

**How Effective Are Social Distancing Policies?
Evidence on the Fight Against COVID-19**

by

Ulrich GLOGOWSKY
Emanuel HANSEN
Simeon SCHÄCHTELE

Working Paper No. 2023
November 2020

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

Corresponding author: ulrich.glogowsky@jku.at

How Effective Are Social Distancing Policies?

Evidence on the Fight Against COVID-19 *

Ulrich Glogowsky,[†] Emanuel Hansen,[‡] Simeon Schächtele[§]

First Version: May 13, 2020

To fight the spread of COVID-19, many countries implemented social distancing policies. This is the first paper that examines the effects of the German social distancing policies on behavior and the epidemic's spread. Exploiting the staggered timing of COVID-19 outbreaks in extended event-study models, we find that the policies heavily reduced mobility and contagion. In comparison to a no-social-distancing benchmark, within three weeks, the policies avoided 84% of the potential COVID-19 cases (point estimate: 499.3K) and 66% of the potential fatalities (5.4K). The policies' relative effects were smaller for individuals above 60 and in rural areas.

JEL codes: I18, H12, I12

Keywords: COVID-19, Coronavirus, social distancing policies, policy evaluation, mobility, fatalities

Acknowledgments: We have benefited from comments by and discussions with Felix Bierbrauer, Tobias Hauck, Erik Hornung, Andreas Peichl, Johannes Rincke, Dominik Sachs, Paul Schempp, Sebastian Sieglöcher, Marc Stoeckli, and Anthony Strittmatter. David Gstrein, Marvin Immesberger, and Vincent Weber provided excellent research assistance. We also thank Ute Rexroth and Muna Abu Sin from the Robert Koch Institut for providing the data on conducted COVID-19 tests. The opinions expressed in this article are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank (IDB), its Board of Directors, or the countries they represent. **Funding:** Ulrich Glogowsky received financial support by the CRC TRR 190, Rationality and Competition. Emanuel Hansen received financial support by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC2126/1 – 390838866.

Declarations of interest: none.

[†]Department of Economics, Johannes Kepler University Linz.

[‡]Center for Macroeconomic Research, University of Cologne.

[§]Inter-American Development Bank.

1 INTRODUCTION

Since its outbreak in Wuhan, the SARS-CoV-2 virus causing the respiratory disease COVID-19 has spread across the globe (Lu et al., 2020; Wang, Horby, Hayden, & Gao, 2020; Zhu et al., 2020). To prevent human-to-human transmission of the virus, many governments have adopted social distancing (SD) policies. For example, more than 190 countries have implemented nationwide school closures.¹ These and similar policies aimed at reducing interpersonal contacts to dissipate the epidemic and, ultimately, save lives.

In this paper, we evaluate the effectiveness of the German social distancing policies in the fight against COVID-19. We offer two contributions. First, we provide a comprehensive analysis of the policies. We do not only estimate their impact on confirmed COVID-19 cases but also on fatalities. Additionally, we investigate if the policies affected certain socio-demographic groups more than others, and we use cell phone data to link the policies to changes in mobility behavior. Second, we propose a flexible quasi-experimental strategy that can be applied to many settings. At the core, it exploits variation in the spread of COVID-19 at the subnational level.

As for the policy variation, we focus on nationwide SD policies that the German federal and state governments jointly enacted in mid-March 2020. The most significant pieces of this policy response were Chancellor Merkel’s televised appeal for voluntary social distancing (March 12), the closure of schools, childcare facilities, and retail stores (March 16), and the implementation of a national contact ban (March 23).

It is challenging to identify the effects of nationwide SD policies. The impact of such policies is the difference in an outcome of interest (e.g., confirmed cases) between states of the world with and without them. After the policy interventions, however, we cannot observe how the outcome would have developed without the policies. We tackle this problem with an extended event-study approach that exploits variation at the level of German districts (NUTS-3 regions; comparable to US counties). Some districts experienced a COVID-19 outbreak several weeks before the policies took effect; others were not yet affected. Hence, we

¹See <https://en.unesco.org/covid19/educationresponse> for current numbers.

can compare how the outcomes developed after a local outbreak without SD policies (former districts) and with SD policies (latter districts). This comparison identifies the policies' effects if, in the counterfactual state without policy interventions, the outcome would have evolved similarly in both types of districts. We cannot test this assumption directly, but we verify its plausibility.²

Three features render our approach and setting especially suited to estimate the policies' effects. First, focusing on within-country variation lowers the potential for bias from heterogeneity in institutions, measurement, and populations. Particularly, all German districts faced the same policies, and identical testing and reporting rules.³ Second, the German data are sufficiently granular for quantitative impact analysis: 401 districts with varying local outbreak dates report cases and fatalities to one federal agency. Third, the data quality is arguably high, and the expected share of undetected cases is lower than in most other countries (Stafford, 2020).⁴

Figure 1 gives a graphical account of our key results. It shows how the SD policies affected the number of cases (Subfigure 1A) and fatalities (Subfigure 1B) within three weeks of implementation. According to our estimates, the policies avoided 499.3 thousand cases (95% CI: 389.4K-634.1K) and 5.4 thousand fatalities (95% CI: 3.0K-8.7K) until April 2. Put differently, the policies prevented around 84% of the confirmed cases and 66% of the fatalities that would have occurred without policy interventions by that time. The heterogeneity analysis shows that the policies' relative effects were smaller for (a) the oldest age group (60+) and (b) in rural areas.

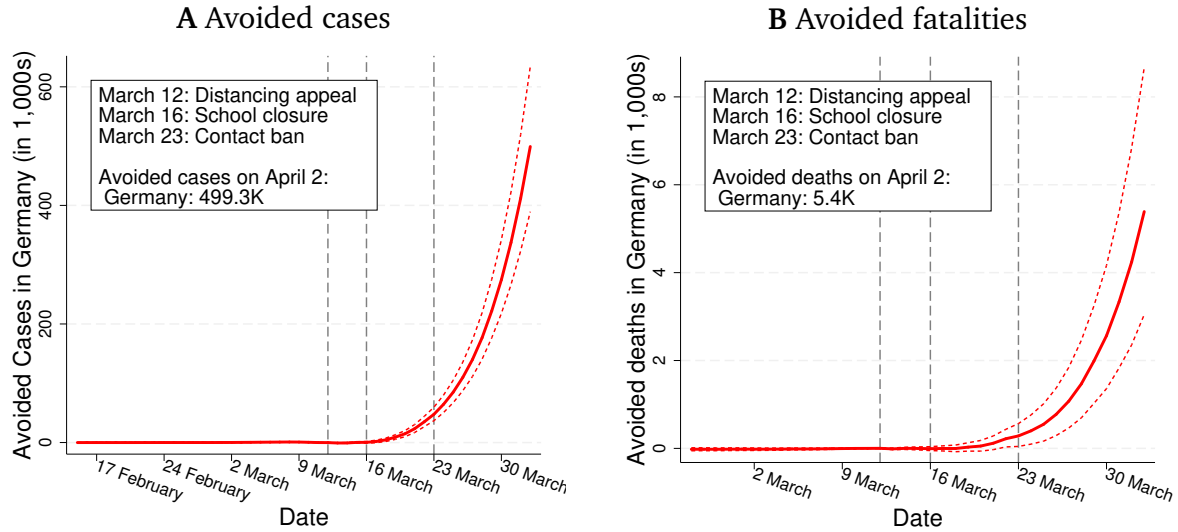
Furthermore, our analysis of cell phone data implies that, without SD policies, citizens would not have limited their social contacts to the same extent: According to our estimates, individuals reduced their mobility by about 30.7% with SD policies. Without them, they

²One study employs a similar empirical strategy to examine the role of social capital in the spread of COVID-19 (Bartscher, Seitz, Siegloch, Michaela, & Werhöfer, 2020). In contrast to us, they do not focus on the impact of policies.

³There is some variation at the local level (e.g., a state announced school closures one day earlier). Even if such variation would have discernible effects, it would only bias our estimates when correlated with the timing of the pandemic. To control for this possibility, we also estimate specifications that interact the epidemic time with the calendar time. The long-run SD effects remain almost identical.

⁴The main reasons why epidemiologists expect a low share of undetected cases in Germany are low case fatality rates and extensive testing. As of April 22, Germany conducted 2.07 million tests (2.6% of the population).

Figure 1: The effectiveness of social distancing policies in Germany



Notes: This figure summarizes the main take-away message of the paper. Subfigure 1A depicts our estimates on the number of cases avoided by social distancing policies. By contrast, Subfigure 1B focuses on avoided fatalities. The three vertical lines mark the Chancellor’s appeal for social distancing (March 12), the nationwide school closures (March 16), and the nationwide contact bans (March 23). The dashed lines represent 95% confidence intervals based on district-level-clustered standard errors.

only would have lowered it by 3.9%. This suggests that the citizens limited their social contacts as intended by the German authorities.

By providing the first comprehensive evaluation of the German SD policies, we add to an ongoing public and scientific debate on whether SD policies contained the virus. In Germany, for example, many citizens consider the lockdown measures appropriate to fight the pandemic.⁵ Other individuals, including German or US citizens, stage large-scale protests against them and question their effectiveness. The scientific debate on the policies’ effectiveness is not settled, either. Particularly, researchers have argued that the prevalent model-based evaluations of SD policies (see, e.g., Chinazzi et al., 2020; Dehning et al., 2020; Donsimoni, Glawion, Plachter, & Waelde, 2020; Ferguson et al., 2020; Ferretti et al., 2020; Greenstone & Nigam, 2020; Hellewell et al., 2020; Maier & Brockmann, 2020; Tian et al., 2020) suffer from methodological issues.⁶ A recent suggestion to evaluate SD policies

⁵For a recent German survey, see https://www.politikpanel.uni-freiburg.de/docs/Auswertung_PPD_Corona_Umfrage_Mai2020.pdf.

⁶It has been argued that epidemiological models (a) are frequently weakly identified as they fit many parameters to a single time series (Atkeson, 2020; Korolev, 2020), (b) rely on too restrictive assumptions (Kuhbandner, Homburg, Hockertz, & Walach, 2020), and (c) often have limited predictive ability (Roda, Varughese, Han, & Li, 2020). A recent paper summarizes and discusses some of these issues (Fernández-Villaverde & Jones, 2020).

is to use quasi-experimental (instead of model-based) methods (Goodman-Bacon & Marcus, 2020). In this spirit, we propose a flexible and widely-applicable quasi-experimental approach that exploits district-level variation in the spread of COVID-19. We then apply this method to provide a broad analysis of SD policies, including their effects on individual behavior, confirmed cases, and fatalities. Hereby, we contribute to a small but growing literature that exploits quasi-experimental techniques to evaluate the effectiveness of non-pharmaceutical interventions (see, e.g., Courtemanche, Garuccio, Le, Pinkston, & Yelowitz, 2020; Fang, Wang, & Yang, 2020; Juranek & Zoutman, 2020). Compared to our study, the corresponding papers have different focuses: They study fewer or other outcomes, employ different identification approaches, and investigate other policies.⁷

The paper’s structure is as follows: Section 2 describes the institutional background and our estimation approach. Section 3 contains the results for mobility (Subsection 3.1), cases and deaths (Subsection 3.2), each with a description of the relevant data. The results section also features our heterogeneity analyses (Subsection 3.3) and robustness checks (Subsection 3.4). Section 4 concludes.

2 METHODS

2.1 COVID-19 and Social Distancing in Germany

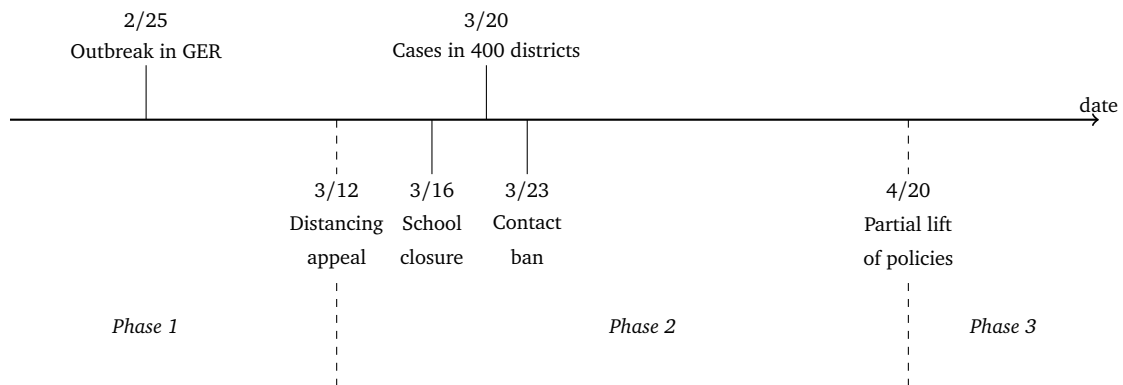
COVID-19 Outbreak: In Germany, COVID-19 spread after the detection of two cases on February 25, 2020.⁸ In the following weeks, the infection propagated to the entire country. On March 20, there were confirmed infections in all but one of the 401 German districts. Figure A7 in the Online Appendix shows the distribution of district-specific outbreak dates.⁹

⁷For example, one study investigates the effectiveness of travel restrictions (Fang et al., 2020), another one studies the effect of SD policies on hospitalizations and cases (Juranek & Zoutman, 2020), and a third paper examines the impacts of shelter-in-place orders on cases in the US (Courtemanche et al., 2020). Methodologically, these papers exploit policy variation across countries or regions in difference-in-difference models.

⁸An earlier outbreak detected on January 27 had been completely contained.

⁹We define the local outbreak date as the first day when ten cases had occurred within two weeks. Figures A8 and A9 consider other outbreak definitions.

Figure 2: Timing of social distancing policies in Germany



Notes: This figure shows the timing of social distancing policies in Germany. On top, it also indicates two key dates of the COVID-19 spread.

Social Distancing Policies: We classify the policy response into three phases (see Figure 2). In the first phase, starting with the detection of the first COVID-19 cases, the German authorities only took limited actions: They put infected persons under quarantine, recommended intensified hygiene practices to the public, and canceled large events with more than thousand participants around March 9.

In the second phase, the German federal and state governments agreed on simultaneous, nationwide SD policies to limit social contacts of the entire population. The phase began on March 12, when Chancellor Merkel appealed to all citizens to avoid social contacts whenever possible. Between March 13-15, the state governments announced the closure of schools, childcare facilities, and most retail stores starting on March 16. On March 22, they declared a strict contact ban: From March 23, meeting more than one person from outside the household was prohibited and keeping a minimum distance of 1.5 meters was required. Our analysis identifies the combined effect of these nationwide policies. Data on internet search behavior suggests that citizens did not anticipate these interventions (see Figures A15-A18).

The third phase started on April 20, when the authorities gradually relaxed the policies.

COVID-19 Testing: Official guidelines determine who qualifies for a COVID-19 test. During the study period, patients with flu-like symptoms were tested if they had been in contact with a person diagnosed with COVID-19 or in a high-risk area. This general rule applied in all federal states and remained almost unchanged.¹⁰

2.2 Estimation Method

Because it is impossible to observe the scenarios with and without SD policies simultaneously, one cannot estimate the policies' effects by directly comparing outcomes between both states. Instead, we need to find a way to approximate the latter, counterfactual scenario. To that end, we present an extended event-study model (Borusyak & Jaravel, 2017; Kleven, Landais, & Sogaard, 2019; Schmidheiny & Siegloch, 2019). Figure A12 presents the model's idea graphically and also highlights its fit to the data.

We denote by Y_{it} an outcome of interest in district i at date t . For each district i , $epitime_{it}$ denotes the “epidemic time” at date t (i.e., the number of days since the local outbreak of COVID-19). The event-study regression model reads:

$$Y_{it} = \sum_{k \in \{\dots, -1, 0, 1, \dots\}} \alpha_k \cdot 1[k = epitime_{it}] + \sum_{j \neq \text{March 11}} \beta_j \cdot 1[j = t] + \varepsilon_{it}. \quad (1)$$

The first sum contains a separate dummy variable for each day before, at, and after the district-specific outbreak. The second sum contains a set of dummies for each day before and after the nationwide SD policies started. We omit the dummy for March 11, the day before Merkel's appeal. The α coefficients account for how the outcome would have developed without the SD policies, while the β coefficients capture the policies' effects (henceforth, SD effects). To see why β captures the SD effects, note that the outcome with SD policies in district i at t is Y_{it} . The predicted outcome without SD is $\hat{Y}_{it} = \hat{\alpha}_k$, where $k = epitime_{it}$ indicates the epidemic time in district i . Consequently, an estimate of the SD effect at $t \geq \text{March 12}$ is $\hat{\beta}_t = E_i[Y_{it}] - E_i[\hat{Y}_{it}]$. We can estimate $\hat{\alpha}$ and $\hat{\beta}$ separately because, conditional on the calendar date, there is variation in the districts' epidemic times.

¹⁰After the virus had spread all over Europe, the authorities dropped the high-risk criterion on March 24.

Two aspects of our approach are essential to note. First, the identifying assumption (known as “parallel trends”) is that the expected outcome without SD policies would have been $E_i[\hat{Y}_{it}]$ for $t \geq \text{March 12}$. Intuitively, this assumption implies that, without SD policies, the outcomes in districts with outbreaks after the intervention would have developed similarly as in districts with outbreaks before the intervention. We can conduct plausibility checks. Before the implementation of the policies, the estimated SD effects $\hat{\beta}_t$ should not differ from zero. Furthermore, prior to the intervention, the outcomes in districts with earlier and later outbreaks should be similar conditional on epidemic time. Both checks suggest that the assumptions hold (see Figures 3 and Table A1).¹¹

Second, the further away date t moves from the implementation date, March 12, the fewer no-SD observations we can use to estimate the counterfactual $E_i[\hat{Y}_{it}]$ and, in consequence, the treatment effect β_t .¹² Until April 2, there is, on average, at least one no-SD control observation per district i . Therefore, our analysis focuses on the policy effects up to this date. For interested readers, we nevertheless provide estimates from an extended analysis until the policies’ relaxation on April 19 in the Online Appendix.

3 RESULTS

3.1 Mobility

Did citizens limit their human-to-human interactions due to the SD policies? One approach to studying this question is to examine if the policies changed mobility (Tian et al., 2020). The rationale is simple: Individuals who travel less or stay at home cut back the personal contacts outside their household.

¹¹To explicitly allow that the outbreak and fatality patterns change over time, we estimate further specifications that interact the epidemic time with the calendar time. The long-run SD effects remain almost identical. This finding is in line with Table A1, which also suggests constant patterns over time.

¹²To see this, note first that, at the date of the policy intervention, the districts are at epidemic times between a lower threshold \underline{k} and an upper threshold \bar{k} . Second, Δ days after the intervention, the districts have progressed to epidemic times between $\underline{k} + \Delta$ and $\bar{k} + \Delta$. Loosely speaking, we estimate the treatment effect by comparing the districts’ outcomes with SD to no-SD outcomes from before the intervention, but with the same epidemic time above $\underline{k} + \Delta$ (and below \bar{k}). Hence, the larger Δ , the smaller is the number of no-SD observations that we can use.

Measuring Mobility: When cell phone users move, their phones switch cell towers to ensure connectivity. From these switches, providers can determine the number of trips starting or ending in a given geographic zone (Oliver et al., 2020). We obtained data on the number of trips at the district level from Teralytics, a business partner of Telefónica. Specifically, for each district i , N_{it} denotes the number of trips on date t in March or April 2020, and $\bar{N}_i(d)$ is the average number of trips on weekday $d \in \{\text{Monday}, \dots, \text{Sunday}\}$ in March 2019. Our mobility measure on date t in district i is:

$$\Delta N_{it} = \left(\frac{N_{it}}{\bar{N}_i(d_t)} - 1 \right) \cdot 100, \quad (2)$$

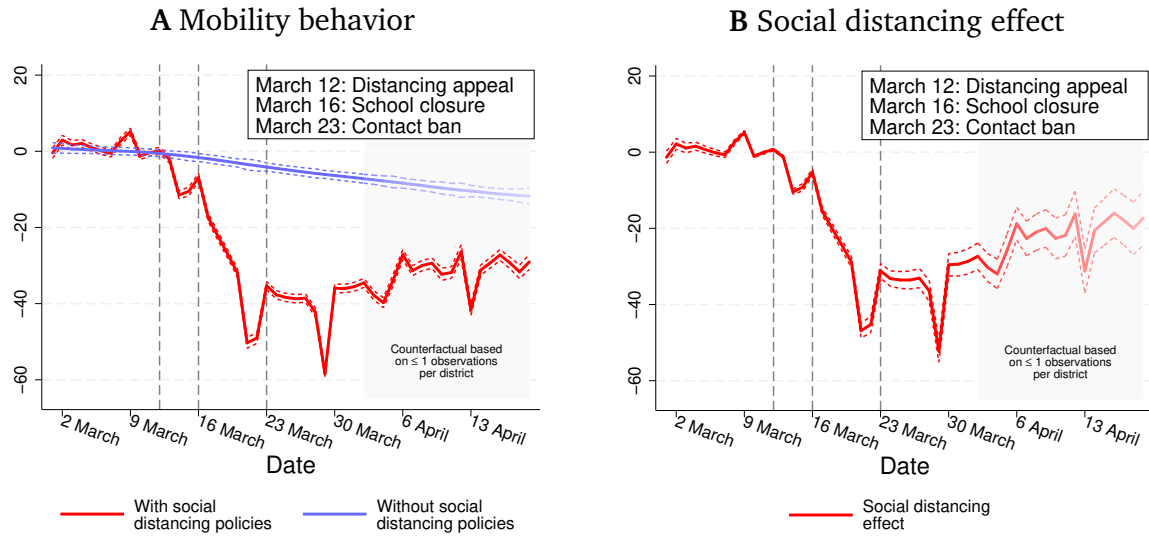
where d_t is the weekday corresponding to date t . Hence, the measure adjusts for weekday-specific mobility patterns. Moreover, it has a simple interpretation: ΔN_{it} measures mobility relative to the number of movements in 2019.

Results: To study the effects of the SD policies on mobility, we use measure (2) as the outcome of model (1). Figure 3 presents our results graphically. Subfigure 3A plots our mobility measure with SD policies (red line) and the estimated behavior without policies (blue line). For each date t , the SD effect corresponds to the vertical difference between these two lines. Subfigure 3B shows this effect measured in percentage points. In both subfigures (and all following figures), the dashed lines represent 95% confidence intervals based on district-level clustered standard errors.

Three observations stand out. First, before the start of the second phase, individuals hardly changed their behavior relative to the baseline year 2019 (see red line in Subfigure 3A). Second, from March 12 on, citizens became less and less mobile.¹³ Shortly after Merkel’s appeal, they traveled slightly less. Mobility decreased more sharply and persistently, however, after the school and business closures on March 16. From March 16 to April 2, individuals traveled, on average, 30% less than in 2019. This reduction is six times larger than the estimated change without SD policies (-3.9%). Third, the effects of SD

¹³Note that we do not find changes in individuals’ mobility before March 12, suggesting that the cancellation of large events around March 9 did not affect this type of behavior.

Figure 3: The social distancing effect on mobility



Notes: This figure shows how social distancing policies affect mobility behavior. Subfigure 3A demonstrates how individuals' behave in a world with (red line) and without (blue line) social distancing policies. The spike on March 29 reflects very bad weather conditions at this particular day. Subfigure 3B shows the social distancing effect measured in percentage points. The three vertical lines mark the Chancellor's appeal for social distancing (March 12), the nationwide school closures (March 16), and the nationwide contact bans (March 23). The dashed lines represent 95% confidence intervals based on district-level-clustered standard errors.

policies on mobility are large over the entire second phase, although they decrease over time (see Subfigure 3B). In sum, Figure 3 suggests that the SD policies reduced mobility considerably and, presumably, also social contacts.

Further Evidence: In the Online Appendix, we provide additional descriptive evidence that Germans became less mobile after the implementation of SD policies. For example, they reduced their trips to workplaces by more than 30% and used public transportation by about 50% less (see Figure A13).

3.2 Cases and Fatalities

Next, we explore if SD policies effectively constrained the spread of COVID-19. Again, we focus on the period until April 2. Figures A1 and A2 provide our estimates for the extended period until April 19.

Measuring COVID-19 Cases and Fatalities: The district-level health offices are legally obliged to report confirmed COVID-19 cases and fatalities to the federal Robert Koch Institute, which collates and publishes these data daily ([Robert Koch-Institut, 2020b](#)). We use the data set provided on April 30.¹⁴ Figures A5 and A6 show descriptive statistics. The data quality is comparatively high. First, the share of undetected cases is expected to be lower than in many other countries ([Stafford, 2020](#)). Second, all COVID-19 cases and fatalities are laboratory-confirmed. Third, all health offices apply the same testing and reporting criteria. Fourth, the data contain information on the day of the first symptoms for most cases and fatalities.¹⁵ Additionally, we gathered state-level data on the numbers of conducted COVID-19 tests for robustness checks.

Results for Confirmed COVID-19 Cases: To study the SD effects on COVID-19 cases, we use the inverse hyperbolic sine (IHS) of the cumulative number of confirmed cases in each district as the outcome in model (1).¹⁶ To simplify the interpretation, Figure 4 presents our estimation results re-transformed to cases (rather than IHS-values). Subfigure 4A displays how the confirmed cases per district evolve with (red line) and without (blue line) SD policies, using a log scale.¹⁷

Importantly, our analysis is based on the day of the first symptoms. Therefore, if the blue line lies above the red one, our estimates imply that the number of individuals suffering from first symptoms at date t would have been higher without SD policies. Subfigure 4B depicts the corresponding effects of the SD policies on a linear scale. To make the timing of the policy effects more easily visible, the figure zooms in on the period until April 2. As the mean incubation period is 5-6 days ([WHO-China Joint Mission, 2020](#)), we do not expect to find significant policy effects before March 16.

The key insights from the Subfigures 4A and 4B are as follows: First, before the closure

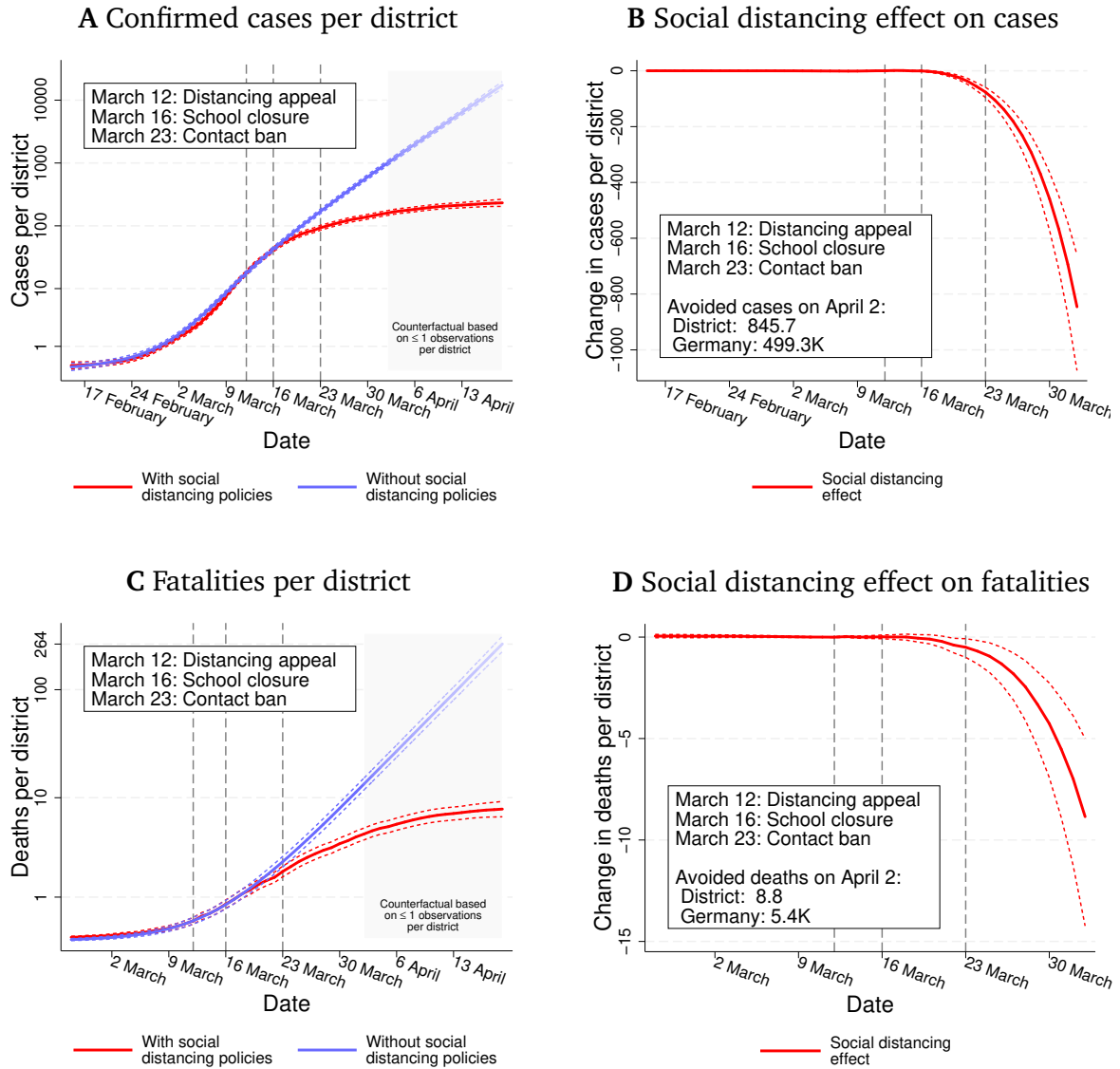
¹⁴Data download: <https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets>.

¹⁵For asymptomatic cases, the day of the first symptoms is set equal to the registration date.

¹⁶The IHS is a commonly used transformation that approximates the natural logarithm for large values. In contrast to the log, it is defined at zero ([Pence, 2006](#)). Hence, it is well-suited to model epidemic growth processes.

¹⁷As the IHS converges to the natural log, both lines depict, approximately, how the geometric means of confirmed cases evolve.

Figure 4: The social distancing effect on cases and fatalities



Notes: This figure shows how social distancing policies affect confirmed COVID-19 cases and fatalities. Subfigure 4A demonstrates the development of cases in a world with (red line) and without (blue line) social distancing policies. It shows IHS means per district. Subfigure 4C presents a similar figure for fatalities. Furthermore, Subfigure 4B (Subfigure 4D) presents the social distancing effect on cases (on fatalities). The three vertical lines mark the Chancellor’s appeal for social distancing (March 12), the nationwide school closures (March 16), and the nationwide contact bans (March 23). The dashed lines represent 95% confidence intervals based on district-level-clustered standard errors.

of schools, the case numbers with and without SD policies match closely (see Subfigure 4A). This finding suggests that our identifying assumption holds. Second, the growth rate of actual cases (red line) starts to diminish a few days after the start of the nationwide policy response, while counterfactual cases (blue line) continue to grow at a similar rate as before. Specifically, the first significant (yet small) SD effects appear on March 18, six days

after Merkel's appeal (see Subfigure 4B). Given the mean incubation period, the timing of the effects is hence in line with the policies' implementation dates. Third, on April 2, our point estimate indicates that the SD policies reduced COVID-19 cases by about 84% or 846 cases per district. Converted to the national level, this estimate indicates that the SD policies prevented 499.3 thousand cases (95% CI: 389.4K-634.1K). The extended analysis suggests that the effects would have continued to grow strongly over time. Fourth, we can also interpret our results in terms of the reproduction number R .¹⁸ After the policies' introduction, R quickly decreased from 2 to below 1. Our estimates suggest that, without SD policies, it would have stayed around 2 until April 2. In summary, the analysis implies that the SD policies effectively contained COVID-19.

Results for Fatalities: Next, we estimate model (1) with the IHS of the cumulative number of fatalities as our outcome. Recall that each fatality is reported together with the day of the patient's disease onset. Hence, we study the effects of the SD policies on the number of (eventually) lethal infections that started on date t . The Subfigures 4C and 4D of Figure 4 present our estimation results for fatalities.¹⁹

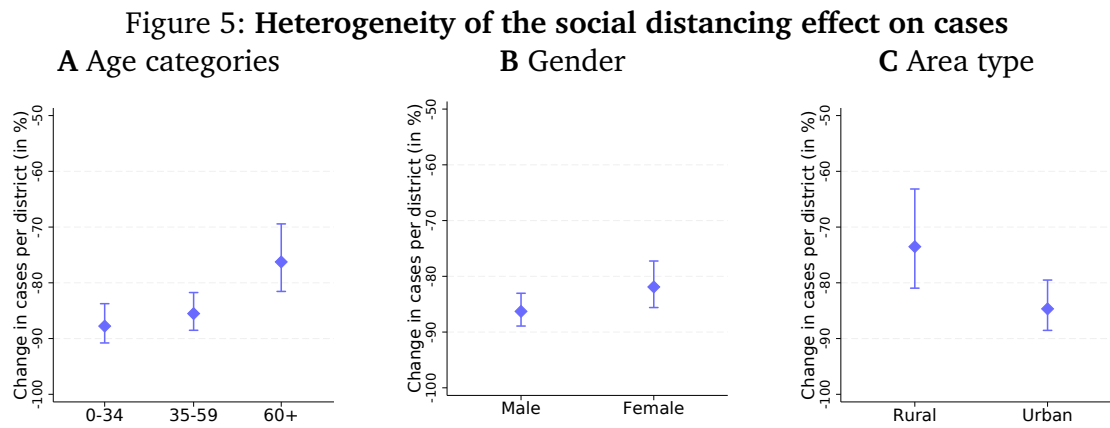
The results are in line with those for confirmed cases: The fatalities in the SD and no-SD scenarios initially follow the same growth path (see blue and red lines in Subfigure 4C). A few days after the implementation of the policies, however, the scenarios diverge: While actual growth in fatalities slows down sharply (red line), the number of counterfactual fatalities continues to grow at a similar rate (blue line). Specifically, the SD effects are significant from March 21 on and increase strongly over time. We estimate that during the period March 11 to April 2, the SD policies decreased fatalities by 66% or 8.8 per district. Transformed to the national level, this estimate suggests that the policies reduced lethal cases with first symptoms until April 2 by 5.4 thousand (95% CI: 3.0K-8.7K). Again, the extended analysis hints at steadily growing effects over time.

¹⁸We use the methodology of the [Robert Koch-Institut \(2020a\)](#) to calculate R .

¹⁹The abscissa starts with February 26, the first day with ten or more eventually lethal infections.

3.3 Heterogeneity Analyses

Evidence is growing that some groups are at higher risk to suffer from severe COVID-19 progressions (Cheng & Shan, 2020; WHO-China Joint Mission, 2020). International data show that hospitalization rates increase above 60 years of age. Furthermore, men seem to be at higher risk than women. Therefore, in the next step, we investigate how the SD effects differ (a) across age groups, (b) by gender, and (c) between urban and rural districts.



Notes: This figure shows how the effect of the nationwide social distancing policies on April 2 differed (a) across age groups, (b) by gender, and (c) between urban and rural districts. The estimates rely on sample splits and show the percentage of confirmed COVID-19 cases avoided by the nationwide policies. The vertical lines represent 95% confidence intervals based on district-level-clustered standard errors.

Figure 5 shows the subgroup-specific SD effects on confirmed cases. The estimates rely on sample splits and indicate the percentages of cases avoided due to the policies until April 2. Three results emerge. First and foremost, we find large effects in all groups. The point estimates range from 76% to 88%. Second, Figure 5 presents evidence for age-group heterogeneity. The policies prevented 88% (182.5K) of the cases among individuals below 35 that would have otherwise occurred, 86% (257.5k) of the cases in the medium-age group, and 76% (76.9k) of the cases among persons of age 60 and above. The difference between the relative effects in the last and the two former groups is significant at the 5% level. The finding is in line with the observation that, after the policy intervention, the share of persons above 60 among all infected persons increased from about 20% (March 11) to 27% (April 2). This age heterogeneity seems plausible: Policies such as school and business closures likely have stronger implications for the working-age population and for

children and their parents than for retirees. Third, we also find somewhat larger relative effects for urban districts than for rural districts and for men than for women. While the former difference is significant at the 10% level, the gender difference is not.

Figure A3 shows the same pattern of subgroup-specific SD effects on fatalities. However, due to lower numbers, we cannot study some groups and the estimation uncertainty is higher. Figure A4 extends the analysis until April 19. Again, the patterns are similar.

3.4 Robustness Analyses

We probe the robustness of our results in various dimensions. First, we explore different definitions of a local outbreak. Second, we run additional regressions in which we control for the number of conducted COVID-19 tests per day. Third, we repeat our analysis using alternative outcome definitions. For example, we drop districts with zero cases or fatalities and estimate models in logarithms. We also apply the $\ln(1 + x)$ transformation to our outcomes. Fourth, we cluster the standard errors on the state level. Figures A10 and A11 report the corresponding results. All conclusions remain substantially unchanged.

4 DISCUSSION

This paper provides evidence on the effects of the German social distancing (SD) policies on (a) individual behavior, (b) confirmed COVID-19 cases, and (c) fatalities. We show that, first, the SD policies affected individuals' mobility. Second, we find that the policies sharply slowed down the spread of the epidemic: According to our estimates, they precluded about 84% (499.3K) of the COVID-19 cases and about 66% (5.4K) of the related fatalities that would have occurred without SD policies within three weeks (until April 2). While large effects emerged across the entire population, the relative effects were smaller for the oldest age group.

From a broader perspective, we have made a step towards quantifying the effects of SD policies. At the same time, we believe that we still need a more comprehensive evaluation of such policies. First, the evidence on confirmed cases may not capture the entire impact

of the policies on the epidemic spread. One reason is that, although we use high-quality data, not all infections are detected. If the data improve over time, our analysis can be repeated. Second, we estimate the number of confirmed COVID-19 cases and fatalities avoided within three weeks after the policies' introduction. In the medium or long run, the picture might change in many ways. On the one hand, some of the avoided infections may have only been shifted to a later time. On the other hand, medical capacities may have been exceeded without SD policies, resulting in even higher numbers of fatalities. Third, our analysis identifies the joint effects of all elements of the policy response. Ideally, future studies find ways to disentangle the effects of appeals for voluntary SD, school closures, and contact bans. To shed light on these and other issues, researchers could adapt the event-study approach to new data and settings, including the policies' removals.

References

- Atkeson, A. (2020). How Deadly is COVID-19? Understanding the Difficulties with Estimation of its Fatality Rate. *NBER Working Paper Series*, 26965.
- Bartscher, A. K., Seitz, S., Siegloch, S., Michaela, S., & Werhöfer, N. (2020). *Social capital and the spread of COVID-19: Insights from European countries*. IZA Discussion Paper 13310. Cold Spring Harbor Laboratory Press. Retrieved from <https://covid-19.iza.org/publications/dp13310/> (<https://www.iza.org/publications/dp/13310>)
- Borusyak, K., & Jaravel, X. (2017). *Revisiting event study designs*. SSRN Working Paper. (DOI: 10.2139/ssrn.2826228) doi: 10.2139/ssrn.2826228
- Cheng, Z., & Shan, S. (2020). 2019 Novel coronavirus: where we are and what we know. *Infection*, 48, 155-163.
- Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., ... Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, 368(6489), 395–400. Retrieved from <https://science.sciencemag.org/content/368/6489/395> doi: 10.1126/science.aba9757
- Courtemanche, B. C., Garuccio, J., Le, A., Pinkston, J., & Yelowitz, A. (2020). Strong social distancing measures in the United States reduced the COVID-19 growth rate. *Health Affairs*, 39(7), 1–8. doi: 10.1377/hlthaff.2020.00608
- Dehning, J., Zierenberg, J., Spitzner, F. P., Wibral, M., Neto, J. P., Wilczek, M., & Priesemann, V. (2020). Inferring change points in the spread of covid-19 reveals the effectiveness of interventions. *Science*. Retrieved from <https://science.sciencemag.org/content/early/2020/05/14/science.abb9789> (eabb9789)
- Donsimoni, J. R., Glawion, R., Plachter, B., & Waelde, K. (2020). *Projecting the spread of COVID19 for Germany*. (medRxiv 2020.03.26.20044214, DOI: 10.1101/2020.03.26.20044214) doi: 10.1101/2020.03.26.20044214
- Fang, H., Wang, L., & Yang, Y. (2020). *Human mobility restrictions and the*

- spread of the novel coronavirus (2019-nCoV) in China*. Cold Spring Harbor Laboratory Press. Retrieved from <https://www.medrxiv.org/content/early/2020/03/26/2020.03.24.20042424> (medRxiv 2020.03.24.20042424. DOI: 10.1101/2020.03.24.20042424) doi: 10.1101/2020.03.24.20042424
- Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., ... Ghani, A. C. (2020). *Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand*. Imperial College COVID-19 Report 9 (Vol. 9) (No. March). Retrieved from <https://doi.org/10.25561/77482> (DOI: 10.25561/77482) doi: 10.25561/77482
- Fernández-Villaverde, J., & Jones, C. I. (2020). Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities. *NBER Working Paper Series*, 27128.
- Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., ... Fraser, C. (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, 368(6491). Retrieved from <https://science.sciencemag.org/content/368/6491/eabb6936> doi: 10.1126/science.abb6936
- Goodman-Bacon, A., & Marcus, J. (2020). Using difference-in-differences to identify causal effects of covid-19 policies. *DIW Berlin Discussion Paper*, 1870.
- Greenstone, M., & Nigam, V. (2020). *Does social distancing matter?* University of Chicago, *BFI Working Paper*. (DOI: 10.2139/ssrn.3561244) doi: 10.2139/ssrn.3561244
- Hellewell, J., Abbott, S., Gimma, A., Bosse, N. I., Jarvis, C. I., Russell, T. W., ... Eggo, R. M. (2020). Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health*, 8(4), e488-e496. doi: 10.1016/S2214-109X(20)30074-7
- Juranek, S., & Zoutman, F. (2020). *The effect of social distancing measures on the demand for intensive care: Evidence on COVID-19 in Scandinavia*. CESifo Working Paper 8262. (URL: https://www.ifo.de/DocDL/cesifo1_wp8262.pdf)
- Kleven, H., Landais, C., & Søgaaard, J. E. (2019, October). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4),

- 181-209. Retrieved from <https://www.aeaweb.org/articles?id=10.1257/app.20180010> doi: 10.1257/app.20180010
- Korolev, I. (2020). *Identification and estimation of the SEIRD epidemic model for COVID-19*, SSRN Working Paper. (DOI: 10.2139/ssrn.3569367)
- Kuhbandner, C., Homburg, S., Hockertz, S., & Walach, H. (2020). Comment on Dehning et al. (Science, 15 May 2020, eabb9789). *SageSubmissions*. doi: <https://doi.org/10.31124/advance.12362645.v1>
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., ... Tan, W. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, 395(10224), 565–574. Retrieved from [http://dx.doi.org/10.1016/S0140-6736\(20\)30251-8](http://dx.doi.org/10.1016/S0140-6736(20)30251-8) doi: 10.1016/S0140-6736(20)30251-8
- Maier, B. F., & Brockmann, D. (2020). Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*, 368(6492), 742–746. Retrieved from <https://science.sciencemag.org/content/368/6492/742> doi: 10.1126/science.abb4557
- Oliver, N., Lepri, B., Sterly, H., Lambiotte, R., Delataille, S., De Nadai, M., ... others (2020). Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Science Advances*. doi: 10.1126/sciadv.abc0764
- Pence, K. M. (2006). The role of wealth transformations: An application to estimating the effect of tax incentives on saving. *The BE Journal of Economic Analysis & Policy*, 5(1).
- Robert Koch-Institut. (2020a). *Erläuterung der Schätzung der zeitlich variierenden Reproduktionszahl R*. (URL: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Projekte_RKI/R-Wert-Erlaeuterung)
- Robert Koch-Institut. (2020b). *Fallzahlen in Deutschland*. Retrieved from https://services7.arcgis.com/m0BPYk0jAyB02ZKk/arcgis/rest/services/RKI{}_COVID19/FeatureServer/0/query?where=1{%}%3D1{%}&outFields=*&outSR=4326{%}&f=json (URL: <https://services7.arcgis.com/m0BPYk0jAyB02ZKk/arcgis/rest/services/>

[RKI_COVID19/FeatureServer/](#))

- Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it Difficult to Accurately Predict the COVID-19 Epidemic? *Infect Dis Model*, 5, 271-281.
- Schmidheiny, K., & Sieglöcher, S. (2019). *On event study designs and distributed-lag models: Equivalence, generalization and practical implications*. CESifo Working Paper 7481. (URL: https://www.cesifo.org/DocDL/cesifo1_wp7481.pdf)
- Stafford, N. (2020). Covid-19: Why Germany's case fatality rate seems so low. *BMJ*, 369. doi: 10.1136/bmj.m1395
- Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M. U. G., ... Dye, C. (2020). An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science*, 368(6491), 638–642. Retrieved from <https://science.sciencemag.org/content/368/6491/638> doi: 10.1126/science.abb6105
- Wang, C., Horby, P. W., Hayden, F. G., & Gao, G. F. (2020). A novel coronavirus outbreak of global health concern. *The Lancet*, 395(10223), 470–473. doi: 10.1016/S0140-6736(20)30185-9
- WHO-China Joint Mission. (2020). *Report of the WHO-China Joint Mission on coronavirus disease 2019 (COVID-19)*. Geneva. (URL: <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>)
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., ... Tan, W. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*, 382(8), 727–733. doi: 10.1056/NEJMoa2001017

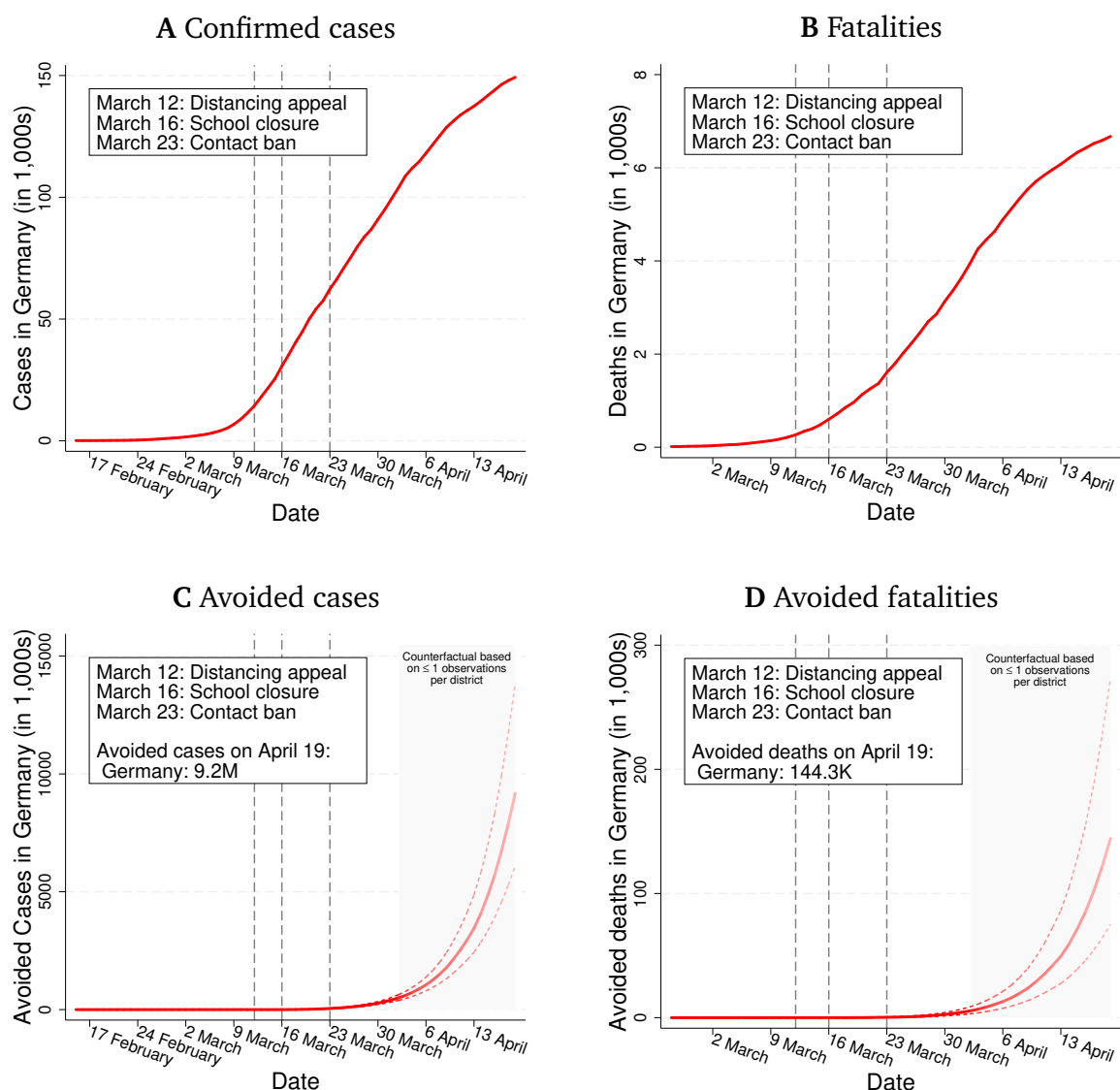
Online Appendix (not for publication)

Table A1: Correlations between outbreaks and progression of dependent variables

| Panel A. Confirmed cases measured in IHS units after: | | | | | | | |
|--|------------------------------|------------------------------|-----------------------------|-----------------------------|-----------------------------|------------------------------|-----------------------------|
| | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days |
| Local outbreak date | 0.002 (0.004) [0.633] | -0.001 (0.004) [0.830] | 0.001 (0.005) [0.903] | 0.007 (0.006) [0.272] | 0.017 (0.009) [0.051] | 0.021 (0.012) [0.081] | 0.016 (0.016) [0.315] |
| Panel B. Fatalities measured in IHS units after: | | | | | | | |
| | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days |
| Local outbreak date | -0.005 (0.007) [0.508] | -0.002 (0.008) [0.814] | 0.001 (0.009) [0.896] | 0.013 (0.011) [0.239] | 0.006 (0.015) [0.675] | -0.003 (0.017) [0.860] | 0.005 (0.022) [0.826] |
| <i>N</i> | 337 | 321 | 294 | 265 | 216 | 178 | 138 |

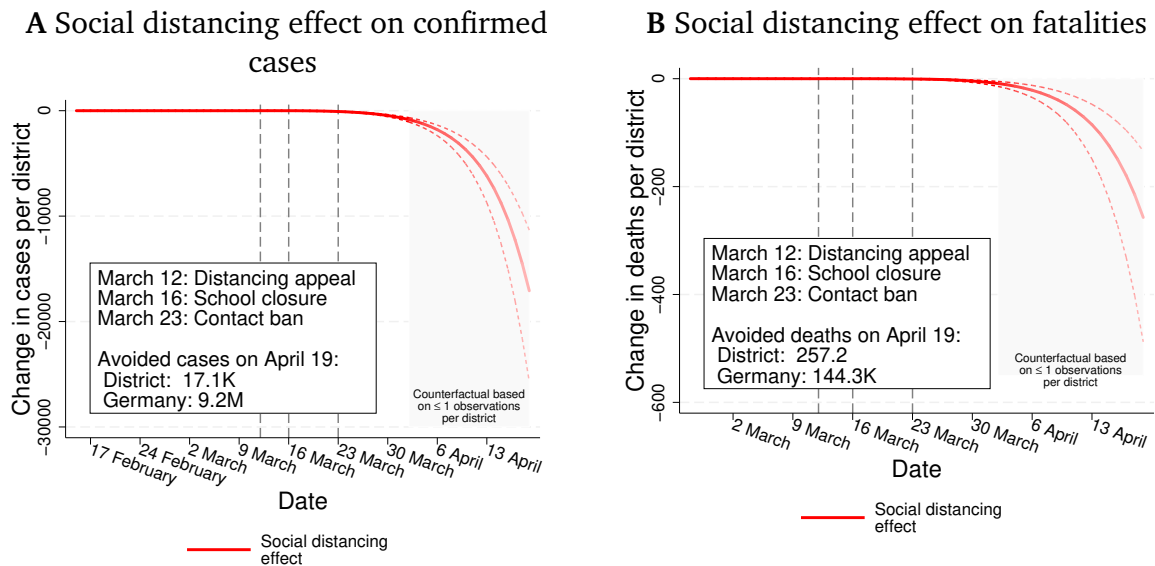
Notes: This table tests if the epidemic spread depends on the date of the local outbreak. The logic is simple: If, for example, districts with earlier outbreaks experienced a more rapid spread of COVID-19, then it is unlikely that these districts constitute a valid counterfactual. The details of the tests are as follows: First, we restrict our sample to observations before the implementation of SD policies. Second, we estimate bivariate OLS regressions of local cumulative COVID-19 cases (Panel A) and deaths (Panel B) X days after the local outbreak on the date of the local outbreak and a constant. Each column displays a different $X \in (1, 2, 3, 4, 5, 6, 7)$. As in the main model, the dependent variables are the inverse hyperbolic sine (IHS) transformation of the local cumulative case or death count. The table presents heteroscedasticity-robust standard errors in parentheses and p-values in brackets. With a Bonferroni correction for multiple testing, no correlation is significant (even at the ten percent level). To see this, note that all p-values exceed the threshold $0.10/7=0.0143$ (multiple testing within a panel) and a fortiori also exceed $0.10/14=0.0071$ (multiple testing across panels).

Figure A1: The epidemic's spread and social distancing policies in Germany (April 19)



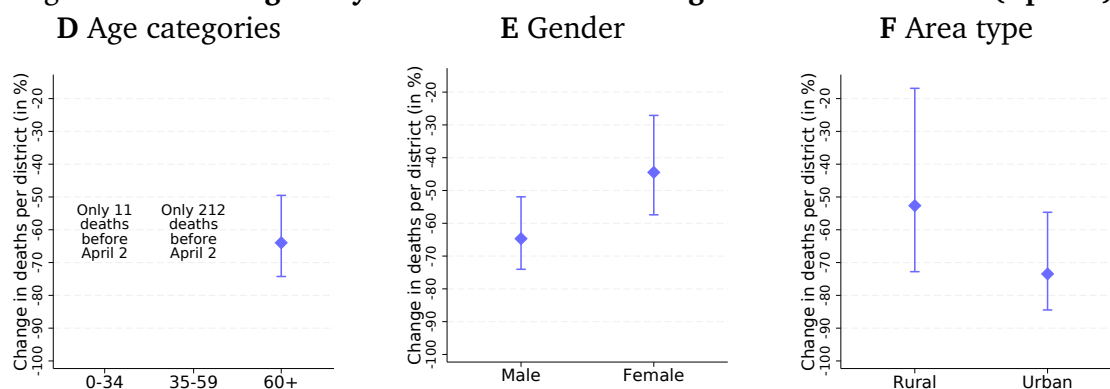
Notes: This figure depicts the epidemic spread and the effect of social distancing policies in Germany until April 19. It shows cumulative counts of confirmed COVID-19 cases (Subfigure A1A) and fatalities (Subfigure A1B) before and after the introduction of social distancing policies. On top of that, Subfigure A1C depicts our estimates on the number of cases avoided by social distancing policies. By contrast, Subfigure A1D focuses on avoided fatalities. The three vertical lines mark the Chancellor's appeal for social distancing (March 12), the nationwide school closures (March 16), and the nationwide contact bans (March 23). The dashed lines represent 95% confidence intervals based on district-level-clustered standard errors.

Figure A2: Social distancing effects in the extended period



Notes: This figure shows social distancing effects until April 19. Subfigure A2A focuses on confirmed cases and Subfigure A2B on fatalities. The three vertical lines mark the Chancellor’s appeal for social distancing (March 12), the nationwide school closures (March 16), and the nationwide contact bans (March 23). The dashed lines represent 95% confidence intervals based on district-level-clustered standard errors.

Figure A3: Heterogeneity of the social distancing effects on fatalities (April 2)



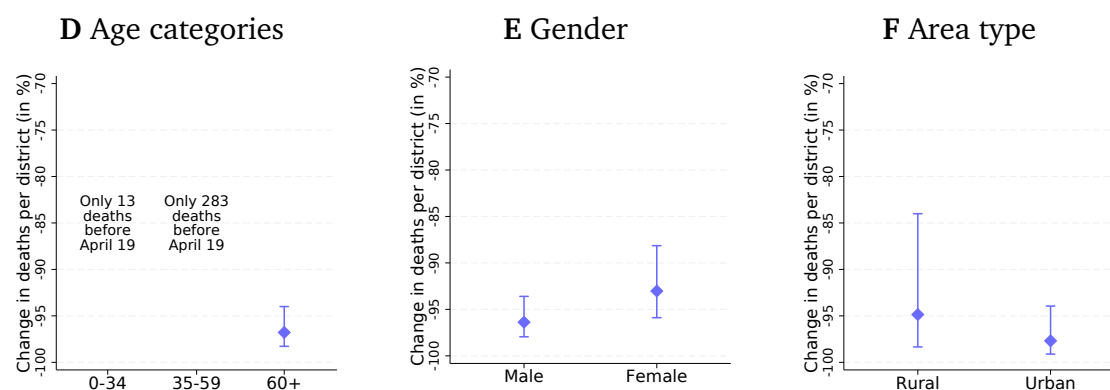
Notes: This figure shows how the effect of the nationwide social distancing policies on April 2 differed (a) across age groups, (b) by gender, and (c) between urban and rural districts. The estimates rely on sample splits and show the percentage of fatalities avoided by the nationwide policies. The vertical lines represent 95% confidence intervals based on district-level-clustered standard errors.

Figure A4: Heterogeneity of the social distancing effects (April 19)

Confirmed cases

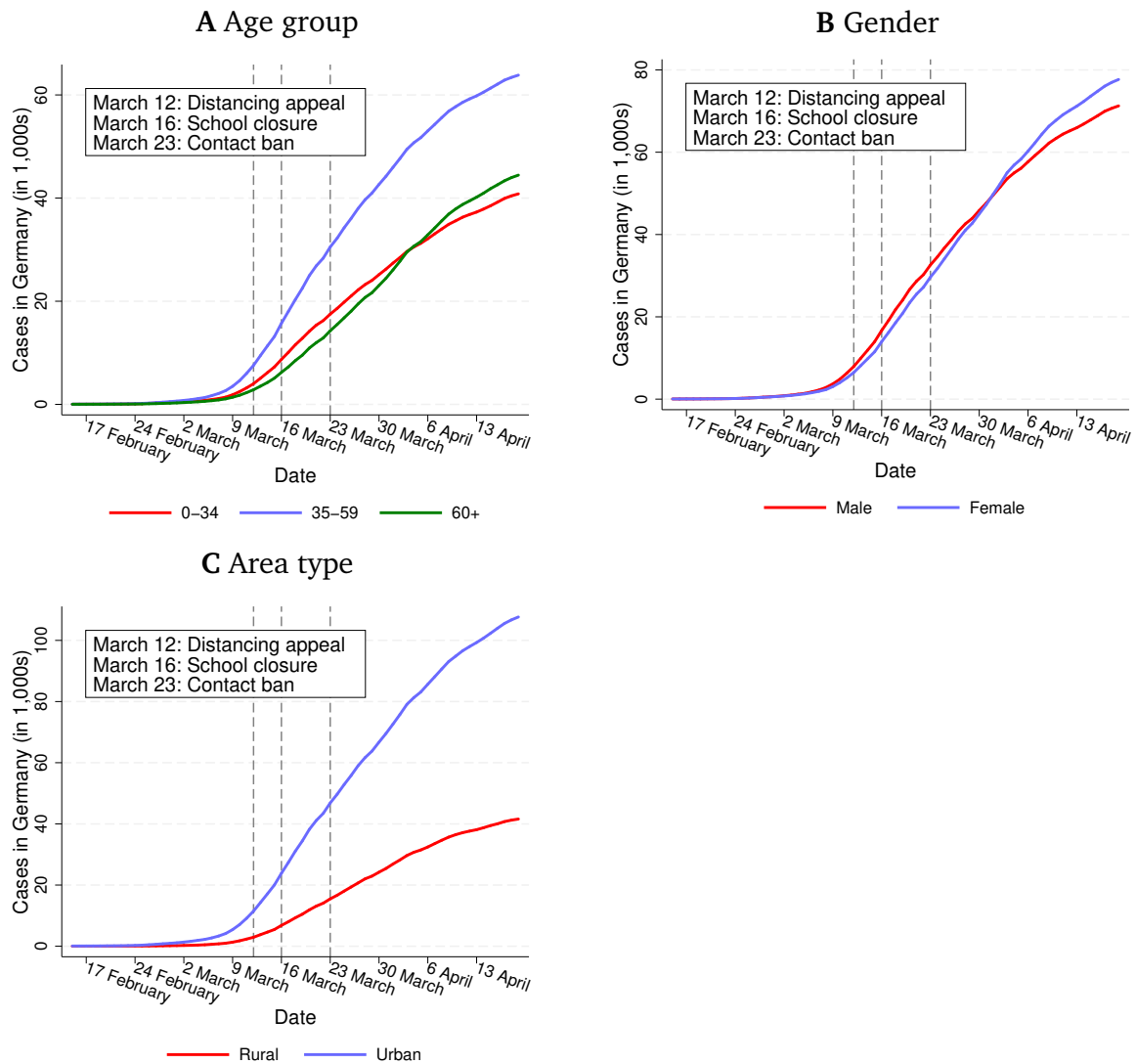


Fatalities



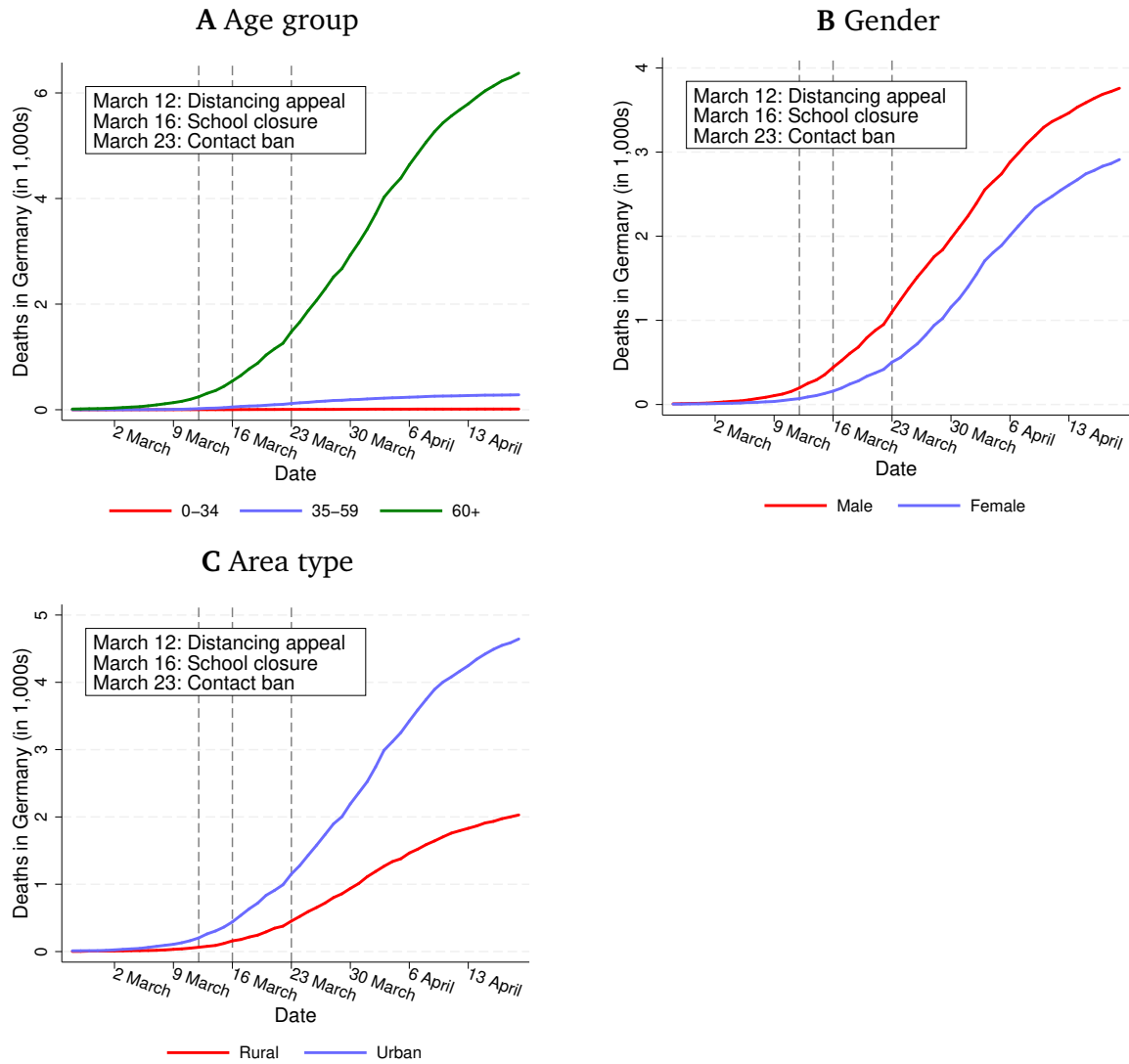
Notes: This figure shows how the effect of the nationwide social distancing policies on April 19 differed (a) across age groups, (b) by gender, and (c) between urban and rural districts. The estimates rely on sample splits and show the percentage of cases (first row) and fatalities (second row) avoided by the nationwide policies. The vertical lines represent 95% confidence intervals based on district-level-clustered standard errors.

Figure A5: Confirmed COVID-19 cases in Germany by groups



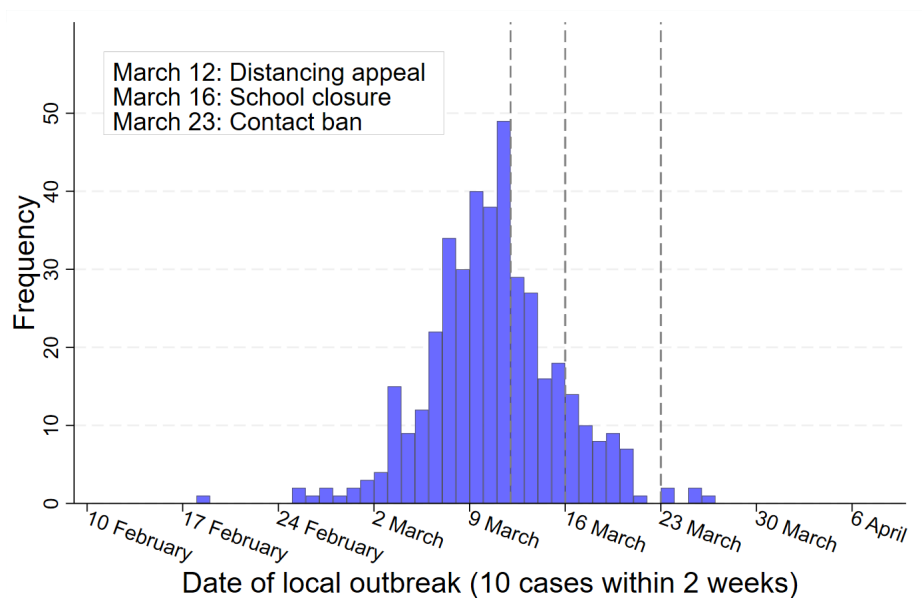
Notes: This figure shows cumulative counts of confirmed COVID-19 cases before and after the introduction of social distancing policies. It decomposes COVID-19 cases (a) across age groups, (b) by gender, and (c) between urban and rural districts.

Figure A6: COVID-19 fatalities in Germany by groups



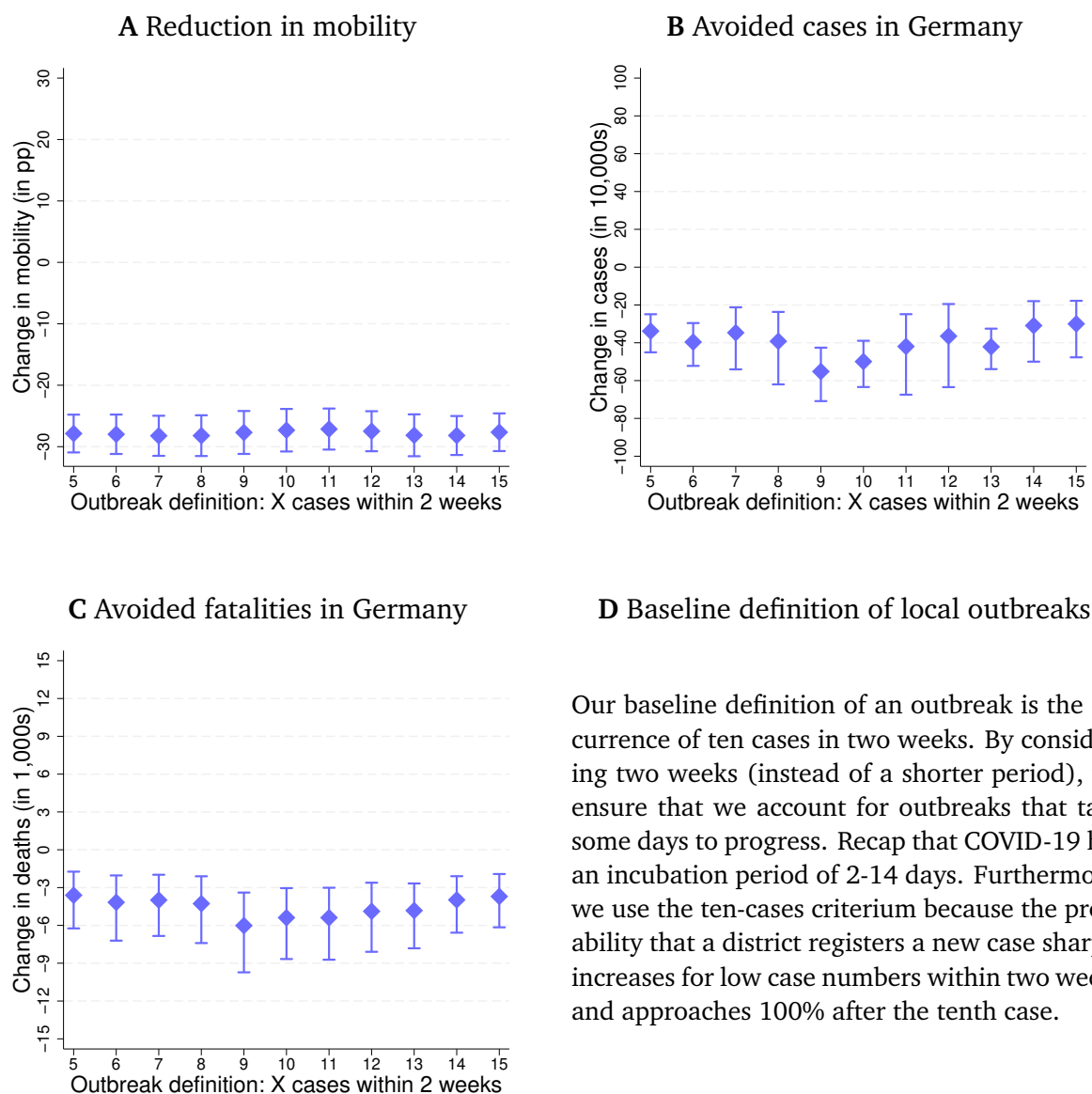
Notes: This figure shows cumulative counts of COVID-19 fatalities before and after the introduction of social distancing policies. It decomposes fatalities (a) across age groups, (b) by gender, and (c) between urban and rural districts.

Figure A7: Distribution of district-specific outbreaks



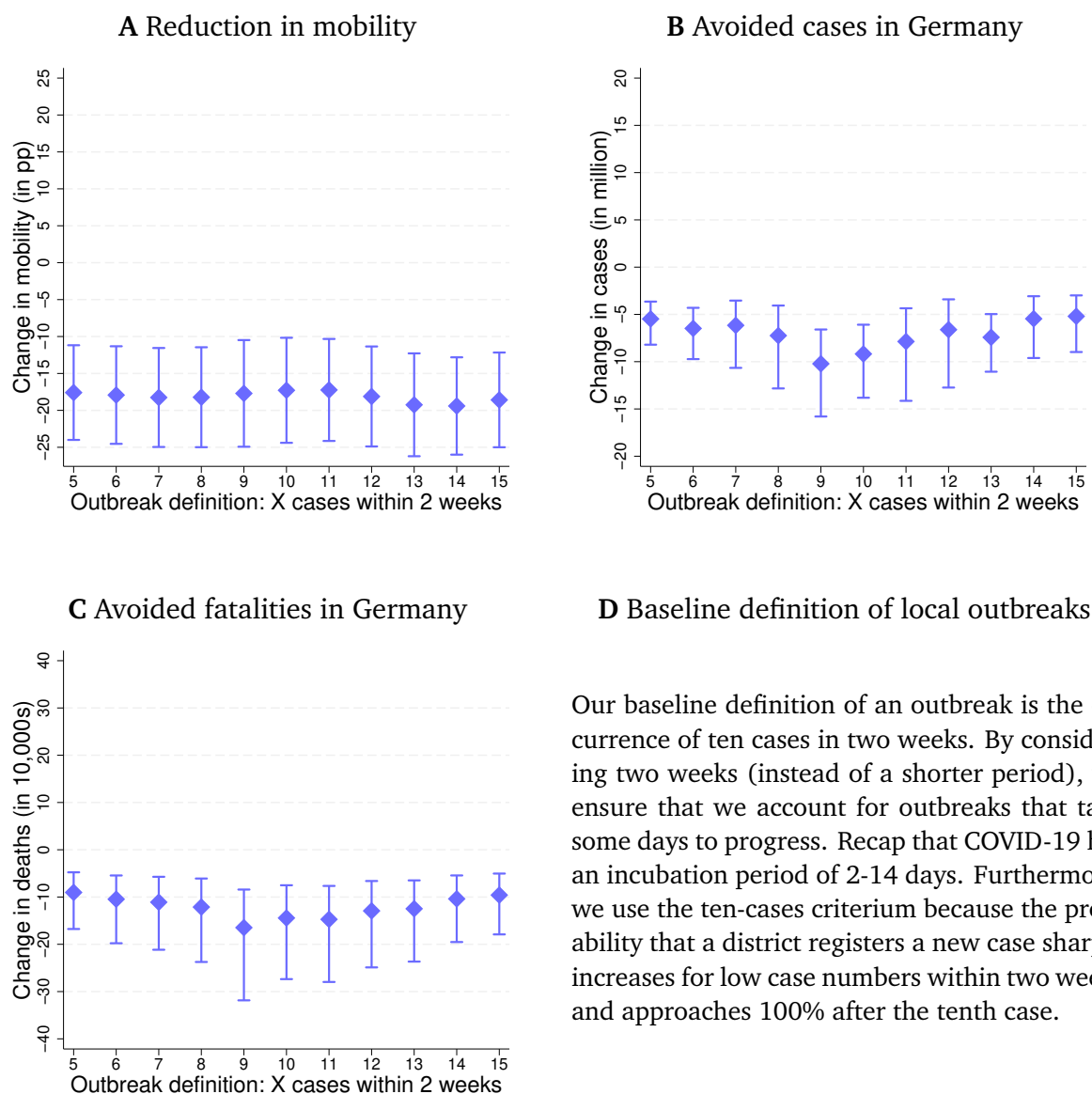
Notes: This figure shows how the district-specific COVID-19 outbreaks are distributed over time. Local outbreak dates are defined as the first day when ten confirmed COVID-19 cases have occurred within the last two weeks (where occurred refers to the date of the first symptoms or, if not available, the registry date of the case).

Figure A8: Social distancing effects under different outbreak definitions (April 2)



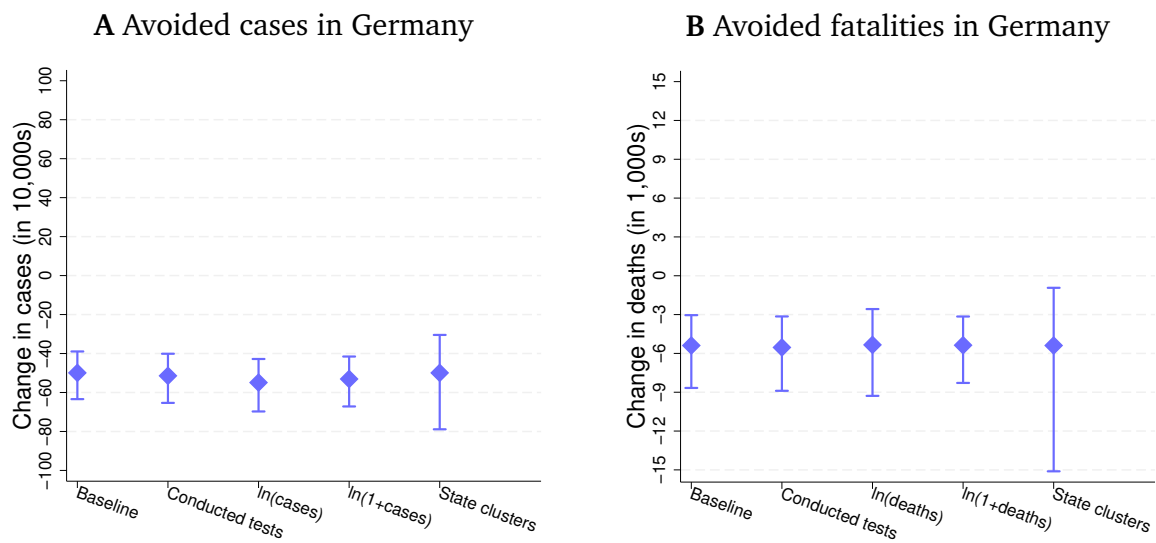
Notes: This figure shows social distancing effects on April 2 for different local outbreak definitions. Particularly, it varies how many district-specific COVID-19 cases within two weeks define an outbreak. The figure presents results for different outcome variables: Subfigure A8A focuses on mobility behavior, Subfigure A8B on the total number of avoided cases, and Subfigure A8C on the total number of avoided fatalities. Subfigure A8D discusses our baseline outbreak definition. The vertical lines represent 95% confidence intervals based on district-level-clustered standard errors.

Figure A9: Social distancing effects under different outbreak definitions (April 19)



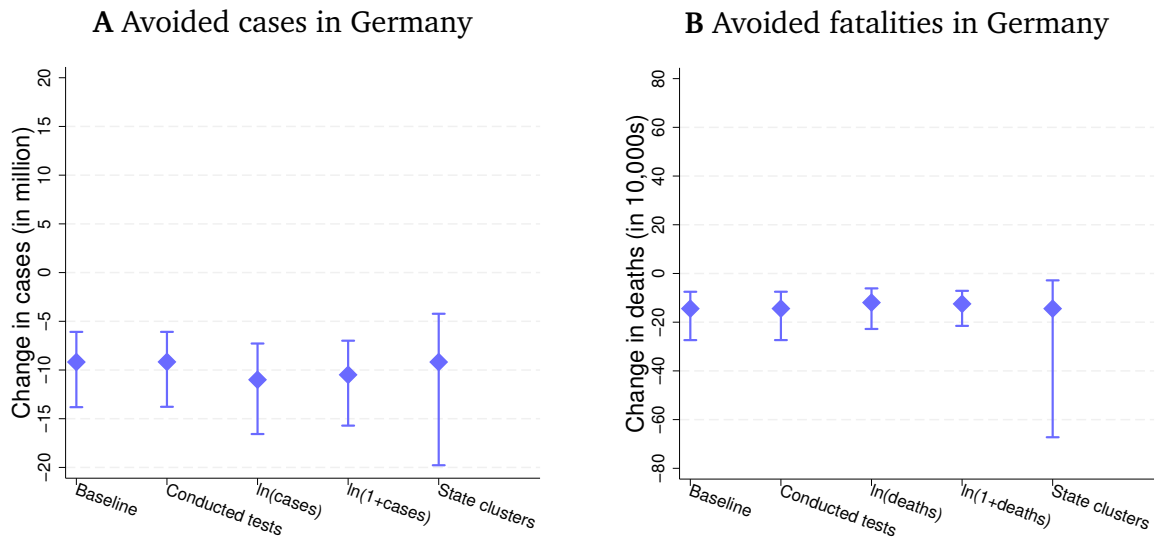
Notes: This figure shows social distancing effects on April 2 for different local outbreak definitions. Particularly, it varies how many district-specific COVID-19 cases within two weeks define an outbreak. Further, the figure presents results for different outcome variables: Subfigure A9A focuses on mobility behavior, Subfigure A9B on the total number of avoided cases, and Subfigure A9C on the total number of avoided fatalities. Subfigure A9D discusses our baseline outbreak definitions. The vertical lines represent 95% confidence intervals based on district-level-clustered standard errors.

Figure A10: Further robustness tests for the social distancing effect (April 2)



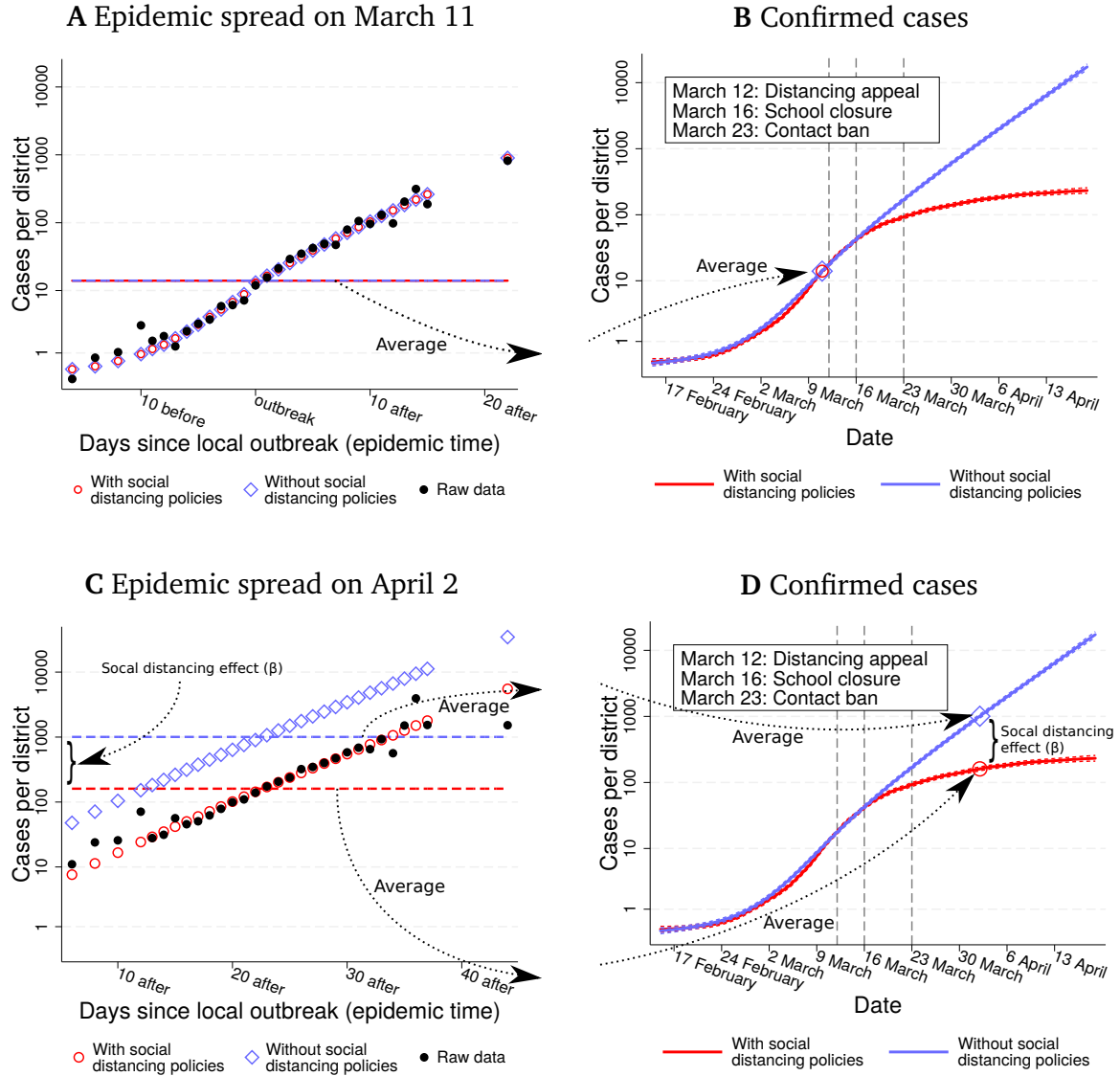
Notes: This figure presents the results of several robustness checks. Subfigure A10A focuses on avoided cases (April 2). Subfigure A10B presents the results for avoided fatalities (April 2). Both subfigures consider five specifications: First, for comparison, they show the baseline specification presented in the main text. The second specification controls for the number of conducted COVID-19 tests per capita. In the third specification, we drop observations with zero counts and take logs of the outcome. The fourth specification keeps observations with zero counts and applies the commonly used $\ln(1 + y)$ outcome transformation. The specifications one to four cluster standard errors on the district level. By contrast, specification five shows state-level-clustered standard errors for the baseline estimate. The vertical lines represent 95% confidence intervals.

Figure A11: **Further robustness tests for the social distancing effect (April 19)**



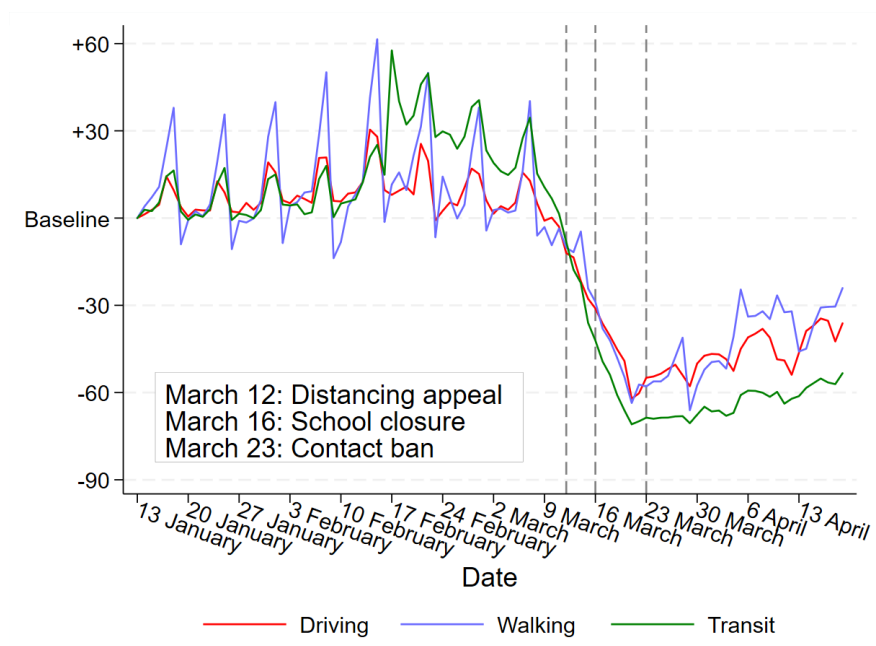
Notes: This figure presents the results of several robustness checks. Subfigure A11A focuses on avoided cases (April 19). Subfigure A11B presents the results for avoided fatalities (April 19). Both subfigures consider five specifications: First, for comparison, they show the baseline specification presented in the main text. The second specification controls for the number of conducted COVID-19 tests per capita. In the third specification, we drop observations with zero counts and take logs of the outcome. The fourth specification keeps observations with zero counts and applies the commonly used $\ln(1 + y)$ outcome transformation. The specifications one to four cluster standard errors on the district level. By contrast, specification five shows state-level-clustered standard errors for the baseline estimate. The vertical lines represent 95% confidence intervals.

Figure A12: Estimating the world without social distancing from the raw data



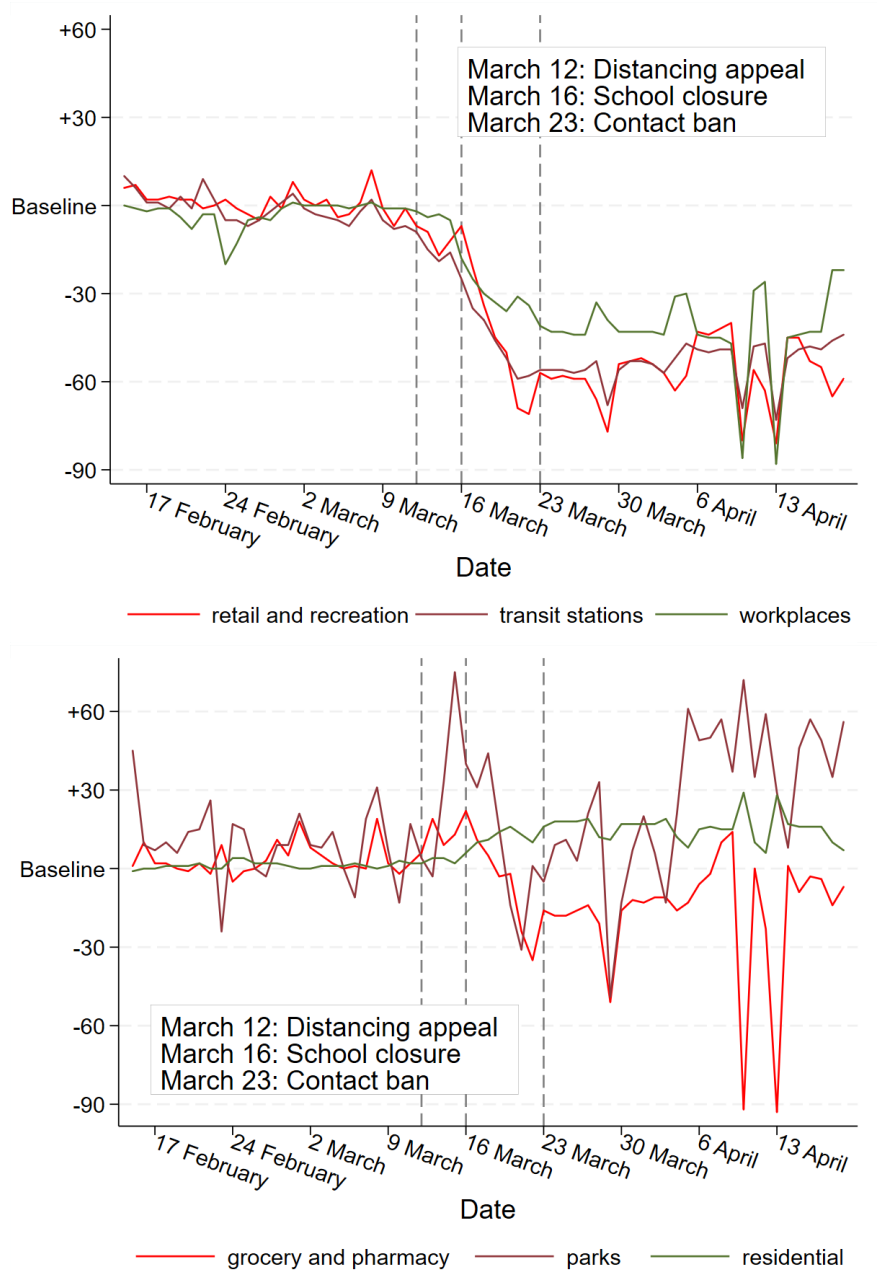
Notes: This figure illustrates our estimation approach graphically. The top (bottom) panel shows how confirmed cases on March 11, 2020 (April 2), enter the estimated evolution of cases. All subfigures represent cases per district (IHS means) on a logarithmic scale. Subfigures A and C show cases conditional on epidemic time. Black circles indicate raw data. Blue diamonds represent the epidemic time dummies estimated from (1) and constitute the no-SD policies prediction. Red circles indicate predicted cases with SD policies. For the omitted base day March 11, blue and red shapes coincide by construction. For April 2, the counterfactual values (blue diamonds) differ from the SD values (red circles) due to the SD effect: It corresponds to the mean deviation of April 2 cases from the cases of districts with the same epidemic times before SD policies. Subfigure A (C) suggests that the model fits the data on March 11 (April 2) very well. Despite the model not imposing constant growth, the predicted values almost lie on a straight line. The averages of the predictions in Panel A (C) (i.e., the horizontal dashed lines) form the per district values on March 11 (April 2) represented in Subfigure B (D).

Figure A13: Apple mobility trends for Germany



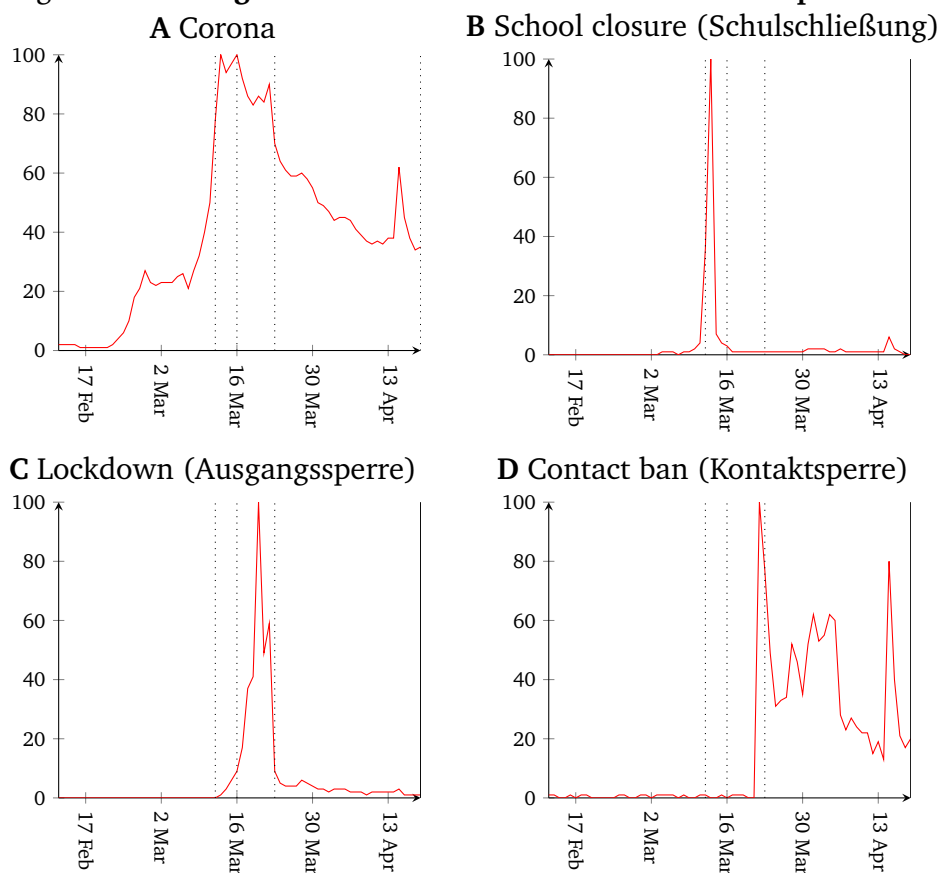
Notes: This figure shows percentage changes in the number of Apple Maps requests for several German cities and districts. The values are measured relative to a baseline date (Jan 13, 2020). The three vertical lines mark the Chancellor’s appeal for social distancing (March 12), the nationwide school closures (March 16), and the nationwide contact bans (March 23). Data: <https://www.apple.com/covid19/mobility>

Figure A14: Google mobility trends for Germany



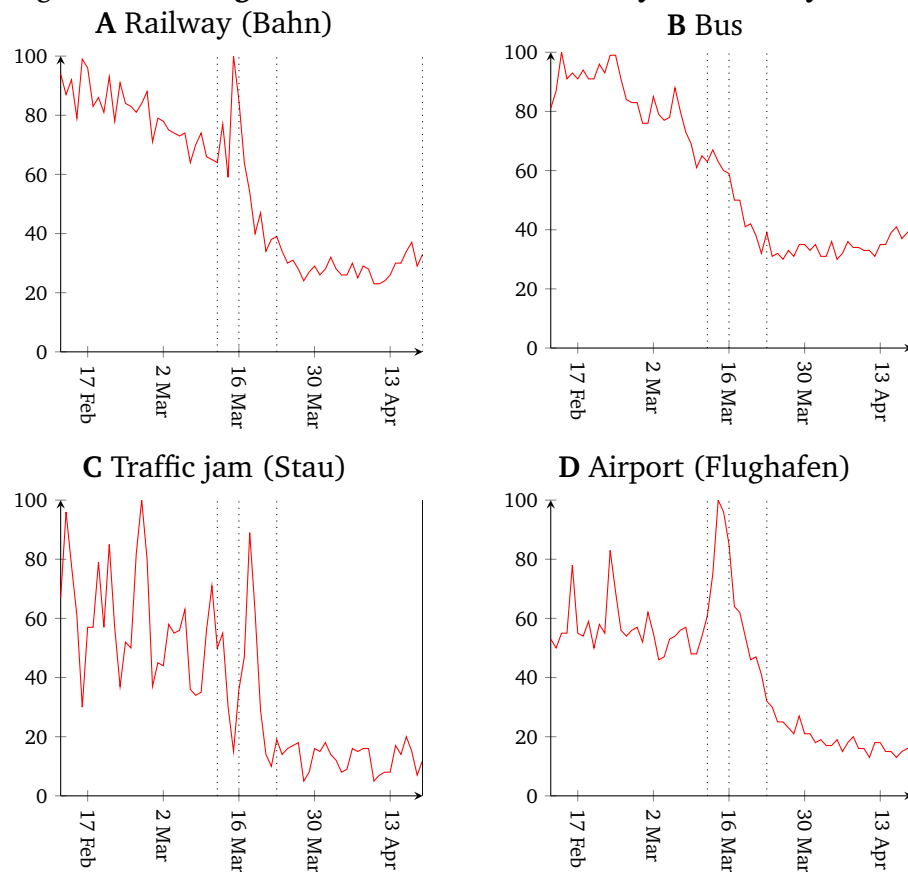
Notes: This figure shows percentage changes in a Google metric based on visits and length of stay at different places in Germany. The baseline is the median value, for the corresponding day of the week, during the 5-week period Jan 3-Feb 6, 2020. The upper figure presents location categories with large reductions in mobility. The bottom figure shows, with more volatility, increases for some places. The downward spikes on April 10 and 13 are due to Easter holidays in Germany. Data: <https://www.google.com/covid19/mobility/>.

Figure A15: Google-search trends for COVID-19 and SD policies



Notes: This figure shows how the relative frequency of Google searches for keywords related to COVID-19 and the German SD policies evolved from February 12 to April 19 in Germany. For each keyword, the maximum frequency of searches on a day is normalized to 100, and the frequencies on all other days are measured relative to this maximum. The three vertical lines mark the Chancellor's appeal for social distancing (March 12), the nationwide school closures (March 16), and the nationwide contact bans (March 23). Data: <https://trends.google.de/>.

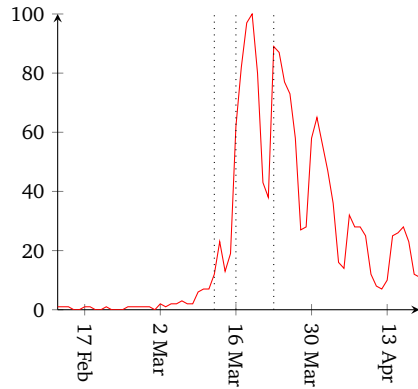
Figure A16: Google-search trends for mobility-related keywords



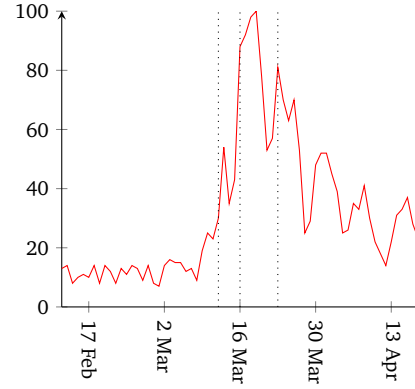
Notes: This figure shows how the relative frequency of Google searches for mobility-related keywords evolved from February 12 to April 19 in Germany. For each keyword, the maximum frequency of searches on a day is normalized to 100, and the frequencies on all other days are measured relative to this maximum. The three vertical lines mark the Chancellor's appeal for social distancing (March 12), the nationwide school closures (March 16), and the nationwide contact bans (March 23). Data: <https://trends.google.de/>.

Figure A17: Google-search trends for work- and catering-related keywords

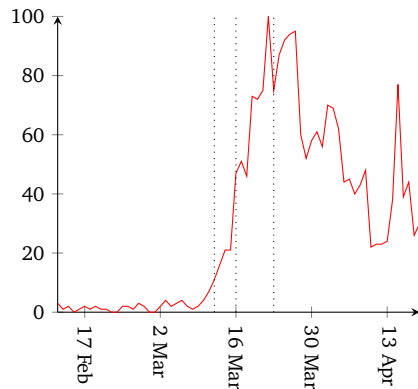
A Short-term working (Kurzarbeit)



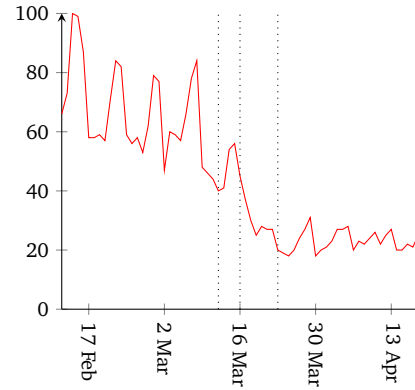
B Home office



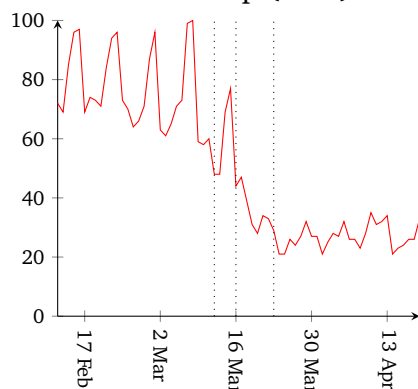
C Video conference (Videokonferenz)



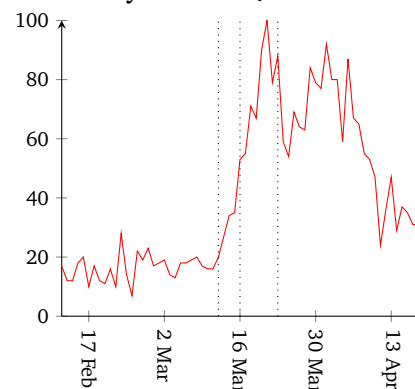
D Restaurant



D Coffee shop (Cafe)

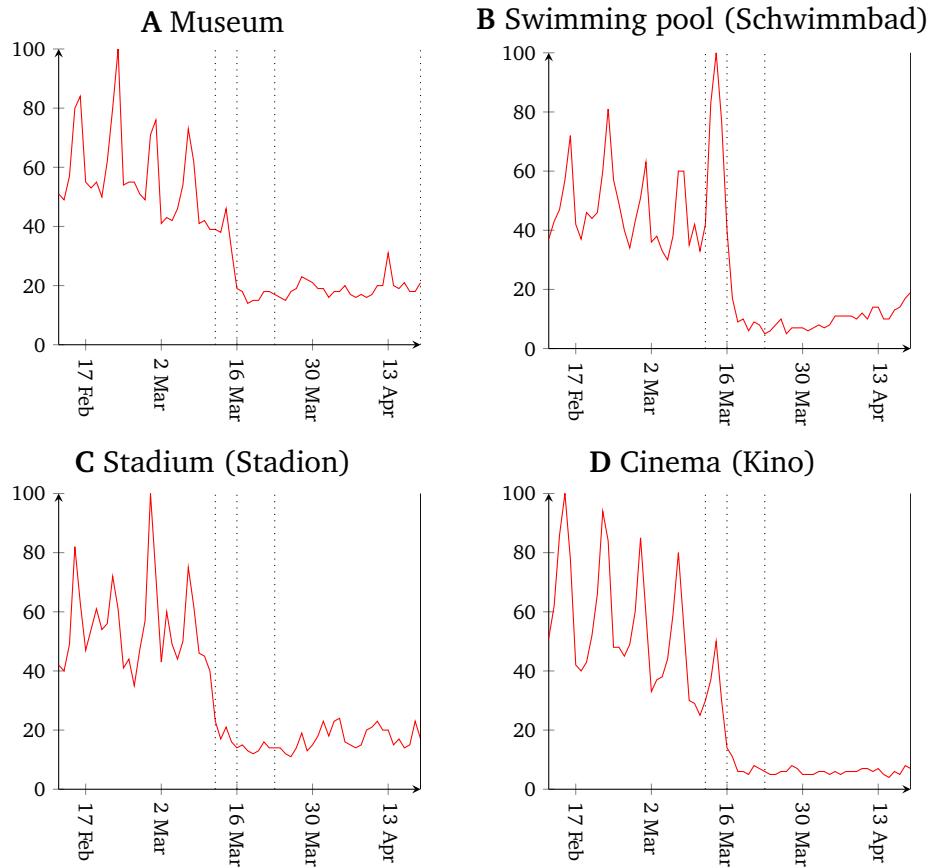


D Delivery service (Lieferdienst)



Notes: This figure shows how the relative frequency of Google searches for work-related and catering-related keywords evolved from February 12 to April 19 in Germany. For each keyword, the maximum frequency of searches on a day is normalized to 100, and the frequencies on all other days are measured relative to this maximum. The three vertical lines mark the Chancellor's appeal for social distancing (March 12), the nationwide school closures (March 16), and the nationwide contact bans (March 23). Data: <https://trends.google.de/>.

Figure A18: Google-search trends for leisure time activities



Notes: This figure shows how the relative frequency of Google searches for keywords related to leisure time activities evolved from February 12 to April 19 in Germany. For each keyword, the maximum frequency of searches on a day is normalized to 100, and the frequencies on all other days are measured relative to this maximum. The three vertical lines mark the Chancellor's appeal for social distancing (March 12), the nationwide school closures (March 16), and the nationwide contact bans (March 23). Data: <https://trends.google.de/>.