

## RESEARCH ARTICLE

# Regional variations in vaccination against COVID-19 in Germany

Verena Bade<sup>1</sup> , Hendrik Schmitz<sup>1,2,3</sup> , Beatrice Baaba Tawiah<sup>4</sup> 

1 Paderborn University, Paderborn, Germany, 2 RWI – Leibniz Institute for Economic Research, Essen, Germany, 3 Leibniz Science Campus Ruhr, Essen, Germany, 4 Munich Research Institute for the Economics of Aging and SHARE Analyses, Munich, Germany

 These authors contributed equally to this work.

\* [hendrik.schmitz@uni-paderborn.de](mailto:hendrik.schmitz@uni-paderborn.de)



## Abstract

Vaccination willingness against COVID-19 is generally perceived as low. Moreover, there is large heterogeneity across and within countries. As a whole, Germany has average vaccination rates compared to other industrialized countries. However, vaccination rates in the 16 different German federal states differ by more than 20 percentage points. We describe variation in vaccination rates on the level of the 400 German counties using data on all vaccinations carried out until December 2022. Around 52–72% of that variation can be explained by regional differences in demographic characteristics, housing, education and political party preferences. We find indications that the remaining part may be due to differences in soft factors such as risk aversion, trust in the German government, trust in science, and beliefs in conspiracy theories regarding the origins of the Corona virus. We conclude that improving the trust in science and the fight against conspiracy theories may possibly be effective tools to improve vaccination rates and effectively fight pandemics.

## OPEN ACCESS

**Citation:** Bade V, Schmitz H, Tawiah BB (2024) Regional variations in vaccination against COVID-19 in Germany. PLoS ONE 19(4): e0296976. <https://doi.org/10.1371/journal.pone.0296976>

**Editor:** Simon Porcher, Universite Paris Pantheon-Assas, FRANCE

**Received:** March 3, 2023

**Accepted:** December 25, 2023

**Published:** April 18, 2024

**Copyright:** © 2024 Bade et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The data underlying the main results presented in the study are available from Mendeley data (<https://data.mendeley.com/drafts/83sb8nxw4f>). Data for the supplementary analyses cannot be made available due to data protection and data security laws. Data are available from the European Social Survey's (ESS) data portal (<https://ess-search.nsd.no/>), and the German Socio-Economic Panel's (SOEP) Data Access/Ethics Committee ([https://www.diw.de/en/diw\\_01.c.601584.en/data\\_access.html](https://www.diw.de/en/diw_01.c.601584.en/data_access.html)) for researchers who meet the criteria for access.

## Introduction

In order to curb the SARS-CoV-2 virus, vaccines were developed as quickly as possible in 2020 and enormous vaccination campaigns started by the end of 2020 in the industrialized countries. In the first months of 2021, undersupply of vaccines was the major problem of these vaccination campaigns. Yet, this changed throughout the year and, later, too low vaccination willingness was considered the major challenge to yield herd immunity [1–3]. Herd immunity, in turn, was among the most important health policy goals in most countries in order to overcome the Corona crisis. However, vaccination willingness against COVID-19 is generally perceived as low [4–6]. While several mutations of the virus made the original vaccinations less effective—but also reduced the severity of the virus [7]—there is little controversy that vaccination still is considered a highly effective tool to reduce mortality and severe illness after a COVID-19 infection [8, 9]. This holds in particular in comparison with other measures such as lockdowns or school closures which have extremely high social costs [10–13].

Understanding the determinants of vaccination status is necessary to improve the success and acceptance of vaccination campaigns—both for future pandemics and also for endemic

**Funding:** HS acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, <https://www.dfg.de/>) with project number 492351283. HS acknowledges support for the publication cost by the Open Access Publication Fund of Paderborn University.

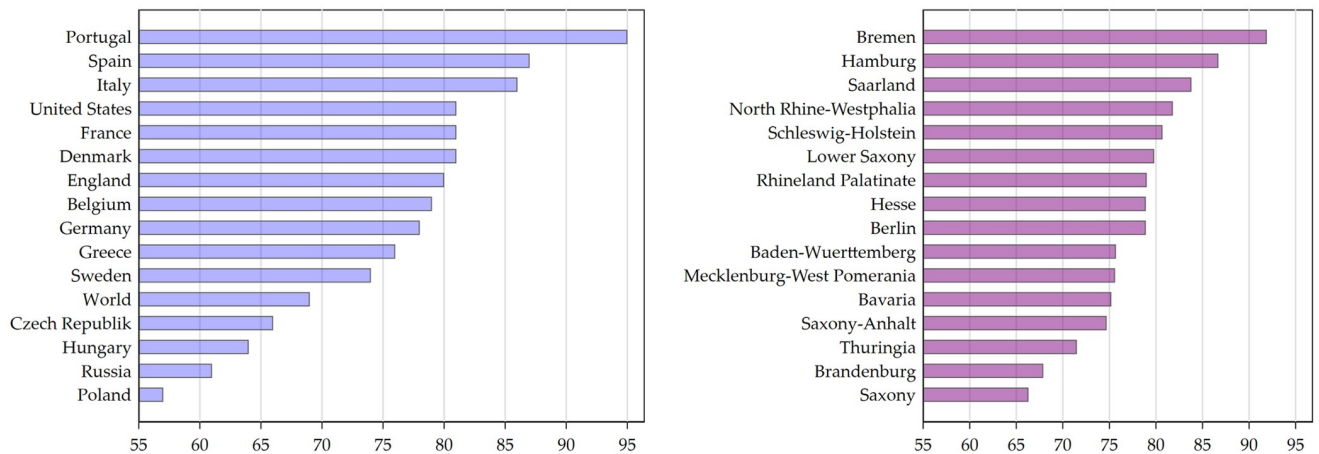
**Competing interests:** The authors have declared that no competing interests exist.

viruses such as influenza and potential mutations of the Corona virus [14]—even more so as vaccination hesitancy is not peculiar to just the COVID-19 vaccines but vaccines in general [15, 16]. Another dimension of vaccination hesitancy is the one towards the COVID-19 vaccine booster shots [17–19], which is even stronger compared to the first two doses of COVID-19 vaccines [20]. One of many potentially fruitful ways to study the determinants of vaccination status is to analyze the determinants of regional differences in vaccination rates. Studying regional variations helps in at least three ways. First, it allows to get a benchmark of which vaccination rates might be achievable in a thought experiment where all regions have the same rates as, say, the region with the highest ones, or, for instance, as those in the upper quartile. It also makes transparent how much room there is for improvement, that is, if all regions are fairly similar or some lack behind by large amounts. Second, it enables to study which observable variables account for the regional differences. In principle—although beyond the scope of this paper—one might figure out what regions with low rates may learn from regions with higher rates in order to increase vaccination rates. Yet, not each and every difference between regions is necessary as sign of inefficiencies. Thus, as a third point, the analysis of regional variations may help to understand which degree of differences is acceptable, for instance, because differences in age structure might imply justified differences in vaccination rates.

As a point of departure of our analysis, the left panel of Fig 1 reports the share of people with at least two Corona vaccinations (typically considered as those with completed immunization) across selected European countries and the USA. Among the European countries, Germany—the country we study in this paper—is somewhere in the middle with about 78% of its population having completed the initial vaccination protocol, i.e., two vaccination doses for most vaccines. However, only looking inside Germany—as the right panel of Fig 1 does—we see that variation within Germany is almost as large as the one across the countries in the left panel. The right panel shows vaccination rates in the 16 German federal states. For instance, while the share of vaccinated people in Saxony is as low as the mean in the Czech Republic (around 66%), the share in the Saarland (84%) is almost as large as the share in leading European states such as Italy and Spain. With more than 20%-point differences across German regions, it seems important to understand what factors might be associated with these differences. In this paper, we analyze correlations of regional variation in Corona vaccination in Germany. We hypothesize that differences in demographic status, levels of education in the population and economic situation could account for some of these differences. We further hypothesize that political preferences and susceptibility to conspiracy theories also add to regional differences in vaccination rates.

## Background

Fig 1 shows considerable regional variations in COVID-19 vaccine uptake. This has also been found in other countries. Most of the studies on the regional variation in COVID-19 vaccine uptake are based on the USA. [21] find more variation across counties than states and [22] find more variation in vaccination levels across postal codes within cities than between cities. A common source of variations in COVID-19 vaccinations in the US is racial disparities. However, the findings are mixed: [22] find higher vaccination rates in areas with a high share of white or Asian people and [23] find lower vaccination rates in regions with a high share of Black people. These findings differ from the results of [24] who find that minorities, defined as all persons except white, non-Hispanic, have a positive impact on vaccination rates, and from the results of [2] who finds that regions with a high share of non-Hispanic white residents have lower vaccination rates. Political affiliation has also been associated with vaccine uptake, where counties with a large proportion of Republican voters have lower vaccine uptake [2, 23].



**Fig 1. Share of people vaccinated against COVID-19.** Vaccinated at least twice as of January 2023. Left panel: Worldwide. Source: Our world in data, <https://ourworldindata.org/covid-vaccinations>; Right panel: German states. Source: Robert Koch Institut, [https://www.rki.de/DE/Content/InfAZ/N/Neuartiges\\_Coronavirus/Daten/Impfquoten-Tab.html](https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Daten/Impfquoten-Tab.html).

<https://doi.org/10.1371/journal.pone.0296976.g001>

Other socio-demographic and socio-economic factors which have also been related with the regional variation in vaccine uptake include the share of population below the poverty line, the share of uninsured persons, the share without internet access and the share of younger persons, where higher shares are associated with lower vaccination rates, and the share of persons with a college degree, the share of older persons and per capita income, where higher levels are associated with higher vaccination rates [2, 21–24].

Some of these factors associated with COVID-19 uptake also correlate with compliance or non-compliance with other COVID-19 preventive measures. Political affiliation is mostly found to be associated with the compliance or non-compliance with COVID-19 preventive measures, where Democrats in the USA are more welcoming to measures such as social distancing [25–27]. Poverty [28] and the share of younger population [29] have been associated with lower compliance with social distancing.

For Germany, [30] find considerable variations in vaccination rates between federal states: states in West Germany have higher vaccination rates than those in East Germany. They also find that support for the populist right-wing political party, Alternative for Germany (AfD), is correlated with lower vaccination rates at the state level. Among the studies done on regional variations in COVID-19 vaccinations, only a few, such as [23] and [24], have tried to explain the variations. All in all, with an adjusted  $R^2$  of about 76–79% these studies can explain a considerable share of variation in vaccination rates in the USA by observable characteristics. Our study contributes to the literature by bringing in a new perspective from not only a different country but a country with a different health care system.

Our paper also adds to a growing field of health economic research on regional variations in the health care market. Strong regional variations have been found in many dimensions across different countries. Differences have been found in health [31], health expenditures [32–38], access to healthcare [39], medical services utilization [40–45], and productivity in the health care sector [46]. These differences may be linked to the variation in vaccine uptake. [23] find that counties with more health facilities have higher vaccination rates. They argue that a lack of these health facilities may have resulted in the delay or loss of opportunity to vaccinate in some regions.

We study regional variations in the share of at least double-vaccinated persons in Germany using administrative data on all vaccinated doses in all 400 German counties until December

2022. Using a host of potentially relevant socio-economic variables (guided from determinants found in the previous literature) we can explain about 52 per cent of regional variations in vaccination on county level and 72 per cent on federal state level. Major explanatory factors are demographic characteristics, levels of education, and further preferences proxied by political party vote shares. To get an idea of what might account for the remaining part, we describe the federal state level distribution in selected “soft” factors and document that individuals in federal states with lower vaccination rates, on average, report to be less risk averse, have less trust in the German government and scientists, and are more likely to believe in conspiracy theories regarding the origins of the Corona virus. While our study design is not able to provide any causal evidence and, thus, does not allow to immediately derive policy recommendations, it adds to our knowledge on the picture of regional variations in health care.

## Methods

### Vaccination rates

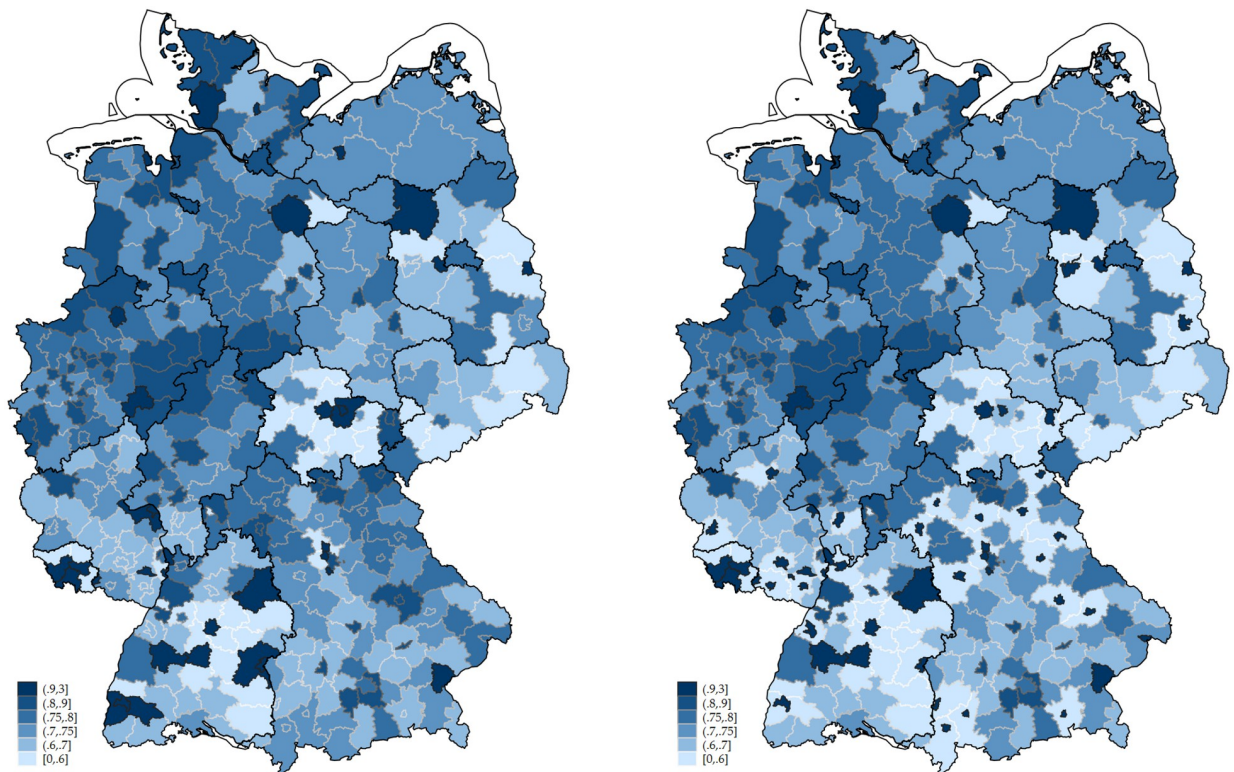
Our main dependent variable is the share of individuals with at least two vaccinations against the Corona virus. It is based on freely accessible and continuously updated data from the German Robert Koch Institute (RKI) [47] that publishes data on the universe of daily vaccinated doses per age group on the level of the 400 German counties. German counties can either be rural areas (*Landkreis*,  $N = 293$ ) or single larger cities (*Kreisfreie Stadt*,  $N = 107$ ). The RKI vaccination data are assigned to the county where the vaccination took place and not where the vaccinated individual lives. In principle, individuals can freely choose where to get vaccinated and it may be possible that they get vaccinated in neighbouring counties. In particular, it seems to be frequently done that individuals in rural areas get vaccinations in the next bigger cities where larger vaccination centers had been installed.

Fig 2 shows the number of second corona vaccinations—usually defined as “complete vaccination protocol”—per inhabitant of the corresponding county until end of December 2022, calculated as follows:

$$\text{Share of second vaccination per inhabitant}_c = \frac{SV_c}{I_c}$$

where  $c$  is the county,  $SV_c$  is the sum of second vaccination doses carried out in county  $c$  until end of 2022, and  $I_c$  is the number of inhabitants living in county  $c$ . The left panel shows the map of the 400 German counties where darker colors imply larger shares of vaccinated people and lighter ones imply lower shares. A striking pattern are the small dark dots, mainly in the south of Germany. These are larger cities completely surrounded by rural counties. Often, individuals from these rural counties got their vaccination in the larger cities, leading to an overestimation of actual vaccination rates in these cities and an underestimation in the rural areas. To account for this, we manipulate the data in the following way: for the 36 cases of cities completely surrounded by one rural county, we assign both the same vaccination rate of the sum of vaccination doses divided by the sum of inhabitants in city and rural county. S1 Table in the Supporting Information lists all these counties. The resulting regional distribution of the share is shown in the right panel of Fig 2. In our baseline analysis, we will work with this outcome variable, that is, based on partly merging certain counties. Robustness checks with the original variable, however, yield results of similar magnitude, implying that crossing county borders do not completely drive the results.

The mean share of individuals with two Covid vaccinations is 0.75, close to the German average reported in Fig 2, the standard deviation is 0.15. Five counties (cities) have a share larger than one while the five counties with lowest numbers have rates between 0.34 and 0.41.



**Fig 2. Share of 2nd vaccination per inhabitant.** Left panel: Original data; Right panel: Cities merged with rural counties. Data from the Robert Koch Institute. Number of second Corona vaccinations per county until December 2022, divided by inhabitants per county. The left panel uses the original data from 400 counties. The right panel presents average numbers for 36 pairs of cities surrounded by rural counties, see text for explanations. The basemap bases on data from GeoBasis-DE / BKG (2022), dl-de/by-2-0 ([www.govdata.de/dl-de/by-2-0](http://www.govdata.de/dl-de/by-2-0)), <https://gdz.bkg.bund.de/index.php/default/digitale-geodaten/verwaltungsgebiete/verwaltungsgebiete-1-5-000-000-stand-31-12-vg5000-12-31.html>.

<https://doi.org/10.1371/journal.pone.0296976.g002>

In robustness checks we will exclude these 10 counties. The map reveals strong variations in Corona vaccinations in Germany, both across federal states (marked by the bold borders), thereby repeating the pattern from Fig 1, as well as within states. North-West Germany has a significantly higher number of counties with larger vaccination rates than the other parts of Germany. Particularly the eastern states (Thuringia, Saxony, Brandenburg) and southern states (Baden-Wuerttemberg, Bavaria) have many counties with less than 70% vaccination coverage.

### Pool of variables on the county level

We obtain county level demographic and socio-economic variables from the INKAR database [48]. INKAR is an interactive online atlas of the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR), containing about 600 indicators, which enable rural-urban comparisons and analyses over the last decades. INKAR provides current regional statistics on socially important topics such as education, demography, job market, economy, accommodation, transportation, and the environment.

For the regressions, we consider those variables that are associated with vaccine uptake or hesitancy. As variations in the age structure could influence the variation in vaccinations given that older age groups were initially given preference over others when the vaccines were made available, we take into account differences in the demography by considering the age structure

and also include the share of male persons, the share of non-natives, the share of people in need of long-term care, life-expectancy and fertility rate.

Previous exposure of counties to the COVID-19 pandemic might affect the willingness to vaccinate. Thus, we account for the cumulative sum of COVID-19 infections per county until December 2020, that is, at the time vaccination campaigns started. Moreover, since the spread of the COVID-19 virus is through the air, the living situation is likely to have an impact on its spread which in turn could have an influence on vaccination uptake. We, therefore, include variables that refer to the housing conditions such as the share of residential buildings, the different apartment sizes based on the number of rooms, per capita living space and the share of apartments in residential buildings.

Further potentially contributing factors to the rate of vaccine uptake are the location and its economy, education, political affiliation and healthcare accessibility. Thus, we consider the rurality, the distance to centers, net commuting and cadastral area as well as gross earnings, taxes, debts, GDP, unemployment rate and the share of people receiving minimum social security benefits, as well as the number of pupils, university students, and different educational levels of employees. In addition, we include the share of voter turnout and the votes for the different political parties and medical supply by taking into account the number of hospital beds and physicians, and the distance to the nearest pharmacy. All variables are listed in [Table 1](#) below.

### Further variables measured on the federal state level for supplementary analyses

On the individual level, different aspects and types of trust were identified as determinants of vaccination [49, 50]. Therefore, in supplementary analyses described below, we also consider four questions from the 10th wave of the European Social Survey (ESS) [51] and the German Socio-Economic Panel (SOEP) [52]—the most prominent and long-running German representative household data set. SOEP is based on about 22,000 participants and the latest wave was carried out in 2020. The German part of the ESS covers around 8,700 participants, interviews regarding the questions below took place between 2021 and 2022. Both data sets are on the individual level but only representative for the federal state, not the county level. Therefore, we use these individual level data but present averages on federal state level as suggestive complementary evidence.

We consider four questions:

- How do you see yourself: Are you generally a person who is very willing to take risks or do you try to avoid taking risks? 0 (not at all willing to take risks) to 10 (very willing to take risks). (Source: SOEP)
- Please tell me on a score of 0–10 how much you personally trust each of the institutions I read out. 0 means you do not trust an institution at all, and 10 means you have complete trust: Germany's parliament? (Source: ESS)
- Please tell me on a score of 0–10 how much you personally trust each of the institutions I read out. 0 means you do not trust an institution at all, and 10 means you have complete trust: Scientists. (Source: ESS)
- Please tell me how much you agree or disagree with the following statement: Coronavirus is the result of deliberate and concealed efforts of some government or organization. Disagree strongly (1)—Agree strongly (5) (Source: ESS, answer categories recoded by the authors)

Table 1. Descriptives.

Variable	Description	Mean	St. dev	Min	Max
<i>Demographics</i>					
Age < 3	Share age under 3 in %	2.76	0.29	1.85	3.60
3 ≤ Age < 6	Share age 3 to under 6 in %	2.78	0.22	2.18	3.36
6 ≤ Age < 18	Share age 6 to under 18 in %	10.71	0.83	7.58	13.38
18 ≤ Age < 30	Share age 18 to under 30 in %	13.13	3.03	6.63	23.60
30 ≤ Age < 50	Share age 30 to under 50 in %	24.43	1.68	20.66	32.27
50 ≤ Age < 65	Share age 50 to under 65 in %	23.64	2.10	16.84	28.47
65 ≤ Age	Share age 65 and older in %	22.56	2.96	15.59	32.71
Women	Share of women in %	50.58	0.63	48.72	52.49
Non-native	Share of non-natives in %	10.82	5.40	2.11	36.56
Life-expectancy	Average life-expectancy of a newborn	81.15	0.97	78.28	83.92
Fertility	Total fertility rate	1.62	0.15	1.16	2.04
LTC	Share of people in need of long-term care (LTC) per 100 inhabitants	5.15	1.23	2.23	9.57
<i>Housing</i>					
COVID cases	Cumulative number of COVID cases per people until December 2020	0.02	0.01	0.00	0.05
Residential buildings	Share residential buildings with 3 or more apartments in %	16.69	10.41	3.90	49.48
1 or 2 rooms	Share apartments with 1 and 2 rooms in %	10.76	5.38	3.00	32.00
5 or more rooms	Share apartments with 5 and more rooms in %	45.35	13.44	13.58	70.34
Living space	Living space per inhabitant in m <sup>2</sup>	48.92	4.93	35.94	69.49
Appartm. residential buildings	Share apartments in residential buildings in %	45.91	19.27	13.96	89.41
<i>Area</i>					
Rurality	Share with population density below 150 people per km <sup>2</sup>	29.40	30.05	0.00	100.00
Distance major center	Avg. driving time to the nearest major center in min	23.12	16.23	0.00	69.61
Distance medium center	Avg. driving time to the nearest medium or major center in min	7.23	5.73	0.00	36.34
Net commuting	Net commuting per 100 employees (subject to social insurance) at workplace	-10.14	28.81	-143.96	61.58
Cadastral area	Cadastral area in km <sup>2</sup>	893.68	724.00	35.70	5495.60
<i>Economy</i>					
Gross earnings	Gross monthly earnings of employees in euros	2863.44	409.05	2244.47	5196.66
Municipal tax capacity	Municipal tax capacity in euros per inhabitant	947.34	276.42	495.01	2818.92
Business tax	Business tax in euros per inhabitant	569.97	299.05	198.99	2637.91
Sales tax	Sales tax in euros per inhabitant	90.97	37.56	34.47	309.93
Municipal debt	Municipal debt in euros per inhabitant	1490.23	1455.17	0.00	9811.02
GDP	Gross domestic product in 1,000 euros per inhabitant	38.54	16.97	16.61	188.29
Unemployment rate	Share unemployed in %	4.69	2.10	1.35	12.83
Beneficiaries SGB II	Share beneficiaries according to SGB (Sozialgesetzbuch) II aged under 65 in %	7.14	4.11	1.16	24.11
Min. social security benefits	Share receiving minimum social security benefits in %	7.27	3.64	2.13	21.98
<i>Education</i>					
Pupils	Pupils per 100 inhabitants	10.03	1.43	5.82	17.26
Students	University students per 1.000 inhabitants	28.51	52.14	0.00	379.57
No degree	Share employees without completed education in %	12.04	2.92	5.03	20.98
Academic degree	Share employees with academic education in %	12.78	5.56	5.29	35.69
Experts	Share employees who are experts in their field in %	10.70	3.87	5.14	29.34
<i>Politics</i>					
Voter turnout	Share of voters among those eligible to vote in % (2017)	75.85	3.71	64.10	84.40
CDU/CSU	Share votes for CDU/CSU (Christlich Demokratische/Soziale Union) in % (2017)	34.27	6.02	21.61	53.79
SPD	Share votes for SPD (Sozialdemokratische Partei Deutschlands) in % (2017)	20.14	6.35	7.80	38.25
Green party	Share votes for Green party in % (2017)	8.09	3.68	2.07	23.31

(Continued)

Table 1. (Continued)

Variable	Description	Mean	St. dev	Min	Max
AfD	Share votes for AfD (Alternative für Deutschland) in % (2017)	13.38	5.33	4.94	35.46
FDP	Share votes for FDP (Freie Demokratische Partei) in % (2017)	10.11	2.48	5.11	17.65
Left-wing party	Share votes for Left-wing party in % (2017)	8.80	4.52	3.58	22.90
Other parties	Share votes for the sum of the other parties in % (2017)	5.20	1.81	2.01	12.58
<i>Supply medical services</i>					
Hospital beds	Hospital beds per 1,000 inhabitants	6.29	3.88	0.00	28.66
Physicians	Physicians per 10,000 inhabitants	14.51	4.33	7.55	30.76
Distance pharmacy	Population-weighted linear distance to the nearest pharmacy in m	1503.60	775.87	347.00	3819.00

Number of observations: 400. Source: Own calculations based on INKAR data.

<https://doi.org/10.1371/journal.pone.0296976.t001>

## Data analysis

### Basic approach

To fix ideas, consider the following linear regression model:

$$y_c = \beta_0 + \beta_1 X_{1c} + \beta_2 X_{2c} + \beta_3 X_{3c} + \beta_4 X_{4c} + \beta_5 X_{5c} + \beta_6 X_{6c} + u_c \tag{1}$$

where  $c$  stands for the county. Further let:

- $y_c$  = COVID-19 vaccinations per inhabitant,
- $X_{1c}$  = Vector of demographic characteristics,
- $X_{2c}$  = Vector of variables reflecting previous Corona cases and housing conditions,
- $X_{3c}$  = Vector of variables reflecting area and local economy,
- $X_{4c}$  = Vector of educational levels,
- $X_{5c}$  = Vector of political preferences,
- $X_{6c}$  = Vector of supply of medical services,
- $u_c$  = error term.

In Section *Pool of variables on the county level* and Table 1 below, we list a number of variables for these six blocks. In total, we have a pool of 48 potential variables that may enter the regression equation. Since we do not aim to perform a causal analysis but to maximize explanatory power in a prediction sense, we use a data-driven approach to select variables from the pool of  $K = 48$  variables. The approach is in the spirit of [53], here adapted to increase the explanatory power of the model. To be specific, we apply the following pre-processing procedure:

1. Start with  $K$  bivariate regressions of  $y_c$  on each of the variables from the pool of potential variables. Store the  $K$  values of adjusted  $R^2$ .
2. Take the variable that leads to the largest adjusted  $R^2$  and put it into the set  $X_{lin}$ . Let  $L$  be the number of variables in  $X_{lin}$ . After this step, we have  $L = 1$ .
3. Perform  $K - L$  regressions of  $y_c$  on  $X_{lin}$  and each of the remaining  $K - L$  variables not in  $X_{lin}$ . Store the  $K - L$  adjusted  $R^2$ .
4. Take the variable that leads to the largest increase in the adjusted  $R^2$  compared to the regression from the previous step. Put in into the set  $X_{lin}$ .
5. Repeat the steps 3 and 4 until no further variable increases the adjusted  $R^2$ .



6. Generate all  $H$  possible bivariate interactions of the variables in  $X_{lin}$  after step 5. This is an additional pool of potential variables.
7. Repeat the procedure from steps 3 to 5. Now, start with  $H$  regressions of  $y_c$  on  $X_{lin}$  and one of the  $H$  interactions. Proceed as before where you put selected interactions into the vector  $X_{int}$  and regress  $y_c$  on  $X_{lin}$  and  $X_{int}$  until no further interaction increases the adjusted  $R^2$ .

The regressions then use the vectors  $X_{lin}$  and  $X_{int}$ . Both are composed of selected variables from the above-mentioned blocks  $X_{1c}$  to  $X_{6c}$ . Allowing for many possible variables and their interaction means to have more potential variables than observations. In the data science literature, this is a common problem. Our solution to this, outlined above, is the “forward stepwise selection” discussed in Chapter 6 of [54].

### Spatial interdependencies

Note that this subsection is taken almost word by word from [38]. The error term  $u_c$  formulated in Eq (1) is probably not distributed independently over the individual counties. Rather, exogenous shocks in one county can be assumed to have knock-on effects in neighbouring counties. It is therefore likely that the error terms are correlated with each other spatially. To take this into account, we formulate the error term as

$$u_c = \rho \sum_{j=1}^{400} m_{cj} u_j + \varepsilon_c \quad (2)$$

where  $m_{cj}$  represents the weighting factor for the effect of a shock in county  $j$  on county  $c$ . Formulated more compactly, we get  $u = \rho Mu + \varepsilon$  where  $u$  is the vector of all error terms (dimension  $400 \times 1$ ),  $M$  is the weighting matrix ( $400 \times 400$ ) and  $\varepsilon$  the vector of errors whose elements are assumed not to correlate with each other. The lower the spatial distance between two counties, the stronger the correlation of the error terms should be. In the definition of the weighting matrix this is achieved by the fact that the weights for adjacent counties are greater than for more distant ones. For  $M$ , we use an *inverse distance matrix*, in which the inverse of the distances between the county centroids are used as elements and the main diagonal contains zeros. The shape files of the counties from the Federal Agency for Cartography and Geodesy (as of 1.1.2022) were used to calculate the distances and transferred into Stata format with the help of the Stata programme `shp2dta`. We follow [55] and normalize the elements with the eigenvalue standardisation, in which every element is divided by the largest module of the eigenvalues of  $M$  (see [56], for a discussion of the normalisation procedures).

Furthermore, it can be assumed that the vaccination rates themselves are also subject to spatial interdependencies. It is conceivable, for example, that local vaccination initiatives have positive spill over effects from one county to another. Moreover, the issue of differences between vaccination location and home county may lead to negative spatial interdependence in the outcome variable. In a more general model in the tradition of [57]’s spatial model or the expanded spatial autoregressive model with spatial autoregressive disturbances (SARAR, [58]), the following variant is estimated:

$$y = \lambda Wy + X\beta + u \quad (3)$$

Here  $W$  is again a weighting matrix. We define the matrix  $W$  as a *contiguity matrix* where only the contingent counties are taken into account and the weights for all others are set at 0. To estimate the parameters  $\beta$ ,  $\rho$  und  $\lambda$  we use the Stata module `sppack`, which contains both the estimator for the model formulated above (`spreg`) and the programme for creating the weighting matrix (see [59]).

To be specific, we run the following spatial regression model:

$$y_c = \lambda W y + \beta_0 + \beta_1 X_{c,lin} + \beta_2 X_{c,int} + u_c \quad (4)$$

where

$$u_c = \rho \sum_{j=1}^{400} m_{cj} u_j + \varepsilon_c$$

This spatial regression model is also used for the variable selection described above.

## A compact approach to present the results

For an easier understanding of how much of the differences between federal states can be explained, we then turn to a more compact approach. We perform a step-wise regression. We start with a regression on a constant only, go on with a regression on a constant and variables from vector  $X_{1c}$ , then add  $X_{2c}$ , and so on. All regressions follow the spatial-econometric approach of Eq (2) and the variable selection procedure. However, when using block  $X_{1c}$  only, only variables from this block are in the pool of potential variables. The same applies when block  $X_{1c}$  and  $X_{2c}$  are used and so on. After each regression, we store the residuals and calculate the federal state averages of the residuals. The standard deviation of these average residuals should be smaller the larger the explanatory power of our model is. The results are presented in Table 3.

## Results

### Baseline results

Table 1 provides the descriptive statistics of the pool of variables on the level of the 400 counties. Table 2 shows the regression results of the model with the highest predictive power according to the spatial regression model outlined in Section *Spatial interdependencies* and the variable selection procedure outlined in Section *Basic approach*. The table includes the chosen variables and interactions. Both  $\rho$  and  $\lambda$  are negative. This most likely reflects the definition of the outcome variable, capturing vaccinations where they took place, not where individuals live. This possibly induces a negative correlation in vaccination rates across neighboring counties. The spatial model takes this into account. However,  $\rho$  is not significantly different from zero and  $\lambda$  is small in magnitude. All in all, spatial correlations do not seem to play a large role in this application.

Note that Table 2 reports coefficients but not marginal effects of certain characteristics. The many interactions make the interpretations of the coefficients complicated. However, many of the baseline coefficients (though not all) go into the expected directions. For instance, the share of individuals over 65 years and previous exposure to the COVID pandemic (measured by COVID cases) have a positive baseline coefficient, while the share of voters for right populist parties (AfD) has a negative coefficient. A higher share of individuals with academic degree goes along with more vaccinations, indicating a positive relationship between education and vaccination. The share of women is positively associated with vaccination rates but the relationship is not statistically significant. All in all, with an  $R^2$  of 0.52 we can explain more than half of the variation on county level using variables on the socio-demographic structure.

Table 3 reports the results of the approach outlined in Section *A compact approach to present the results* where we stepwise include variables in order to understand their share of explained variation in Corona vaccinations. A regression of the COVID vaccinations on only demographics has an  $R^2$  of 19% and can reduce the standard deviation of the residuals by 48%

**Table 2. Regression results.**

	Coef.	std. err
Age < 3	0.166	(0.31)
6 ≤ Age < 18	-1.939***	(0.75)
65 ≤ Age	0.146***	(0.05)
Women	0.160	(0.15)
Life-expectancy	-0.225**	(0.10)
COVID cases	2.857**	(1.17)
Rurality	-0.113***	(0.03)
Distance major center	0.029	(0.04)
Distance medium center	0.658***	(0.18)
Cadastral area	0.000**	(0.00)
Gross earnings	0.005***	(0.00)
Sales tax	-0.014**	(0.01)
Municipal debt	-0.000	(0.00)
Pupils	1.140***	(0.43)
Students	-0.039*	(0.02)
Academic degree	0.660***	(0.25)
Experts	-0.784**	(0.35)
AfD	-0.656***	(0.21)
AfD × Experts	-0.004**	(0.00)
AfD × Rurality	0.000**	(0.00)
AfD × 6 ≤ Age < 18	0.011***	(0.00)
AfD × Students	0.000*	(0.00)
AfD × Academic degree	0.004***	(0.00)
AfD × Municipal debt	-0.000	(0.00)
AfD × Women	0.011***	(0.00)
AfD × Sales tax	0.000***	(0.00)
AfD × Gross earnings	-0.000**	(0.00)
Experts × Cadastral area	-0.000	(0.00)
Experts × Pupils	0.002	(0.00)
Experts × Age < 3	0.131*	(0.07)
Experts × Academic degree	-0.000	(0.00)
Experts × Municipal debt	-0.000***	(0.00)
Experts × 65 ≤ Age	0.021***	(0.01)
Experts × COVID cases	0.001	(0.00)
Experts × Gross earnings	0.000	(0.00)
Life-expectancy × Rurality	0.001***	(0.00)
Life-expectancy × Distance medium center	-0.006***	(0.00)
Life-expectancy × 6 ≤ Age < 18	0.029***	(0.01)
Life-expectancy × Municipal debt	0.000	(0.00)
Life-expectancy × Gross earnings	-0.000*	(0.00)
Distance major center × Rurality	0.000	(0.00)
Distance major center × Cadastral area	0.000	(0.00)
Distance major center × Women	-0.001	(0.00)
Distance major center × COVID cases	0.002***	(0.00)
Rurality × 6 ≤ Age < 18	0.001	(0.00)
Rurality × Age < 3	-0.002	(0.00)
Rurality × Academic degree	-0.000**	(0.00)

(Continued)

Table 2. (Continued)

	Coef.	std. err
Rurality × Municipal debt	-0.000	(0.00)
Rurality × 65 ≤ Age	0.000	(0.00)
Rurality × COVID cases	-0.001**	(0.00)
Rurality × Gross earnings	0.000***	(0.00)
Distance medium center × Pupils	-0.003	(0.00)
Distance medium center × 6 ≤ Age < 18	-0.003	(0.00)
Distance medium center × Students	0.000***	(0.00)
Distance medium center × 65 ≤ Age	-0.003**	(0.00)
Distance medium center × Gross earnings	-0.000**	(0.00)
Cadastral area × 6 ≤ Age < 18	-0.000***	(0.00)
Cadastral area × Age < 3	0.000**	(0.00)
Cadastral area × Students	-0.000*	(0.00)
Cadastral area × COVID cases	0.000	(0.00)
Cadastral area × Sales tax	0.000	(0.00)
Pupils × Students	0.000*	(0.00)
Pupils × Municipal debt	0.000***	(0.00)
Pupils × Women	-0.024***	(0.01)
Pupils × 65 ≤ Age	0.001	(0.00)
6 ≤ Age < 18 × Municipal debt	-0.000**	(0.00)
6 ≤ Age < 18 × COVID cases	-0.105***	(0.02)
6 ≤ Age < 18 × Sales tax	0.001**	(0.00)
6 ≤ Age < 18 × Gross earnings	-0.000*	(0.00)
Age < 3 × Students	-0.001	(0.00)
Age < 3 × Academic degree	-0.150***	(0.05)
Age < 3 × Sales tax	0.002	(0.00)
Students × Municipal debt	0.000***	(0.00)
Students × Women	0.001**	(0.00)
Students × COVID cases	-0.001***	(0.00)
Students × Sales tax	0.000	(0.00)
Academic degree × Women	0.002	(0.00)
Academic degree × 65 ≤ Age	-0.018***	(0.01)
Municipal debt × COVID cases	0.000***	(0.00)
Municipal debt × Gross earnings	0.000*	(0.00)
Women × COVID cases	-0.026	(0.02)
65 ≤ Age × COVID cases	-0.017***	(0.01)
65 ≤ Age × Gross earnings	-0.000***	(0.00)
COVID cases × Sales tax	-0.001*	(0.00)
Sales tax × Gross earnings	-0.000***	(0.00)
$\lambda$	-0.063**	(0.03)
$\rho$	-0.173	(0.17)
$R^2$	0.52	

Number of observations: 400.

\*  $p < 0.10$ ,

\*\*  $p < 0.05$ ,

\*\*\*  $p < 0.01$

<https://doi.org/10.1371/journal.pone.0296976.t002>

**Table 3. Reduction in variation.**

Adjustment	R <sup>2</sup> (%)	State level	
		Std. dev.	Reduction (%)
None	0	.036	0
... + Demographics	19	.019	48
... + Cases and Housing	24	.016	54
... + Area and Economy	27	.013	65
... + Education	48	.01	72
... + Politics	52	.01	72
... + Supply medical services	52	.01	72

<https://doi.org/10.1371/journal.pone.0296976.t003>

at the federal state level. Previous COVID exposure and housing increases the share to 54%. Variables for area and economy increase it further to 65%. Education, politics and supply of medical services do not add much, once the other variables are controlled for, but ultimately 52% of the variation on county level and 72% on federal state level can be explained by our observable characteristics. Of course, these single numbers also depend on the ordering of the blocks but the total number after all blocks are included is invariant to the ordering.

### Robustness checks

Table 4 reports results from robustness checks. These follow the same estimation protocol as above but do not split up the results by different blocks of variables and merely show results for all blocks  $X_{1c}$  to  $X_{6c}$ . We first run the regressions with 390 observations where the ten counties with largest and smallest vaccination rates are dropped. We cannot rule out that these are still strongly affected by differences in vaccination location and place of residence. Moreover, we use the original RKI data without merging numbers for 36 county-pairs (see Section *Data*). The results are basically the same when we drop the 10 outliers. Using the original data, we receive a very large  $R^2$  of 84. However, we tend to have more trust in the baseline specification.

### Potential remaining factors measured at the federal state level

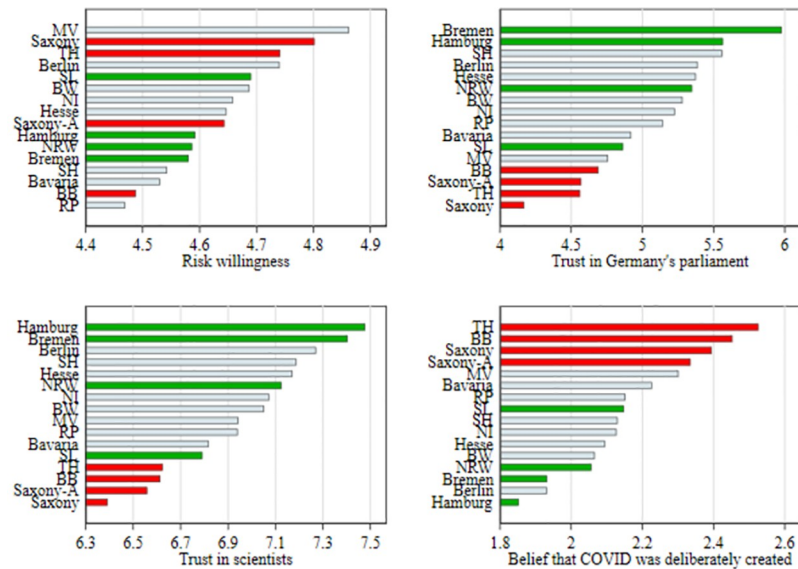
The variables used in the previous analysis cannot fully explain the regional variation in differences in COVID-19 vaccination. Therefore, this subsection provides examples of variables that can be expected to be associated with vaccination but for which data are not available at county level.

Fig 3 reports the sample means in the different federal states of the four answers concerning risk willingness, trust in Germany's parliament, trust in scientists and the belief that COVID was deliberately created. For the sake of presentation, we color the number of the four states with highest vaccination rates (see Fig 1) in green and the four states with the lowest vaccination rates in red. Specifically, the four states with lowest vaccination rates also form a

**Table 4. Robustness checks.**

Sample	R <sup>2</sup> (%)	Reduction (%) Federal state level
Baseline	52	72
Without 10 outliers	50	78
Original data	84	79

<https://doi.org/10.1371/journal.pone.0296976.t004>



**Fig 3. Federal state-level differences in risk preferences and selected statements.** Sample means by federal states presented. MV stands for Mecklenburg-West Pomerania, TH = Thuringia, BB = Brandenburg, RP = Rhineland Palatinate, SL = Saarland, NI = Lower Saxony, SH = Schleswig-Holstein, Saxony-A. = Saxony-Anhalt, NRW = North Rhine-Westphalia, BW = Baden-Wuerttemberg. Upper left panel: *Source:* SOEP; Other panels: *Source:* ESS.

<https://doi.org/10.1371/journal.pone.0296976.g003>

consistent group of highest or lowest ranked states in these four panels (with one exception). In these states, individuals tend to be more risk loving, have less trust in the German parliament and scientists in general, and are more likely to believe that the Corona virus was deliberately created. Bavaria and Mecklenburg-West Pomerania, the states that rank 5 and 6 from the bottom of vaccination status, also broadly rank in this range in the four answers. It is less consistent with the states with highest vaccination rates. By and large, however, they rank on the other extreme also in these four questions. Hence, Fig 3 offers additional explanatory factors for regional variations in vaccination rates.

## Discussion

In terms of Corona vaccination rates, Germany ranks only in the midfield among the industrialized countries. This is partly due to some regions within Germany that have very low vaccination rates. In this study we sought to understand the determinants of the regional variation in Corona vaccination in Germany. It turned out that differences in demographic status are able to explain a considerable share of this regional variation. For instance, regions with a higher share of individuals older than 65 have higher vaccination rates. All in all, differences in demographics alone account for 48 percent of the variation in Corona vaccination on the federal states level. Arguably, differences in vaccination rates due to differences in demographics seem to be justified, for instance given that older individuals belong to the high-risk group and should have higher vaccination rates. Thus, parts of the regional differences observed in Figs 1 and 2 are something policy does not need to worry about or even address. Yet, this refers to only half of the regional variations.

Using a host of other characteristics together with a highly flexible spatial-econometric model, we came to the result that around 45% of the county-level differences and 72% of the federal state-level differences can be explained by observable characteristics. Important factors

other than demographics are previous exposure to Corona in the county, the economic situation and educational status. A number of studies have reported positive correlations of education and vaccination on the individual level (e.g., [14]), thus, this does not come as a surprise. Our explained share is lower than in studies from the USA, where, based on the adjusted  $R^2$ , about 76–79% could be explained [23, 24]. Apparently, our results are not directly comparable to these studies as they use other data, other methods and refer to a different health care system. Nevertheless, an important share of variation in vaccination rates among German counties is left unexplained by basic “hard” factors. Thus, in the second part of the analysis we turned to “soft” factors.

We found suggestive evidence that preferences like risk attitudes and trust might be important explanatory factors for these differences. In federal states with low vaccination rates, individuals in two representative surveys were also more likely to report higher levels of willingness to take risks and lower levels of trust in science. While differences in vaccination due to risk preferences seem justified (when being unvaccinated is just an expression of the preference to take risks), policy should take mistrust in science seriously.

That a person’s trust in institutions or authorities is an important factor in their intention to vaccinate was found in many studies [49, 60, 61]). For [62], trust as part of social capital is a plausible channel through which the Communist past of the East German countries affects vaccination willingness. The negative correlation between belief in conspiracy theories and vaccination willingness has also already been confirmed in empirical studies on the individual level. For Germany, [63] find that parents who have high levels of agreement with vaccine conspiracy theories show a lower willingness to get vaccinated against the corona virus. [64] find for Turkey and the UK that people who believe in a natural origin of the Corona virus are more likely to be vaccinated against the virus.

Trust and belief in conspiracy theories are also important factors in predicting other kinds of preventive behavior: People with a higher trust in government are more likely to comply to the recommended health behaviors such as handwashing and self quarantine [65] and people who believe in conspiracy theories related to COVID are less likely to adhere to measures recommended to contain the Corona virus such as wearing masks. Moreover, they are less likely to intend to get vaccinated [66]. It should be noted, however, that we are not aware of studies that explain regional variations in other preventive measures. Thus, on the one hand, our results are not directly comparable. On the other hand, the strong correlation between vaccination and other preventive measures allows the prediction that regional variations in other measures may share similarities with regional variations in vaccination.

Thus, improving the trust in science and the government as well as the fight against conspiracy theories in general might not only be important in itself to safeguard democracy, but moreover be effective tools to improve vaccination rates. This includes the transparent and impartial discussion and evaluation of all measures taken to fight the COVID pandemic (such as vaccination campaigns, compulsory vaccination in the health care sector, compulsory mask wearing, school closures and lockdowns). An examination with an open mind and the courage to tell unpleasant truths (e.g., where scientist made recommendations regarding the fight of the Corona crisis that turned out to be wrong in hindsight) may be a way to increase trust in science again. This will be particularly relevant for future pandemics.

There are several limitations to our study. Our study design is not able to provide any causal evidence and, thus, does not allow to immediately derive clear policy recommendations and all conclusions derived here must be taken with caution. The main data limitation of this study is that vaccination rates are measured in the county where vaccinations are carried out and not where individuals live. This may lead to underestimations in some areas and overestimations in others. A large part of differences in vaccination rates remains unexplained. Access to more

and different variables at the county level might further reduce the unexplained share. Finally, more granular data on a level below county level should be helpful, also to avoid potential problems of ecological fallacy. To date these data do not exist for Germany.

## Conclusion

High vaccination rates have repeatedly been stressed by many scholars as a decisive factor to fight pandemics such as the COVID-19 pandemic. However, vaccination rates differ strongly across and within countries where some regions have vaccination rates that are generally perceived too low. In this study we analyzed regional variations of COVID-19 vaccination and their determinants in Germany. A highly flexible spatial-econometric model came to the result that around 52% of the county-level differences and 72% of the federal state-level differences can be explained by different county structure such as demographics, housing, economy, education, and party preferences. Around 50% of the differences on federal state level are due to differences in demographics. Thus, arguably, up to 50% of the regional variations may be regarded as justified. However, in future pandemics, policymakers may want to reduce the other 50%. We find suggestive evidence that soft factors on preferences like risk attitudes and trust might be important additional explanatory factors. Based on our findings we recommend improving the trust in science and the fight against conspiracy theories as possibly effective tools to improve vaccination rates and effectively fight pandemics.

## Supporting information

**S1 Table. Counties assigned the average vaccination rate.**

(PDF)

## Author Contributions

**Conceptualization:** Hendrik Schmitz.

**Data curation:** Verena Bade, Hendrik Schmitz.

**Formal analysis:** Verena Bade, Hendrik Schmitz.

**Funding acquisition:** Hendrik Schmitz.

**Investigation:** Verena Bade, Hendrik Schmitz.

**Methodology:** Hendrik Schmitz.

**Project administration:** Hendrik Schmitz.

**Supervision:** Hendrik Schmitz.

**Visualization:** Verena Bade, Hendrik Schmitz.

**Writing – original draft:** Verena Bade, Hendrik Schmitz, Beatrice Baaba Tawiah.

**Writing – review & editing:** Verena Bade, Hendrik Schmitz, Beatrice Baaba Tawiah.

## References

1. Campos-Mercade P, Meier AN, Schneider FH, Meier S, Pope D, Wengström E. Monetary incentives increase COVID-19 vaccinations. *Science*. 2021; 374(6569):879–82. Available from: <https://www.science.org/doi/abs/10.1126/science.abm0475>. PMID: 34618594
2. Albrecht D. Vaccination, politics and COVID-19 impacts. *BMC Public Health*. 2022; 22(1):96. Available from: <https://doi.org/10.1186/s12889-021-12432-x>. PMID: 35031053



3. Barber A, West J. Conditional cash lotteries increase COVID-19 vaccination rates. *Journal of Health Economics*. 2022; 81:102578. Available from: <https://www.sciencedirect.com/science/article/pii/S0167629621001636>. <https://doi.org/10.1016/j.jhealeco.2021.102578> PMID: 34986437
4. Nehal KR, Steendam LM, Campos Ponce M, van der Hoeven M, Smit GSA. Worldwide vaccination willingness for COVID-19: A systematic review and meta-analysis. *Vaccines*. 2021; 9(10). Available from: <https://www.mdpi.com/2076-393X/9/10/1071>. <https://doi.org/10.3390/vaccines9101071> PMID: 34696179
5. Mewhirter J, Sagir M, Sanders R. Towards a predictive model of COVID-19 vaccine hesitancy among American adults. *Vaccine*. 2022; 40(12):1783–9. Available from: <https://www.sciencedirect.com/science/article/pii/S0264410X22001347>. <https://doi.org/10.1016/j.vaccine.2022.02.011> PMID: 35164989
6. Sterl S, Stelzmann D, Luettschwager N, Gerhold L. COVID-19 vaccination status in Germany: Factors and reasons for not being vaccinated (yet). *Frontiers in Public Health*. 2023; 11:1070272. Available from: <https://www.frontiersin.org/articles/10.3389/fpubh.2023.1070272/full>. PMID: 36860382
7. del Rio C, Malani PN, Omer SB. Confronting the Delta Variant of SARS-CoV-2, Summer 2021. *JAMA*. 2021 09; 326(11):1001–2. Available from: <https://doi.org/10.1001/jama.2021.14811>. PMID: 34406361
8. Nasreen S, Chung H, He S, Brown KA, Gubbay JB, Buchan SA, et al. Effectiveness of COVID-19 vaccines against symptomatic SARS-CoV-2 infection and severe outcomes with variants of concern in Ontario. *Nature Microbiology*. 2022; 7(3):379–85. Available from: <https://www.nature.com/articles/s41564-021-01053-0>. <https://doi.org/10.1038/s41564-021-01053-0> PMID: 35132198
9. Nordström P, Ballin M, Nordström A. Risk of infection, hospitalisation, and death up to 9 months after a second dose of COVID-19 vaccine: a retrospective, total population cohort study in Sweden. *The Lancet*. 2022; 399(10327):814–23. Available from: <https://www.sciencedirect.com/science/article/pii/S0140673622000897>. PMID: 35131043
10. Besley T, Stern N. The economics of lockdown. *Fiscal Studies*. 2020; 41(3):493–513. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/1475-5890.12246>. PMID: 33362309
11. Jessop J. The UK lockdown and the economic value of human life. *Economic Affairs*. 2020; 40(2):138. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7361259/>. <https://doi.org/10.1111/ecaf.12417>
12. Bayham J, Fenichel EP. Impact of school closures for COVID-19 on the US health-care workforce and net mortality: a modelling study. *The Lancet Public Health*. 2020; 5(5):e271–8. [https://doi.org/10.1016/S2468-2667\(20\)30082-7](https://doi.org/10.1016/S2468-2667(20)30082-7) PMID: 32251626
13. Viner RM, Russell SJ, Croker H, Packer J, Ward J, Stansfield C, et al. School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. *The Lancet Child & Adolescent Health*. 2020; 4(5):397–404. Available from: [https://www.thelancet.com/journals/launch/article/PIIS2352-4642\(20\)30095-X/fulltext](https://www.thelancet.com/journals/launch/article/PIIS2352-4642(20)30095-X/fulltext). [https://doi.org/10.1016/S2352-4642\(20\)30095-X](https://doi.org/10.1016/S2352-4642(20)30095-X) PMID: 32272089
14. Monsees D, Schmitz H. The Effect of Compulsory Schooling on Vaccination against COVID and Influenza. *Ruhr Economic Papers* 1011. 2023.
15. de Figueiredo A, Simas C, Karafillakis E, Paterson P, Larson HJ. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. *The Lancet*. 2020; 396(10255):898–908. Available from: <https://www.sciencedirect.com/science/article/pii/S0140673620315580>. [https://doi.org/10.1016/S0140-6736\(20\)31558-0](https://doi.org/10.1016/S0140-6736(20)31558-0) PMID: 32919524
16. Mayrhofer T, Schmitz H. Prudence and Prevention. *Ruhr Economic Papers*. 2020;(863). Available from: [https://www.rwi-essen.de/media/content/pages/publikationen/ruhr-economic-papers/rep\\_20\\_863.pdf](https://www.rwi-essen.de/media/content/pages/publikationen/ruhr-economic-papers/rep_20_863.pdf).
17. Lounis M, Bencherit D, Rais MA, Riad A. COVID-19 vaccine booster hesitancy (VBH) and its drivers in Algeria: national cross-sectional survey-based study. *Vaccines*. 2022; 10(4). Available from: <https://www.mdpi.com/2076-393X/10/4/621>. <https://doi.org/10.3390/vaccines10040621> PMID: 35455371
18. Yadete T, Batra K, Netski DM, Antonio S, Patros MJ, Bester JC. Assessing Acceptability of COVID-19 Vaccine Booster Dose among Adult Americans: A Cross-Sectional Study. *Vaccines*. 2021; 9(12). Available from: <https://www.mdpi.com/2076-393X/9/12/1424>. <https://doi.org/10.3390/vaccines9121424> PMID: 34960170
19. Yoshida M, Kobashi Y, Kawamura T, Shimazu Y, Nishikawa Y, Omata F, et al. Factors Associated with COVID-19 Vaccine Booster Hesitancy: A Retrospective Cohort Study, Fukushima Vaccination Community Survey. *Vaccines*. 2022; 10(4). Available from: <https://www.mdpi.com/2076-393X/10/4/515>. <https://doi.org/10.3390/vaccines10040515> PMID: 35455264
20. Shah A, Coiado OC. COVID-19 vaccine and booster hesitation around the world: A literature review. *Frontiers in Medicine*. 2023; 9:3994. Available from: [https://www.frontiersin.org/articles/10.3389/fmed.2022.1054557/full?&utm\\_source=Email\\_to\\_authors\\_&utm\\_medium=Email&utm\\_content=T1\\_11.5e1\\_](https://www.frontiersin.org/articles/10.3389/fmed.2022.1054557/full?&utm_source=Email_to_authors_&utm_medium=Email&utm_content=T1_11.5e1_)

- [author&utm\\_campaign=Email\\_publication&field=&journalName=Frontiers\\_in\\_Medicine&id=1054557. https://doi.org/10.3389/fmed.2022.1054557](https://doi.org/10.3389/fmed.2022.1054557) PMID: 36714110
21. Mishra A, Sutermeister S, Smittenaar P, Stewart N, Sgaier SK. COVID-19 Vaccine Coverage Index: Identifying barriers to COVID-19 vaccine uptake across U.S. counties. medRxiv. 2021. Available from: <https://www.medrxiv.org/content/early/2021/06/22/2021.06.17.21259116>.
  22. DiRago NV, Li M, Tom T, Schupmann W, Carrillo Y, Carey CM, et al. COVID-19 vaccine rollouts and the reproduction of urban spatial inequality: Disparities within large US cities in March and April 2021 by racial/ethnic and socioeconomic composition. *Journal of Urban Health*. 2022; 99(2):191–207. Available from: <https://doi.org/10.1007/s11524-021-00589-0>. PMID: 35118595
  23. Pallathadka A, Chang H, Han D. What explains spatial variations of COVID-19 vaccine hesitancy?: A social-ecological-technological systems approach. *Environmental Research: Health*. 2022; 1(1):011001.
  24. Mollalo A, Tatar M. Spatial modeling of COVID-19 vaccine hesitancy in the United States. *International Journal of Environmental Research and Public Health*. 2021; 18(18). Available from: <https://www.mdpi.com/1660-4601/18/18/9488>. <https://doi.org/10.3390/ijerph18189488> PMID: 34574416
  25. Allcott H, Boxell L, Conway J, Gentzkow M, Thaler M, Yang D. Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics*. 2020; 191:104254. Available from: <https://www.sciencedirect.com/science/article/pii/S0047272720301183>. <https://doi.org/10.1016/j.jpubeco.2020.104254> PMID: 32836504
  26. Baccini L, Brodeur A. Explaining Governors' Response to the COVID-19 Pandemic in the United States. *American Politics Research*. 2021; 49(2):215–20. Available from: <https://doi.org/10.1177/1532673X20973453>.
  27. Gerace A, Rigney G, Anderson JR. Predicting attitudes towards easing COVID-19 restrictions in the United States of America: The role of health concerns, demographic, political, and individual difference factors. *Plos one*. 2022; 17(2):e0263128. Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0263128>. PMID: 35196316
  28. Wright AL, Sonin K, Driscoll J, Wilson J. Poverty and economic dislocation reduce compliance with COVID-19 shelter-in-place protocols. *Journal of Economic Behavior & Organization*. 2020; 180:544–54. Available from: <https://www.sciencedirect.com/science/article/pii/S0167268120303760>. <https://doi.org/10.1016/j.jebo.2020.10.008> PMID: 33100443
  29. Engle S, Stromme J, Zhou A. Staying at home: Mobility effects of COVID-19; 2020. No. 3565703. Available from: <https://ssrn.com/abstract=3565703>.
  30. Desson Z, Kauer L, Otten T, Peters JW, Paolucci F. Finding the way forward: COVID-19 vaccination progress in Germany, Austria and Switzerland. *Health Policy and Technology*. 2022; 11(2):100584. The COVID-19 Pandemic: Vaccination Strategies and Global Health Policies. Available from: <https://www.sciencedirect.com/science/article/pii/S2211883721001076>. <https://doi.org/10.1016/j.hlpt.2021.100584> PMID: 34868834
  31. Cullen MR, Cummins C, Fuchs VR. Geographic and racial variation in premature mortality in the US: analyzing the disparities. *PLoS One*. 2012; 7(4):e32930. Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0032930>. PMID: 22529892
  32. Doyle JJ. Returns to local-area health care spending: Evidence from health shocks to patients far from home. *American Economic Journal: Applied Economics*. 2011 July; 3(3):221–43. Available from: <https://www.aeaweb.org/articles?id=10.1257/app.3.3.221>. PMID: 23853699
  33. Rettenmaier AJ, Wang Z. Regional variations in medical spending and utilization: a longitudinal analysis of US Medicare population. *Health Economics*. 2012; 21(2):67–82. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.1700>. PMID: 22223553
  34. Reich O, Weins C, Schusterschitz C, Thöni M. Exploring the disparities of regional health care expenditures in Switzerland: some empirical evidence. *The European Journal of Health Economics*. 2012; 13(2):193–202. Available from: <https://doi.org/10.1007/s10198-011-0299-x>. PMID: 21290253
  35. Prieto DC, Lago-Peñas S. Decomposing the determinants of health care expenditure: the case of Spain. *The European Journal of Health Economics*. 2012; 13(1):19–27. Available from: <https://doi.org/10.1007/s10198-010-0276-9>. PMID: 20853126
  36. Panopoulou E, Pantelidis T. Cross-state disparities in US health care expenditures. *Health Economics*. 2013; 22(4):451–65. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.2816>. PMID: 22473657
  37. Sheiner L. Why the geographic variation in health care spending cannot tell us much about the efficiency or quality of our health care system. *Brookings Papers on Economic Activity*. 2014; 2014(2):1–72. <https://doi.org/10.1353/eca.2014.0012>

38. Göppfarth D, Kopetsch T, Schmitz H. Determinants of regional variation in health expenditures in Germany. *Health Economics*. 2016; 25(7):801–15. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.3183>. PMID: 25962986
39. Gusmano MK, Weisz D, Rodwin VG, Lang J, Qian M, Bocquier A, et al. Disparities in access to health care in three French regions. *Health Policy*. 2014; 114(1):31–40. *Geographic Variation in Health Care—40 Years of “Small-Area Variation”*. Available from: <https://www.sciencedirect.com/science/article/pii/S0168851013002005>. <https://doi.org/10.1016/j.healthpol.2013.07.011> PMID: 23927846
40. Fuchs VR. More variation in use of care, more flat-of-the-curve medicine. *Health Affairs*. 2004; 23 (Suppl2):104–7. Available from: <https://doi.org/10.1377/hlthaff.var.104>. PMID: 15471787
41. Fuchs VR, McClellan MB, Skinner JS. Area differences in utilization of medical care and mortality among US elderly. In: Wise DA, editor. *Perspectives on the Economics of Aging*. University of Chicago Press; 2004. p. 367–414. Available from: <http://www.nber.org/chapters/c10349>.
42. Augurzky B, Kopetsch T, Schmitz H. What accounts for the regional differences in the utilisation of hospitals in Germany? *The European Journal of Health Economics*. 2013; 14(4):615–27. Available from: <https://doi.org/10.1007/s10198-012-0407-6>. PMID: 22760519
43. Kopetsch T, Schmitz H. Regional variation in the utilisation of ambulatory services in Germany. *Health Economics*. 2014; 23(12):1481–92. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.3001>. PMID: 24123586
44. Tu JV, Chu A, Maclagan L, Austin PC, Johnston S, Ko DT, et al. Regional variations in ambulatory care and incidence of cardiovascular events. *CMAJ*. 2017; 189(13):E494–501. Available from: <https://www.cmaj.ca/lookup/doi/10.1503/cmaj.170116>. PMID: 28385894
45. Salm M, Wübker A. Sources of regional variation in healthcare utilization in Germany. *Journal of Health Economics*. 2020; 69:102271. Available from: <https://www.sciencedirect.com/science/article/pii/S016762961830866X>. <https://doi.org/10.1016/j.jhealeco.2019.102271> PMID: 31874377
46. Bojke C, Castelli A, Street A, Ward P, Laudicella M. Regional variation in the productivity of the English National Health Service. *Health Economics*. 2013; 22(2):194–211. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.2794>. PMID: 22334404
47. Robert Koch-Institut. COVID-19-Impfungen in Deutschland. Github; 2022. Available from: [https://github.com/robert-koch-institut/COVID-19-Impfungen\\_in\\_Deutschland](https://github.com/robert-koch-institut/COVID-19-Impfungen_in_Deutschland).
48. BBSR Bonn. Bundesinstitut für Bau-, Stadt- und Raumforschung, [www.bbsr.bund.de](http://www.bbsr.bund.de); 2022. Available from: <https://www.inkar.de/>.
49. Grüner S, Krüger F. The intention to be vaccinated against COVID-19: stated preferences before vaccines were available. *Applied Economics Letters*. 2021; 28(21):1847–51. Available from: <https://doi.org/10.1080/13504851.2020.1854445>.
50. El-Far Cardo A, Kraus T, Kaifie A. Factors That Shape People’s Attitudes towards the COVID-19 Pandemic in Germany—The Influence of Media, Politics and Personal Characteristics. *International Journal of Environmental Research and Public Health*. 2021; 18(15). Available from: <https://www.mdpi.com/1660-4601/18/15/7772>.
51. ESS. ESS Round 10: European Social Survey Round 10 Data. Data file edition 1.2. Sikt—Norwegian Agency for Shared Services in Education and Research, Norway—Data Archive and distributor of ESS data for ESS ERIC.; 2020.
52. Goebel J, Grabka MM, Liebig S, Kroh M, Richter D, Schröder C, et al. The German Socio-Economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik*. 2019; 239(2):345–60. Available from: <https://doi.org/10.1515/jbnst-2018-0022> [cited 2022-07-20].
53. Imbens GW. Matching Methods in Practice: Three Examples. *Journal of Human Resources*. 2015; 50 (2):373–419. <https://doi.org/10.3368/jhr.50.2.373>
54. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning—With Applications in R*. Springer, New York, Heidelberg, Dordrecht, London.
55. Felder S, Tauchmann H. Federal state differentials in the efficiency of health production in Germany: An artifact of spatial dependence? *European Journal of Health Economics*. 2013; 14(1):21–39. <https://doi.org/10.1007/s10198-011-0345-8> PMID: 21901473
56. Kelejian HH, Prucha IR. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*. 2010; 157(2):53–67. <https://doi.org/10.1016/j.jeconom.2009.10.025> PMID: 20577573
57. Cliff AD, Ord JK. *Spatial Autocorrelation*. London: Pion; 1973.
58. Anselin L, Florax R. Small sample properties of tests for spatial dependence in regression models: Some further results. In: Anselin L, Florax R, editors. *New Directions in Spatial Econometrics*. New York: Springer; 1995. p. 75–95.

59. Drukker DM, Peng H, Prucha I, Raciborski R. SPPACK: Stata module for cross-section spatial-autoregressive models; 2011. Statistical Software Components, Boston College Department of Economics. Available from: <http://ideas.repec.org/c/boc/bocode/s457245.html>.
60. Bughin J, Cincera M, Peters K, Reykowska D, Zyszkiewicz M, Ohme R. Make it or break it: Vaccination intent at the time of Covid-19. *COVID Economics*. 2021. Available from: <https://cepr.org/node/390707>.
61. Dabla-Norris E, Khan H, Lima F, Sollaci A. Who doesn't want to be vaccinated? Determinants of vaccine hesitancy during COVID-19. *COVID Economics*. 2021. Available from: <https://cepr.org/node/390764>.
62. Pronkina EB, Berniell I, Fawaz Y, Laferrere A, Mira P. The COVID-19 curtain: Can past communist regimes explain the vaccination divide in Europe? *Social Science and Medicine*. 2023; 115759(321). Available from: <https://doi.org/10.1016/j.socscimed.2023.115759>. PMID: 36774703
63. Altenbuchner A, Haug S, Schnell R, Scharf A, Weber K. Impfbereitschaft von Eltern mit einem COVID-19-Vakzin. Die Rolle von Elternschaft und Geschlecht. *Pädiatrie & Pädologie*. 2021;(5):230–4. Available from: <https://doi.org/10.1007/s00608-021-00925-2>.
64. Salali GD, Uysal MS. COVID-19 vaccine hesitancy is associated with beliefs on the origin of the novel coronavirus un the UK and Turkey. *Psychological Medicine*. 2022; 15(52):3750–2. Available from: <https://doi.org/10.1017/S0033291720004067>.
65. Han Q, Zheng B, Cristea M, Agostini M, Belanger JJ, Gützkow B, et al. Trust in government regarding COVID-19 and its associations with preventive health behaviour and prosocial behaviour during the pandemic: a cross-sectional and longitudinal study. *Psychological Medicine*. 2021; 1(53):149–59. Available from: <https://doi.org/10.1017/S0033291721001306>. PMID: 33769242
66. Romer D, Jamieson KH. Conspiracy theories as barriers to controlling the spread of COVID-19 in the U. S. *Social Science and Medicine*. 2020;(263):113356. Available from: <https://doi.org/10.1016/j.socscimed.2020.113356>. PMID: 32967786