# Syndemic tendencies?
# A spatial perspective on COVID-19 and socioeconomic deprivation in Germany.

Alexander Busch
Matrikelnummer 4052247
Keplerstr. 66
69120 Heidelberg

Submitted to the Department of Sociology in partial
fulfilment of the requirements for the degree of

Bachelor of Arts in Sociology

University of Heidelberg
First supervisor: Prof. Dr. Thomas Klein
Second supervisor: apl. Prof. Dr. Johannes Stauder

# Contents

# Declaration of Originality

[**English**]

"I hereby certify that I have written this thesis independently and without unauthorised outside help and that all passages taken verbatim or in spirit from publications in this thesis have been individually identified with reference to the source."

[**Deutsch**]

"Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig und ohne unerlaubte fremde Hilfe verfasst habe und dass alle wörtlich oder sinngemäß aus Veröffentlichungen entnommenen Stellen dieser Arbeit unter Quellenangabe einzeln kenntlich gemacht sind."

A. Busch

Alexander Busch, Heidelberg 13.06.2021

# Images



# Tables

# Abbreviations

- Case Fatality Ratio = CFR
- Corona Virus Disease 2019 = COVID-19
- Federal Agency for Cartography and Geodesy = BKG
- Federal Institute for Research on Building, Urban Affairs and Spatial Development = BBSR
- German Index of Socioeconomic Deprivation = GISD
- Human Development Index = HDI
- Incidence Rate = IR
- Nomenclature des unités territoriales statistiques = NUTS
- per capita = p.c.
- Robert Koch Institute = RKI
- Severe Acute Respiratory Syndrome Coronavirus 2 = SARS-CoV-2
- World Health Organization = WHO

# Abstract


**[English]**

This study explores interactions of the COVID-19 pandemic in Germany with social deprivation in the light of the current 'syndemic'-debate. The relationships between incidence rates (IR) and case fatality ratios (CFR) with socioeconomic variables were assessed. SAR and OLS regression models on county scale (N=401) were fitted for monthly estimations from February 2020 to April 2021. The main findings are that IR were negatively associated with social deprivation in the beginning of the pandemic, but positively during the second and third wave. For counties with a higher share of foreign nationals, as well as counties in the historic east, the deprivation effect during the second and third wave is magnified. No such effect was found for CFR. It is concluded that IR are syndemic on the county scale, while CFR may not be. Furthermore, mistrust in the government (measured by votes for the far-right AfD party) was connected to higher IR during the second and third wave.

**[Deutsch]**

Diese Studie untersucht die Wechselwirkungen der COVID-19-Pandemie in Deutschland mit sozialer Deprivation im Lichte der aktuellen 'Syndemie'-Debatte. Die Beziehung zwischen sozioökonomischen Variablen und Inzidenzraten (IR) und Case Fatality Ratios (CFR) wurden untersucht. SAR- und OLS-Regressionsmodelle auf Kreisskala (N=401) wurden monatlich von Februar 2020 bis April 2021 geschätzt. Die wichtigsten Ergebnisse sind, dass IR zu Beginn der Pandemie negativ mit sozialer Benachteiligung assoziiert waren, aber positiv während der 2. und 3. Welle. Für Landkreise mit einem höheren Anteil an Ausländer_innen sowie für Landkreise in Ostdeutschland ist der Deprivationseffekt während der 2. und 3. Welle verstärkt. Für CFR wurde kein solcher Effekt gefunden. Es wird gefolgert, dass IR auf der Kreisebene syndemisch sind, die CFR nicht. Außerdem war das Misstrauen in die Regierung (gemessen an den Stimmen für die rechtsextreme Partei AfD) mit höheren IR während der 2. und 3. Welle verbunden.


**Keywords** COVID-19, Germany, social deprivation, GISD, German Index of Socioeconomic Deprivation, syndemic, ecological, spatial, temporal, SAR, spatial econometrics

# 1 Introduction

Is COVID-19 a social equaliser? Is it true that the virus does not discriminate between poor and rich? Or are people with little means stronger affected by the disease? These questions are at the heart of the international 'syndemic debate' that arose primarily in the Lancet in autumn 2020 (Horton 2020; Bambra et al. 2020; Gravlee 2020; Kenyon 2020; Fronteira et al. 2021) [1]. While many spatial socioeconomic studies on COVID-19 highlight the uneven distribution of the virus and links between infections, deaths, social and economic artifacts, this debate has not been settled yet.

A country rather unaffected by these questions and this string of research is Germany: Just three country-wide socio-spatial peer-reviewed studies on COVID-19 (Scarpone et al. 2020; Wachtler et al. 2020b; Hoebel et al. 2021), two working papers (Ehlert 2020; Plümper and Neunmayer 2020), and some journalistic attention to inequalities within certain cities (Norddeutscher Rundfunk 2021; Kaiser 2021; infas360 2021) have been published. All publications except for one study were published in summer 2020, one in January 2021 (Hoebel et al. 2021). Whether the pandemic in Germany is embedded in a larger syndemic has not been discussed yet.

This study connects the national COVID-19 records provided by the RKI with the German Index of Socioeconomic Deprivation (Kroll et al. 2018) and selected indicators from the INKAR-data base (BBSR 2021). Cases and deaths of the virus were aggregated for each of the 401 German counties, monthly, from the first case in Germany in January 2020 through 30 April 2021. Due to strong spatial autocorrelation, spatial regression models (SAR) were fitted when a model involved incedence rates (IR), while regular OLS regressions could be used for relationships involving case-fatality-ratios (CFR). All models were fitted and compared on a monthly bases.

The goal of this study is twofold: First, to contribute to the understanding of the role of socioeconomic factors in the pandemic via the syndemic framework of Singer (1994). Second, to take an intertemporal perspective and analyse how the impact of these factors changed during the whole first year of COVID-19 in Germany.

The analysis shows that the pandemic in Germany in regard to its social impact

[1] What constitutes the sociological concept of a 'Syndemic', a 'synergistic co-occurring epidemic', is defined and further explained in the theoretical section of this study.

has two phases: During the first wave of cases until summer 2020, IR in counties with less socially deprived population were higher. Afterwards, the pandemic moved on to socially deprived counties, especially to those with a higher share of foreign nationals and those in Eastern Germany. For infections with the virus, this paper presents evidence of syndemic tendencies by the definition of Mendenhall and Singer (2019) and Singer (1994). Furthermore, it is shown that trust in government was associated with lower IR after the first wave, and that doctor density was associated with a higher CFR during the second and third wave.

# 2 Literature and Theoretical Discourse

After a short excursus on socioeconomic status in previous pandemics, this chapter progresses to individual risk factors for severe cases of COVID-19, discusses previous ecological studies on the pandemic, and finally expands on the international debate on whether this pandemic may fulfill the criteria of a previously given definition of the syndemic-concept.

For contextualising statements on COVID-19 in Germany (e.g. references to the first, second, or third wave of infections), several maps and graphs are included in the appendix to illustrate the course of the pandemic in Germany (appendix 8, 9, 6, 7).

## 2.1 The socioeconomics of previous pandemics

Since John Snow's seminal works on the outbreak of Cholera in London (Snow 1849), the social and geographical determinants of pandemics have been widely studied. The epidemiological focus on public health distinguishes this field from clinical medicine and constitute it as a neighbouring discipline of macro-sociology. Pandemics accompany civilizations throughout history, so it is not surprising that the current spread of a virus can be compared to its many predecessors. This is just a very brief overview of the most important findings on past global pandemics. As the COVID-19 pandemic is still occurring and thus cannot be studied in its totality, a historical perspective yields many insights.

The 'Spanish Flu' is often regarded as the pandemic most similar to the current situation: It affected Western countries with a rather developed healthcare system and spread quickly throughout the world, infecting approximately 500 million people and killing more than 50 million (Niall and Mueller 2002). A meta-study by Sharma et al. (2021) compared the Spanish Flu with COVID-19 and found the world unprepared for the sudden surge in demand for healthcare both times. Mortality was shown to be higher for socioeconomically deprived individuals in Oslo/Kristiania (Mamelund 2006) and Sidney (McCracken and Curson 2003), although the debate over this overall socioeconomic dimension is not settled and results might differ between countries (Mamelund 2006). Looking at the dynamics of this pandemic, Mamelund (2018) found evidence for a change of the role of the socioeconomic status in mortality in the Norwegian city of Bergen: First, people with a low status had a higher chance of contracting the virus, while later, people with high status were the bearers of the disease.

In contrast to this historical example, the 2009 H1N1 'swine flu' pandemic may also serve as example of pandemic mechanisms in contemporary times. The effect of socioeconomic status differed between studies: In the US (Massachusetts), people of high socioeconomic status were more likely to be in ICU (Placzek and Madoff 2014), while the opposite was observed in Canada (Lowcock et al. 2012) and Spain (Mayoral et al. 2013). Other more recent outbreaks of viruses such as SARS (2002), MERS (2012), avian influenza (e.g. 1997 in Hong Kong, 2015 in the US, 2016 and 2020 in Germany) are not fit for comparison on socioeconomic factors. They were regionally contained and did not affect countries similar to Germany on a substantial scale.

In a yet unpublished literature review and meta-analysis, Mamelund et al. (2020) compare studies on the impact of socioeconomic status on influenza pandemics. They only found studies on either the 1918 or the 2009 pandemics. Overall, low socioeconomic status was associated with a more severe illness, the effect on mortality however was unclear.

## 2.2 Individual Traits and COVID-19

The current pandemic follows to some degree the lines of its predecessors. To represent the current stage of research on individual traits (in contrast to grand sociological interactions), this section is mainly based on systematic overviews and meta studies on risk factors of COVID-19. They mainly focus on morbidity, not on transmission.

Men are more likely to be hospitalised and have a higher risk of dying from the virus in comparison to women (Li et al. 2021; Gao et al. 2021; van Gerwen et al. 2021). These studies do not differentiate between sex, gender, and gender identity, so differences may be explained by both biological (e.g. hormone differences) and 'lifestyle' differences (e.g. smoking, type of work) (Gao et al. 2021; Wenham et al. 2020). On the other hand, in most countries, health-care workers (excluding doctors) are predominantly female and are thus more likely to be exposed to the virus (Wenham et al. 2020).

Diabetes, hypertension (high blood pressure), and chronic heart/renal/liver/lung diseases have been identified as medical risk factors of COVID-19 mortality and severity (Li et al. 2021; Gao et al. 2021). So has been obesity (Gao et al. 2021; Malik et al. 2020) and cancer (Gao et al. 2021).

Once infected, old people are more likely to die due to the prevalence of medical risk factors, a weaker immune defense, and a higher viral load (Gao et al. 2021). In a meta analysis of medical studies on COVID-19, Li et al. (2021) found that survivors of the virus were on average almost 20 years younger than non-survivors. Additionally, the reported infection rates for children are lower than those of adults (Mehta et al. 2020). However, this may not stem from actual lower infection rates, but rather from the fact that children are more likely to be asymptomatic or have mild cases (ibid.).

The role of ethnicity in morbidity is unclear. Especially studies in the UK and the US find a much higher risk of hospitalisation, severity, and mortality for Black and other non-white ethnicities (Gao et al. 2021). However, some studies report that ethnicity has no effect after accounting for socioeconomic factors and lifestyle variables, while others still find significant influence (ibid.).

Gao et al. (2021) list several 'lifestyle' risk factors for COVID-19: Some diets (Mediterranean, Ketogenic, fermented vegetables) may lower the chances of dying of the virus, while some occupations such as working in healthcare increase these chances. Active and former smokers also have suffer more severe (Gülsen et al. 2020) and have a higher mortality (Gao et al. 2021).

Studies on air pollution also find evidence that exposure to pollutants such as $PM_{2.5}$ and $NO_2$ may increase the spread and lethality of the novel Corona virus (Copat et al. 2020; Bourdrel et al. 2021).

## 2.3 Ecological Studies of the current Pandemic

The interest in the spatial and structural impact of the COVID-19-pandemic has been immense. This section will thus not provide a complete overview of this topic, but rather a focused summary of findings concerning Germany and similar countries.

### 2.3.1 Ecological Research on COVID-19 world-wide (especially OECD)

Wachtler et al. (2020a) present an overview of 46 ecological studies that focuses on socioeconomic aspects (excluding race) published before 15 June 2020. They generally find that in the US and the UK, strong inequalities exist and poor areas are more affected, both in mortality and in infections. On the other hand, one paper on Germany and one paper on Italy (both not peer-reviewed) reported a positive relationship between cases and employment rates.

Mackey et al. (2021) review 52 papers (published before 1 September 2020, grey literature before 2 November 2020) on the spatial relationship of the pandemic with race in the US and find that the African American/Black and Hispanic population have higher case numbers and mortality, but a similar CFR. Thus, the higher mortality might stem from more infections in the first place.

Another overview by Fatima et al. (2021) on 38 papers (published before 1 October 2020) focuses on methodology, geographical clustering, and to a lesser degree on socioeconomics. They find that most research was conducted in and on China, Brazil, and the US. Of the 38 papers included, 13 were categorised as 'spatial epidemiological modeling' and thus relevant for this chapter. These studies confirm the results of the other two reviews, namely that the virus affected underprivileged areas more. Some of the papers are identical to those reviewed by Mackey et al. (2021) and Wachtler et al. (2020a).

Although these three overviews cover most of the relevant international research, some new papers have not been included or have emerged since. Also, these overviews did not report on the temporal dimension of relationships, but only on statics.

The COVID-19 CFR in the US was slightly higher in rural counties during the whole year of 2020 (Iyanda et al. 2021). Iyanda et al. (2021) and Mourad et al. (2020) also found a positive relationship between the 2020 share of Trump voters and incidence rates. In a working paper using county data from several European countries until May 2020, Bartscher et al. (2020) report that voter turnout correlated negatively with cases, which they interpreted as the effect of social capital.

Studies comparing whole countries show that incidences correlated positively with GDP (Sorci et al. 2020; Gangemi et al. 2020), HDI (Shabazi and Khazaei 2020) and international travel (Gangemi et al. 2020). Deaths were correlated positively with age (Gangemi et al. 2020; Sorci et al. 2020) and international travel (Pana et al. 2021). Death rates were correlated negatively with healthcare capacities (Sorci et al. 2020) (Khan et al. 2020) (all of these studies rely on data older than June 2020). Although the country-scale is the crudest aggregation level for the pandemic, it is interesting to see that the results from the mesa-level (county and post-code analysis) for the beginning of the pandemic align with the country scale. However, it should be noted that these studies aggregate and compare cases in countries that might not have reached the same stage of the pandemic. In some, COVID-19 might have just begun spreading, while it has diffused into all areas and social classes in other countries.

For data until 11 May 2020, Gaudart et al. (2021) found no relationship between incidence or mortality rates and economic or healthcare indicators in France. Zeitoun et al. (2020) also found no consistent impact of local healthcare on mortality for data until 11 April 2020 in France. Ginsburgh et al. (2021) on the other hand show that inequality is a driving factor for mortality in France for data until 3 September 2020. For Spain until 23 May 2020, García (2021) found that GDP, aeroplane passengers, nursing home beds, and public health expenditure were positively associated with the mortality rate, while urbanity, island location, and doctors per capita were negatively associated. All three studies relied on OLS regressions, which might not account for the spatial dimension of the data, as later discussed in the chapter on methodology.

### 2.3.2 Ecological Research on COVID-19 in Germany

There are four peer-reviewed and published studies on spatial differences of COVID-19 in Germany, as well as two preprints and some journalistic or 'grey' inquiries. In comparison to other countries such as the US, the UK, or China, this is a very low number.

The earliest study conducted in May by Plümper and Neunmayer (2020) identifies two phases of the pandemic: An initial phase of higher IR in wealthy counties, and a second phase after national lockdown (around 20 March), during which underprivileged counties had higher IR. In the first phase, welfare claimants p.c. were negatively associated with cases, but positively in the second phase, and vice versa for academics and income. Their explanation is twofold: By chance, the virus initially spread from Austria to southern Germany, which is a wealthy region. Second, those that brought the virus were mostly ski-tourists and travelers, therefore mostly upper class citizens. The authors model is a simple OLS regression including previous cases, to control for path dependency. This does not to take spatial autocorrelation into account, however, strong autocorrelation should be assumed and violates central assumptions of OLS regressions. It should be noted that the study by Plümper and Neunmayer (2020) is no published or peer-reviewed paper, but a working paper.

Wachtler et al. (2020b) also finds two phases of the pandemic until 15 June 2020: For the data until mid-April, deprived counties had lower case numbers, but starting in mid-May, no difference was observable. In the German south (Bavaria, Baden-Württemberg), this reversal was more pronounced and lead to higher cases for deprived counties. The authors conclude that as the pandemic began in the south, this reversal may also happen in the north. Their measure of deprivation was the respective quintiles of the German Index of Socioeconomic Deprivation (GISD) by Kroll et al. (2017).

In the most recent spatial analysis, the authors of Wachtler et al. (2020b) repeated their study for the second wave and found that deprived counties were more heavily affected by infections (Hoebel et al. 2021). Hoebel et al. (2021) use data from 31 August 2020 to 10 January 2021. Especially at the end of the second wave, deprived regions had higher incidence rates than less deprived. However, in contrast to the first paper, only three groups were compared (lowest quintile, highest quintile, middle quintiles). Why the authors changed their methodology is not reported.

Using machine-learning algorithms, Scarpone et al. (2020) show that for the total cases until 31 March a strong north-south difference in cases is observable. Almost the whole of the INKAR-data base with its 376 variables was tested to search for strong predictors of age adjusted case rates, as well as additional information on structural features. Church density, population density, and foreign guests had a positive impact on incidence rates, whereas unemployment, and voter participation rate had a negative association. Although this study uses socioeconomic data, it is mostly devoid of sociological theory.

For the first wave (until 15 June 2020) Ehlert (2020) confirms that well-off, urban areas were more affected. Spatial regression models (SAR, SEM and SAC) are used to estimate the influence of the variables on case numbers. The models find a negative relationship between cases and unemployment, population over 75, physicians and pupils per population. A positive relationship is found for academics p.c., mean age, life expectancy, and population density. However, they only regard the aggregated data, without accounting for possible changes in effect-size and direction during these six months. As their paper is neither peer-reviewed nor published, these results should be interpreted cautiously.

Although Schröder et al. (2020) did not publish an ecological study, their findings are also important for this paper sine they rely on a German panel survey (socioeconomic Panel, SOEP). They report that the access to home-office is unevenly distributed, as it strongly correlates with education and income. In another study, Wahrendorf et al. (2021) found that unemployed people were vastly overrepresented in COVID-19 related hospital admissions. Ahlers et al. (2021) find that especially low-paid workers fear contracting the Corona virus at their work place.

Finally, there are some non-academic sources showing that on a city-district-scale, deprived districts were heavier affected in early 2021. Media-attention arose after cities such as Cologne (Stadt Köln 2021b), Mannheim (Stadt Mannheim 2021), München, and Hamburg (Stadt Hamburg 2021) started to publish finer-grained data of case rates in spring 2021. Several newspapers and blogs (e.g., Kaiser 2021; infas360 2021; Norddeutscher Rundfunk 2021; Effern et al. 2021) reported on low numbers in rich districts and high rates in densely populated 'problem districts'. As a response, cities such as Mannheim and Cologne started special vaccination campaigns in poorer districts (Stadt Köln 2021a).

To summarise: What these aforementioned studies have in common is their reliance on geospatial data on the NUTS-3 level.[2] Most of the studies also utilize the INKAR-data base. Only one study was published after the first wave, all others were published between May and August 2020. They generally find that socioeconomic relationships with COVID-19 are not stable over time, but move from well-off counties to more deprived ones

[2]Nomenclature des unités territoriales statistiques, a referencing system of the European Union. In Germany, NUTS-1 refers to states (e.g. Baden-Württemberg), NUTS-2 to government regions (e.g. administrative region Karlsruhe), and NUTS-3 to counties (e.g. Heidelberg).

(Plümper and Neunmayer 2020; Wachtler et al. 2020b; Hoebel et al. 2021). Data released by cities suggests that during spring 2021, the pandemic is stratified within counties as well, for which the official data by the RKI does not account for (Kaiser 2021; infas360 2021; Norddeutscher Rundfunk 2021; Effern et al. 2021).

## 2.4 Syndemic Framework

The medical anthropologist Merrill Singer and colleagues coined the term 'syndemic' as a portmanteau of 'synergistic co-occurring epidemics' (Singer 1994). A syndemic is classified by at least two diseases that interact biologically, socially and/or psychologically[3] in a larger social logic[4] (Mendenhall and Singer 2019): Originally, Singer described the AIDS-pandemic of the 1980s and 1990s as part of an 'inner city syndemic', a health crisis in which already vulnerable people (urban young men of colour who have sex with other men) were not only hit by the HIV-virus, but by a drug-pandemic, and other health problems (Singer 1994). This perspective allows to understand a pandemic not only as a health crisis, but as a cultural phenomenon and thus object to sociological studies.

Richard Horton, Editor-in-chief of The Lancet, opined in September 2020 that COVID-19 is a global syndemic (Horton 2020). Some scholars agree with this claim or had previously stated similar beliefs (Bambra et al. 2020; Gravlee 2020; Kenyon 2020; Fronteira et al. 2021), while others refused this as too generalising (Mendenhall 2020). Proponents argue that prevalent pandemics (e.g. obesity, drugs), systematic racism, social/economic inequality, unequal access to healthcare, and diverging living conditions created communities especially vulnerable for the virus and its (economic) consequences. They argue that the pandemic and its long term effects may deepen inequalities around the world and within countries (Fronteira et al. 2021; Gravlee 2020; Bambra et al. 2020).

What is particularly interesting about this debate is that these claims rely on ecological studies. This marks a turn in the history of syndemic research: As highlighted by Trasi et al. (2018), syndemic research has previously been either qualitative/ethnographic, or individualistic and thereby often ignorant of structural features of social differences. Perhaps, as the timeliness and circumstances of the COVID-19 pandemic does not allow for

[3]For example co-morbidity as a biological interaction or stigma as a social interaction.
[4]For example systematic oppression, colonialism, patriarchy.

immediate qualitative research, ecological perspectives are able to serve as first indicators on how the current pandemic interacts with the social realities it encounters. This opens a unique opportunity to enrich this debate in epidemiology with a macro-sociological perspective.

While the syndemic framework plays a vital role in North American and British discourse, this perspective has been missing in Germany. However, structural inequalities in income, location (especially east-west, urban-rural), and the persistence of stratified health pandemics in general (Lampert et al. 2016) and in specific COVID-19 risk areas such as obesity (Mader et al. 2020) and smoking (Tönnies et al. 2021) could mean that COVID-19 may integrate itself into a syndemic system in Germany as well. Current studies show that access to protection against the pandemic (for example: remote working from home) (Schröder et al. 2020) and hospital admissions due to COVID-19 (Wahrendorf et al. 2021) are subject to socioeconomic class differences in Germany and thereby already hint at syndemic tendencies.

# 3 Research Gap and Hypotheses

There are several shortcomings of the current state of research: First of all, the pandemic is constantly evolving, and there is no single up-to-date study with consistent design spanning the whole of the second wave and beyond. Although the research team at the RKI (Wachtler et al. 2020b; Hoebel et al. 2021) conducted a very similar study in June 2020 and January 2021, they slightly change their categories on deprived counties (from five categories in June to three in January) and only published a descriptive report. They did not show, e.g. how much an increase of the IR was associated with an increase of the deprivation index. Second, the German studies that rely on models neglect important details. As mentioned above, the two studies by Wachtler et al. (2020b) and Hoebel et al. (2021) did not construct a model at all. Plümper and Neunmayer (2020) did not account for spatial autocorrelation. Some studies have a faulty interpretation of significance levels (especially: Plümper and Neunmayer 2020; Ehlert 2020), which might have adversely influenced model building previous to publication. Additionally, the temporal dimension is at times ignored (Ehlert 2020; Scarpone et al. 2020), although it may be sensible to assume that the relationships between cases and socioeconomic factors develop over time. Finally, no German study is explicitly grounded in social theory. In this regard Scarpone et al. (2020) represent an extreme as they tested hundreds of variables without further explaining their selection, but other studies did not expand beyond a simple assumption that poor people may be more likely to contract the virus or die from it as well.

What is missing is thus a coherent and model-grounded approach that considers spatial dependencies, the temporal dimension, and grounds empirical research in a theoretical framework. That is precisely the aim of this study.

## 3.1 Hypotheses

In accordance to the identified research gap, two main research hypotheses on the course and impact of COVID-19 in Germany can be derived. These hypotheses are both on IR and CFR. Furthermore, there are two exploratory hypotheses that are rather grounded in empirical research from other countries than theory. The main goal is to test whether the reported evidence of socioeconomic differences in the pandemic may be interpreted as syndemic in Germany.

The definition of syndemics by Mendenhall and Singer (2019) mentions three criteria:

1. At least two epidemics coexists in a given population.
2. These epidemics interact biologically, socially, and/or psychologically.
3. This happens in a larger structural context of oppression.

The mere existence of at least two pandemics in Germany does not need to be proven by this paper, as for example a smoking (World Health Organization 2010) and an obesity epidemic (Swinburn et al. 2020) are well documented. The focus of this paper is rather to gather evidence for the second and third criteria: To uncover interactions between social status and IR and CFR, and to place these interactions into a broader context of structural differences.

Following the previous research in Germany, a change in impact (IR and CFR) of the pandemic over time is expected, namely a move from the rich south to the less wealthy northern and eastern parts of the country. This shift has been explained by other researchers as the diffusion from the initial outbreak into the broader society. If the syndemic framework can be applied, this diffusion is not random but follows socioeconomic categories. This shift should be noticeable once the pandemic has moved from few hot spots to a general incidence throughout the country and should persist throughout the second and third wave (Hypothesis I).

In addition to the social distribution, a syndemic is characterised by its embeddedness in a larger social logic. This paper focuses on two structurally disadvantaged groups: Foreigners and people living in the old GDR (German Democratic Republic). For the primary, the language barrier, lower education, and systematic racism are just some of the obvious structural disadvantages. For the secondary, these are unemployment, lower wages, brain drain, and the aging population. While the first group is rather found in urban areas, the second is more rural. In addition to the general trend in hypothesis I, an additional positive interaction effect is expected for the IR and CFR in counties with a higher proportion of these groups after the first wave (Hypothesis II.1 and II.2).

Furthermore, there are two exploratory hypotheses based on empirical findings in other countries. First, a negative relationship between trust in government and IR is expected. If a larger share of people does not believe in the existence or severity of COVID-19, or does not think that the government acts in the best interest of the people, this might increase rule-breaking and thus the IR. Trust in the government during the pandemic is measured by vote share for the extremist party AfD, which has positioned itself as the parliamentary voice of both virus-deniers and extreme critics of government regulations during the pandemic (Alternative für Deutschland 2021) (Hypothesis E1).

A second exploratory hypothesis concerns healthcare capacities: The findings of the effect of those has been heterogeneous on the international scale. However, with declining hospitals capacities in Germany (DESTATIS 2021), it is of greatest interest to see whether heterogeneities in these may explain part of the CFR. Besides hospital capacities, the population per doctor is used as a second indicator of local healthcare capacities that might be more sensible for less severe cases (Hypothesis E2).

Main Hypotheses

I. For the first wave of the pandemic (until August 2020), counties with lower socioeconomic deprivation were more affected (higher IR and CFR). With the second wave (September 2020 until February 2021) and third wave (starting March 2021), this dynamics shifted and deprived counties were more affected.

II.1 The effect of Hypothesis I is stronger for counties with a larger share of foreigners.

II.2 The effect of Hypothesis I is stronger for counties within the former German Democratic Republic.

Exploratory Hypotheses

E1. The longer the pandemic dwelt, counties with less trust in the government exhibited a higher IR.

E2. Stronger local healthcare capacities resulted in a lower CFR.

Additionally, two robustness checks will be performed: The effect of a state-dummy-system and of a lagged CFR on the relationship in hypothesis I will be calculated. The first allows to see whether state-policies had an impact on the relationship assessed. As especially during the early phases of the pandemic, these regional policies differed, an explanatory power for IR and CFR may be assumed. Second, since there is a time interval between contraction and death, the common definition of the CFR may lead to overestimation in times of rising case numbers and underestimation in times of declining case numbers. To test whether this can be seen in the data, a lagged CFR variable which uses deaths of one month and cases of the previous month is calculated and used for the regression analysis performed in hypothesis I.

# 4 Method

## 4.1 Data

The COVID-19-cases in Germany were obtained from the official data base of the RKI for the time starting with the first confirmed case in January 2020 until 31 May 2021 (Robert Koch Institut 2021). The RKI reports state and county of residence, age group, gender, and reported date of case or death. Thus, the place and date of infection are not reported. While the number of deaths by the virus is probably accurate, the case count should be higher as there is an unknown amount of unrecorded cases. As the ratio between recorded and unrecorded cases probably changes throughout the pandemic, no attempt is made to correct this in the RKI data. This would create a source for false interpretations and render this study incomparable to others. The same applies to the discrepancy between location of infection and of residence.

Second, socioeconomic data from the INKAR-data base by the BBSR was used (BBSR 2021). Most of the variables stem from 2017. The focus on this study lies on how the pandemic affects counties stratified by their socioeconomic characteristics, not how these change during the pandemic. Thus, in an ideal case, the data should be from 2019, as close to the outbreak of the pandemic as possible. 2017 is very close to this ideal case and as relative differences within Germany are the main focus, country-wide trends are not of importance. Within-country changes in these two years may confound the results, though this is expected to be of negligible influence.

As the two data sets mentioned above use different geographic referencing systems for counties (RKI: 'Kreisschlüssel', a German referencing system; INKAR: NUTS-3, a European referencing system), a third data set was needed to merge the two: The data set for counties by population and population density (DESTATIS 2018) by the Federal Statistical Office of Germany.

Although the original data from the German Index of Socioeconomic Deprivation (GISD) by Kroll et al. (2017) stems from the INKAR data base, the already calculated index from the 2018 revision was obtained separately (Kroll et al. 2018).

In order to create the maps and to estimate the spatial regression model, shapefiles of the German NUTS-1 and NUTS-3 regions were obtained from the Bundesamt für Kartographie und Geodäsie (BKG 2020).

## 4.2 Variables

### 4.2.1 Dependent Variables

The Incidence Rate (IR) during a pandemic is defined as number of new recorded cases within a specified time period in a specified population (Rychethnik et al. 2004). For this study, the IR was calculated for each German county and calendar month (excluding cases in January, which were included in February due to low numbers). The IR were standardised to cases per 100,000, as has been convention in Germany.

$$Incidence\ Rate\ (IR, in\ cases\ per\ 100,000) = \frac{Number\ of\ cases\ of\ COVID-19}{Population} \times 100,000 \quad (1)$$

The Case Fatality Ratio (CFR) was calculated in accordance to the definition of the World Health Organization (Organization 2020). It is defined as the proportion of people with a diagnosis that die from the disease in question (ibid.). It differs from the Infection Fatality Ratio (IFR), which relates the deaths to all infections. It also is different to the general mortality, that merely states how many people as a share of the population die from an illness.

$$Case\ Fatality\ Ratio\ (CFR, in\%) = \frac{Number\ of\ deaths\ of\ COVID-19}{Number\ of\ confirmed\ cases\ of\ disease} \times 100 \quad (2)$$

A sample of independent variables is shown in table 1, with the complete table in the appendix as table 3.

Table 1: Dependent Variables by county, selection

| Statistic | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| CFR | 401 | 0.03 | 0.01 | 0.01 | 0.1 |
| IR | 401 | 4,440.4 | 1,412.0 | 1,054.8 | 9,453.2 |
| aggrCases | 401 | 9,270.0 | 11,985.7 | 838 | 179,702 |
| aggrDeaths | 401 | 224.6 | 260.0 | 9 | 3,529 |
| CFR20.04 | 401 | 0.1 | 0.04 | 0.0 | 0.3 |
| CFR21.04 | 401 | 0.01 | 0.01 | 0.00 | 0.00 |
| IR20.04 | 401 | 111.4 | 97.4 | 7.2 | 841.3 |
| IR21.04 | 401 | 686.4 | 259.4 | 143.0 | 1,663.4 |

#### 4.2.2 Independent Variables

As a measure for deprivation, the GISD introduced by Kroll et al. (2017) is employed in its 2018 version (Kroll et al. 2018) (for a map of GISD values in Germany see appendix 10). The GISD is an index constructed for depicting socioeconomic deprivation in a public health context and has been able to explain substantial regional differences in life expectancy and smoking (Kroll et al. 2017). The index is additive and consists of three equally weighted categories with two to three items each: Education (school leavers without certificate, employed at place of residence with university degree), occupation (Unemployed, gross income and wage, employment quota), and income (debtor quota, net household income, tax revenue). The index is standardised to $[0, 1]$, with 0 as the county with the lowest deprivation (Kreisfreie Stadt Erlangen) and 1 for the one with the highest deprivation (Landkreis Mansfeld-Südharz). As the data relies on the INKAR and the oldest item is from 2014, the GISD actually measures the relative deprivation as of 2014. As there has been one merging of counties in Germany since 2014, the observation for 'Landkreis Osterode am Harz' has been deleted and its GISD has been included in the GISD of the 'Landkreis Göttingen' (via weighted average by population). As the GISD is a relative index, this change might distort the index on a very small scale.

On a critical note, the GISD does not show interactions of indicators. Its macroscopic nature cannot differentiate between two counties with an identical share of unemployed and in debt people, in which in one of the counties all debtors are unemployed, while in the other, these shares are disjoint. Thus, the actual deprivation in the first county may be higher. This is a form of the ecological fallacy.

The variable for hospital beds per 1,000 people had two missing values for the counties Fürth and Sömmerda, these were replaced by the state averages.

A complete summary of the independent variables is depicted in appendix table 5 (excluding variables such as state or binaries, e.g. east/west). However, it is important to note that the summary statistics are not weighted by population. Thus, for example the mean population density in Germany is not 533.7 people per square kilometre, but close to 232 (BBSR 2019). This is just the average of the 401 German counties.

## 4.3 Method

### 4.3.1 General Specifications

As the main hypotheses predict a change in relationship between the independent and dependent variables, a singular model that only calculates one coefficient for the whole pandemic is of no use. Instead, for each calendar month a model was fitted. Afterwards, the resulting coefficients are compared.

It may be that the actual number of cases is higher than the recorded cases. However, as estimates of the actual number of cases are controversial and change depending on the stage within the pandemic (e.g. due to the changing availability of tests), an assumption will be made for simplification: The recorded cases are treated as totality. This makes prerequisites for significance levels obsolete, which would not be fulfilled anyhow as recorded cases are not randomly drawn from population. For example, it is plausible that recorded cases are more likely severe, thus treating them as random sample would overestimate the share of older people or those with preexisting conditions. This point will be revisited in the discussion.

The IR was logarithmised in order to compute monthly effects as percentages and compare months with differing average IR. Independent variables were logarithmised when they were not already expressed as percentages (e.g. unemployment) to make effect sizes comparable. As the logarithm of zero is not defined, a consistent strategy for zeros was formulated: Following common practice (Bellégo and Pape 2019), a small positive number was added to each zero. For IR, this value was 0.1. In words: one case per one million people. Calculations with other values were run to confirm that this arbitrary chosen number does not influence the results by a substantial degree.

Due to the definition of CFR, if no cases were recorded in a given month in a county, the CFR is not defined. In order to make regressions comparable, these undefined values were recoded to zeros. Overall, 38 of 5614 CFR ($\#counties \times \#months = 401 \times 14$) were recoded this way. This is preferable to trying to omit missing values completely, as spatial regressions rely on values of neighboring regions.

All models controlled for age structure and gender (im)balance of the county, as these factors have been shown to impact mortality and transmissions of COVID-19 on the individual level.

All models and images were created within RStudio (RStudio Team 2021) in the statistical programming language R (R Core Team 2021). For a complete overview of packages and the link to the code-repository see section 8 (Code Repository and Data).

#### 4.3.2 Regression Model

Ordinary Least Square (OLS) Regressions are the general purpose tool of social research. However, they rely on an assumption that is violated in a pandemic: OLS regressions assume that values are identically independently distributed, thus that an increase of $Y_i$ has no direct influence on $Y_j$ with $i \neq j$. During a pandemic, it seems plausible that an increase in the IR in one county has a positive impact on the IR in its neighbors: People frequently move across county borders to socialise, work, or go to school and thereby transmit the virus. The more people are infected in county $i$, the higher the likelihood that infected people spread the disease to regions $\neg i$ nearby. For the CFR, it seems less obvious why it should behave this way, but if underlying patterns (unemployment, other pandemics) are spatially autocorrelated, the CFR will also be regionally clustered.

The common measure for spatial autocorrelation is Moran's I, which takes a value between $-1$ and 1 for variables in at least two dimensions. $-1$ corresponds to perfect dispersion, $+1$ to perfect spatial autocorrelation, and close to 0 for a random arrangement [5]. As depicted in table 2, spatial autocorrelation can indeed be assumed for IR, with the potential exception of the CFR.

Table 2: Moran's I, selection of dependent variables

| Variable | Moran's I |
|---|---|
| IR total | 0.71 |
| IR April 2020 | 0.64 |
| IR April 2021 | 0.58 |
| CFR total | 0.33 |
| CFR April 2020 | 0.01 |
| CFR April 2021 | 0.28 |

A model that can consider the spatial dimension of the data as well as keeping the simplicity of OLS regressions is the Simultaneous Autoregressive model (SAR) (this description of SAR-models mainly follows: LeSage 2008). It is essentially a regression with a 'spatial lag' that has also been employed in epidemiological modelling (e.g., Ehlert 2020). A model of a data set with $n$ regions can be specified as follows:

$$Y = \rho WY + X\beta + \epsilon \tag{3}$$

Similar to an OLS regression, $Y$ is the dependent variable (e.g. IR), $X$ the independent variable (e.g. unemployment), $\beta$ the regression coefficient, and $\epsilon$ the error term. The expression $\rho WY$ captures the autocorrelation between dependent variables. $\rho$ is an estimated coefficient of the strength of the spatial autocorrelation with a range of $[-1, +1]$.

[5] A graphical example for perfect dispersion and perfect spatial autocorrelation is included in the appendix as figure 4.

$W$ is a $n \times n$ matrix that contains all possible pairings of all $n$ regions (region 1 with region 1, region 1 with region 2, region 1 with region 3, etc.). Thus, $Y_i$ is dependent on every other $Y_j$. In practice, a bilateral connection in $W$ is either binary (e.g. 1 if neighbor, 0 if not), or a standardised binary (1/(*number neighbors*) if neighbor, 0 if not). The latter allows that the term $\rho WY$ calculates the weighted average $Y$ of all neighbors. In this paper, $W$ was row-standardised. In contrast to OLS regressions, the coefficients of the SAR model cannot be estimated via ordinary least squares. The R package used for this analysis (Bivand et al. 2013) employs maximum likelihood estimation, although other techniques are possible.

A major drawback of the SAR model is that its coefficient cannot be interpreted directly as marginal effects. As the relationship $Y \backsim X$ is influenced by $n^n$ bilateral connections, a change in $X$ creates a spillover-effect throughout the whole model. A measure to interpret effects in a SAR model is to distinguish direct and indirect effects, as discussed by LeSage and Pace (2014). In this specification, $Average\ Total\ effect = Average\ Direct\ effect + Average\ Indirect\ effect$. The Average Direct effect can be understood as the average change in a specific $Y_i$ by a change in $X_i$, considering all spillover-effects and feed-back-loops from other regions. For example, how a higher share of unemployed in one specific region would change the IR in this region. The Average Indirect effect can be understood as the average change in a specific $Y_i$ if all other $X_j$ change, considering all spillover-effects and feed-back-loops from other regions. For example, how an increase in unemployment in all regions except $i$ impacts the IR in a specific region $i$. The Average Total Effect as a sum of these two effects can be interpreted akin to an OLS coefficient: As the effect on $Y$ by a marginal change in $X$.

# 5 Results

The comparison of the average total effects show that for IR, hypothesis I cannot be rejected (figure 1, for a full overview of effects tables, see appendix table 6). There is a clear trend from higher IR in less deprived regions during the first wave (February to May), to a summer without clear direction (June to September), towards higher IR in deprived regions during the second and third wave (October to May 2021). The size of the effect was highest in March 2020 (negative) and February 2021 (positive). In March 2020, the model depicted a difference between the highest and lowest county by GISD (0 to 1) of −2.64%. In February 2021, this gap was +1.78%. This is a small effect: For all of March 2020 with an average monthly county-IR of 86, this maximal predicted difference was equivalent to less than 3 cases per 100,000. In February 2021 this gap was slightly over 5 cases per 100,000.

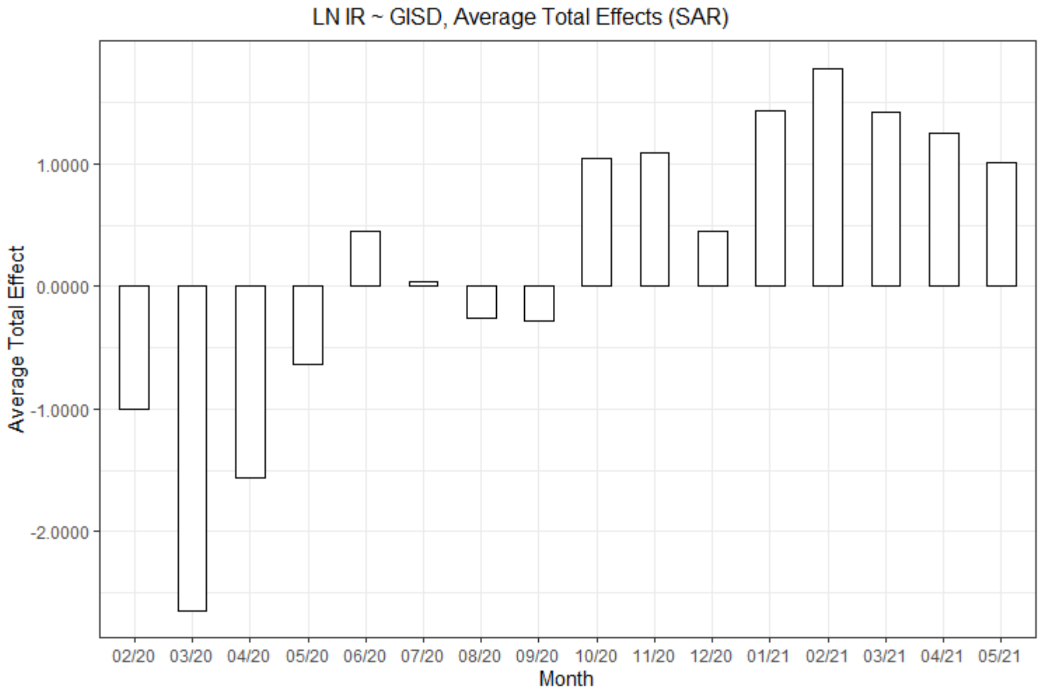


Figure 1: Hypothesis I: IR ∽ GISD

However, the expected effect cannot be observed in the CFR (figure 2, for a full overview of OLS and SAR effects, see appendix table 6). Mortality does not seem to be systematically associated with GISD in a temporal perspective. The difference between OLS and SAR regression was small, as was to be expected due to the Moran's I values. However, the effect of the GISD on the CFR was larger than its effect on the IR. Since the effect did not behave as predicted, the separate monthly effects should not be interpreted.

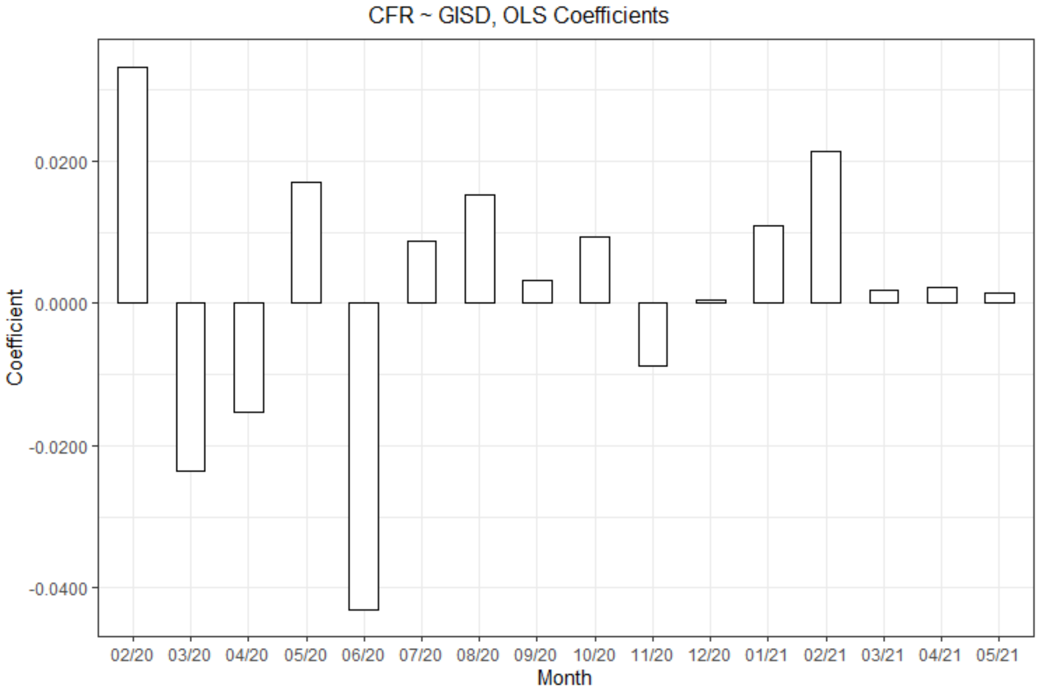


Figure 2: Hypothesis I: CFR ∽ GISD

As the GISD is an additive index, three of its core components (unemployment, median income, workers with academic education) were further analysed to asses whether the direction of the effect is uniform in its components or heterogeneous. This analysis shows that the trend, including the change of direction, of the relationship between IR and GISD is most similar to the income effect, while the turn of the relationship of IR with unemployment (negative to positive) and academic education (positive to negative) started as early as May 2020 (see appendix figures 11, 12, 13). For the relationship with the CFR, these components also showed no homogeneous direction (see appendix figures 14, 15, 16).

Testing hypothesis II.1, the interaction of the share of people with a foreign nationality and the GISD on IR is similar to the relationship of GISD and IR: While deprived areas with a higher share of foreign nationals had a lower IR from February to April 2020, the trend reverses in May 2020 and remains positive (for a table with the total average effects of GISD, foreign nationality and the interaction, see appendix 7). During August and September 2020, the interaction effect dropped, otherwise, a slow upward trend during the second and third can be recognised. The effect size is moderate: Given a maximum GISD of 1, an increase of the percentage of foreign nationals by 1 percentage points increases IR by circa 0.10% after summer 2020 (figure 3).

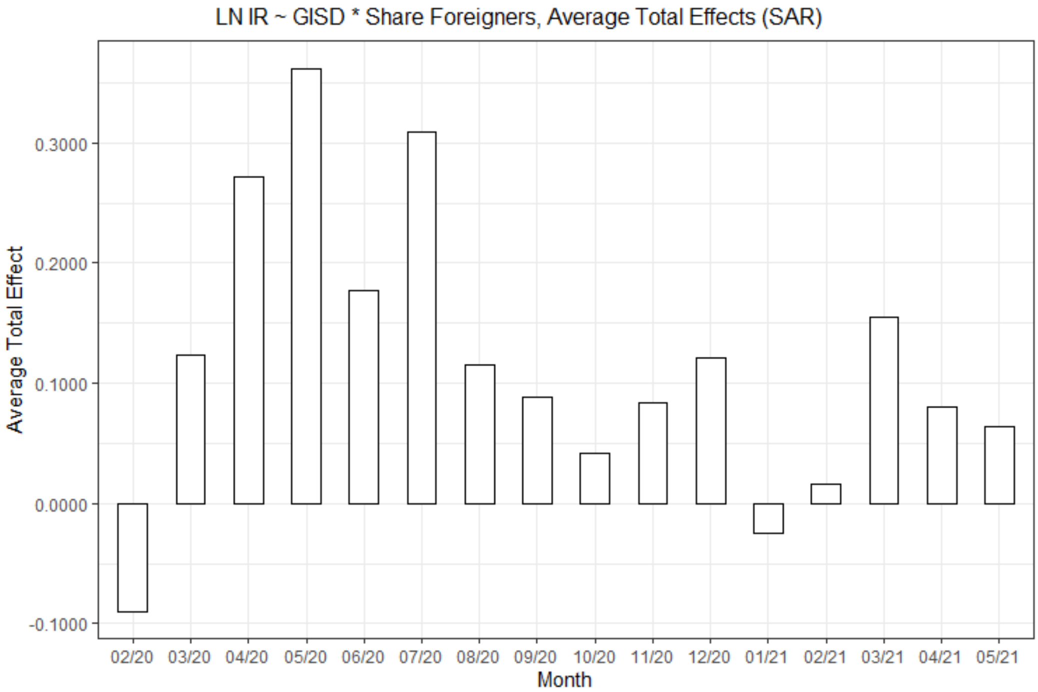


Figure 3: Hypothesis II.1: IR $\backsim$ GISD $\times$ share of foreign nationals

The interaction effect of GISD and share of foreign nationals shows a different pattern when regressed against the CFR: The interaction was strongest at the beginning of the first wave, subsequently shrunk and showed no great effect of consistent direction after June 2020 (for a table with the total average effects of GISD, foreign nationality and the interaction, see appendix 8). Considering the size of the CFR, the effect size started strong and became very small after June 2020. During February 2020, with a GISD of 1, an increase of the percentage of foreign nationals by 1 percentage points increases CFR by circa 0.001. As the CFR was in February on overage 0.002, this is a strong impact. But while the CFR rose and stabilised at between 0.01 and 0.04 during the second wave, the interaction had no relevant size afterwards.

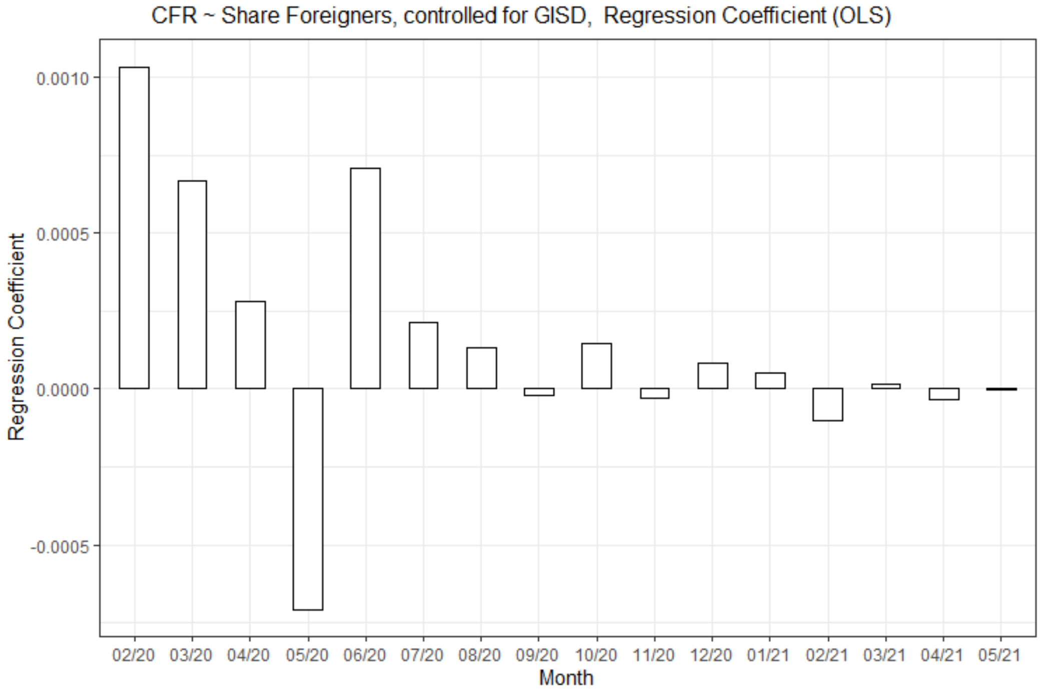


Figure 4: Hypothesis II.1: CFR $\backsim$ GISD $\times$ share of foreign nationals

The interaction of eastern counties and GISD tested for hypothesis II.2 also follows the patterns described in the previous subsections: While deprived eastern counties had lower IR from February to May 2020, this trend reverses in June 2020 and continues with few exceptions until April 2021 (for a table with the total average effects of GISD, the east dummy and the interaction, see appendix **??**). For a county with the maximum GISD of 1, additionally being located in the former GDR increased IR by $-2.86\%$ in March 2021 and $+1.76\%$ in February 2021.

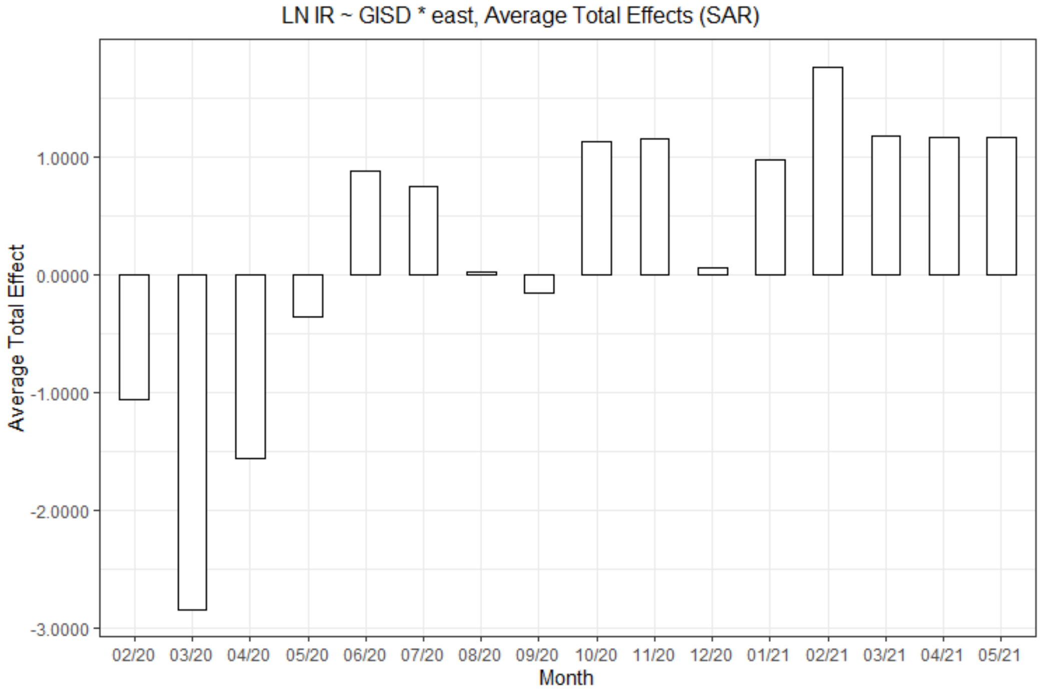


Figure 5: Hypothesis II.2: IR $\backsim$ GISD $\times$ east dummy

As with the previous indicators, the relationship with the CFR was without temporal order, but strong in some months (see appendix figure 17, for a table with the total average effects of GISD, the east dummy and the interaction, see appendix **??**). As the predicted course is not followed, and the possibility of other explanations for this pattern may have been missed, this interaction is not interpreted.

With the exception of July 2020, the first wave and the subsequent summer 2020 saw no uniform or grand impact of the share of votes the AfD received in 2017 on the IR (see appendix figure 18). However, from September 2020, the share of votes had a strong relationship with the IR, peaking in December 2020 at +0.10% per additional percentage point for the AfD. Thus, it may be inferred that trust in government had a strong impact on risky behaviour and thus IR.

For assessing the impact of local healthcare on CFR, the variables 'People per Doctor' and 'Hospital beds per 1,000 people' were used. Although in some months the effect of hospital beds was strong on the CFR, as with most variables, the overall pathway was unpredictable (see appendix figure 20). Thus, as with previous predictors of CFR, these explorations yields no insights into components of the CFR. On the other hand, the effect of population per doctor had the predicted impact, with the exception of February 2020.

However, this effect was rather small (see appendix figure 19).

The robustness checks with a lagged CFR and state dummies indicate that both have no strong impact on the results. While the relationship in hypothesis I with a lagged CFR or state dummies for policy differences still follows no clear path over the months (graphic 22, graphic 23), state dummies also neither strengthened nor weakened the effect of hypothesis I for IR (graphic 21). For the lagged CFR, the first two months had to be omitted for scarcity of cases and deaths.

# 6 Discussion

The results show that the virus was more prevalent in socially less deprived counties during the first wave. They also show that this effect has reversed during the early phase of the pandemic and that more deprived counties were more affected during the second and third wave. Furthermore, there is evidence that both counties with more people with foreign nationalities as well as counties in the old GDR have a stronger deprivation effect on IR. This supports the claim that the COVID-19 pandemic is syndemic in transmission: The propensity to take public transportation, working on-site, and living in crowded apartments are potential social reasons why counties with more people with lower socio-economic status have a higher IR. Additionally, the effect for foreign nationals and the East Germany confirm that structural disadvantages in IR are present.

On the other hand, the CFR does not seem to be systemically related to either social deprivation or the stage of the pandemic. An obvious explanation is that there were not 'enough' severe cases for a strong, socially stratified impact of deaths. Another argument could be that the German healthcare and social system protected those infected more or less equally, so that mortality on a county scale shows little evidence of a syndemic.

It seems surprising that healthcare variables had no consistent effect on CFR. However, this result is similar to Gaudart et al. (2021) for France during the first wave. An explanation may be that German hospitals never reached their full capacities for a considerable period of time during the previous year. As for doctors, the rather small effect in line with the effect of the GISD on IR may show that to some extent, the varying availability of the first line of healthcare may have caused a higher CFR in deprived counties. As the second effect was very small, it could be possible that the German healthcare system may have proven to be an equaliser in Germany during the pandemic.

As there are no studies of similar design and topicality, the effect sizes cannot be compared directly. However, the findings are generally en par with other studies on Germany, as they also documented the shift from less to more deprived counties (Plümper and Neunmayer 2020; Wachtler et al. 2020b; Hoebel et al. 2021). It should be noted that for the GISD, this change in direction occurs later in comparison to the indicators (e.g. welfare claimants) used by other studies. A small to non-existent effect of healthcare capacities on CFR has also been observed in France (Zeitoun et al. 2020; Gaudart et al. 2021). Additionally, the effect of political variables on IR has been observed in Europe (voter turnout) (Bartscher et al. 2020) and the US (share of votes for Donald Trump) (Mourad et al. 2020; Iyanda et al. 2021). On the other hand, the relationship between IR and GISD is much weaker in comparison to the study of Wachtler et al. (2020b), who found the difference between the highest quintile and the lowest quintile in IR was between 169 and 184 for data accumulated until mid-June 2020. One difference might be different approaches to age-standardisation, or the use of a spatial model for this study.

The greatest weakness to a study of this design is the ecological fallacy: As cases and deaths are registered at the county level, the actual micro-reality could be entirely opposed to what is inferred from the data. Just because areas with high social deprivation have higher IR during the second and third wave, it is still possible that the infected individuals are of highest socioeconomic status. Thus, studies of this design can only describe aggregated effects that can hint at individual mechanisms. However, current media publications on city districts find similar results as this study for the IR during the third wave (e.g., Kaiser 2021; infas360 2021; Norddeutscher Rundfunk 2021; Effern et al. 2021). Hopefully, data that allows for research on the socioeconomic status of individuals during the pandemic will be made available to confirm or reject the ecological findings.

A methodological confound of the study might be that relying on reported cases causes an overestimation of severe cases. There are certainly unreported cases, for each of the observed months. As data for a complete picture is lacking, this is a weakness that cannot be addressed to a satisfactory extent.

What neither this study nor other studies consider is the impact of neighbouring countries. This should be addressed by further studies, as this could mitigate the east-west differences, as cases in Czech were very high during most of the second and third wave. On the other hand, the SAR model already indirectly considers this effect when it takes neighbouring counties into account that also lie at the country border.

Another possible confound is that IR can be path-dependent: The current IR may be positively correlated to the IR of the previous month. On the other hand, this path dependency is not fully random but higher cases in the past may be due to social factors as well. Controlling for these, as done by Plümper and Neunmayer (2020), may filter out an important channel that is shaped by social factors and therefore of interest.

INKAR data is slightly outdated. Especially the GISD with data from 2014 is five years past the optimal date for an independent variable in this study. However, it may be argued that most developments in Germany during these years leave the relative social deprivation between counties intact.

Relevant for the long run effects of the COVID-19 pandemic is how well vaccines are received. As rich people are more likely to be vaccinated against the seasonal flu (Nagata et al. 2013; Jain et al. 2017), such an inequality in COVID-19-vaccinations would lead to unequal long-term effects. This study cannot examine these effects, as the vaccine rollout is still taking place and no detailed data on socioeconomic status of treated people is available. Another long-term issue is described as 'long COVID', chronic fatigue, that lasts after people are cured (Mahase 2020). If the load of the pandemic is carried by certain people and groups, as this study suggests, long COVID and unequal vaccination rates may further deepen the stratified impact of the pandemic in the years to come.

Other areas for further studies could be whether closeness to the sea exerts a negative impacts on infection rates. In the IR map (appendix 8), ocean-counties seem to fare better than other counties. Also, a meta-analysis of the newspaper articles regarding inner-city differences and their relationship to socioeconomic status seems pressing. All in all, there is a great need for an academic – and social – discourse on how COVID-19 shapes the German society and which groups are disproportionately affected. These studies to come need better data, the RKI should consider releasing its data on the postcode, or at least municipality level.

This study uncovers the impacts of social deprivation on the course of COVID-19 during the first year of the pandemic in Germany. It demonstrates that in regard to infections, Germany might have been and might still be in a syndemic.

# 7 Appendix

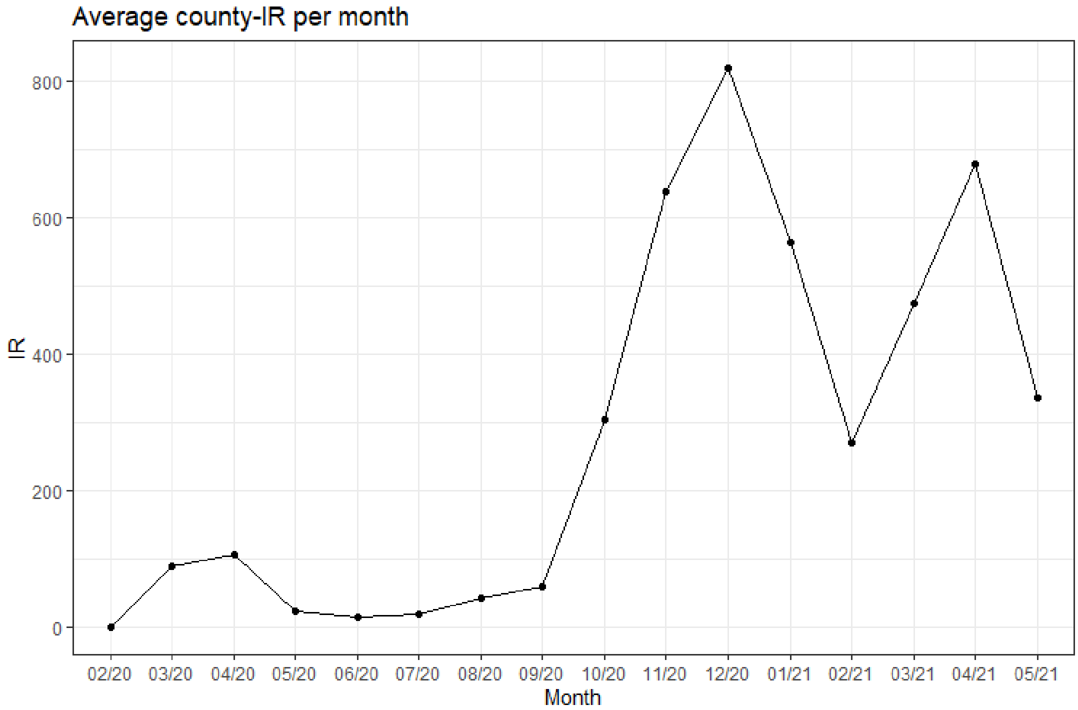


Figure 6: German average county-IR per Month

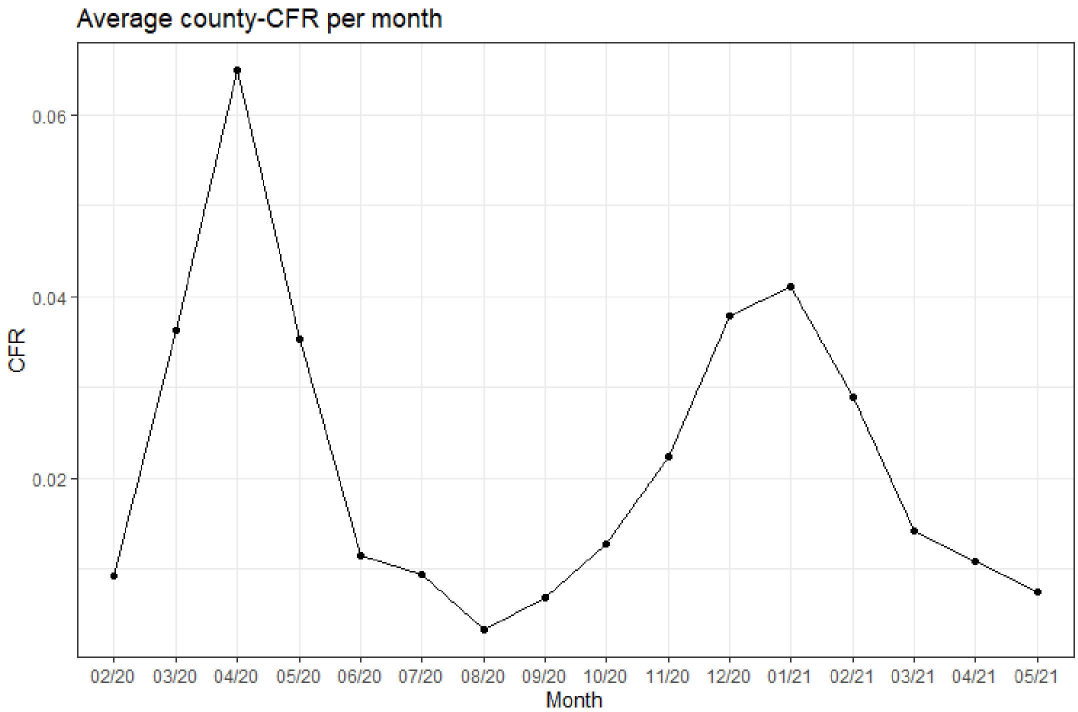


Figure 7: German average county-CFR per Month

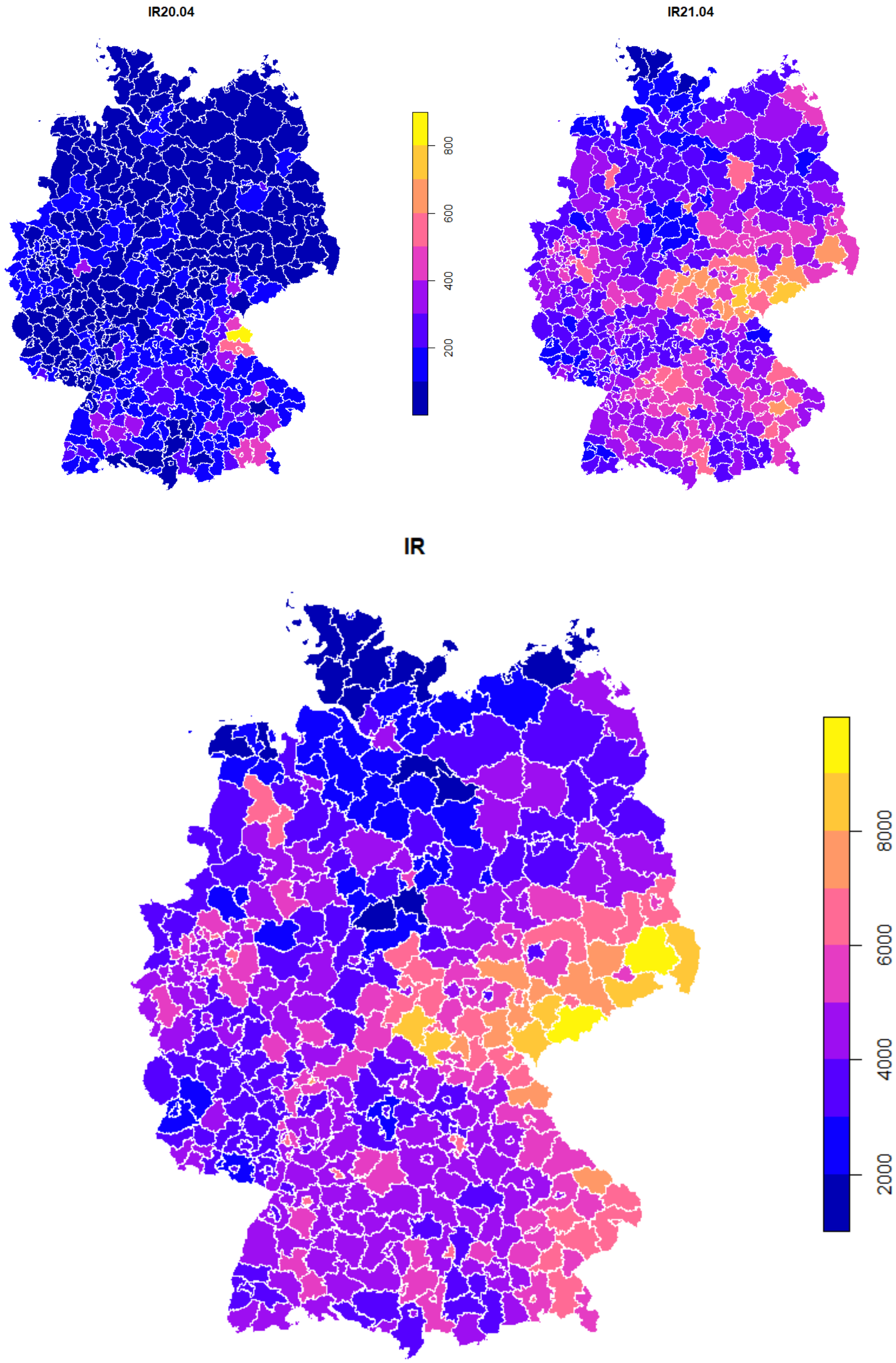


Figure 8: Map: German IR in April 2020; April 2021; and in total

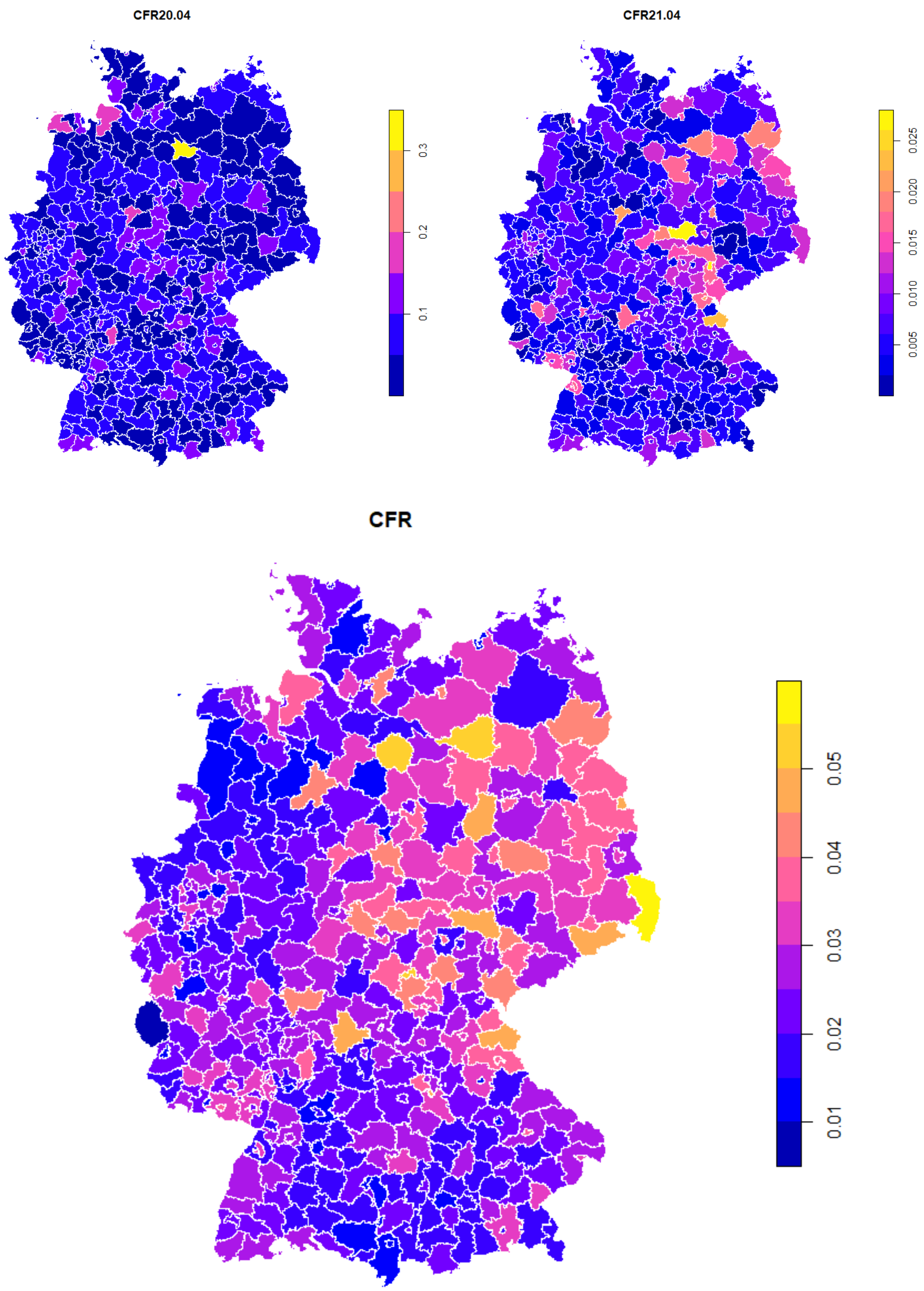


Figure 9: Map: German CFR in April 2020; April 2021; and in total

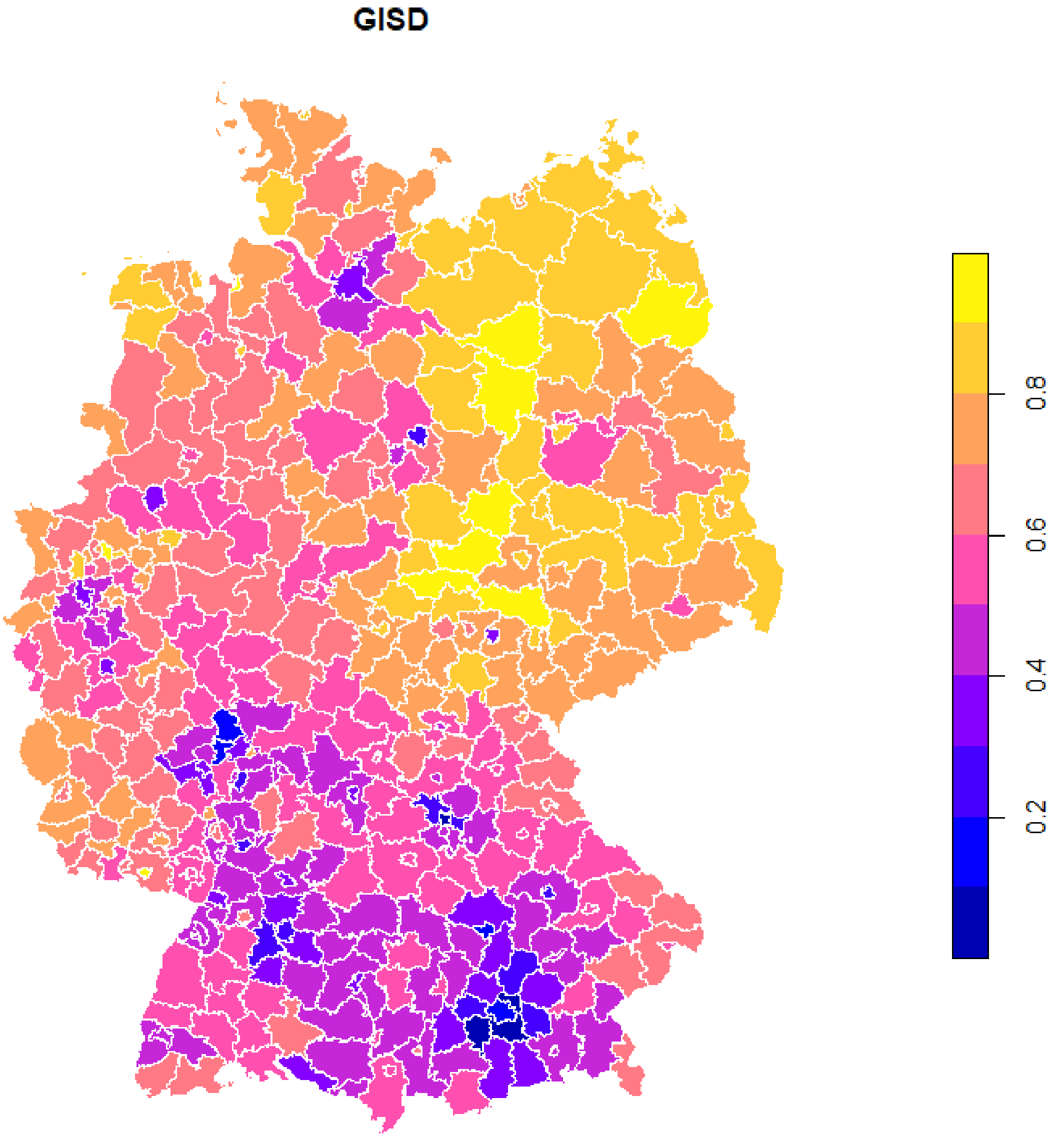


Figure 10: Map: GISD

Table 3: Dependent Variables by county, full table, rounded to second decimal

| Statistic | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| CFR | 401 | 0.03 | 0.01 | 0.01 | 0.06 |
| IR | 401 | 4,440.36 | 1,411.96 | 1,054.82 | 9,453.18 |
| aggrCases | 401 | 9,270.00 | 11,985.72 | 838 | 179,702 |
| aggrDeaths | 401 | 224.62 | 259.99 | 9 | 3,529 |
| CFR20.02 | 401 | 0.002 | 0.05 | 0.00 | 1 |
| CFR20.03 | 401 | 0.03 | 0.03 | 0.00 | 0.29 |
| CFR20.04 | 401 | 0.06 | 0.04 | 0.00 | 0.33 |
| CFR20.05 | 401 | 0.04 | 0.09 | 0.00 | 1 |
| CFR20.06 | 401 | 0.02 | 0.08 | 0.00 | 1 |
| CFR20.07 | 401 | 0.01 | 0.04 | 0.00 | 0.00 |
| CFR20.08 | 401 | 0.004 | 0.01 | 0.00 | 0.00 |
| CFR20.09 | 401 | 0.01 | 0.02 | 0.00 | 0.00 |
| CFR20.10 | 401 | 0.01 | 0.01 | 0.00 | 0.00 |
| CFR20.11 | 401 | 0.02 | 0.01 | 0.00 | 0.08 |
| CFR20.12 | 401 | 0.04 | 0.02 | 0.01 | 0.10 |
| CFR21.01 | 401 | 0.04 | 0.02 | 0.00 | 0.12 |
| CFR21.02 | 401 | 0.03 | 0.02 | 0.00 | 0.09 |
| CFR21.03 | 401 | 0.01 | 0.01 | 0.00 | 0.05 |
| CFR21.04 | 401 | 0.01 | 0.01 | 0.00 | 0.00 |
| CFR21.05 | 401 | 0.01 | 0.01 | 0.00 | 0.00 |
| IR20.02 | 401 | 0.12 | 1.01 | 0.00 | 20 |
| IR20.03 | 401 | 85.99 | 69.13 | 9.18 | 691.00 |
| IR20.04 | 401 | 111.40 | 97.35 | 7.18 | 841.33 |
| IR20.05 | 401 | 23.12 | 26.50 | 0.00 | 245.57 |
| IR20.06 | 401 | 12.01 | 27.52 | 0.00 | 468.30 |
| IR20.07 | 401 | 15.58 | 20.23 | 0.00 | 300.36 |
| IR20.08 | 401 | 35.40 | 27.93 | 0.87 | 342.97 |
| IR20.09 | 401 | 49.51 | 34.45 | 2.39 | 220.85 |
| IR20.10 | 401 | 263.64 | 138.56 | 42.75 | 859.77 |
| IR20.11 | 401 | 593.02 | 262.81 | 72.99 | 1,538.66 |
| IR20.12 | 401 | 822.85 | 413.84 | 120.16 | 2,649.89 |
| IR21.01 | 401 | 590.33 | 280.76 | 147.43 | 1,577.55 |
| IR21.02 | 401 | 285.88 | 154.34 | 54.05 | 1,292.34 |
| IR21.03 | 401 | 486.34 | 243.63 | 97.94 | 2,151.61 |
| IR21.04 | 401 | 686.38 | 259.45 | 142.98 | 1,663.40 |
| IR21.05 | 401 | 338.34 | 142.66 | 71.51 | 881.15 |

Table 4: Graphical Example, Moran's I

| | | | |
|---|---|---|---|
| ■ | ★ | ■ | ★ |
| ★ | ■ | ★ | ■ |
| ■ | ★ | ■ | ★ |
| ★ | ■ | ★ | ■ |

(a) Moran's $I = -1$

| | | | |
|---|---|---|---|
| ■ | ■ | ★ | ★ |
| ■ | ■ | ★ | ★ |
| ■ | ■ | ★ | ★ |
| ■ | ■ | ★ | ★ |

(b) Moran's $I = +1$

Table 5: Independent Variables by county

| Statistic | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| GISD | 401 | 0.6 | 0.2 | 0.0 | 1.0 |
| unemployment | 401 | 5.4 | 2.4 | 1.5 | 14.0 |
| medInc | 401 | 3,064.9 | 451.1 | 2,183 | 4,635 |
| workersAcadem | 401 | 12.0 | 5.2 | 4.8 | 33.2 |
| popDensity | 401 | 736.3 | 178.3 | 328 | 1,364 |
| population | 401 | 207,030.5 | 243,880.3 | 34,209 | 3,644,826 |
| shareWomen | 401 | 50.6 | 0.6 | 48.4 | 52.7 |
| avgAge | 401 | 44.5 | 2.0 | 39.8 | 50.2 |
| shareForeign | 401 | 10.0 | 5.1 | 1.9 | 35.0 |
| AfD | 401 | 13.4 | 5.3 | 4.9 | 35.5 |
| hospBeds | 401 | 6.4 | 3.9 | 0.0 | 29.6 |
| popPerDoc | 401 | 533.7 | 702.7 | 36 | 4,686 |

Table 6: Hypothesis 1: IR $\backsim$ GISD, CFR $\backsim$ GISD

| | IR | | CFR | | |
|---|---|---|---|---|---|
| Month | ATE | $\rho$ | ATE | $\rho$ | $\beta$ |
| 02/2020 | -1.0001 | 0.2355 | 0.0328 | -0.0151 | 0.0333 |
| 03/2020 | -2.6428 | 0.5991 | -0.0287 | 0.1940 | -0.0236 |
| 04/2020 | -1.5589 | 0.6725 | -0.0147 | -0.0199 | -0.0148 |
| 05/2020 | -0.6243 | 0.5633 | 0.0108 | -0.1143 | 0.0157 |
| 06/2020 | 0.4590 | 0.4371 | -0.0430 | -0.0090 | -0.0430 |
| 07/2020 | 0.0343 | 0.3351 | 0.0078 | -0.1046 | 0.0093 |
| 08/2020 | 0.2590 | 0.4104 | 0.0159 | 0.0100 | 0.0157 |
| 09/2020 | -0.2857 | 0.4082 | 0.0019 | -0.0954 | 0.0027 |
| 10/2020 | 1.0311 | 0.7096 | 0.0098 | 0.1270 | 0.0092 |
| 11/2020 | 1.0829 | 0.7546 | -0.0089 | 0.0797 | -0.0088 |
| 12/2020 | 0.4466 | 0.8260 | -0.0003 | 0.1523 | 0.0002 |
| 01/2021 | 1.4121 | 0.6818 | 0.0106 | 0.1079 | 0.0107 |
| 02/2021 | 1.7732 | 0.6169 | 0.0211 | 0.2062 | 0.0191 |
| 03/2021 | 1.4211 | 0.7007 | 0.0001 | 0.1008 | 0.0004 |
| 04/2021 | 1.2812 | 0.7239 | 0.0006 | 0.2340 | 0.0008 |
| 05/2021 | 1.2812 | 0.7239 | 0.0006 | 0.2340 | 0.0008 |

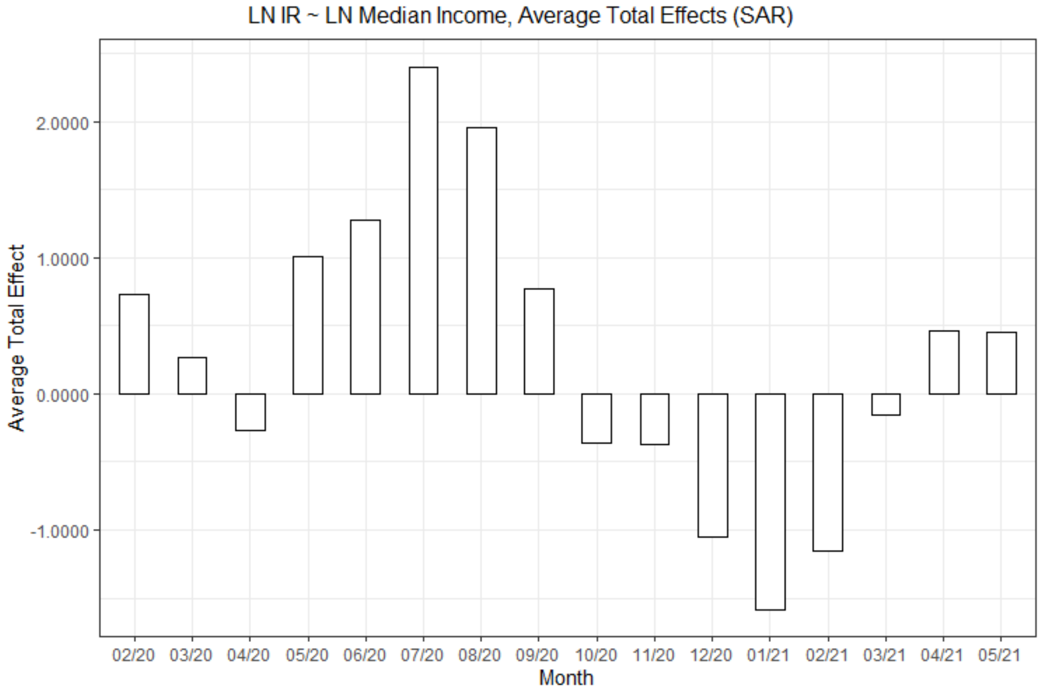


Figure 11: Hypothesis 1: IR ∽ LN Median Income

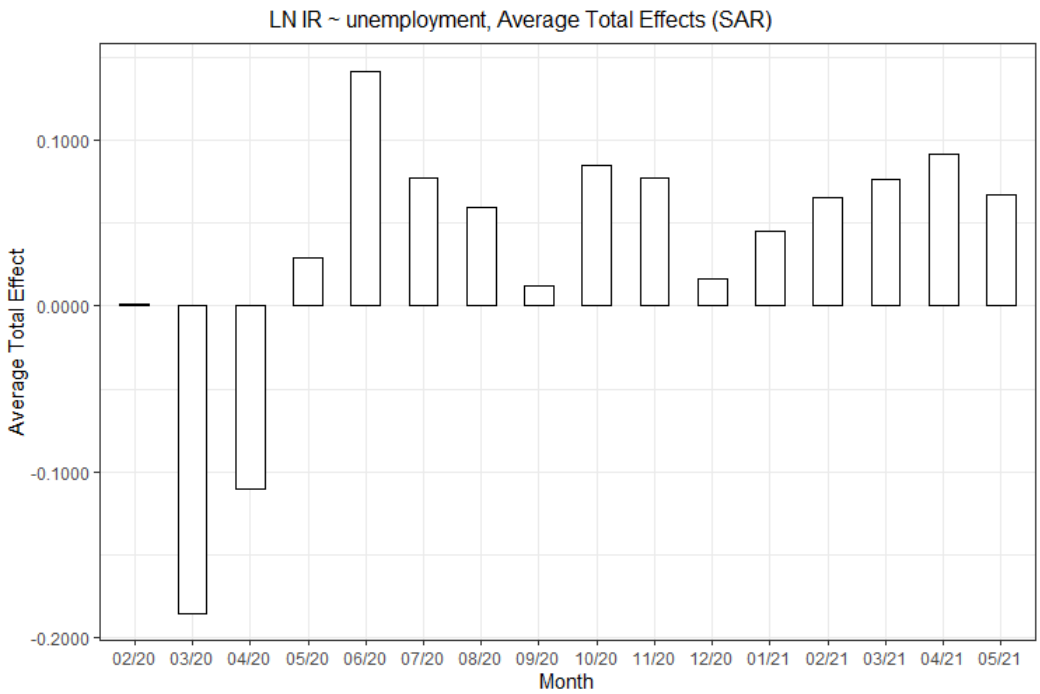


Figure 12: Hypothesis 1: IR ∽ Unemployment

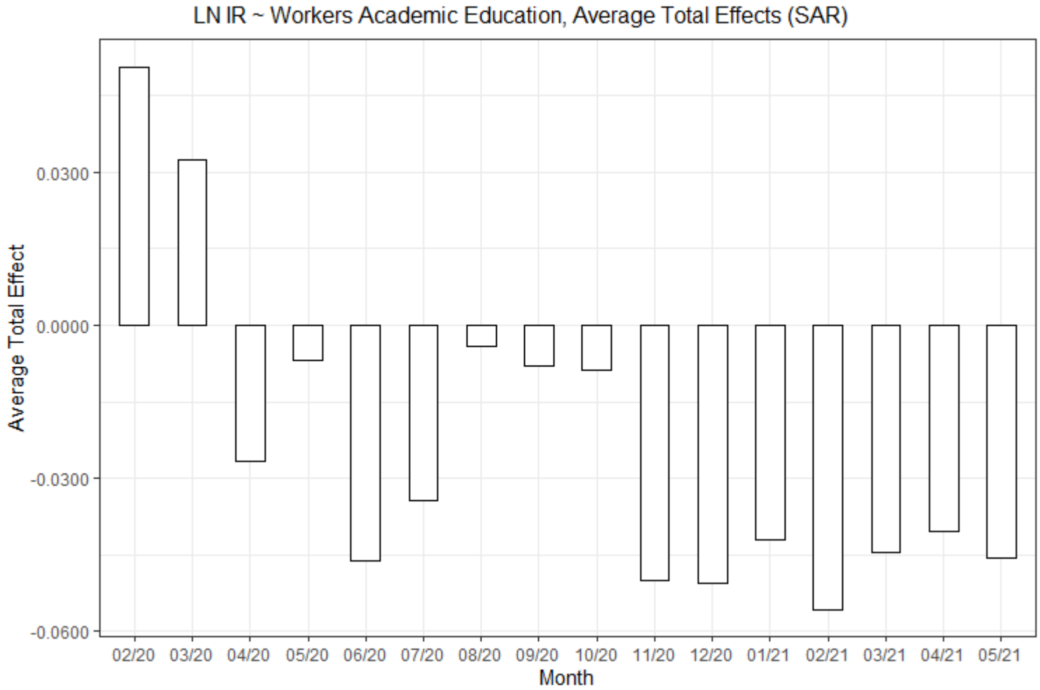


Figure 13: Hypothesis 1: IR ∽ Workers with Academic Education

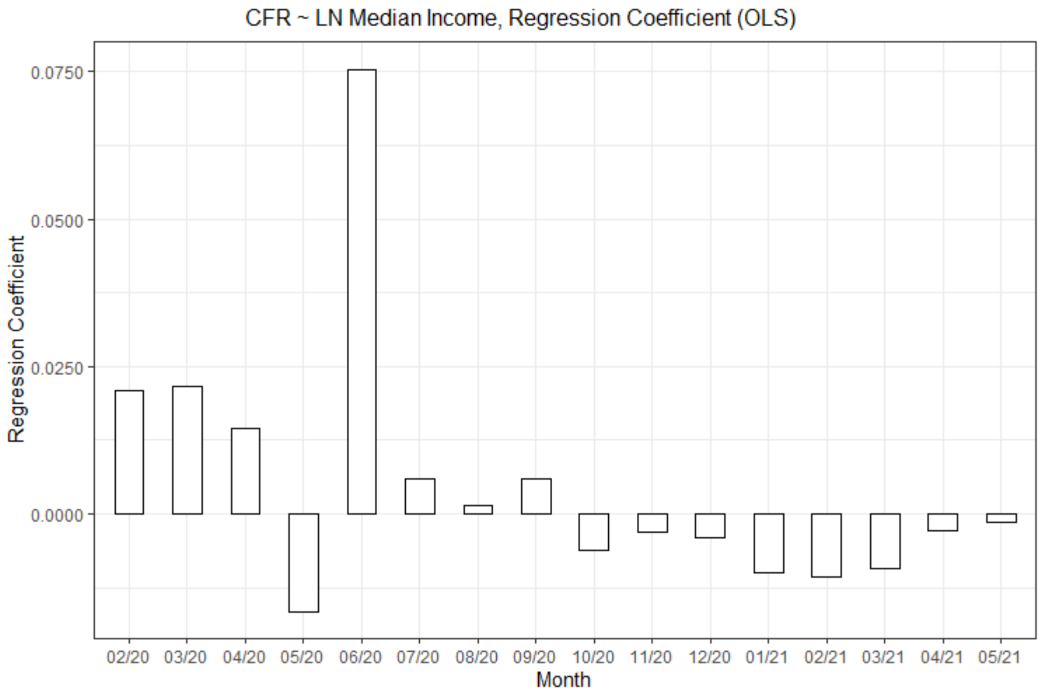


Figure 14: Hypothesis 1: CFR ∽ LN Median Income

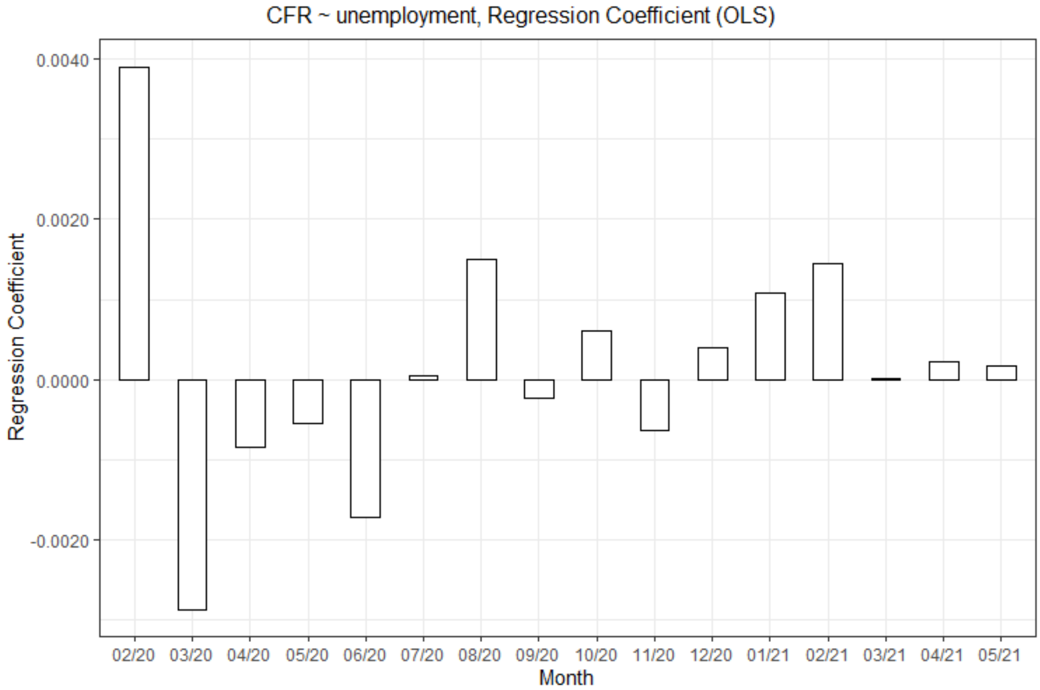


Figure 15: Hypothesis 1: CFR ∽ Unemployment

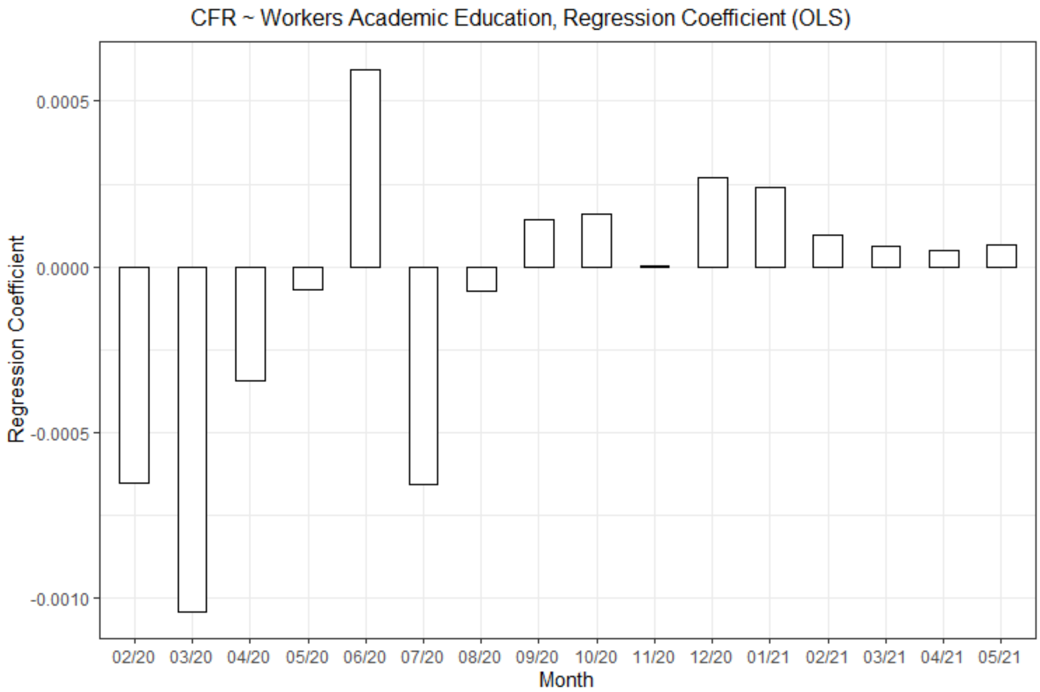


Figure 16: Hypothesis 1: CFR ∽ Workers with Academic Education

Table 7: Hypothesis 2.1: LN IR ∽ GISD*Share foreign nationals

| | Month | ATE GISD*Share for. nat. | GISD | share foreign national | rho |
|---|---|---|---|---|---|
| 1 | 02/20 | −0.090 | 0.180 | 0.076 | 0.225 |
| 2 | 03/20 | 0.124 | −4.344 | −0.133 | 0.627 |
| 3 | 04/20 | 0.271 | −5.016 | −0.153 | 0.649 |
| 4 | 05/20 | 0.362 | −5.199 | −0.153 | 0.536 |
| 5 | 06/20 | 0.177 | −1.732 | 0.005 | 0.393 |
| 6 | 07/20 | 0.309 | −3.842 | −0.069 | 0.220 |
| 7 | 08/20 | 0.115 | −1.827 | 0.027 | 0.188 |
| 8 | 09/20 | 0.088 | −1.467 | −0.0002 | 0.340 |
| 9 | 10/20 | 0.042 | 0.317 | 0.053 | 0.656 |
| 10 | 11/20 | 0.084 | −0.116 | 0.015 | 0.728 |
| 11 | 12/20 | 0.121 | −1.052 | −0.001 | 0.819 |
| 12 | 01/21 | −0.025 | 1.848 | 0.045 | 0.696 |
| 13 | 02/21 | 0.016 | 1.672 | 0.026 | 0.634 |
| 14 | 03/21 | 0.155 | −0.482 | −0.039 | 0.705 |
| 15 | 04/21 | 0.081 | 0.282 | 0.026 | 0.720 |
| 16 | 05/21 | 0.064 | 0.167 | 0.033 | 0.652 |

Table 8: Hypothesis 2.1: CFR ∽ GISD*Share foreign nationals

| | Month | ATE GISD*Share for. nat. | GISD | share foreign national |
|---|---|---|---|---|
| 1 | 02/20 | −0.078 | 0.146 | 0.070 |
| 2 | 03/20 | 0.092 | −3.224 | −0.057 |
| 3 | 04/20 | 0.204 | −4.260 | −0.100 |
| 4 | 05/20 | 0.250 | −4.000 | −0.096 |
| 5 | 06/20 | 0.141 | −1.521 | 0.006 |
| 6 | 07/20 | 0.255 | −3.236 | −0.041 |
| 7 | 08/20 | 0.101 | −1.752 | 0.031 |
| 8 | 09/20 | 0.087 | −1.763 | −0.003 |
| 9 | 10/20 | 0.054 | −0.777 | 0.020 |
| 10 | 11/20 | 0.097 | −1.616 | −0.023 |
| 11 | 12/20 | 0.094 | −1.585 | −0.039 |
| 12 | 01/21 | 0.034 | −0.039 | −0.021 |
| 13 | 02/21 | 0.042 | 0.331 | −0.021 |
| 14 | 03/21 | 0.087 | −0.675 | −0.043 |
| 15 | 04/21 | 0.069 | −0.871 | −0.020 |
| 16 | 05/21 | 0.064 | −0.931 | −0.006 |

Table 9: Hypothesis 2.2: LN IR $\backsim$ GISD*East

| | Month | GISD*East | GISD | East | rho |
|---|---|---|---|---|---|
| 1 | 02/20 | 1.716 | −1.173 | −1.132 | 0.256 |
| 2 | 03/20 | −3.541 | −2.640 | 2.985 | 0.614 |
| 3 | 04/20 | −4.119 | −1.290 | 3.012 | 0.676 |
| 4 | 05/20 | −5.146 | −0.023 | 3.198 | 0.559 |
| 5 | 06/20 | 0.790 | 0.823 | −1.460 | 0.423 |
| 6 | 07/20 | −1.785 | 0.870 | −0.244 | 0.277 |
| 7 | 08/20 | 0.318 | −0.002 | −0.963 | 0.336 |
| 8 | 09/20 | 1.047 | −0.226 | −1.065 | 0.396 |
| 9 | 10/20 | 1.890 | 1.015 | −1.570 | 0.711 |
| 10 | 11/20 | 1.843 | 1.044 | −1.474 | 0.757 |
| 11 | 12/20 | 3.336 | −0.152 | −1.783 | 0.813 |
| 12 | 01/21 | 1.662 | 0.894 | −0.658 | 0.645 |
| 13 | 02/21 | 0.849 | 1.723 | −0.601 | 0.621 |
| 14 | 03/21 | 2.114 | 1.043 | −1.193 | 0.685 |
| 15 | 04/21 | 4.203 | 0.881 | −2.965 | 0.715 |
| 16 | 05/21 | 2.379 | 1.013 | −1.977 | 0.680 |

Table 10: Hypothesis 2.2: CFR $\backsim$ GISD*East

| | Month | GISD*East | GISD | East |
|---|---|---|---|---|
| 1 | 02/20 | 1.433 | −0.996 | −0.989 |
| 2 | 03/20 | −1.020 | −1.981 | 0.740 |
| 3 | 04/20 | −1.365 | −1.498 | 0.802 |
| 4 | 05/20 | −2.520 | −0.464 | 1.355 |
| 5 | 06/20 | 0.417 | 0.481 | −0.950 |
| 6 | 07/20 | −0.953 | 0.685 | −0.774 |
| 7 | 08/20 | 0.510 | −0.170 | −1.214 |
| 8 | 09/20 | 0.586 | −0.567 | −0.793 |
| 9 | 10/20 | 0.283 | −0.066 | −0.382 |
| 10 | 11/20 | 0.296 | −0.468 | −0.136 |
| 11 | 12/20 | 0.468 | −0.753 | 0.304 |
| 12 | 01/21 | 0.197 | 0.062 | 0.522 |
| 13 | 02/21 | −0.240 | 0.750 | 0.438 |
| 14 | 03/21 | 0.673 | 0.140 | 0.005 |
| 15 | 04/21 | 1.436 | −0.291 | −0.658 |
| 16 | 05/21 | 0.826 | −0.285 | −0.440 |

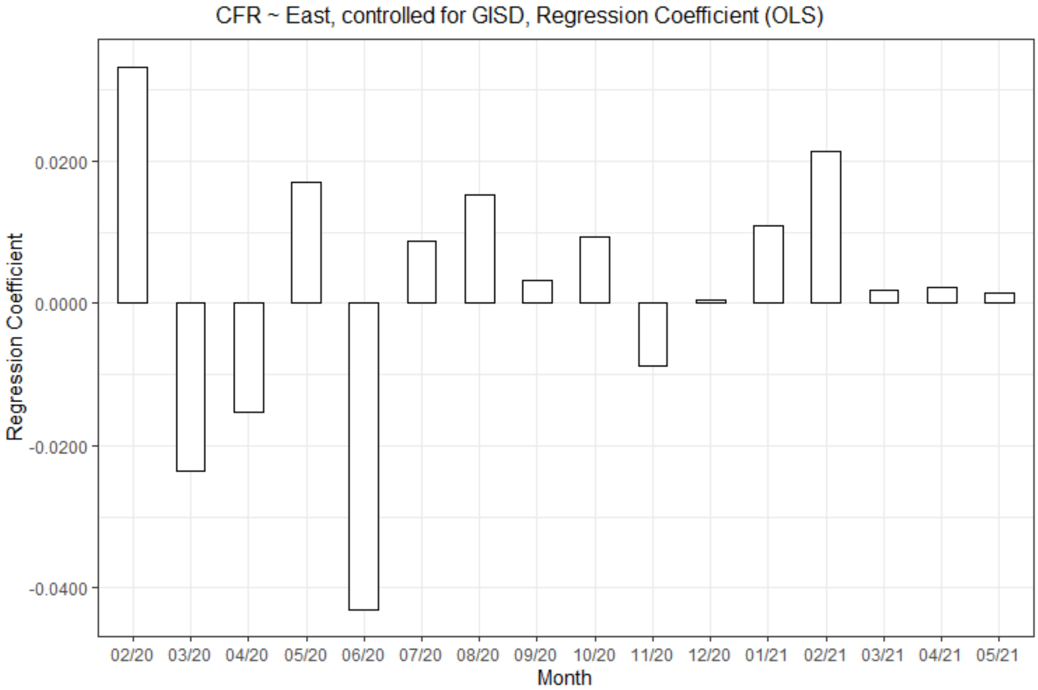


Figure 17: Hypothesis II.2: CFR $\backsim$ GISD $\times$ east dummy

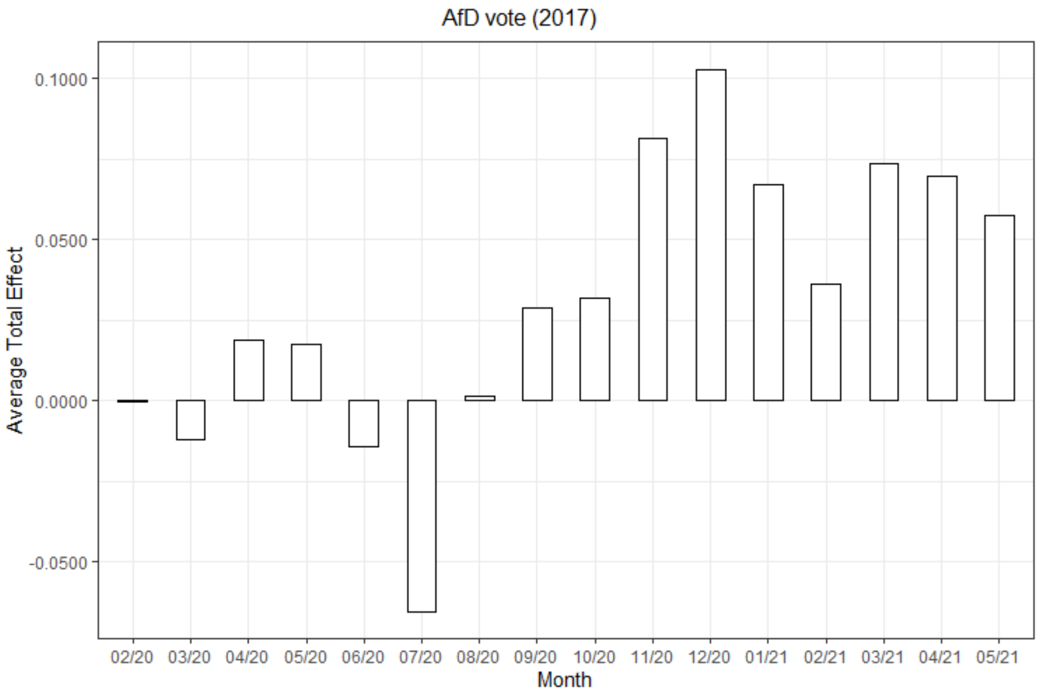


Figure 18: Hypothesis E1: IR $\backsim$ AfD

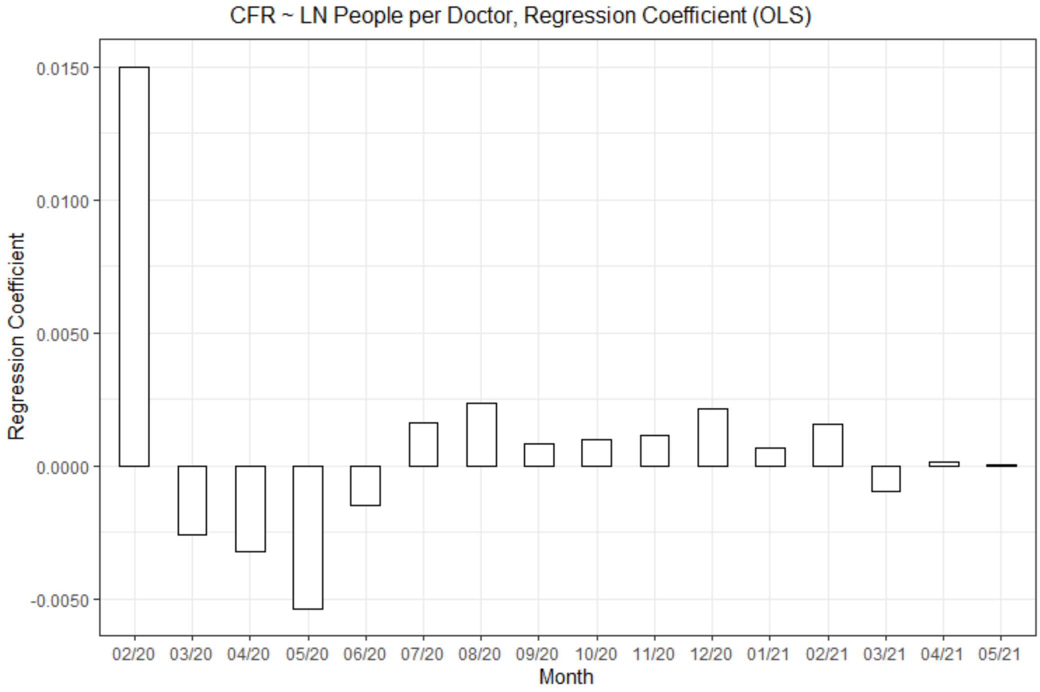


Figure 19: Hypothesis E2: CFR ∽ Population per Doctor

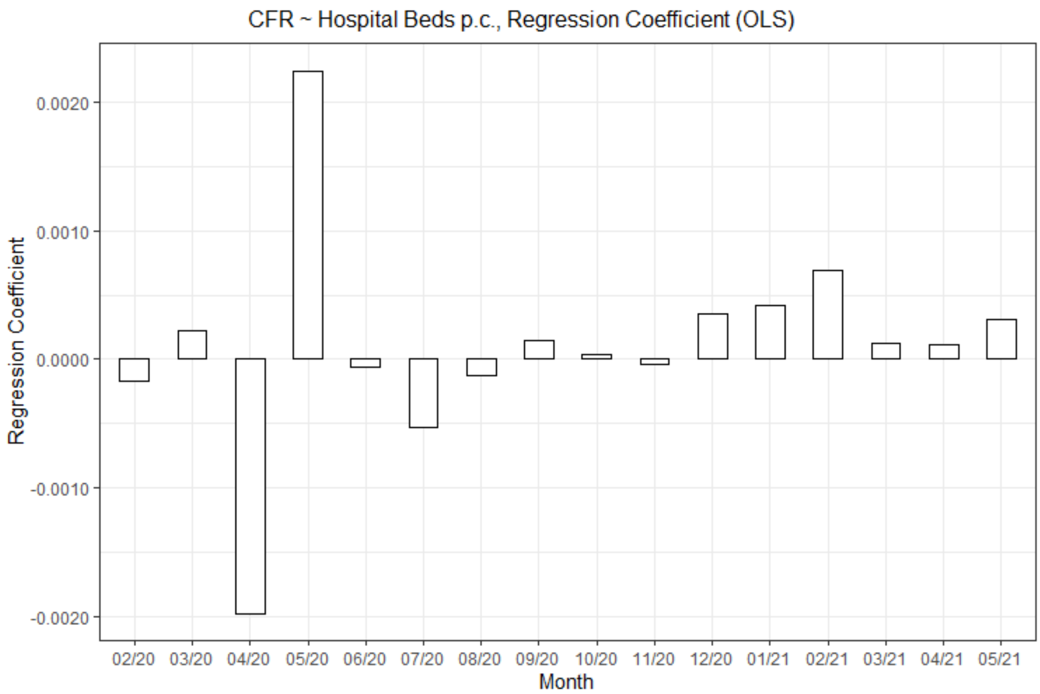


Figure 20: Hypothesis E2: CFR ∽ Hospital Beds per 1,000

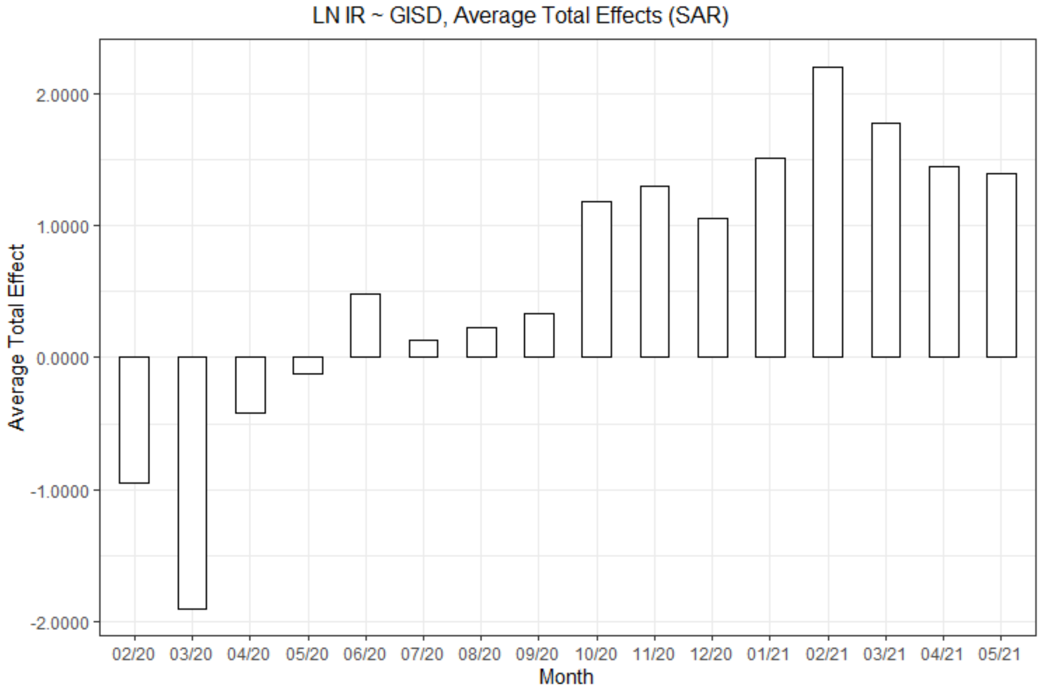


Figure 21: Robustness Check: Hypothesis 1, IR, with state dummies

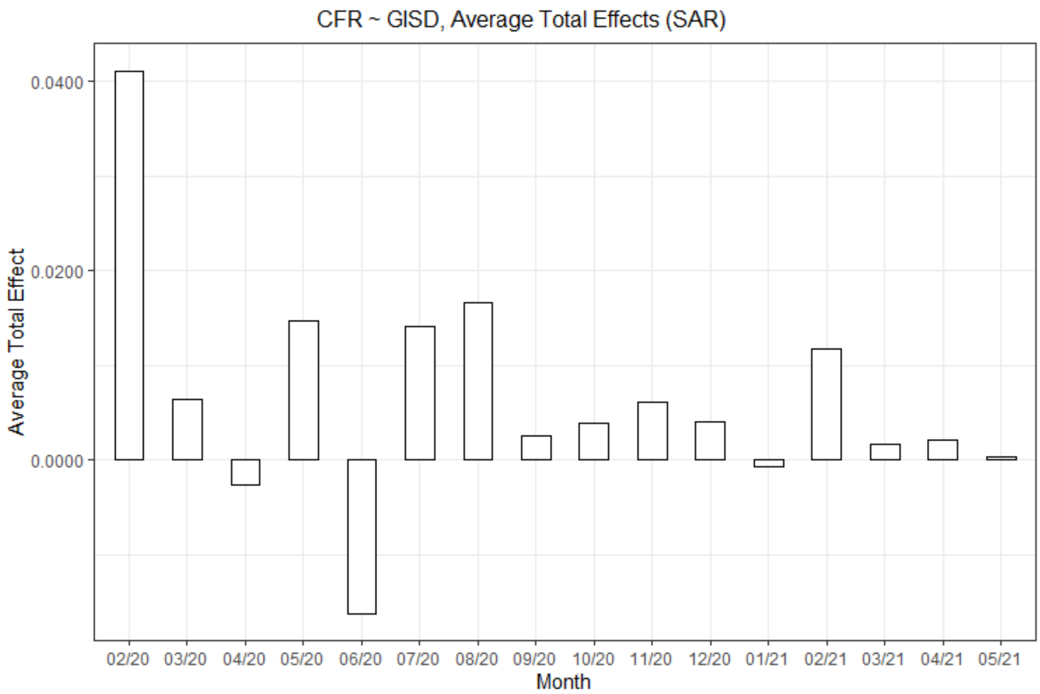


Figure 22: Robustness Check: Hypothesis 1, CFR, with state dummies

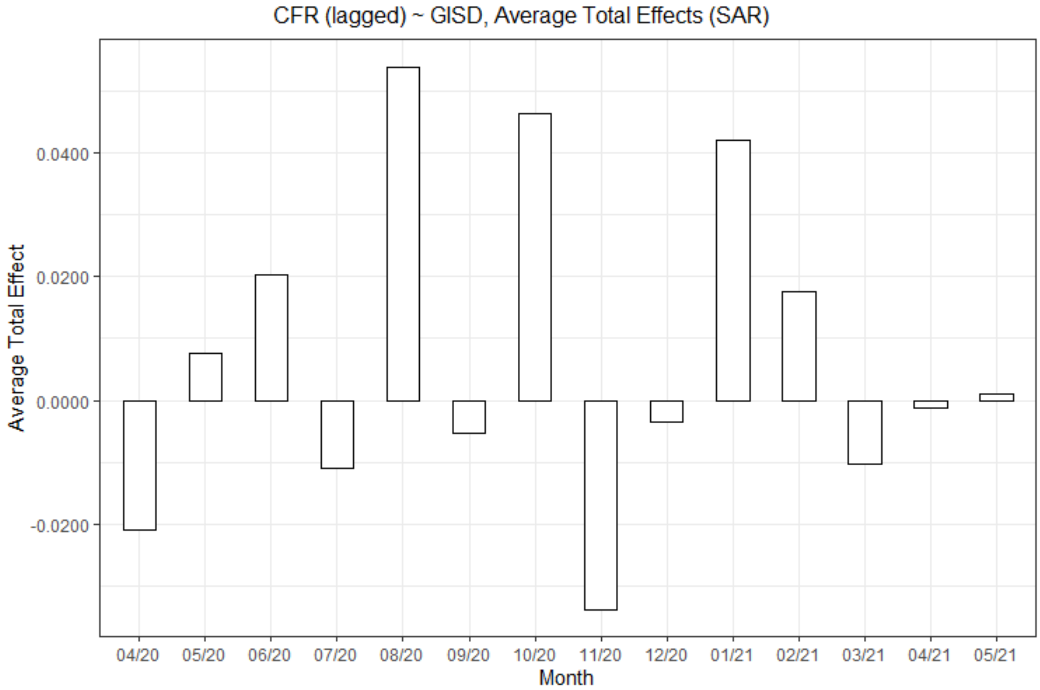


Figure 23: Robustness Check: Hypothesis 1, lagged CFR

# 8 Code Repository and Data

All data and software used for this paper are publicly available and free of charge. Only open source software was used for calculations and image creation.

All calculations, models, and images were created within RStudio (RStudio Team 2021) in the statistical programming language R (R Core Team 2021). Packages used: plyr (Wickham 2011), dplyr (Wickham et al. 2021), ggplot2 (Wickham 2016), ggpubr (Kassambara 2020), purrr (Henry and Wickham 2020), readxl (Wickham and Bryan 2019), spdep (Bivand et al. 2013), sf (Pebesma 2018), stringr (Wickham 2019), and stargazer (Hlavac 2018).

The complete code of this study is available as repository at GitHub: Click on this sentence for the GitHub-Repository containing all R scripts for calculations and image creation. Alternatively, visit: https://github.com/hieronymusBusch/BA-Sociology/

# 9 References

Ahlers, E., Mierich, S., and Zucco, A. (2021). *Homeoffice. Was wir aus der Zeit der Pandemie für die zukünftige Gestaltung von Homeoffice lernen können.* Wirtschafts- und Sozialwissenschaftliches Institut, Düsseldorf, Germany.

Alternative für Deutschland (2021). Corona-resolution. `https://www.afd.de/bundesparteitag-in-dresden-beschliesst-corona-resolution/`. [Online; last access May 17, 2021].

Bambra, C., Riordan, R., Ford, J., and Matthews, F. (2020). The covid-19 pandemic and health inequalities. *Journal of Epidemiology & Community Health*, 74(11):964–968.

Bartscher, A. K., Seitz, S., Siegloch, S., Slotwinski, M., and Wehrhöfer, N. (2020). Social Capital and the Spread of Covid-19: Insights from European Countries . *CESifo working papers*, page 8346.

BBSR (2019). Bevölkerungsdichte. `https://www.deutschlandatlas.bund.de/DE/Karten/Wo-wir-leben/006-Bevoelkerungsdichte.html`. [Online; last access May 11, 2021].

BBSR (2021). INKAR-Daten, erhoben aus der laufenden Raumbeobachtung, basierend auf dem Zensus 2011 BRD: Altered by Alexander Busch. `https://www.inkar.de/`. [Online; last access May 4, 2021].

Bellégo, C. and Pape, L.-D. (2019). Dealing with Logs and Zeros in Regression Models. *Série des Documents de Travail*, 2019-13.

Bivand, R. S., Pebesma, E., and Gomez-Rubio, V. (2013). *Applied spatial data analysis with R, Second edition.* Springer, NY.

BKG (2020). Digitales Landschaftsmodell 1:250 000 (Ebenen): Altered by Alexander Busch. `https://gdz.bkg.bund.de/index.php/default/catalog/product/view/id/773/s/nuts-gebiete-1-250-000-stand-01-01-nuts250-01-01/category/8/?___store=default`. [Online; last access May 4, 2021].

Bourdrel, T., Annesi-Maesano, I., Alahmad, B., Maesano, C. N., and Blind, M. A. (2021). The impact of outdoor air pollution on covid-19: a review of evidence from in vitro, animal, and human studies. *European Respiratory Journal*, 31(159):200242.

Copat, C., Christaldi, A., Fiore, M., Grasso, A., Zuccarello, P., Signorelli, S. S., Conti, G. O., and Ferrante, M. (2020). The role of air pollution (pm and no2) in covid-19 spread and lethality: A systematic review. *Environmental Research*, 191:110129.

DESTATIS (2018). Kreisfreie Städte und Landkreise nach Fläche, Bevölkerung und Bevölkerungsdichte am 31.12.2018: Altered by Alexander Busch. https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/Standardtabellen/04_KreiseVorjahr.html;jsessionid=8484FA1B86F77062B91D6B2B726EE9AC.live741. [Online; last access May 4, 2021].

DESTATIS (2021). Krankenhäuser, Betten, Patienten: Deutschland, Jahre. https://www-genesis.destatis.de/genesis/online?operation=abruftabelleBearbeiten&levelindex=1&levelid=1621176891342&auswahloperation=abruftabelleAuspraegungAuswaehlen&auswahlverzeichnis=ordnungsstruktur&auswahlziel=werteabruf&code=23111-0001&auswahltext=&werteabruf=starten#abreadcrumb. [Online; last access May 20, 2021].

Effern, H., Kel, E., and Kronewiter, T. (2021). Pandemie in München: Welche Stadtviertel besonders von Corona betroffen sind. https://www.sueddeutsche.de/muenchen/muenchen-coronavirus-stadtteile-brennpunkte-1.5286288. [Online; last access May 10, 2021].

Ehlert, A. (2020). The socio-economic determinants of COVID-19: A spatial analysis of German county level data . *Preprint*.

Fatima, M., O'Keefe, K. J., Wei, W., Arshad, S., and Gruebner, O. (2021). Geospatial Analysis of COVID-19: A Scoping Review. *International journal of environmental research and public health*, 18(5):2336.

Fronteira, I., Sidat, M., Magalhães, J. P., de Barros, F. P. C., Delgado, A. P., Correia, T., Daniel-Ribeiro, C. T., and Ferrinho, P. (2021). The sars-cov-2 pandemic: A syndemic perspective. *One Health*, 12:100228.

Gangemi, S., Billeci, L., and Tonacci, A. (2020). Rich at risk: socio-economic drivers of covid-19 pandemic spread. *Clinical and Molecular Allergy*, 18(12).

Gao, Y.-d., Ding, M., Dong, X., Zhang, J.-j., Kursat Azkur, A., Azkur, D., Gan, H., Sun, Y.-l., Fu, W., Li, W., Liang, H.-l., Cao, Y.-y., Yan, Q., Cao, C., Gao, H.-y., Brüggen, M.-C., van de Veen, W., Sokolowska, M., Akdis, M., and Akdis, C. A. (2021). Risk factors for severe and critically ill covid-19 patients: A review. *Allergy*, 76(2):428–455.

García, C. N. (2021). Socioeconomic, demographic and healthcare determinants of the COVID-19 pandemic: an ecological study of Spain . *BMC Public Health*, 21(1):606.

Gaudart, J., Landier, J., Heiart, L., Legendre, E., Lehot, L., Bendiane, M. K., Chiche, L., Petitjean, A., Mosnier, E., Samadoouloucou, F. K., Demongeot, J., Piarroux, R., and Rebaudet, S. (2021). Factors associated with the spatial heterogeneity of the first wave of COVID-19 in France: a nationwide geo-epidemiological study. *Lancet Public Health*, 6(4):e222–e231.

Ginsburgh, V., Magerman, G., and Natali, I. (2021). COVID-19 and the role of inequality in French regional departments . *The European Journal of Health Economics*, 22(2):311–327.

Gravlee, C. C. (2020). Systemic racism, chronic health inequities, and COVID-19: A syndemic in the making? *American journal of human biology*, 32(5).

Gülsen, A., Yigitbas, B. A., Uslu, B., Drömann, D., and Kilinc, O. (2020). The effect of smoking on covid-19 symptom severity: Systematic review and meta-analysis. *Pulmonary Medicine*, page 2020:7590207.

Henry, L. and Wickham, H. (2020). *purrr: Functional Programming Tools*. R package version 0.3.4.

Hlavac, M. (2018). *stargazer: Well-Formatted Regression and Summary Statistics Tables*. Central European Labour Studies Institute (CELSI), Bratislava, Slovakia. R package version 5.2.2.

Hoebel, J., Michalski, N., Wachtler, B., Diercke, M., Neuhauser, H., Wieler, L. H., and Hövener, C. (2021). Sozioökonomische Unterschiede im Infektionsrisiko während der zweiten SARS-CoV-2-Welle in Deutschland. *Dtsch Arztebl International*, 118(15):269–270.

Horton, R. C. (2020). Covid-19 is not a pandemic. *The Lancet*, 396(10255):874.

infas360 (2021). Der Virus und sein Veedel. `https://blog.infas360.de/2021/04/10/der-virus-und-sein-veedel/`. [Online; last access May 4, 2021].

Iyanda, Ayodeji, E., Boakye, K. A., Lu, Y., and Oppong, J. R. (2021). Spatial Disparities in Coronavirus Incidence and Mortality in the United States: An Ecological Analysis as of May 2020. *Journal of Racial and Ethnic Health Disparities*.

Jain, A., van Hoek, A. J., Boocia, D., and Thomas, S. L. (2017). Lower vaccine uptake amongst older individuals living alone: A systematic review and meta-analysis of social determinants of vaccine uptake. *Vaccine*, 35(18):2315–2328.

Kaiser, O. (2021). Pandemie legt soziale Ungerechtigkeit "schonungslos" offen. `https://www.rnz.de/nachrichten/mannheim_artikel,-mannheim-pandemie-legt-soziale-ungerechtigkeit-schonungslos-offen-_arid,662872.html`. [Online; last access May 4, 2021].

Kassambara, A. (2020). *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.4.0.

Kenyon, C. (2020). Syndemic responses to covid-19 should include an ecological dimension. *The Lancet*, 396(10264):1730–1731.

Khan, J. R., Awan, N., Islam, M. M., and Muurlink, O. (2020). Healthcare capacity, health expenditure, and civil society as predictors of covid-19 case fatalities: A global analysis. *Frontiers in Public Health*, 8:347.

Kroll, L. E., Schumann, M., Hoebel, J., and Lampert, T. (2017). Regional health differences – developing a socioeconomic deprivation index for Germany. *Journal of Health Monitoring*, 2(2):98–114.

Kroll, L. E., Schumann, M., Hoebel, J., and Lampert, T. (2018). GISD - The German Index of Socioeconomic Deprivation. Data Revision 2018. `https://github.com/lekroll/GISD/tree/master/Revisions/2018`. [Online; last access May 18, 2021].

Lampert, T., Richter, M., Schneider, S., Spallek, J., and Dragano, N. (2016). Soziale Ungleichheit und Gesundheit. *Bundesgesundheitsblatt*, 59:153–165.

LeSage, J. P. (2008). An introduction to spatial econometrics. *Revue d'économie industrielle*, 123(3):19–44.

LeSage, J. P. and Pace, R. K. (2014). Interpreting spatial econometric models. In Fischer, M. M. and Nijkamp, P., editors, *Handbook of Regional Science*, pages 1535–1552. Springer Berlin Heidelberg, Berlin, Heidelberg.

Li, J., Huang, D. Q., Zou, B., Yang, H., Hui, W. Z., Rui, F., Yee, N. T. S., Liu, C., Nerurkar, S. N., Kai, J. C. Y., Teng, M. L. P., Li, X., Zeng, H., Borghi, J. A., Henry, L., Cheung, R., and Nguyen, M. H. (2021). Epidemiology of covid-19: A systematic review and meta-analysis of clinical characteristics, risk factors, and outcomes. *Journal of Medical Virology*, 93(3):1449–1458.

Lowcock, E. C., Rosella, L. C., Foisy, J., McGeer, A., and Crowcroft, N. (2012). Effect of race/ethnicity and socioeconomic status on pandemic H1N1-related outcomes in Massachusetts. *American journal of public health*, 102(8):e51–e58.

Mackey, K., Ayers, C. K., Kondo, K. K., Saha, S., Advani, S. M., Young, S., Spencer, H., Rusek, M., Anderson, J., Veazie, S., Smith, M., and Kansagara, D. (2021). Racial and Ethnic Disparities in COVID-19-Related Infections, Hospitalizations, and Deaths : A Systematic Review . *Annals of internal medicine*, 174(4):362–373.

Mader, S., Rubach, M., Schaecke, W., Röger, C., Feldhoffer, I., and Thalmeier, E.-M. (2020). Healthy nutrition in germany: a survey analysis of social causes, obesity and socioeconomic status. *Public Health Nutrition*, 23(12):2109–2123.

Mahase, E. (2020). Covid-19: What do we know about "long covid"? *BMJ*, 370.

Malik, V. S., Ravindra, K., Attri, S. V., Bhadada, S. K., and Singh, M. (2020). Higher body mass index is an important risk factor in covid-19 patients: a systematic review and meta-analysis. *Environmental science and pollution research international*, 27(33):42115–42123.

Mamelund, S.-E. (2006). A socially neutral disease? Individual social class, household wealth and mortality from Spanish influenza in two socially contrasting parishes in Kristiania 1918–19. *Social Science & Medicine*, 62(4):923–940.

Mamelund, S.-E. (2018). 1918 pandemic morbidity: The first wave hits the poor, the second wave hits the rich. *Influenza and other respiratory viruses*, 12(3):307–313.

Mamelund, S.-E., Shelley-Egan, C., and Rogeberg, O. (2020). The association between socioeconomic status and pandemic influenza: systematic review and meta-analysis. *Preprint*.

Mayoral, J. M., Alonso, J., Garín, O., Herrador, Z., Astray, J., Baricot, M., Castilla, J., Cantón, R., Castro, A., Delgado-Rodríguez, M., Ferri, A., Godoy, P., Gónzález-Candelas, F., Martín, V., Pumarola, T., Quintana, J. M., Soldevila, N., Tamames, S., and Domínguez, A. (2013). Social factors related to the clinical severity of influenza cases in Spain during the A (H1N1) 2009 virus pandemic. *BMC public health*, 13(1):118–118.

McCracken, K. and Curson, P. (2003). Flu downunder: A demographic and geographic analysis of the 1919 pandemic in sydney, australia. In Phillips, H. and Killingray, D., editors, *The Spanish influenza Pandemic of 1918–19. New perspectives*, chapter 10, pages 110–131. Routledge Social History of Medicine series, London.

Mehta, N. S., Mytton, O. T., Mullins, E. W. S., Fowler, T. A., Falconer, C. L., Murphy, O. B., Langenberg, C., Jayatunga, W. J. P., Eddy, D. H., and Nguyen-Van-Tam, J. S. (2020). SARS-CoV-2 (COVID-19): What Do We Know About Children? A Systematic Review. *Clinical Infectious Diseases*, 71(9):2469–2479.

Mendenhall, E. (2020). The covid-19 syndemic is not global: context matters. *The Lancet*, 396(10264):1731.

Mendenhall, E. and Singer, M. (2019). What constitutes a syndemic? Methods, contexts, and framing from 2019. *Current Opinion in HIV and AIDS*, 15:213–217.

Mourad, A., Turner, N. A., Baker, A. W., Okeke, N. L., Narayanasamy, S., Rolfe, R., Engemann, J. J., Cox, G. M., and Stout, J. E. (2020). Social Disadvantage, Politics, and Severe Acute Respiratory Syndrome Coronavirus 2 Trends: A County-level Analysis of United States Data . *Clinical Infectious Diseases*, page ciaa1374.

Nagata, J. M., Hernández-Ramos, I., Kurup, A. S., Albrecht, D., Vivas-Torrealba, C., and Franco-Paredes, C. (2013). Social determinants of health and seasonal influenza vaccination in adults $\geq$ 65 years: a systematic review of qualitative and quantitative data. *BMC Public Health*, 13:388 (2013).

Niall, P. A. S. J. and Mueller, J. (2002). Updating the Accounts: Global Mortality of the 1918-1920 "Spanish" Influenza Pandemic. *Bulletin of the History of Medicine*, 76(1):105–115.

Norddeutscher Rundfunk (2021). Corona-Zahlen: Erhebliche Unterschiede in Hamburgs Stadtteilen. `https://www.ndr.de/nachrichten/hamburg/Corona-Zahlen-Erhebliche-Unterschiede-in-Hamburgs-Stadtteilen,stadtteile628.html`. [Online; last access May 4, 2021].

Organization, W. H. (2020). Estimating mortality from covid-19: scientific brief. `https://www.who.int/news-room/commentaries/detail/estimating-mortality-from-covid-19`. [Online; last access May 11, 2021].

Pana, T. A., Bhattacharya, S., Gable, D. T., Pasdar, Z., Szlachetka, W. A., Perdomo-Lampignano, J. A., Ewers, K. D., McLernon, D. J., and Myint, P. K. (2021). Country-level determinants of the severity of the first global wave of the COVID-19 pandemic: an ecological study . *BMJ Open*, 11:e042034.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1):439–446.

Placzek, H. and Madoff, L. (2014). Effect of race/ethnicity and socioeconomic status on pandemic H1N1-related outcomes in Massachusetts. *American journal of public health*, 104(1):e31–e38.

Plümper, T. and Neunmayer, E. (2020). The COVID-19 Pandemic Predominantly Hits Poor Neighborhoods, or does it? Evidence from Germany. *Preprint*.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Robert Koch Institut (2021). RKI Corona Landkreise: Altered by Alexander Busch. `https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/dd4580c810204019a7b8eb3e0b329dd6_0?page=386`. [Online; last access May 4, 2021].

RStudio Team (2021). *RStudio: Integrated Development Environment for R*. RStudio, PBC., Boston, MA.

Rychethnik, L., Hawe, P., Waters, E., Barratt, A., and Frommer, M. (2004). A glossary for evidence based public health. *Journal of Epidemiology and Community Health*.

Scarpone, C., Brinkmann, S. t., Große, T., Sonnenwald, D., Fuchs, M., and Walker, B. B. (2020). A multimethod approach for county-scale geospatial analysis of emerging infectious diseases: a cross-sectional case study of COVID-19 incidence in Germany. *International Journal of Health Geographics*, 19:32.

Schröder, C., Entringer, T., Goebel, J., Grabka, M. M., Graeber, D., Kröger, H., Kroh, M., Kühne, S., Liebig, S., Schupp, J., Seebauer, J., and Zinn, S. (2020). Vor dem Covid-19-Virus sind nicht alle Erwerbstätigen gleich . *DIW aktuell*, 41:8.

Shabazi, F. and Khazaei, S. (2020). Socio-economic inequality in global incidence and mortality rates from coronavirus disease 2019: an ecological study. *New Mircobes and New infections*, 38:100762.

Sharma, A., Ghosh, D., Divekar, N., Gore, M., Gochhait, S., and Shireshi, S. (2021). Comparing the socio-economic implications of the 1918 spanish flu and the covid-19 pandemic in india: A systematic review of literature. *International Social Science Journal*.

Singer, M. (1994). AIDS and the health crisis of the U.S. urban poor; the perspective of critical medical anthropology . *Social science and medicine*, 39(7):931–948.

Snow, J. (1849). *On the Mode of Communication of Cholera*. Wilson and Ogilvy, London, United Kingdom.

Sorci, G., Faivre, B., and Morand, S. (2020). Explaining among-country variation in COVID-19 case fatality rate. *Scientific Report*, 10:18909.

Stadt Hamburg (2021). Entwicklung der Zahl der positiv auf COVID-19 getesteten Personen nach Bezirken. `https://www.hamburg.de/corona-zahlen/`. [Online; last access May 4, 2021].

Stadt Köln (2021a). Corona-Virus - Impfungen. `https://www.stadt-koeln.de/artikel/70610/index.html`. [Online; last access May 11, 2021].

Stadt Köln (2021b). Fallzahlen nach Stadtteilen. `https://stadt-koeln.maps.arcgis.com/apps/webappviewer/index.html?id=515374f75b2f4c08bbf73e05ae1239e8`. [Online; last access May 4, 2021].

Stadt Mannheim (2021). Fallzahlen COVID-19 Mannheim Stadtbezirke. `https://www.mannheim.de/de/informationen-zu-corona/aktuelle-situation-in-mannheim/inzidenzzahl`. [Online; last access May 4, 2021].

Swinburn, B. A., Sacks, G., Hall, K. D., McPherson, K., Finegood, D. T., Moodie, M. L., and Gortmaker, S. L. (2020). The global obesity pandemic: shaped by global drivers and local environments. *Lancet*, 378(9793):804–814.

Trasi, A. C., Mendenhall, E., Trostle, J. A., and Kawachi, I. (2018). Co-occurring epidemics, syndemics, and population health. *Lancet*, 389(10072):978–982.

Tönnies, T., Pohlabeln, H., Eichler, M., Zeeb, H., and Brand, T. (2021). Relative and absolute socioeconomic inequality in smoking: time trends in Germany from 1995 to 2013. *Annals of Epidemiology*, 53:89–94.e2.

van Gerwen, M., Alsen, M., Little, C., Barlow, J., Genden, E., Naymagon, L., and Tremblay, D. (2021). Risk factors and outcomes of COVID-19 in New York City; a retrospective cohort study. *Journal of Medical Virology*, 93(2):907–915.

Wachtler, B., Michalski, N., Nowossadeck, E., Diercke, M., Wahrendorf, M., Santos-Hövener, C., Lampert, T., and Hoebel, J. (2020a). Socioeconomic inequalities and COVID-19 – A review of the current international literature . *Journal of Health Monitoring*, 5(S7).

Wachtler, B., Michalski, N., Nowossadeck, E., Diercke, M., Wahrendorf, M., Santos-Hövener, C., Lampert, T., and Hoebel, J. (2020b). Sozioökonomische Ungleichheit im Infektionsrisiko mit SARS-CoV-2 – Erste Ergebnisse einer Analyse der Meldedaten für Deutschland . *Journal of Health Monitoring*, 5(S7).

Wahrendorf, M., Rupprecht, C., Dortmann, O., Scheider, M., and Dragano, N. (2021). Erhöhtes Risiko eines COVID-19-bedingten Krankenhausaufenthaltes für Arbeitslose: Eine Analyse von Krankenkassendaten von 1,28 Mio. Versicherten in Deutschland. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 64.

Wenham, C., Smith, J., and Morgan, R. (2020). COVID-19: the gendered impacts of the outbreak. *The Lancet*, 395(10227):846–848.

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, 40(1):1–29.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham, H. (2019). *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.4.0.

Wickham, H. and Bryan, J. (2019). *readxl: Read Excel Files*. R package version 1.3.1.

Wickham, H., François, R., Henry, L., and Müller, K. (2021). *dplyr: A Grammar of Data Manipulation*. R package version 1.0.3.

World Health Organization (2010). *WHO Report on the Global Tobacco Epidemic, 2019*. World Health Organization, Geneva, Switzerland.

Zeitoun, J.-D., Faron, M., and Lefevre, J. H. (2020). Impact of the local care environment and social characteristics on aggregated hospital fatality rate from COVID-19 in France: a nationwide observational study . *Public Health*, 189:104–109.