SLS [6] | 2SLS [7] |
|---|---|---|---|---|---|---|---|
| Local Catholic share | .091*** (0.026) | .084*** (0.024) | .140*** (0.050) | .101** (0.046) | .100** (0.046) | .090* (0.049) | .088* (0.052) |
| Predicted social interaction index | – | .042** (0.017) | – | .030* (0.017) | – | – | – |
| Live with parents | – | – | – | – | .344* (0.216) | – | .358* (0.213) |
| Regularly attend religious services | – | – | – | – | – | .120 (0.156) | .133 (0.149) |
| County social interaction controls | No | No | No | No | No | No | Yes |
| County economic & demographic controls | No | No | Yes | Yes | Yes | Yes | Yes |
| *R*-squared | .090 | .153 | .304 | .306 | .300 | .286 | .324 |
| First-stage result | – | – | 1.440*** (0.181) | 1.440*** (0.181) | 1.440*** (0.181) | 1.440*** (0.181) | 1.440*** (0.181) |
| Observations | 118 | 118 | 118 | 118 | 118 | 118 | 118 |

Source: Robert Koch Institute (RKI); Federal Statistical Office of Germany; Google Maps; European Values Study 2017 (EVS).
Notes: Results from Pesaran and Zhou (2018) two-stage approach. In the first stage, the logged number of COVID-19 cases per 100,000 local population is regressed on a full set of county fixed effects, lagged number of cases and time trends. In the second stage, the dependant variable is the mean residual obtained from the fixed effects estimation in the first stage. The local (logged) Catholic share is calculated relative to the county's total population. County controls include the share of people over than 65 years old, the share of foreign-born, GDP per capita, population density, the share of those completed secondary education, the average number of nights spent by tourists, the number of hospital beds per capita. County social interaction controls include the shares of those considering family as very important in their lives, those considering friends as very important in their lives, those living together with their parents and/or their parents-in-law, those who completely trust their family, those who completely trust people they personally know, those who live in households with maximum 2 members, those who regularly attend religious services, those who consider religion as very important in their lives, and those who believe that children should provide care for their parents. The predicted social interaction index was constructed from a principal component analysis using the social interaction controls. Regressions are weighted by the size of the county population. In columns [3] to [7], the local Catholic share is instrumented with the (logged) distance between the county's capital and Wittenberg. Estimation sample is restricted to Western German counties in which respondents were surveyed for the 2017 EVS wave. Robust standard errors in parentheses. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

estimation sample only to North Rhine-Westphalia counties who exhibit a considerable variation in being predominantly Catholic or not. Columns 5–6 in Table 2 report the results. Conditional on local demographics and time trends, the estimated coefficient of either the logged Catholic share or the dummy variable is positive, significant and higher in magnitude. Moreover, as the virus circulated locally at the beginning of the pandemic, we further restricted our sample by excluding extreme counties, such as Heinsberg, from our sample of North Rhine-Westphalia counties but the results are similar to those in columns 5–6 of Table 2.[18]

The baseline results in columns 1–4 of Table 2 are robust to a series of sensitivity checks. These are reported in Appendix Table A.4. Controlling for day-of-week fixed effects and their interaction with state indicators to account for differences in working, commuting and leisure patterns across states did not bring any notable differences in the parameters of interest. A second check was to exclude Bavarian counties from the estimation sample due to their quite high share of Catholics. Again, the positive relationship is confirmed. Conditional on all observables and state-specific trends, the estimated coefficients of our Catholic county measures remained statistically significant. The same conclusions held when all counties from both Bavaria and Baden-Württemberg, a neighbouring state with an also large share of Catholics and a large number of predominantly Catholic counties in it, were excluded.

To address concerns regarding proximity to the Northern Italian border, an area that was hit the hardest at the pandemic outbreak in Europe, we calculated the fastest highway driving distance (in kilometres) between the major city of each Western German county and Milan, Italy, using Google Maps.[19] Although this does not entirely rule out any geographic heterogeneity, it mitigates any concerns related to how infections were scattered across continental Europe. As seen in Table A.4, including the logged fastest highway driving distance, in kilometres, from Milan, did not affect the baseline coefficients. Driving distance from Milan could not explain any of the variation in the data either, as the associated point estimate was quite noisy and insignificant. In fact, closeness to borders in general, did not seem to offer a convincing explanation for the observed link. Excluding all counties right on any border area left the results unaffected. We also tested the robustness of the baseline evidence by including weather-related variables, such as temperature and precipitation for 136 available counties taken from the Climate Data Centre (*Deutscher Wetterdienst*), that have been argued to be related to the number of cases, however, the same conclusions were reached.[20]

[18] The results for this test are reported, among others, in Appendix Table A.4.

[19] Large airports are primarily thought of as the primary entryways into a country, and air travel as an important contributor in spreading infectious diseases, we considered land travel based on recent evidence (Christidis and Christodoulou, 2020; Pluemper and Neumayer, 2020).

[20] The same conclusions hold when the logged cumulative number of virus-related incidences per 100,000 inhabitants is used as the dependent variable. Results are not reported here but they are available upon request.

Another concern could be that there might be some other groups, mostly concentrated in urban areas, such as Muslims, that were reportedly more exposed to the virus. Our test for this was to exclude cities with more than 200,000 population, in which such groups are more likely to be concentrated, but again the positive relationship was confirmed. This last test also helps in excluding the large number of non-religious individuals concentrated in urban areas, reducing potential biases from non-random migration and rendering the remaining counties more comparable. Finally, we performed a placebo test by replacing the local share of Catholics in the county with the share of Protestants; as well as the dummy indicating a Catholic county with similarly constructed one indicating a Protestant county. As seen in Table A.4, the effect of Protestants was not significant and negative, providing some further reassurance about the direction and magnitude of our empirical observation. Moreover, controlling for both Catholic and Protestant regressors, confirms the positive relationship between the Catholic variable and the COVID-19 incidence at the county level. In models controlling for both religious groups the coefficients of the Protestant variables are close to zero and not significant. In this last specification the reference category is the share of other/non-religious. The results imply that given the share of other/non-religious, a higher share of Catholics is positively associated with COVID-19 incidence, whereas the share of Protestants is not.

### *5.2. Controlling for county fixed effects*

A limitation of the results discussed so far is that they were conditional only to state level fixed effects, otherwise the coefficient of interest could not be estimated. However, it needs to be ruled out that the local share of Catholics is not correlated with unobserved factors that influence the COVID-19 incidence at the county level. This is a common problem in the COVID-19 literature because local economic and demographic covariates do not vary at the daily level while the virus incidence does. In fact, they are time invariant in this context.

To address this issue, we use the fixed-effects filtered (FEF) estimator suggested by Pesaran and Zhou (2018). This allows to uncover the effect of a time-invariant covariate when using large $N$, small $T$ data. In the first stage, a FE estimator is applied to Eq. (3) and the residual COVID-19 incidence is averaged over the period for each county. In the second stage, Eq. (4), the mean residual is regressed on the local Catholics share and other characteristics using the cross-sectional sample of Western German counties.

The results from the second stage are reported in Table 3. In Panel A, the outcome is the mean residual obtained from applying a FE estimator to the total period in the first stage. The share of Catholics is positively and significantly affecting both the residual number of infections and deaths per 100,000 population, and this holds after controlling for demographic and economic characteristics. This is confirmed by both OLS and 2SLS estimates, where the local share of Catholics in each county was instrumented by the (logged) distance between each county's capital and the town of Wittenberg as suggested by Becker and Woessmann (2009). In the case of the 2SLS estimates, the first-stage result is particularly strong, suggesting that the current religious composition is strongly related to the historical one, an observation similar to Spenkuch (2017). In Panel B, the outcome is the mean residual when the estimation sample in the first stage is restricted to the period before the lockdown, in March 22, 2020. The effect of the local Catholics share in the county remains positive and significant on both cases and deaths after controlling for the usual set of county characteristics. The same conclusions are reached when the first stage estimation is conditioned on the period during the lockdown. In the second stage, the link between COVID-19 incidence and Catholics share is positive and significant according to both OLS and 2SLS estimates. Finally, in Panels D and E, we use cases and related deaths for those below and above 60 years old, respectively, in the first stage. The RKI data report the age for each COVID-19 case or fatality, so we were able to test for age-related non-linearities. The results in the second stage confirm a positive relationship for both age groups and, as expected, the estimated coefficients of the local Catholic share are higher for those above 60 years old, especially when considering the number of deaths per 100,000 population. This is worth mentioning, because even though non-Catholic counties have a slightly higher share of elderly, the number of infections and related deaths among the more vulnerable increases with the local share of Catholics.

### *5.3. Differences in individual behaviour*

The county-level analysis so far, established a positive link between COVID-19 incidence and the prevalence of Catholics in Western Germany. Behind this observed relationship, differences in individual behaviours and preferences could play a role. More specifically, stronger family ties and tighter social networks, could provide a possible channel that partially explains the aggregate results.

Religiosity is rather low in contemporary Western Germany with only 12 percent of EVS respondents reporting that religion is very important in their lives and 8 percent attending church regularly. Nevertheless, it may be the case that some shared values or preferences may have become part of culture outside of religion. Historically, it has been shown that Protestants and non-Catholics have developed more market orientated and individualistic behaviours, whereas Catholics are characterised by rent-seeking behaviours through established social structures (Becker and Pascali, 2019; Ekelund et al., 2002). Catholics have been traditionally organising their social life through the Church and much less through independent institutions, which implies that Church has been a strong focal point of their social activity (Satyanath et al., 2017). Moreover, these differences mostly reflect on social dimensions rather than economic dimensions or work ethics. Furthermore, Protestants view social interaction more as an enforcement and control mechanism over peers and society (Arruñada, 2009). Hence, more frequent social interaction, and stronger social and family ties as an intrinsic behaviour are more likely to be observed among Catholics. Our argument is that historic religious composition and the associated behaviours have manifested as a local way of life. Even though we have shown that contemporary religious composition closely resembles the historic one, religiosity in Western Germany is declining.[21] However, some of the shared values may remain and persist at the local level (Spenkuch, 2017), especially ones that are not directly related to religious activities such as the strength of social and family ties. This could imply that individuals living in these areas are likely to adapt to this local way of life regardless of religious affiliation. For these reasons we believe that social and family ties are more pronounced among Catholics and in predominantly Catholic regions.

We rely on individual-level data from the 2018 ESS and the 2017 EVS waves for Western Germany to test whether these differences can partially explain the county-level relationship. The ESS contains questions on religious denomination, the frequency of social interactions, on the perceived frequency of such interaction compared to one's peers, and also questions about the number of individuals with whom one feels comfortable discussing private and individual matters, and the number of individuals each respondent confides in. The EVS also contains information on religion denomination as well as a set of questions that proxy attitudes and behaviours, i.e. importance of family and friends in an individual's life, level of trust towards members of the family, and likelihood to reside with parents (or parents-in-law) in the same household. Moreover, the EVS asks respondents another series of questions on perceptions and beliefs that can affect behaviour in ways other than social contacts. Using this information, we constructed indicators on whether a respondent believe it is justifiable to avoid paying taxes, accept bribe, avoid paying fare in public transport, and whether they have confidence in their government.

Then we used those behavioural variables as outcomes in regressions where the variable of interest was a binary variable equal to one if the respondent was a Catholic and zero otherwise, alongside a series of demographic and other characteristics (i.e. age, gender, subjective health status, employment status, household size, household income, age of youngest household member, city size) and state fixed effects. Table 4 reports the results. In Panel A, we consider social interaction outcomes from the EVS data. The results are supportive of a positive relationship between belonging to the Catholic denomination and frequency of social interactions. When the outcome is the frequency of meeting other people, the coefficient of the Catholic dummy is barely not significant at the 10 percent level ($t$-statistic $= 1.67$), but the other two are, indicating that Catholics tend to have more frequent social interactions and exhibit stronger ties relative to non-Catholics. In Panel B, we turn to EVS data in order to see whether there is any systematic differentiation with respect to outcomes related to social and family ties. In line with Arruñada (2009), the results clearly indicate that Catholics, relative to non-Catholics, regard their family and friends as being very important, they trust their family members more, and they are more likely to reside in the same household with their parents or parents-in-law. This provides further support that the higher incidence of COVID-19 infections and related fatalities in predominantly Catholic counties, could be partially attributed to the existence of stronger social and family ties among Catholics. For robustness, we proceed with excluding all other religious groups from our specifications and only focussing on Catholics, Protestants and non-religious.[22] These three groups account for roughly 92 percent of all EVS respondents in Western Germany. The results are presented in Panel C of Table 4 and our findings remain largely unchanged.

To further strengthen our proposed channel, we employed another set of individual-level regressions using individual EVS data. The idea was to examine whether the observed differences in the COVID-19 incidence in Catholic and non-Catholic counties in the early pandemic are due to differences in behaviours other than through social contacts. Specifically, we test whether Catholics have a higher propensity to justify cheating behaviour in avoiding taxes, accepting bribes and avoiding fares in public transport. Further, we examine for differences in their confidence towards the government; even though recent research suggests that in most European countries confidence in governments and their implemented measures have increased during the early stages of the pandemic and that individuals tend to prioritise health (Bol et al., 2020; Hargreaves Heap et al., 2020). This test should be indicative of whether the higher number of cases and fatalities in predominantly Catholic counties was due to an overall riskier behaviour or disobedience towards the authorities and the mitigation measures they introduced. This would likely have a differential impact on the number of COVID-19 cases even before the mitigation measures become effective. Panel D in Table 4 reports the results.

Broadly speaking, the data reveal that adherence to rules is very high in Western Germany. Moreover, Table 4 does not suggest a higher propensity for Catholics to disobey rules, relative to non-Catholics. If anything, Catholics are less likely to tolerate even the milder offence (avoiding fare) and exhibit marginally higher confidence in the government. Therefore, the higher count of COVID-19 cases during the early pandemic phase seems unlikely to have resulted from differences in risky behaviours and civil disobedience between Catholics and non-Catholics. Based on the evidence from Panels A and B, stronger social and family networks seem more relevant.

An issue here is that unobserved heterogeneity may bias the results. This is often the case with cross-sectional analyses of survey data. To partially correct for this, we follow the methodology proposed by Oster (2019). The idea is to simultaneously observe how the coefficient of interest co-moves with the $R^2$ when including the full set of controls. This allows to approximate the selection on unobservables relative to the selection on observables and obtain an unbiased coefficient of the Catholic indicator. We read the latter as a "conservative" estimate obtained under the standard assumptions that (a) unobservables matter as much as observables, and (b) that the maximum variance that can be explained is 30 percent higher than the one explained when only controlling for observables, i.e.

[21] It is worth noting that Spenkuch (2017) argues that Protestants are twice as likely to denounce their denomination. This implies that Catholics may feel more strongly tied to these shared values.

[22] Having already shown, in Appendix Table A.4, that infections do not vary with the local share of Protestants, Appendix Table A.5 further narrows down the comparison between Catholics and Protestants alone using individual-level data. Both the obtained OLS and unbiased parameter estimates suggest that Catholics seem to have more social interactions as compared to Catholics alone, e.g. they meet other people more than several times per week and they show greater trust towards their family members. In particular the unbiased coefficients point towards this direction in four out of six different outcomes we consider.

$R^2_{max} = 1.3 \times R^2$ (Oster, 2019). This method has found numerous applications in the empirical literature, e.g. Alesina et al. (2016), Satyanath et al. (2017), Tabellini (2020). Column 2 in Table 4 displays the obtained unbiased parameters for the Catholic indicator.

For all the outcome variables considered in Panels A, B, C and D of Table 4, the unbiased parameter estimate of the Catholic dummy suggests that if omitted variables were as important as the included observables, the bias would be against the estimated parameters in column 1. However, the estimated coefficients in column 1, Panel A, remain rather stable implying that Catholics are different in terms of social interactions relative to non-Catholics. The same applies to the outcomes considered in Panel B. Catholics appear to have stronger family and social ties relative to everybody else. Among others, Catholics are more likely to live in the same household with their parents or parents-in-law. This provides some reassurance that the uncovered differences in behaviour and preferences between Catholics and non-Catholics are less likely to be driven by unobserved heterogeneity. Regarding the outcomes considered in Panel C, it seems that accounting for unobservables moves the unbiased estimates in the first three columns closer to zero. This further suggests that being a Catholic does not imply any differences in justifying cheating behaviour and complying with rules. When considering trust in the government as outcome, the coefficient remains remarkably stable indicating that Catholics exhibit indeed more confidence in the government relative to non-Catholics. Therefore, any observed differences in the initial number of cases and fatalities could be due to stronger social and family ties, rather due to riskier behaviour and non-compliance.

Within the context of this paper, a more direct test about how the more frequent social interactions of Catholics, relative to non-Catholics, can partially account for the higher number of COVID-19 cases in counties where the local share of Catholics is higher, would require controlling for variation in such behaviour at the county level, i.e. when estimating Eq. (4). Originally this was not possible as both the ESS and the EVS datasets provide only state-level identifiers. After a licence agreement, county level identifiers for the 2017 EVS data were provided to us, allowing the aggregation of social interaction measures at the county level. However, surveyed individuals were sampled in a subset of 118 Western German counties. Therefore, our hypothesis was directly tested at the county level but under this restriction. Table 5 reports the results. First, we re-estimated Eq. (4) in order to check if the results hold for the subset of counties for which EVS information is available. Column 1 shows that the estimated coefficient is lower from the one reported in Table 3 (column 1, Panel A), however significant at the 1 percent level. The same was true when the local share of Catholics at the county level was instrumented with the distance between the county's capital and the town of Wittenberg, Germany (column 3) as suggested by Becker and Woessmann (2009). The first stage result was strong, and the estimated 2SLS coefficient of the local Catholic share is comparable with the one reported in Table 3 (Panel A, column 3).

To directly test whether increased social interactions among Catholics can explain (some of) the relationship between COVID-19 incidence and the local share of Catholics at the county level, we controlled for a social interaction index, in column 2. This index was constructed through a principal component analysis (PCA) that reduced the dimensionality of the EVS variables proxying social interaction preferences and perceptions. Principal components were constructed as linear combinations of variables that were indicating (a) whether family is very important in life or not, (b) whether friends are very important in life or not, (c) if someone lives with their parents and/or their parents-in-law, (d) if someone attends religious services more frequently than once per month, (e) if someone lives in a small household, i.e. up to two members, or not, and (f) if someone agrees that children should provide care for their parents. The variables were collapsed at the county level and standardised before the principal components were identified. Appendix Fig. A.3 displays the variance in the data explained by each component. The first component explains more than 27 percent of the total variation in the data. When the second and the third components were additionally considered, approximately 70 percent of the total variation in the data is explained. For the first component, the most important variable was the one indicating cohabitation with parents and/or parents-in-law. For the second and third components, the most important variables were whether children should provide care for their parents, and frequency of attending religious services, respectively. For all three components, their means are significantly higher, at the 1 percent level, in counties where the share of Catholics is larger than the share of Protestants, and in counties where the local Catholics share is higher than the respective state average, compared to counties where it is lower.

In column 2 of Table 5, we estimated Eq. (4) using OLS and controlling additionally for the predicted social interaction index (the first principal component). This reduces the magnitude of the coefficient associated with the local Catholic share at the county level, implying that the local share of Catholics is strongly correlated with our measure of social interactions. Even if other factors associated with the historic geographic dispersion may play a role, we find this strong association with social and family ties. Moreover, the relationship between the index and the number of COVID-19 cases per 100,000 county population is positive and significant at the 5 percent. This is also the case in column 4 reporting the 2SLS estimates of the local Catholic share. Compared to column 3, the 2SLS estimate in column 4 is reduced by nearly 28 percent when models control for the predicted social interaction. Instead of the constructed index, columns 5–7 control for the averages of the variables that were used to construct it instead. In all cases, the estimated coefficient of the local Catholic share at the county level is reduced, and the social interaction variables have the expected sign (column 7 controls for all variables used in the principal component construction). Therefore, although based on a limited number of available observations, this direct test provides some support in our argument about more frequent and closer social interactions, cohabitation with parents etc., observed among Catholics, playing a role in explaining the positive relationship we observed at the county level.

Moreover, this relationship does not seem to be due to other behavioural differences between Catholics and non-Catholics. Looking again at the Apple mobility data in Fig. 2, there is a significant drop in mobility after the lockdown, however, there is not any notable differentiation on the basis of religion. A concern with those data could be that they refer to searched for directions to specific destinations via certain transportation means. This cannot be formally tested; however, it is plausible to assume that individuals would not necessarily look up the addresses of their family members and close friends before visiting them prior to the lockdown. Moreover, the Federal Statistical Office provided commuter mobility data at the county level gathered by Teralytics AG, a private service provider, for 30 million mobile phones in Western Germany. The data refer to individuals crossing county borders during peak commuting hours between Monday and Thursday each week. The data are provided as monthly averages between January 2020 and May 2020, compared to the respective month of the previous year. We used those estimates as independent variables in regressions separately for each month as well as for the total period. As commuting could be an important factor in spreading the virus, this would indicate any related differences between Catholics and non-Catholics. However, the estimated coefficient for the local Catholics share is not significant, in Appendix Table A.6. Hence, there is no differential change in commuter mobility compared to the previous year depending on the local share of Catholics in the county. Therefore, the higher COVID-19 incidence in Catholic counties is more likely to be due to stronger family ties, cohabitation with parents, and closer social circles, rather than higher mobility of any sort before or after the lockdown.

## 6. Conclusions

The context within which social preferences are shaped is a crucial determinant of societal and individual behaviour. Such differences are difficult to be observed between otherwise comparable individuals and groups of people. It is often the case that they are unobserved and it is burdensome to isolate their potential impact on individual or group-level outcomes and relationships (Alesina and Giuliano, 2015). Nevertheless, as they can lead to differential outcomes it is of high importance to take them under consideration. Such heterogeneities need to be taken into account not only in epidemiological modelling and parameterisation, but also to design optimal policy responses. Moreover, they do not only exist between, but also within countries and often go beyond standard demographic and socioeconomic characteristics, e.g. age, income or health status.

This is particularly important when facing crises such as the COVID-19 pandemic, which has already claimed many lives both directly and indirectly (Vandoros, 2020). Being able to explain diverging trajectories of countries in the early stages could provide invaluable insight to policymakers to design mitigation measures and policies around pandemics (Platteau and Verardi, 2020). Previous attempts to highlight the importance of culture and how it can determine social interactions are marred by unobserved heterogeneity and it is difficult to disentangle the number of factors that need to be taken into account. In this paper we used data from Western Germany in early 2020 to highlight the relevance of cultural differences in shaping social interactions on the incidence of the disease and the associated mortality in the initial phase of the COVID-19 pandemic. The starting point of our motivation was the observation that the incidence of COVID-19 infections and associated fatalities during the early pandemic phase was strikingly similar to how Catholics are scattered throughout the Western German counties. After validating that Catholic and non-Catholic counties are quite similar in terms of county-level characteristics, mobility patterns, and pre-pandemic mortality rates, we confirmed our empirical observation in a regression-based framework that accounted for differences in county-level covariates, county-level fixed effects and time trends. Our results survived a wide series of robustness checks and they were also confirmed using the predicted Catholic incidence under an instrumental variables framework. COVID-19 cases per 100,000 population were higher in predominantly Catholic counties in Western Germany. It followed that the associated mortality, as well as overall mortality, was higher in those counties.

Our main argument here is that the higher number of COVID-19 cases and related deaths in predominantly Catholic counties can be attributed to the fact that Catholics seem to exhibit a higher degree of social interactions. The historical geographic dispersion of Catholics across Western Germany closely resembles the contemporary one. As a result, some shared values and preferences may have firmly established themselves at the local level. To test its relevance, we first worked at the individual level, showing that Catholics are systematically different from non-Catholics in terms of social interaction frequency, religious services attendance, cohabitation with parents and parents-in-law and a series of other indicators capturing social and family ties. Next, we used the individual-level data to construct a social interaction index, that was able to explain part of the observed positive relationship at the county level.

Our results are important and indicate that, especially in societies with stronger social and family ties, virus outbreaks should be managed carefully, promptly and in a targeted manner by the authorities in order to avoid rapid spread and, consequently, higher death tolls. It is this insight that will enable policymakers to better respond to public health crises with the potential to upend society, the economy and the political landscape.

## Acknowledgments

We would like to thank the Editors of the European Economic Review and two anonymous referees for their comments and suggestions that significantly improved our work. We would also like to thank Maria Petrova, Dimitrios Xefteris, Konstantinos Matakos, Evgenia Passari, Seyhun Orcan Sakalli, Charitini Stavropoulou, Mujaheed Shaikh, Ivan Zilic and Agne Suziedelyte for their valuable input. Finally we would like to thank participants at the Quantitative Political Economy (QPE) seminar of the Department of Political Economy, King's College London and participants at the Conference in Research in Economic Theory and Econometrics (CRETE) 2021 for comments and suggestions.

## Appendix

Fig. A1, Fig. A2, Fig. A3, Table A1, Table A2, Table A3, Table A4, Table A5, Table A6, Table A7.

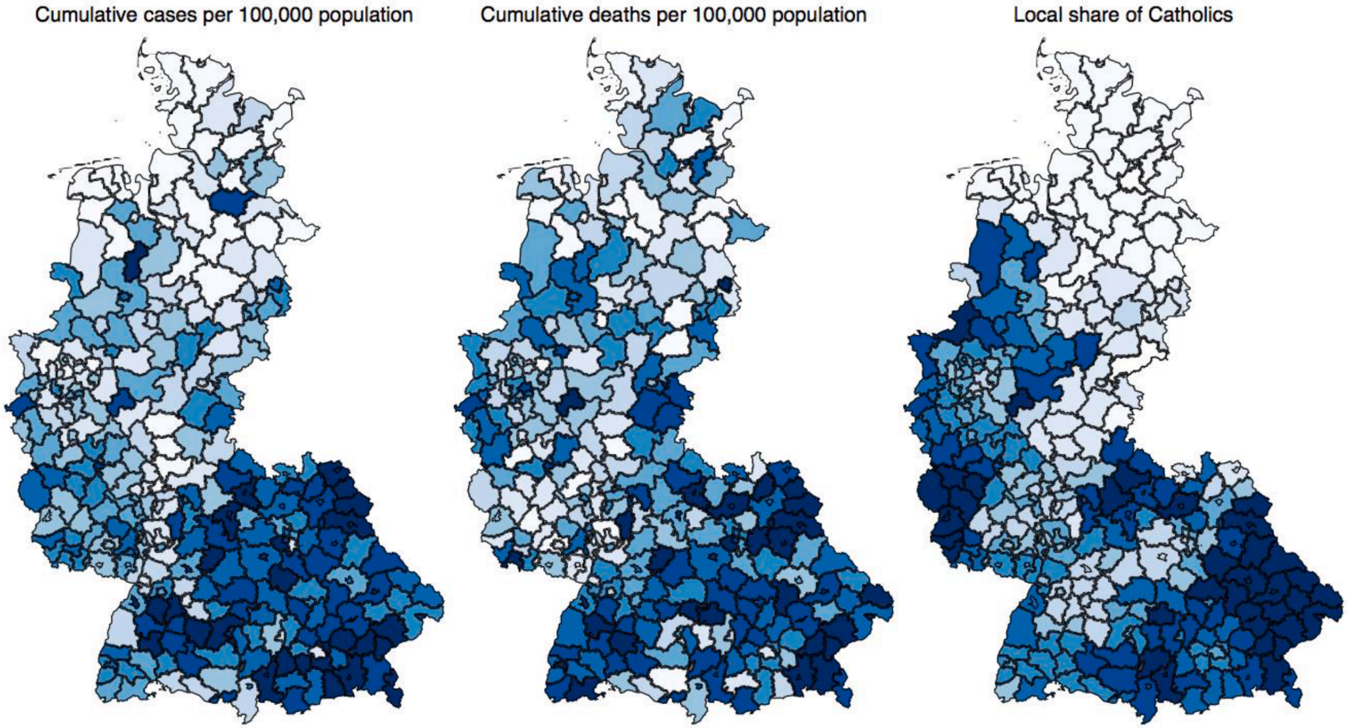


**Fig. A.1. COVID-19 incidence and local share of Catholics at the county level in Western Germany.** Source: Robert Koch Institute (RKI); Federal Statistical Office of Germany. Notes: Data on COVID-19 cases and deaths are as of April 20, 2020. Data on the local share of Catholics (at the county level) relative to the total county population refer to 2018. The correlation coefficient between the share of Catholics at the county level and the number of cumulative cases and cumulative deaths per 100,000 local population is 0.416 and 0.315, respectively.

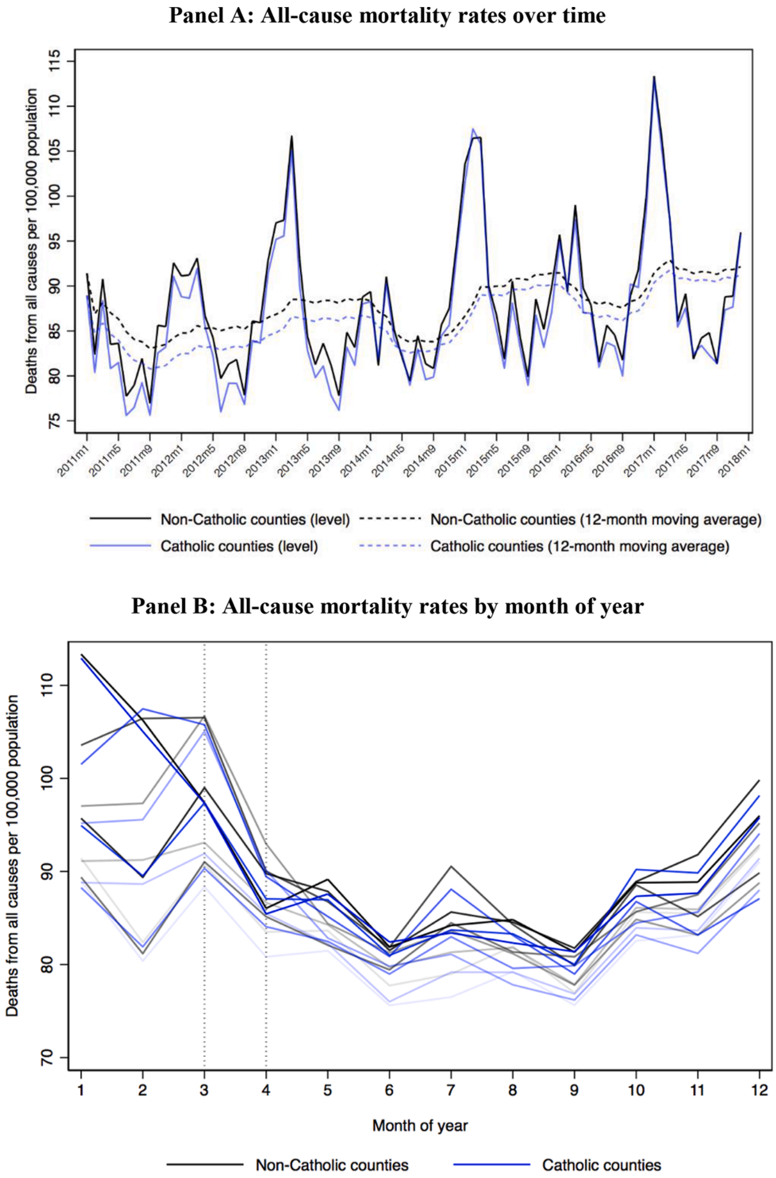


**Fig. A.2. All-cause mortality rate trends in Catholic and non-Catholic counties in Western Germany. Panel A: All-cause mortality rates over time. Panel B: All-cause mortality rates by month of year.** Source: Federal Statistical Office of Germany. Notes: All-cause mortality rate is defined as the total count of deaths per 100,000 local population. Catholic (non-Catholic) counties are defined as those where the local share of Catholics is higher (lower) than their respective state average. In Panel B, each line represents a different year during the 2011–2017 period for Catholic counties (blue lines) and for non-Catholic counties (black lines).

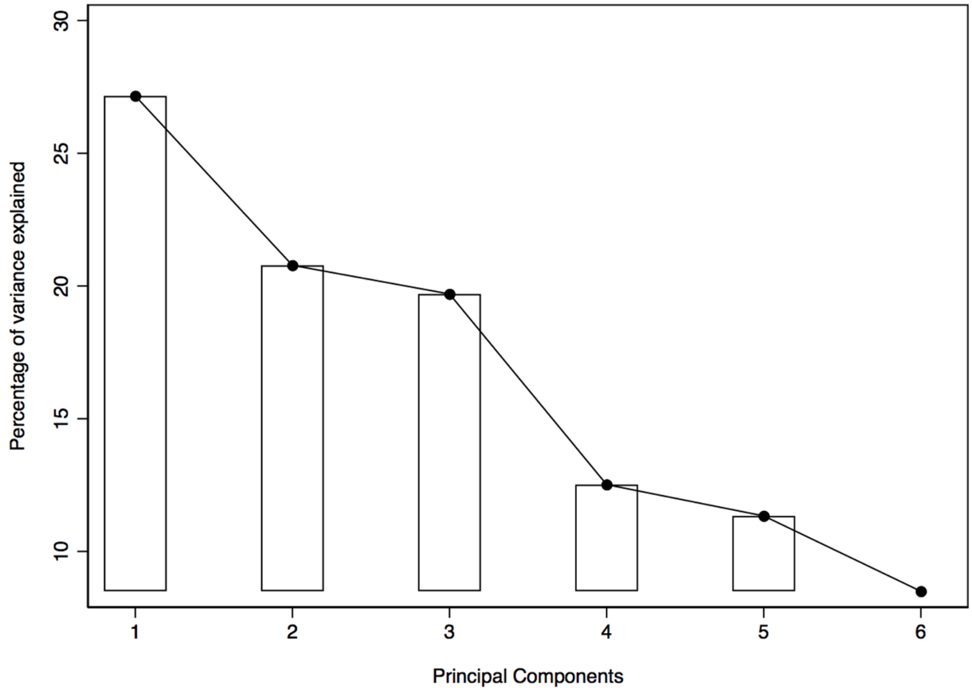


**Fig. A.3. Variation in the EVS data captured by each principal component.** Source: European Values Study 2017 (EVS). Notes: Principal components were constructed as linear combinations of the following variables indicating: (a) whether family is very important in life or not, (b) whether friends are very important in life or not, (c) if someone lives with their parents and/or their parents-in-law, (d) if someone attends religious services more frequently than once per month, (e) if someone lives in a small household, i.e. up to two members, or not, and (f) if someone agrees that children should provide care for their parents.

**Table A.1**
Descriptive statistics for county-level variables.

| | Catholic counties ($N$ = *147*) | | Non-Catholic counties ($N$ = *177*) | | Mean difference |
|---|---|---|---|---|---|
| | Mean [1] | Standard deviation [2] | Mean [3] | Standard deviation [4] | *t*-test (*[1]-[3])* [5] |
| COVID-19 cases per 100,000 population | 2.23 | 0.21 | 1.88 | 0.20 | 0.35** |
| COVID-19 deaths per 100,000 population | 0.10 | 0.03 | 0.06 | 0.02 | 0.04* |
| Cumulative COVID-19 cases per 100,000 population | 24.28 | 1.16 | 20.52 | 1.28 | 3.77** |
| Cumulative COVID-19 deaths per 100,000 population | 0.68 | 0.12 | 0.42 | 0.07 | 0.34** |
| Share (%) of Catholics | 55.5 | 18.9 | 25.3 | 13.7 | 30.2*** |
| GDP per capita (euros) | 70,892.0 | 11,810.5 | 71,248.4 | 13,561.0 | −356.3 |
| Population density | 371.6 | 460.4 | 757.5 | 844.6 | −385.9*** |
| Share (%) of people over 65 years old | 20.99 | 1.88 | 21.54 | 2.28 | −0.54** |
| Share (%) of foreign-born | 10.39 | 3.85 | 12.10 | 5.28 | −1.71** |
| Share (%) with completed secondary education | 29.81 | 9.38 | 33.78 | 8.16 | −3.96*** |
| Number of hospital beds per capita | 5.72 | 3.70 | 6.54 | 4.27 | −0.82* |
| Number of nights spent by tourists | 2.63 | 0.79 | 2.48 | 0.90 | 0.15 |

Source: Robert Koch Institute (RKI); Federal Statistics Office of Germany.
Notes: Statistics on COVID-19 cases and related deaths per population are calculated before the national lockdown (March 22, 2020). Catholic (non-Catholic) counties are defined as those where the local share of Catholics is higher (lower) than their respective state average. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

**Table A.2**
Descriptive statistics for individual-level variables.

| | Catholic individuals | | Non-Catholic individuals | | Mean difference |
|---|---|---|---|---|---|
| | Mean [1] | Standard deviation [2] | Mean [3] | Standard deviation [4] | *t*-test (*[1]–[3])* [5] |
| | **Panel A: European Social Survey** | | | | |
| Age (in years) | 53.7 | 18.2 | 52.9 | 18.5 | 0.77 |
| Female (%) | 47.8 | 50.0 | 46.1 | 49.9 | 1.71 |
| Good/very good health status (%) | 64.7 | 47.8 | 64.3 | 47.9 | 0.33 |
| Completed secondary education (%) | 50.1 | 50.1 | 51.8 | 50.0 | −1.67 |
| In paid employment (%) | 95.6 | 20.5 | 94.6 | 22.6 | 0.99 |
| Household size (number of members) | 2.6 | 1.2 | 2.6 | 1.2 | −0.05 |
| Upper half of total household income distribution (%) | 74.6 | 43.4 | 75.6 | 42.3 | −0.96 |
| Meet other people more than several times a week (%) | 40.6 | 49.2 | 37.6 | 48.5 | 3.00* |
| Meet other people more frequently relative to peers (%) | 64.8 | 47.8 | 60.6 | 48.9 | 4.20** |
| People discussing intimate and private matters with (number) | 3.7 | 1.2 | 3.5 | 1.3 | 0.20** |
| | **Panel B: European Values** Study | | | | |
| Age (in years) | 51.3 | 18.4 | 50.2 | 18.8 | 1.15* |
| Female (%) | 54.0 | 49.9 | 54.5 | 49.8 | −0.54 |
| Good/very good health status (%) | 60.3 | 48.9 | 59.5 | 49.1 | 0.78 |
| Completed secondary education (%) | 75.7 | 42.9 | 72.4 | 44.7 | 3.26** |
| In paid employment (%) | 57.6 | 49.4 | 56.5 | 49.6 | 1.10 |
| Household size (number of members) | 2.6 | 1.2 | 2.7 | 1.4 | −0.11** |
| Age of youngest household member (in years) | 31.8 | 24.3 | 29.1 | 24.3 | 1.76* |
| Upper half of total household income distribution (%) | 63.9 | 48.0 | 57.5 | 49.4 | 6.43*** |
| Living in city with ≥20,000 population | 51.6 | 50.0 | 52.9 | 50.0 | −1.28 |
| Family is important/very important in life (%) | 89.4 | 30.8 | 89.1 | 31.0 | 0.30 |
| Friends are important/very important in life (%) | 49.6 | 50.0 | 47.0 | 49.9 | 2.57* |
| Trust family members (%) | 85.2 | 35.5 | 83.8 | 36.8 | 1.34 |
| Living with parents and/or parents-in-law (%) | 14.5 | 35.1 | 14.0 | 34.8 | 0.40 |
| Attending religious services more than once a month (%) | 30.3 | 46.0 | 24.8 | 43.2 | 5.50*** |
| Religion is quite important in life (%) | 52.4 | 49.9 | 48.0 | 50.0 | 4.40** |
| Children should provide long-term care for parents (%) | 46.7 | 49.9 | 44.9 | 49.8 | 1.78* |
| Never justified to avoid taxes (%) | 87.1 | 33.4 | 85.8 | 34.9 | 1.40 |
| Never justified to accept bribery (%) | 92.4 | 26.6 | 91.0 | 28.6 | 1.31 |
| Never justified to avoid fare in public transport (%) | 76.7 | 42.3 | 73.7 | 44.0 | 2.87 |
| High confidence in government (%) | 41.4 | 49.3 | 40.4 | 49.1 | 0.93 |

Source: European Social Survey (2018); European Values Study (2017).
Notes: Catholics and non-Catholics are distinguished on the basis of self-reported religious affiliation in the ESS and EVS. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

**Table A.3**
Difference in mortality rates from all causes between Catholic and non-Catholic counties in Western Germany during the pre-pandemic period (2011–2017).

| | [1] | [2] | [3] | [4] |
|---|---|---|---|---|
| Local Catholic share (log) | .044 (0.054) | .018 (0.053) | – | – |
| Catholic county indicator (0/1) | – | – | .004 (0.054) | .007 (0.049) |
| County controls | No | Yes | No | Yes |
| County fixed effects | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes |
| $R$-squared | .429 | .531 | .428 | .529 |
| Observations | 2254 | 2226 | 2256 | 2.228 |
| Counties | 322 | 322 | 323 | 323 |

Source: Federal Statistics Office of Germany.
Notes: OLS estimates. The local Catholic share is calculated relative to the county's total population. Catholic (non-Catholic) counties are defined as those where the local share of Catholics is higher (lower) than their respective state average. County controls include the share of people over than 65 years old, the share of foreign-born, GDP per capita, population density, the share of those completed secondary education, the average number of nights spent by tourists, the number of hospital beds per capita. All variables are in logs unless mentioned otherwise. Standard errors in parentheses are clustered at the county level. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

**Table A.4**
Relationship between COVID-19 incidence and share of Catholics in Western German counties: Sensitivity checks.

| Explanatory variable of interest: | Share of county population | | Binary indicator (0/1) | |
|---|---|---|---|---|
| Outcome (logged count per 100,000 population): | Cases | Deaths | Cases | Deaths |
| Specification / estimation sample: | [1] | [2] | [3] | [4] |
| Controlling for day-of-week fixed effects | .060** (0.026) | .013*** (0.005) | .066*** (0.023) | .011** (0.005) |
| Controlling for day-of-week X state fixed effects | .060** (0.026) | .013** (0.005) | .066*** (0.023) | .011** (0.005) |
| Excluding Bavarian counties | .049** (0.025) | .010* (0.006) | .046** (0.022) | .008 (0.006) |
| Excluding Bavarian & Baden-Württemberg counties | .078*** (0.025) | .011** (0.006) | .070*** (0.022) | .011* (0.006) |
| Excluding counties next to any border area | .067*** (0.025) | .011** (0.006) | .080*** (0.021) | .010** (0.005) |
| Excluding counties close to any border area | .059** (0.026) | .006 (0.006) | .074*** (0.022) | .005 (0.005) |
| Controlling for temperature | .086** (0.043) | .012 (0.009) | .089** (0.040) | .010 (0.010) |
| Controlling for precipitation | .097** (0.045) | .014 (0.009) | .103** (0.043) | .009 (0.010) |
| Controlling for driving distance from Milan, Italy | .088*** (0.031) | .017*** (0.006) | .090*** (0.026) | .013** (0.005) |
| North Rhine-Westphalia counties excluding Heinsberg | .086** (0.037) | .019** (0.008) | .094*** (0.032) | .015** (0.007) |
| Excluding cities with more than 200,000 population | .051** (0.024) | .016** (0.006) | .046** (0.023) | .012* (0.006) |
| Replacing Catholic with Protestant indicators | −0.032 (0.023) | −0.009 (0.006) | −0.028 (0.024) | −0.007 (0.005) |
| Controlling for local Protestant share | .069* (0.038) | .014* (0.008) | .072*** (0.027) | .009 (0.006) |

Source: Robert Koch Institute (RKI); Federal Statistical Office of Germany; Google Maps; Climate Data Centre.
Notes: OLS estimates. Each cell reports results from a separate regression. Outcomes are expressed as the logged count of cases (or deaths) per 100,000 local population. In columns 1–2 the local (logged) Catholic share is calculated relative to the county's total population. In columns 3–4 Catholic (non-Catholic) counties are defined as those where the local share of Catholics is higher (lower) than their respective state average. County controls include the share of people over than 65 years old, the share of foreign-born, GDP per capita, the share of those completed secondary education, the average number of nights spent by tourists, the number of hospital beds per capita. We exclude counties that border on other countries and in a second specification all those that are within 20 km from a border. Temperature is measured in Celsius and Precipitation in centimetres and refer to the monthly average recorded in a weather station in a particular county. Specification also include state fixed effects and state-specific time trends. Regressions are weighted by the size of the county population. Standard errors in parentheses are clustered at the county level. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

**Table A.5**
Differences in behavioural indicators between Catholics and Protestants in Western Germany.

| | Catholic parameter estimate [1] | Unbiased Catholic parameter estimate [2] | Individual controls [3] | State fixed effects [4] | *R*-squared [5] | Observations [6] |
|---|---|---|---|---|---|---|
| | **Panel A: Social interaction outcomes (ESS data)** | | | | | |
| Meet other people more than several times a week | .094** (0.047) | .112 | Yes | Yes | .068 | 535 |
| Meet other people more frequently relative to peers | −0.007 (0.047) | −0.003 | Yes | Yes | .069 | 535 |
| Discuss intimate and private matters with other people | .106 (0.121) | .112 | Yes | Yes | .089 | 535 |
| | **Panel B: Social and family ties outcomes (EVS data)** | | | | | |
| Family important in life | .018 (0.011) | .018 | Yes | Yes | .076 | 1928 |
| Friends important in life | .038 (0.026) | .041 | Yes | Yes | .054 | 1921 |
| Trust towards family members | .054* (0.028) | .054 | Yes | Yes | .049 | 1402 |
| Reside with parents or parents-in-law | −0.003 (0.008) | −0.002 | Yes | Yes | .257 | 1950 |

Source: European Social Survey 2018 (ESS); European Values Study 2017 (EVS).
Notes: OLS estimates. All outcomes are binary. Each row refers to a separate regression with a different outcome variable. Individual-level controls include perceived health status, gender, age, household income, education level, employment status, age of youngest household member, as well as city size. The unbiased Catholic parameter estimate has been obtained using the methodology proposed by Oster (2019). Standard errors in parentheses are clustered at the state level. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

**Table A.6**
Falsification test: Commuting mobility differences in Western German counties, January 2020–May 2020.

| | Total period [1] | January [2] | February [3] | March [4] | April [5] | May [6] |
|---|---|---|---|---|---|---|
| Local Catholic share (log) | −0.110 (0.502) | −0.044 (0.712) | −1.010 (0.736) | .698 (0.558) | −0.349 (0.603) | .156 (0.562) |
| R-squared | .873 | .159 | .159 | .432 | .522 | .487 |
| Catholic county indicator (0/1) | .127 (0.467) | .517 (0.657) | −0.663 (0.623) | .774 (0.527) | −0.031 (0.550) | .037 (0.502) |
| *R*-squared | .873 | .162 | .157 | .434 | .521 | .487 |
| County controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Month fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| State fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1550 | 310 | 310 | 310 | 310 | 310 |

Source: Federal Statistical Office of Germany; Teralytics AG.
Notes: OLS estimates. The local Catholic share is calculated relative to the county's total population. Catholic (non-Catholic) counties are defined as those where the local share of Catholics is higher (lower) than their respective state average. County controls include the share of people over than 65 years old, the share of foreign-born, GDP per capita, population density, the share of those completed secondary education, the average number of nights spent by tourists, the number of hospital beds per capita, and the number of vehicles per capita. All variables are in logs unless mentioned otherwise. Standard errors in parentheses are clustered at the county level. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

**Table A.7**
COVID-19 transmission rate in Catholic and non-Catholic counties.

| | [1] | [2] | [3] | [4] |
|---|---|---|---|---|
| Within county ($\alpha$) × Catholic county (0/1) | .070*** (0.025) | .040** (0.016) | .068*** (0.025) | .039** (0.017) |
| Within county ($\alpha$) | .169*** (0.034) | .206*** (0.020) | .165*** (0.034) | .203*** (0.020) |
| *R*-squared | .104 | .163 | .150 | .169 |
| Daily time trend | Yes | Yes | Yes | Yes |
| County controls | No | Yes | Yes | Yes |
| State fixed effects | Yes | Yes | Yes | Yes |
| State linear trends | No | No | Yes | Yes |
| Observations | 17,568 | 17,061 | 17,568 | 17,061 |
| Number of counties | 324 | 314 | 324 | 314 |

Source: Robert Koch Institute (RKI); Federal Statistical Office of Germany.
Notes: OLS estimates. Outcomes are expressed as the logged count of cases per 100,000 local population. Catholic (non-Catholic) counties are defined as those where the local share of Catholics is higher (lower) than their respective state average. County controls include the share of people over than 65 years old, the share of foreign-born, GDP per capita, the share of those completed secondary education, the average number of nights spent by tourists, the number of hospital beds per capita. Regressions are weighted by the size of the county population. Standard errors in parentheses are clustered at the county level. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

### *A.1. A short note on the Standard Inflammatory Response (SIR) model*

Standard epidemiological modelling relies on variants of the Standard Inflammatory Response (SIR) model. The idea behind SIR models is to predict the course of a pandemic over time based on the fractions of susceptible ($S$), infected ($I$), and recovered ($R$) individuals within a population, with $S+I+R=1$. An important feature of such models is that they can be modified to estimate not only the within-county transmission rate, i.e. parameter $\alpha$ in Eq. (2), but also differences in transmission rates based on some county-level observables. The seminal work of Adda (2016) popularised this sort of model in the economics literature by examining how policies, such as school closures, and economic interactions affect virus transmission. For example, if policies to mitigate the pandemic impact were enacted in some counties but not in others, their efficiency could be assessed by testing for differences in virus transmission rates between counties. This idea can be expanded to observe the differential impact of any kind of "treatment". Therefore, to make this adaptation in our setting, Eq. (2) was modified to include an interaction between the term capturing the within-county spread of the virus and the Catholic county binary indicator. Although our paper focuses on differences in the observed incidence between Catholic and non-Catholic counties, documenting a difference in the spread of the virus using a modification of a SIR model adds credence to our results. A positive and statistical significant estimate of the parameter associated with this interaction term would indicate that the transmission rate was significantly higher in Catholic counties during the early pandemic. The results are in Table A.6. Our outcome variable is the logged count of COVID-19 cases per 100,000 population, and the models are estimated using OLS. In columns 1–4, specifications include combinations of county controls, state fixed effects and state-specific time trends. As in Table 2, past infections are significant drivers of current within-county COVID-19 incidence. Moreover, and in line with our results about Catholics exhibiting more frequent social interactions, the transmission rate is higher in predominantly Catholic counties, as indicated by the estimated coefficient of the interaction term.

## References

Adda, J., 2016. Economic activity and the spread of viral diseases: evidence from high frequency data. Q. J. Econ. 131 (2), 891–941.
Alesina, A., Giuliano, P., 2015. Culture and Institutions. J. Econ. Lit. 53 (4), 898–944.
Alesina, A., Harnoss, J., Rapoport, H., 2016. Birthplace diversity and economic prosperity. J. Econ. Growth 21, 101–138.
Alfaro, L., Faia, E., Lamersdorf, N., & Saidi, F. (2020). Social interactions in pandemics: fear, altruism and reciprocity. *NBER* Working Paper 27134.
Arruñada, B., 2009. Protestants and Catholics: similar work ethic, different social ethic. Econ. J. 120 (547), 890–918.
Bargain, O., Aminjonov, U., 2020. Trust and compliance to public health policies in times of COVID-19. J. Public Econ. 192, 1–13.
Bartscher, A.K., Seitz, S., Siegloch, S., Slotwinski, M., WehrhÖfer, N., 2021. Social capital and the spread of COVID-19: insights from European countries. J. Health. Econ. 80, 102531.
Bayer, C., & Kuhn, M. (2020). Intergenerational ties and case fatality rates: a cross-country analysis. *IZA* Discussion Paper 13114.
Becker, G., 1996. Preferences and values. *Accounting for Taste*, Gary Becker. Harvard University Press, Cambridge.
Becker, S.O., Mergele, L., Woessmann, L., 2020. The separation and reunification of Germany: rethinking a natural experiment interpretation of the enduring effects of Communism. J. Econ. Perspect. 34 (2), 143–171.
Becker, S.O., Pascali, L., 2019. Religion, division of labor, and conflict: anti-Semitism in Germany over 600 years. Am. Econ. Rev. 109 (5), 1764–1804.
Becker, S.O., Woessmann, L., 2009. Was Weber wrong? A human capital theory of protestant economic history. Q. J. Econ. 124 (2), 531–596.
Becker, S.O., Woessmann, L., 2018. Social Cohesion, Religious Beliefs, and the Effect of Protestantism on Suicide. Rev. Econ. Stat. 100 (3), 377–391.
Belloc, M., Buonanno, P., Drago, F., Galbiati, R., Pinotti, P., 2020. Cross-country correlation analysis for research on COVID-19. VoxEU.
Bol, D., Giani, M., Blais, A., Loewen, P.J., 2020. The effect of COVID-19 lockdowns on political support: some good news for democracy? Eur. J. Polit. Res. 60 (2), 497–505.
Borgonovi, F., Andrieu, E., 2020. Bowling together by bowling alone: social capital and COVID-19. Soc. Sci. Med. 265, 1–12.
Brodeur, A., Grigoryeva, I., Kattan, L., 2021. Stay-at-home orders, social distancing and trust. J. Popul. Econ. 34, 1321–1354.
Christidis, P., Christodoulou, A., 2020. The predictive capacity of air travel patterns during the global spread of the COVID-19 pandemic: risk, uncertainty and randomness. Int. J. Environ. Res. Public Health 17 (10), 3356.
Delanghe, J., Speeckaert, M., De Buyzere, M., 2020. The host's angiotensin-converting enzyme polymorphism may explain epidemiological findings in COVID-19 infections. Clin. Chim. Acta 505, 192–193.
Deopa, N., Fortunato, P., 2020. Coronagraben. Culture and social distancing in times of COVID-19. *COVID Econ.*, CEPR Press 39, 186–216.
Deslandes, A., Berti, V., Tandjaoui-Lambotte, Y., Alloui, C., et al., 2020. SARS-Cov-2 was already spreading in France in late December 2019. Int. J. Antimicrob. Agents 55 (6), 1–5.
Ding, W., Levine, R., Lin, C., & Xie, W. (2020). Social distancing and social capital: why U.S. counties respond differently to COVID-19. *NBER* Working Paper 27393.
Dowd, J., Rotondi, V., Andriano, L., Brazel, D., et al., 2020. Demographic science aids in understanding the spread and fatality rates of COVID-19. Proc. Natl. Acad. Sci. 117, 9696–9698.
Durante, R., Guiso, L., Gulino, G., 2021. Asocial capital: civic structure and social distancing during COVID-19. J. Popul. Econ. 194, 104342.
Ekelund, R., Hébert, R., Tollison, R., 2002. An economic analysis of the Protestant Reformation. J. Polit. Econ. 110 (3), 646–671.
Georganas, S., Velias, A., Vandoros, S., 2021. On the measurement of disease prevalence. *COVID Econ.*, CEPR Press 69, 109–139.
Glaeser, E., Glendon, S., 1998. Incentives, predestination and free will. Econ Inq 36 (3), 429–443.
Gudbjartsson, D.F., Helgason, A., Jonsson, H., Magnusson, et al., 2020. Spread of SARS-CoV-2 in the Icelandic population. N. Engl. J. Med. 382, 2302–2315.
Guiso, L., Sapienza, P., Zingales, L., 2006. Does culture affect economic outcomes? J. Econ. Perspect. 20 (2), 23–48.
Hargreaves Heap, S.P., Koop, C., Matakos, K., Unan, A., & Weber, N. (2020). COVID-19 and people's health-wealth preferences: information effects and policy implications. *QPE* Working Paper 2020-5.
Iannaccone, L., 1998. Introduction to the economics of religion. J. Econ. Lit. 36 (3), 1465–1495.
Lauer, S.A., Grantz, K.H., Bi, Q., Jones, F.K., et al., 2020. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann. Intern. Med. 172 (9), 577–582.
Liu, Y., Gu, Z., Xia, S., Shi, B., et al., 2020. What are the underlying transmission patterns of COVID-19 outbreak? An age-specific social contact characterization. EclinicalMedicine 22 (100354), 1–10.
Mazzonna, F., 2020. Cultural differences in COVID-19 spread and policy compliance: evidence from Switzerland. COVID Econ., CEPR Press 33, 163–185.
Oster, E., 2019. Unobservable selection and coefficient stability: theory and evidence. J. Bus. Econom. Statist. 37 (2), 187–204.
Pesaran, M.H., Zhou, Q., 2018. Estimation of time-invariant effects in static panel data models. Econom. Rev. 37 (10), 1137–1171.
Platteau, J., Verardi, V., 2020. How to exit COVID-19 lockdowns: culture matters. COVID Econ., CEPR Press 23, 1–57.
Plümper, T., Neumayer, E., 2020. The pandemic predominantly hits poor neighbourhoods? SARS-CoV-2 infections and COVID-19 fatalities in German districts. Eur. J. Public Health 30 (6), 1176–1180.
Remland, M.S., Jones, T.S., Brinkman, H., 1995. Interpersonal distance, body orientation, and touch: effects of culture, gender, and age. J. Soc. Psychol. 135 (3), 281–297.
Sa, F., 2020. Socioeconomic determinants of COVID-19 infections and mortality: evidence from England and Wales. COVID Econ., CEPR Press 22, 47–58.
Satyanath, S., Voigtländer, N., Voth, H.J., 2017. Bowling for fascism: social capital and the rise of the Nazi Party. J. Polit. Econ. 125 (2), 478–526.
Sorokowska, A., Sorokowski, P., Hilpert, P., Cantanero, K., et al., 2017. Preferred interpersonal distances: a global comparison. J. Cross. Cult. Psychol. 48 (4), 577–592.
Spenkuch, J.L., Tillmann, P., 2017. Elite influence? Religion and the electoral success of the Nazis. Am. J. Polit. Sci. 62 (1), 19–36.
Spenkuch, J.L., 2017. Religion and work: micro evidence from contemporary Germany. J. Econ. Behav. Organ. 135, 193–214.
Stafford, N., 2020. COVID-19: why Germany's case fatality rate seems so low. Br. Med. J. 369, m1395.
Tabellini, M., 2020. Gifts of the immigrants, woes of the natives: lessons from the age of mass migration. Rev. Econ. Stud. 87 (1), 454–486.
Van Bavel, J.J., Baicker, K., Boggio, P.S., Capraro, V., et al., 2020. Using social and behavioural science to support COVID-19 pandemic response. Nature Hum. Behav. 4, 460–471.
Vandoros, S., 2020. Excess mortality during the COVID-19 pandemic: early evidence from England and Wales. Soc. Sci. Med. 258, 113101–113104.
Zhao, S., Lin, Q., Ran, J., Musa, S.S., et al., 2020. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak. Int. J. Infect. Dis. 92, 214–217.
EVS (2020): European Values Study 2017: Integrated Dataset – Sensitive Data (EVS2017 Sensitive Data). GESIS Data Archive, Cologne. ZA7501 Data file Version 1.0.0, doi:10.4232/1.13091.
EVS (2020): European Values Study 2017: Integrated Dataset (EVS2017). GESIS Data Archive, Cologne. ZA7500 Data file Version 4.0.0, doi:10.4232/1.13560.
ESS Round 9: European Social Survey Round 9 Data (2018). Data file edition 3.1. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC. doi:10.21338/NSD-ESS9-2018.