Manuscript submitted to *The Econometrics Journal*, pp. 1–18.

The spread of COVID-19 and the BCG vaccine: A natural experiment in reunified Germany

Richard Bluhm[*,†] and Maxim Pinkovskiy[‡]

[*] *Leibniz University Hannover, Institute of Macroeconomics, 30167 Hannover, Germany*
[†] *University of California San Diego, Department of Political Science, La Jolla, CA 92093*
E-mail: bluhm@mak.uni-hannover.de
[‡] *Federal Reserve Bank of New York, Microeconomic Studies Function, New York, NY 10045*
E-mail: maxim.pinkovskiy@ny.frb.org

Summary The "BCG hypothesis" suggests that the Bacillus Calmette-Guérin (BCG) vaccine against tuberculosis limits the severity of COVID-19. We exploit the differential vaccination practices of East Germany and West Germany prior to reunification to test this hypothesis. Using a differences in regression discontinuities (RD-DD) design centred on the end of universal vaccination in the West, we find that differences in COVID-19 severity across cohorts in the East and West are insignificant or have the wrong sign. We document a sharp cross-sectional discontinuity in severity of the disease, which we attribute to limited mobility across the long-gone border.

Keywords: *COVID-19, BCG vaccine, Germany, mobility, SIR model with commuting flows*

1. INTRODUCTION

Since December 2019, the disease caused by the novel coronavirus (COVID-19) has infected over 30 million people worldwide, of whom over one million have died.[1] The pandemic has produced an unprecedented decline in global economic activity as countries enforce social distancing measures to contain the spread of the virus.

While some drugs appear to have positive effects on clinical outcomes,[2] to date, there is no targeted vaccine for COVID-19 that is known to be safe, effective and widely used. This has sparked considerable interest in whether some types of vaccines that are already known to be safe may have positive indirect effects on the spread and severity of COVID-19 infections. Specifically, there is now a lively controversy over whether the Bacillus Calmette-Guérin (BCG) vaccine against tuberculosis also partially protects individuals against COVID-19. Several studies point out that countries with mandatory BCG vaccination tend to have substantially fewer coronavirus cases and deaths per capita than countries without mandatory vaccination, and that the intensity of the epidemic is lower for countries that began vaccinating earlier (Berg et al., 2020; Escobar et al., 2020; Gursel and Gursel, 2020; Hauer et al., 2020; Sharma et al., 2020). Non-specific or off-target effects of live vaccines are not uncommon and have been documented in a variety of settings (Kleinnijenhuis et al., 2015; Chumakov et al., 2020). The BCG vaccine appears to protect against its target, some forms of tuberculosis, for up to 60 years

[1] As of September 28 2020, based on data from coronavirus.jhu.edu.

[2] Dexamethasone—a common steroid—has been shown to reduce the level of respiratory support required in hospitalized patients and reduced COVID-19 related mortality in clinical trials (The RECOVERY Collaborative Group, 2020).

(Aronson et al., 2004) but has also been associated with long-term reductions in all-cause mortality and mortality from respiratory diseases (Rieckmann et al., 2016). Live vaccines appear to elicit a strong response of the immune system which subsequently offers broad protection against unrelated pathogens (Chumakov et al., 2020). Similar 'trained immunity' responses against COVID-19 in BCG vaccinated individuals are therefore plausible (O'Neill and Netea, 2020), but there is a lack of rigorous evidence investigating these non-specific effects. Clinical trials are taking place across the globe which test the effectiveness of the BCG vaccine against COVID-19.[3] These trials are likely to take at least a year while the virus continues to spread at a rapid pace. The WHO cautions that there is currently no evidence that the BCG vaccine protects against the novel coronavirus (Curtis et al., 2020).

In the absence of experimental results, we propose to use the tools of modern applied econometrics to test the hypothesis that BCG may offer long-run protection against COVID-19. We use a regression discontinuity differences-in-differences (RD-DD) analysis of severe coronavirus cases to exploit a natural experiment along the former border between East Germany and West Germany. This border separated the two sides from 1949 until Germany was reunified in October 1990. West Germany phased out a policy of de facto universal BCG vaccination (which began in the 1950s) for the general population starting in 1974, while East Germany strictly enforced a policy of mandatory BCG neonatal vaccination from 1953 until 1990. Moreover, neonatal vaccination was completely interrupted in West Germany in 1975 and 1976 when the only licensed vaccine caused unexpected side effects and had to be withdrawn from the market (Genz, 1977). Residual BCG vaccination of risk groups then ended permanently in reunified Germany in 1998. This gives rise to a natural experiment that creates variation over time and space. Individuals born shortly after 1974 just to the west of the former border experienced a sharp drop in BCG coverage relative to those born shortly before 1974, while their peers just to the east were consistently vaccinated throughout (at rates exceeding 95%).

Our analysis takes advantage of several important features that are present in the German context. First, there is an internal border sharply dividing different vaccination regimes for a limited period of time. Second, the German health ministries of each state report detailed records to the Robert Koch Institute (RKI), which disseminates standardized data on COVID-19 cases by geographic location and by single year of age. The latter allows us to compare individuals in close age cohorts who are also close to each other in physical space, and therefore differ less on unobservable characteristics. This alleviates the concerns raised by Becker et al. (2020) about cross-sectional comparisons and discontinuity designs between East and West Germany (e.g. Alesina and Fuchs-Schündeln, 2007; Campa and Serafinelli, 2019; Goldfayn-Frank and Wohlfart, 2019; Lippmann et al., 2020). While differences in BCG vaccination rates across space (Hauer et al., 2020) or cohorts have been used in related studies (Hamiel et al., 2020), to our knowledge this is the first paper combining geographic *and* age variation in BCG vaccination status for reliable causal inference. A third feature of the German reporting system is that every test result submitted to the RKI indicates a case definition, allowing us to distinguish cases with acute respiratory symptoms (including pneumonia), or patients who have died

[3]For example, the BRACE trial in Australia (`NCT04327206`), the BADAS trial in the United States (`NCT04348370`), or the BCG-CORONA trial in the Netherlands (`NCT04328441`). Clinicaltrials.gov identifiers are provided in parentheses.

from COVID-19, from benign cases without acute symptoms. Fourth, areas of Germany on both sides of the border have been subject to the same state response to the COVID-19 pandemic and have comparable access to medical services, creating a high level of homogeneity in pandemic policy. Finally, Germany publishes detailed data on county-by-county commuter flows, allowing us to investigate a potentially important factor in the transmission of COVID-19.

Our RD-DD results contradict the BCG hypothesis. We find that individuals born just to the east of the former border shortly *after* 1974 are, if anything, more likely to have a reported case of COVID-19 relative to their peers in the west than individuals born just to the east of the border shortly *before* 1974. Our estimates for the potential "effect" of the BCG vaccine consistently have the wrong sign, typically cannot be distinguished from zero, and exclude large effects of the BCG vaccine with the expected sign under the BCG hypothesis. We perform covariate balancing tests to rule out that other outcomes, such as mortality and respiratory hospitalizations, change differentially between East and West Germany for cohorts born after rather than before 1974. Performing a less restrictive differences-in-differences (DD) analysis between all of East and West Germany, as well as between cohorts born before and after 1974, yields similar results. We also obtain comparable results when we focus on severe cases only. Analogous RD-DD and DD analyses in 1990—the other critical date in the history of the BCG vaccine in Germany—deliver even stronger estimates of the differential effect that run counter to the BCG hypothesis, although their covariate balancing tests are not as stable. Taken together, these findings cast serious doubt that the correlations adduced by the literature supporting the BCG hypothesis capture a causal relationship (Berg et al., 2020; Escobar et al., 2020; Gursel and Gursel, 2020; Hauer et al., 2020; Sharma et al., 2020).

Our analysis reveals a puzzle. If we ignore the variation over time, we observe a sharp discontinuity in COVID-19 cases at the former border separating West and East Germany. There are considerably fewer cases and deaths in the former East. If it is not the BCG vaccine, then what explains this jump?[4] Our solution to this puzzle suggests that while the virus does not stop at the long-gone border, people who carry the virus still do. Cross-county commuter flows in Germany are sharply discontinuous at the former border, which reflects a continued lack of connectedness between West German and East German transportation infrastructure. We find that controlling for commuter flows significantly decreases the discontinuities in COVID-19 cases and deaths at the border, while adding some basic demographics removes these discontinuities completely. This echoes the findings of a recent literature on border controls and travel restrictions (e.g., Chinazzi et al., 2020; Eckardt et al., 2020). Going further, we demonstrate that discontinuities in commuter flows can generate discontinuities in COVID-19 cases at the former border without any reference to the BCG hypothesis. We simulate a spatial SIR model of viral spread across German counties where commuter flows act as a transmission channel (as in Wesolowski et al., 2017). We show that even within this very stylized framework, we can obtain a sizable discontinuity in COVID-19 intensity without incorporating anything relating to vaccination into the model. Last but not least, we show that the discontinuities

[4]Prominent newspapers in Germany have noticed that there is much lower COVID-19 prevalence in the former East than in the West but offer only suggestive explanations (e.g., *Die Zeit*, a German weekly, `www.zeit.de/2020/13/coronavirus-ausbreitung-osten-westen-faktoren`, or *Der Tagesspiegel*, a Berlin-based German daily, `www.tagesspiegel.de/politik/mehr-flaeche-mehr-alte-warum-der-osten-weniger-unter-corona-leidet/25796940.html`). Moreover, low COVID-19 mortality in Germany as a whole has been the subject of media interest.

in cases and severe cases weaken over time. In fact, we do not observe significant discontinuities in new cases occurring after the initial outbreak in spring. We interpret this as additional evidence in favor of mobility playing a key role in "seeding" the early distribution of the outbreak, as opposed to the spread being contingent on innate characteristics of the population.

The remainder of the paper is organized as follows. Section 2 describes the county-level data on cases, covariates, and commuter flows. Section 3 outlines our empirical strategy. Section 4 presents the balancing tests and main RD-DD results for cases by age-groups. Section 5 explores mobility as an alternative explanation for the sharp discontinuity in overall cases and presents results from placebo tests with cases simulated from a county-level SIR model with mobility. Section 6 concludes.

2. DATA

The COVID-19 pandemic in Germany can be characterized by two distinct phases: *i)* an initial outbreak in spring which began in late February and lasted about four weeks into the lockdown on 23 March 2020 until about April 26 2020, and *ii)* a slow resurgence of coronavirus cases as travel restrictions were lifted over the summer.[5] For most of the analysis, we focus on the first period between the start of the pandemic in Germany over the peak with more than 6,000 daily cases until new infections dropped again to less than 2,000 cases per day.

Our main dependent variable is the logarithm of one plus the number of cumulative reported COVID-19 cases per million people in a German county (*Kreis*) as of April 26, 2020. Germany reported a total of 158,047 COVID-19 cases and 8,122 deaths by this date. We obtain counts of cumulative and new COVID-19 cases and deaths by German county for every date since January 27 2020 (the start of week 5) from the RKI's *NPGEO Corona Hub 2020*. Figure 1 shows a map of cumulative cases by county as of April 26 2020 (the end of week 17). Given the difficulties in recording asymptomatic cases of COVID-19, it is all but certain that the case counts we employ in this paper are a lower bound for COVID-19 cases in Germany. In the period of observation, testing capacity was high but testing was limited to individuals displaying symptoms, returnees from high risk countries, and those who have been in contact with confirmed cases. We proceed on the necessary assumption that undercounting errors do not vary systematically across age groups and locations.

The RKI also separately reports case data by single-year age groups and week via the platform *SurvStat@RKI 2.0*. By federal law, all cases reported to the RKI have to follow a case definition (*Falldefinition*).[6] This categorization allows us to distinguish *severe* cases which have been confirmed by a PCR test and display acute symptoms (about 63% of all cases)—ranging from respiratory distress over pneumonia to death—from cases which do not display any clinical symptoms or are entirely asymptomatic (26%), and from cases whose clinical status is unknown (11%).[7] The high share of severe cases in overall cases is consistent with the notion that the cumulative case count by and large indicates COVID-

[5] Figure S.3. provides a time-series of daily cases.

[6] For a full description (in German), see `www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Falldefinition.pdf?__blob=publicationFile`.

[7] Figure S.4. includes a figure of the age distribution of cases and severe cases for the territories of the former West and East.

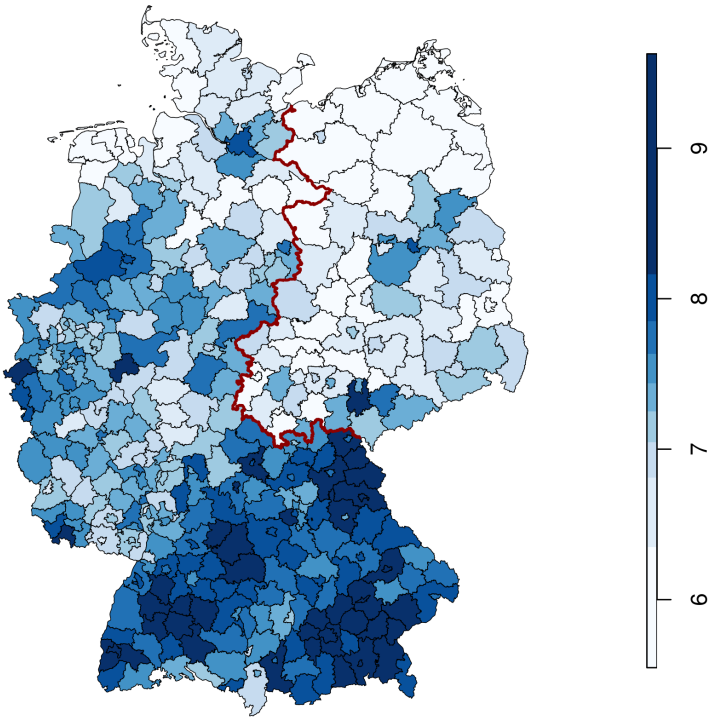

Figure 1. Spatial distribution of COVID-19 cases in Germany

19 severity rather than just incidence. We use both overall and severe cases throughout the analysis. Deaths by single-year age groups and week are not reported by the RKI.

We compile several county-level characteristics: income and demographics, labour market statistics, historical mortality, and commuting flows. Measures of disposable income in 2017, the age distribution in 2018, population density in 2018 and labour market outcomes from the 2011 census are taken from official statistics published by the German federal statistical office (via the *GENESIS* regional database). We also collect data on overall mortality, mortality from selected infectious diseases and mortality from respiratory diseases in 2016 from the same source. Given that the age-profile of the population varies significantly across counties in Germany, we manually age-adjust all mortality figures using the corresponding age distribution of the entire country in 2016. Finally, we use the latest available data on commuting flows published by the Federal Employment Agency from December 2019. The agency regularly releases an origin-destination matrix of commuting flows across German counties. These flows capture about 33 million jobs. Approximately 13 million of the jobs are in a different county than the primary tax residence of the employee.

3. EMPIRICAL STRATEGY

Our study exploits discontinuous changes in vaccination policies across the former border dividing East and West Germany from 1949 until 1990, with a focus on the cessation of widely recommended BCG vaccination after 1974 in the West.[8]

Ignoring variation across cohorts, we can estimate the discontinuity between East and West Germany in COVID-19 intensity, y_c, by running a standard geographic regression discontinuity (RD) design (see e.g. Lalive, 2008; Dell, 2010):

$$y_c = \alpha + \beta \text{EAST}_c + \delta_1 d_c + \delta_2 \left(d_c \times \text{EAST}_c\right) + \lambda_{s(c)} + \varepsilon_c \quad \text{for} \quad |d_c| < b \tag{3.1}$$

where c indexes counties (*Kreise*), EAST_c indicates whether the county was part of East Germany before reunification, d_c is the distance of county c from the former East German border (it is negative if $\text{EAST}_c = 0$ and positive if $\text{EAST}_c = 1$), and $\lambda_{s(c)}$ is a fixed effect for the border segment associated with county c. Border segments are defined as pairs of bordering states (*Bundesländer*). We drop Berlin throughout the analysis and focus on the border separating the two larger countries. Our specification uses an interacted local linear RD polynomial at a variety of plausible bandwidths b (Gelman and Imbens, 2019), ranging from 50 km to 200 km, centred on a value close to a statistically optimal bandwidth (Imbens and Kalyanaraman, 2012).

The coefficient of interest in this design, β, does *not* identify the effect of the BCG vaccine on COVID-19 severity. Instead, it captures a compound treatment effect of discontinuous changes in other variables that are related to being born in East Germany and directly affect COVID-19 transmission. Put simply, East Germany differed from West Germany in many more ways than BCG vaccination (Alesina and Fuchs-Schündeln, 2007; Wolf, 2009; Eder and Halla, 2018; Campa and Serafinelli, 2019; Wyrwich, 2019; Goldfayn-Frank and Wohlfart, 2019; Becker et al., 2020; Lippmann et al., 2020). We only use β as a reference to gauge the size of the baseline discontinuity for all ages or for specific cohorts.

To isolate the effect of BCG vaccination on COVID-19 severity, we estimate the *difference in discontinuities* at the border for cohorts born just before and just after West Germany suspended widespread recommendation of the BCG vaccine in 1974. We call this a regression discontinuity differences-in-differences (RD-DD) design (see e.g., Deshpande, 2016), which we estimate using the following specification:

$$\begin{aligned} y_{c,a} = \; & \alpha_a + \beta \text{EAST}_c + \gamma \text{EAST}_c \times \text{TREATED}_a + \delta_1 d_c + \delta_2 d_c \times \text{EAST}_c + \\ & \delta_3 d_c \times \text{TREATED}_a + \delta_4 d_c \times \text{EAST}_c \times \text{TREATED}_a + \lambda_{s(c),a} + \varepsilon_{c,a} \quad \text{for} \quad |d_c| < b \end{aligned} \tag{3.2}$$

where a indexes age groups (pre- and post-vaccination change), TREATED_a is an indicator for the age group for which the policy experiment created no differential in vaccination status across the border (here it is the age group born before the vaccination policy change), and the intercept and the border segment fixed effects are allowed to vary by age group. The rest of the notation is the same as before. The coefficient of interest, γ, delivers an estimate of the difference in discontinuities across cohorts.

The assumption needed for this specification to recover a causal effect of BCG vaccination on the outcome variable is that any discontinuities across the former border—other than the effect of the vaccine—are constant across cohorts born shortly before

[8] Online Appendix S1 contains a detailed history and timeline of BCG vaccination policies in East and West Germany.

and shortly after the change in vaccination regime. As it is not apparent that anything else happened in 1974 that would affect newborns differently in East and West Germany (e.g. no other vaccines were introduced or withdrawn) we view this as a plausible assumption.[9] Moreover, we can test this assumption by looking at other outcomes for successive cohorts around 1974. Equation 3.2 can also be viewed as nesting a standard differences-in-differences specification, in which we compare outcomes in all of East Germany with those in all of West Germany, and for cohorts born just before or just after 1974. In the framework above, such a specification would set the bandwidth b equal to infinity, and the coefficients on the distance terms and the border segment fixed effects to zero.

4. RESULTS

4.1. Balancing tests

We start by presenting evidence that several other outcomes for cohorts born just before and just after 1974 are similar, suggesting that the RD-DD design is not capturing the effects of differences between these cohorts unrelated to BCG vaccination. Table 1 presents these balancing tests. Unfortunately, we do not have data on alternative outcomes by single year of age. However, we have mortality and hospitalization rates in 5-year age bins derived from administrative data in 2016 and we have labour market indicators in 2011, also in 5-year age bins, from the latest population census.

Table 1 presents the coefficient γ from equation 3.2 when the logarithm of mortality or hospitalization rates or raw unemployment and out-of-labour force rates are used as a dependent variable in the RD-DD design. In each specification, we omit the 5-year age cohort that was born immediately around the 1974 vaccination change in the West as we cannot classify them neatly as treated or untreated. In the case of health outcomes, we omit the 40–44 year old group (since these were measured in 2016) and for the labour market outcomes we omit the 35–39 year old group (as they were measured in 2011). In column 1 we take the 5-year age group immediately older than the omitted group as the "treated" group and the 5-year age group immediately younger than the omitted group as the "untreated" group. For example, in the case of mortality this compares 45–49 year olds to those that are 35–39. In column 2 we take the two 5-year age groups immediately older than the omitted group as the "treated group" and construct the "untreated group" symmetrically (e.g., 45–54 year olds versus 30–39 year olds). All other cohorts are omitted from the analysis. Columns 3 and 4 present analogous balance tests for the cohorts born around 1990.

Nearly all health and labour market outcomes appear to be balanced for the cohorts born around 1974. We only observe a significant difference-in-discontinuities at the 5% significance level between the treated and untreated cohorts across counties for one of the eight outcomes under consideration. Older cohorts in the East appear to be hospitalized more often, no matter if we consider 5-year or 10-year age groups. Reassuringly, this difference-in-discontinuities disappears once we examine hospitalizations from infectious diseases and hospitalizations from respiratory diseases. The differential in overall mortality is only marginally statistically significant at the 10% significance level for the 5-year age groups. Aside from mortality rates for infectious and respiratory diseases (which are low and volatile in these populations) the magnitudes of the differentials between the

[9]The closest policy change is the near synchronous end of universal smallpox vaccination in 1983 in West Germany and 1982 in East Germany (Klein et al., 2012; Klein, 2013).

Table 1. Balance tests

	Policy change			
	1974 West ends universal		1990 East ends mandatory	
	Age interval around policy change			
	5 years	10 years	5 years	10 years
	(1)	(2)	(3)	(4)
Panel A. Log all-cause mortality (2016)				
East × Treated	-.170*	-.030	-.096	-.023
	(.094)	(.113)	(.528)	(.326)
Panel B. Log mortality from infectious diseases (2016)				
East × Treated	.800	.996	-.686*	-1.36**
	(.535)	(.708)	(.369)	(.537)
Panel C. Log mortality from respiratory diseases (2016)				
East × Treated	-.131	.636	.039	1.281**
	(.963)	(1.208)	(.428)	(.526)
Panel D. Log hospitalizations (2016)				
East × Treated	.068***	.091***	-.211***	-.191***
	(.013)	(.016)	(.033)	(.024)
Panel E. Log hospitalizations for infectious diseases (2016)				
East × Treated	-.083	-.080	.046	-.144***
	(.067)	(.052)	(.063)	(.046)
Panel F. Log hospitalizations for respiratory diseases (2016)				
East × Treated	-.022	.054	.153*	.052*
	(.048)	(.063)	(.090)	(.030)
Panel G. Unemployment rate (2011 census)				
East × Treated	-.006	-.001	.011	.011
	(.006)	(.003)	(.010)	(.009)
Panel H. labour force participation rate (2011 census)				
East × Treated	-.008	.003	-.069***	.012
	(.012)	(.011)	(.014)	(.013)

Note: The table reports results from a differences in regression discontinuities specification with an interacted local linear RD polynomial and border segment fixed effects. The RD bandwidth is 100 km. All RD-DD results have 276 observations (two for each county). Standard errors clustered on the state (*Bundesländer*) level are reported in parentheses.

measures are modest as well. However, for cohorts born around 1990, the covariates do not appear to be balanced. This may be because we are dealing with a much younger population in which mortality and hospitalization events are even rarer or because political changes in the 1990s had long-run consequences as the fetal origins hypothesis may suggest (Almond and Currie, 2011). For this reason, we focus on the 1974 experiment

and present results using the end of mandatory BCG vaccination in the East in 1990 only as a robustness check.

4.2. Main results

We start by examining the reference RD specification for overall cases and symptomatic cases, across all ages and for the constituent cohorts of the RD-DD design.

Figure 2*a)* presents local linear estimates of the mean of log(1+cases/million) by distance to the border of the former GDR, with positive (negative) distances indicating locations in former East Germany (West Germany). We find that the local linear estimates are discontinuous at zero, falling from west to east. Otherwise, the conditional expectation function exhibits no apparent discontinuities. At a bandwidth of 100 km we observe a drop of 0.83 log points as one crosses the border from west to east (Panel A in Table 3). Hence, there are more than half (56%) as many cases per capita in a former East German county relative to a West German county just across the border. Figure 2*b)* presents similar local linear estimates for the discontinuity in log symptomatic cases. Here too, crossing the border from west to east entails a 0.75 log point (53%) decrease in the number of severe cases per million residents (Panel B in Table 3). These correlations mirror those documented in the recent literature which typically finds a strong association between BCG status and COVID-19 severity in cross-sectional analyses. While remarkable, these apparent jumps are not causal, as many other confounders also change discontinuously at the former border.

Panels *c)* to *f)* of Figure 2 present the corresponding local linear estimates for the 35–44 age group and the 46–55 age group. These correspond to the cohorts born in the 10 years following and the 10 years preceding the 1974 discontinuation of the recommendation to give the BCG vaccine to most newborns. We see that the jump at the border is present for the older group and is quantitatively similar to the jump for the younger group. As the BCG hypothesis would have predicted little if any discontinuity for the older group (as members of that cohort in both the East and the West received the vaccine), this exercise already provides evidence against the BCG hypothesis. Importantly, using symptomatic cases instead of all reported cases as the COVID-19 intensity measure hardly changes the qualitative conclusion that the discontinuity is similar across the two age groups.

Table 2 formalizes the intuition of the cohort-wise figures and presents estimates of the coefficient γ from specification 3.2 for various definitions of the "treated" and "untreated" groups and outcomes. We start with the 10-year cohorts underlying Figure 2, and then consider narrower age groups, all the way to two years before and after the 1974 policy experiment (46–47 versus 43–44 year olds). We use a 100 km bandwidth throughout, which is somewhat narrower than the average optimal bandwidth for both cohorts (Imbens and Kalyanaraman, 2012). If the discontinuity in COVID-19 cases is caused by the direct long-term effect of BCG vaccination, then we would expect that the discontinuity in cases among individuals born just before 1974 should be smaller than the discontinuity among people born just after 1974. The end of the recommendation to vaccinate in the West together with the temporary cessation would have lowered vaccinations, whereas the East continued mandatory vaccination with no change in 1974. As the sign of the discontinuity in COVID-19 intensity across the border is negative, we would expect the

[9]Table S.2. reports results with optimal bandwidths estimated separately for each cohort. Table S.3. and Table S.4. use other fixed bandwidths.

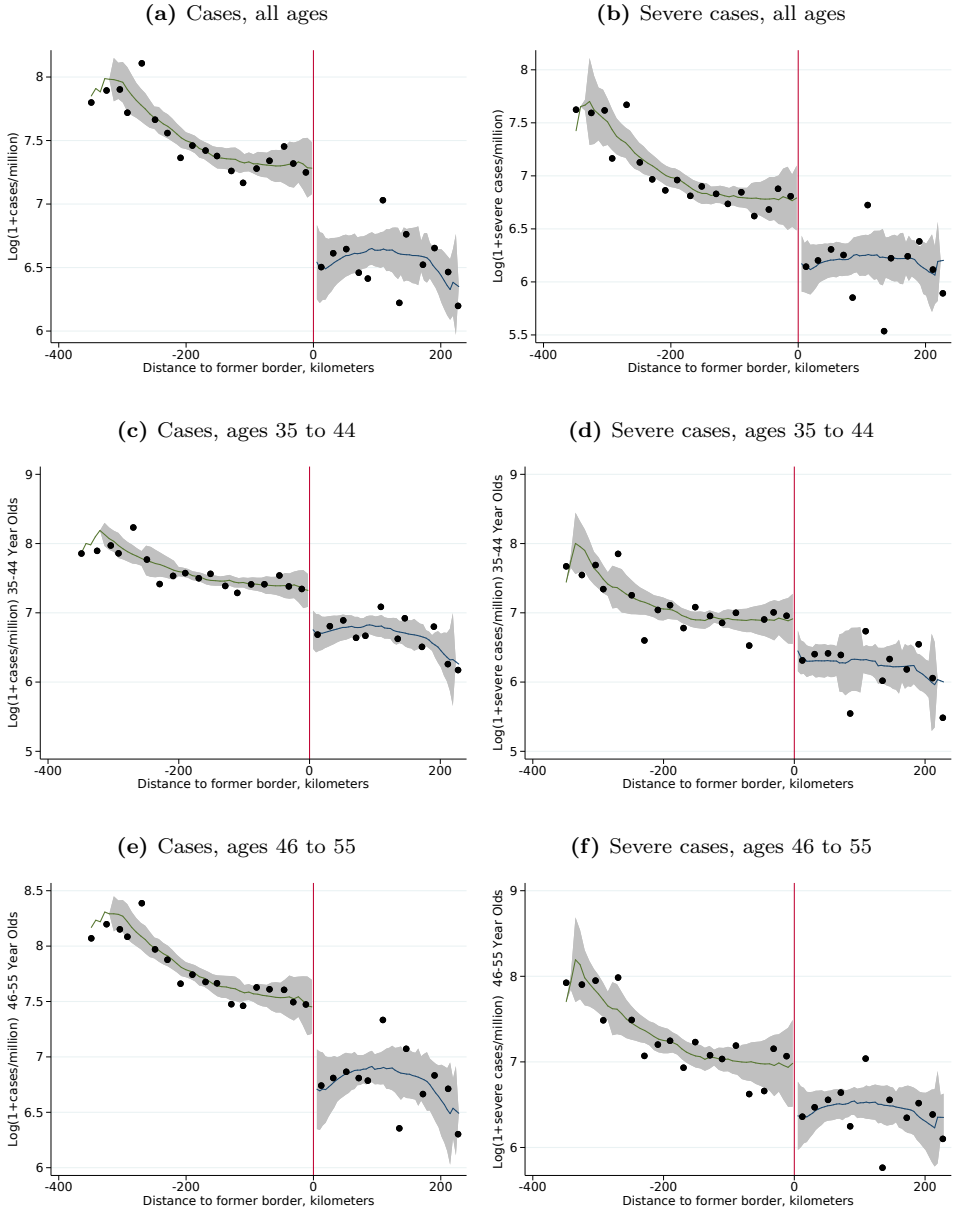

Figure 2. Discontinuities in log(1+cases/million) at former border

coefficient γ, the additional effect on the treated group (those born before 1974), to be positive as it should cancel out the average discontinuity for the population as a whole, or at least counter some portion of the compound treatment at the former border.

Table 2. Differences in regression discontinuities (RD-DD) by age window

	Age interval around policy change				
	10 years	8 years	6 years	4 years	2 years
	(1)	(2)	(3)	(4)	(5)
Panel A. 1974 baseline (cases/million)					
East × Treated	-.114	-.196	-.562**	.343	.066
	(.102)	(.138)	(.223)	(.634)	(.691)
Panel B. 1974 differences-in-differences (cases/million)					
East × Treated	-.090**	-.015	.000	.225	-.160
	(.043)	(.109)	(.151)	(.200)	(.542)
Panel C. 1974 severity measure (symptomatic cases/million)					
East × Treated	-.198	-.264	-.805**	-.347	.147
	(.235)	(.259)	(.324)	(.546)	(.896)
Panel D. 1974 differences-in-differences (symptomatic cases/million)					
East × Treated	.042	.221	.175	.352	-.044
	(.076)	(.182)	(.239)	(.277)	(.559)
Panel E. 1990 East Germany ends mandatory vaccination (cases/million)					
East × Treated	-.603	-1.12*	-1.53*	-2.00*	-2.07***
	(.422)	(.614)	(.847)	(1.082)	(.575)
Panel F. 1990 differences-in-differences (cases/million)					
East × Treated	-.057	-.142	-.238	-.616	-1.01**
	(.113)	(.111)	(.202)	(.420)	(.396)
Panel G. 1990 severity measure (symptomatic cases/million)					
East × Treated	-.370	-.690	-1.37*	-1.24	-1.68*
	(.408)	(.626)	(.720)	(1.063)	(1.023)
Panel H. 1990 differences-in-differences (symptomatic cases/million)					
East × Treated	-.018	-.148	-.204	-.492	-1.22**
	(.167)	(.132)	(.284)	(.533)	(.537)

Note: The table reports results from a differences in regression discontinuities specification with an interacted local linear RD polynomial and border segment fixed effects. The RD bandwidth is 100 km. All RD-DD results have 276 observations and all DD results have 800 observations (two for each county). Standard errors clustered on the state (*Bundesländer*) level are reported in parentheses.

Contrary to the BCG hypothesis, the coefficient γ for overall cases is either negative, or positive but statistically insignificantly different from zero (Panel A of Table 2). When γ is positive, it is, at most, half the magnitude of the baseline discontinuity shown in Figure 2*a)*. With smaller age groups and fewer data, the standard errors on the two positive coefficients become large and contain the magnitude of the population discontinuity

estimate. However, all estimates for wider age groups have modest standard errors that comfortably exclude an estimate of equal or greater magnitude and opposite sign from the population discontinuity. We also present the simple differences-in-differences estimates for the 1974 experiment, in which the whole of West Germany is used as a control for the whole of East Germany (Panel B). We now obtain negative estimates of the coefficient γ for four of the five ways of structuring the treated group, while the lone positive estimate (in column 4) has a large standard error. For all but the narrowest age group definition, we can reject that the policy experiment fully offsets the population discontinuity estimate. In Panels C and D of Table 2 we exploit the unique data reported by the RKI and repeat our analysis using symptomatic cases only. Once again, our estimates of γ are negative, or positive but insignificantly different from zero, inconsistent with the BCG hypothesis.

As a robustness check we examine a second change to the BCG vaccination regime in Germany: the end of mandatory vaccination in the East following reunification in 1990. Now, the "treated" group is the post-1990 cohort because the end of mandatory vaccination in the East reduced the differential in vaccination rates between the East and the West for individuals born just after 1990 relative to individuals born just before. Panels E through H of Table 2 show that the corresponding estimates are consistently negative. If anything, they disagree even more with the BCG hypothesis than the 1974 experiment.

5. ALTERNATIVE EXPLANATIONS FOR THE DISCONTINUITY

If the BCG vaccine does not explain the East-West differential in COVID-19 cases, then what does? We investigate a broad set of additional variables to assess whether they may explain the discontinuity in overall COVID-19 cases and severity across the border. Regardless of the bandwidth used, log population density, log disposable income, the share of the population aged 45–64 and the share older than 64, the date that the first COVID-19 case was recorded, and age-adjusted mortality from all causes, infectious diseases and respiratory diseases all show discontinuities at the former border.[10] This echos the pre-division East-West differences documented elsewhere (e.g. Wolf, 2009; Wyrwich, 2019; Becker et al., 2020) and is precisely why we do not consider the overall discontinuity as a causal estimate, since it may be driven by some of these variables. It is well known that counties just to the east of the border have lower population density, lower consumption, and an older population. We also find that they recorded their first COVID-19 case later and have higher age-adjusted mortality rates than counties just to the west of the border. Nevertheless, these explanations cannot fully account for the East-West difference is COVID-19 intensity. The discontinuity in cases per capita remains after including these variables. The raw correlations within the East and West further suggest that the geography of the early outbreak in Germany was very particular. The virus first spread among affluent and less vulnerable populations.[11]

Mobility is a key driver of the spread of COVID-19 in Germany and across the world (Dehning et al., 2020; Kraemer et al., 2020; Hsiang et al., 2020). Hence, a likely explanation for the discontinuity at the old East Germany border are Germany's regional commuting patterns. Although commuting over long distances is very common in Germany, decades of partition meant that its infrastructure was re-oriented to connect counties

[10] Table S.5. reports the corresponding results.

[11] See RD estimates with controls in Table S.6. and the bivariate OLS regressions in Table S.7.

within the West or East (Santamaria, 2020), with lasting effects on the spatial equilibrium in Germany.[12] If flows from the West usually terminate in the West and flows from the East usually terminate in the East, then cross-border transmission of the virus could be relatively slow.[13] As the epidemic started in the West, it will have had a harder time spreading eastward. The eastward spread was then further interrupted by the nation-wide lock-down on March 22 2020 (Dehning et al., 2020).

We examine the role played by mobility in Table 3. Panels A and B reproduce the benchmark RD specification for a variety of bandwidths centred on 100 km. We also include results for an MSE-optimal bandwidth, as suggested by Imbens and Kalyanaraman (2012), which varies by outcome but is typically close to 100 km. Panel C shows that there is a stark and statistically significant discontinuity in commuter flows across the former border, no matter which bandwidth is considered. Border counties on the eastern side are considerably less likely to receive commuter flows from a West German county than border counties on the western side. Panel D in Table 3 replicates the RD regression in Panel B, but includes the fraction of commuter flows from West Germany as a control. The estimated discontinuities in severe cases fall substantially and are no longer statistically significant for all but the smallest bandwidth. To further account for differences on different sides of the border, Panel F adds a basic set of demographic and income controls to the previous specification. Now all RD estimates are insignificant and of small magnitude. We take this as direct evidence that the discontinuity in commuter flows largely accounts for the discontinuity in severe COVID-19 cases.

To assess the plausibility of this finding, we simulate the epidemic in each county using a canonical SIR model with mobility flows (Bjørnstad and Grenfell, 2008; Wesolowski et al., 2017). This allows us to demonstrate that mobility patterns and the geography of the initial outbreaks can create a counterfactual discontinuity just like the one we observe in the data. In the model, we allow infections to spread along commuting patterns starting from the distribution of COVID-19 cases on February 29 2020 and use the approximate epidemiological characteristics of the outbreak in Germany (e.g., an R_0 of 2.5). We use the observed commuting flows from December 2019 together with county population data to proxy for actual mobility around the time of the outbreak. We simulate the model for 60 periods (days) but stop all commuting flows after 22 days to reflect the nation-wide shutdown. We do not explicitly model social distancing or other local containment measures, apart from the lack of commuting, which implies that the simulation overpredicts total cases.[14]

Panel F in Table 3 shows that the number of cases also discontinuously declines in the simulated data as one crosses from west to east over the former border. With about 0.6 log points, the discontinuity is about 70% of the discontinuity observed for true cases in Panel A. This confirms the results of the RD design with controls for commuter flows and strongly suggests that mobility was a key driver of the geography of the early outbreak. This approach cannot exclude other alternative explanations, and officially registered commuter flows likely do not represent person-to-person movement across Germany perfectly. However, our simulation constructs a situation that shares some essential features

[12]Eder and Halla (2018) document that the sizable East-West population gap in counties close to the former border dates back to 1945–46, when those in the East fled from the Soviet army, rather than from the socialist state created in 1949.

[13]Figure S.1. shows that few counties have flows across the former East-West border of more than one thousand people. The only major destination in former East Germany is Berlin.

[14]Details of the simulation can be found in Online Appendix S2.

Table 3. Regression discontinuities by bandwidth

	The bandwidth is					
	50 km	75 km	100 km	150 km	200 km	IK optimal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Log(1+cases/million)						
East	-.788***	-.708***	-.828***	-.995***	-.899***	-1.03***
	(.230)	(.150)	(.154)	(.196)	(.146)	(.223)
Panel B. Log(1+severe cases/million)						
East	-.743***	-.830***	-.753***	-.805***	-.798***	-.823***
	(.257)	(.193)	(.180)	(.244)	(.159)	(.223)
Panel C. Fraction of incoming flows from West Germany (WGF)						
East	-.354***	-.475***	-.555***	-.669***	-.720***	-.457***
	(.027)	(.026)	(.022)	(.020)	(.021)	(.030)
Panel D. Log(1+severe cases/million), w/ WGF						
East	-.334**	-.312	-.291	-.269	-.255	-.297
	(.155)	(.213)	(.225)	(.248)	(.258)	(.228)
Panel E. Log(1+severe cases/million), w/ demographics, income & WGF						
East	-.262	-.191	-.102	-.140	-.162	-.075
	(.283)	(.278)	(.256)	(.220)	(.200)	(.191)
Panel F. Log(1+simulated cases/million)						
East	-.438*	-.493*	-.588**	-.866***	-.862***	-.493*
	(.227)	(.254)	(.235)	(.245)	(.222)	(.254)
Panel G. Log(1+new cases/million), April 27 to August 30 2020						
East	-.368	-.274	-.237	-.198	-.403**	-.123
	(.314)	(.274)	(.214)	(.236)	(.204)	(.228)
Panel H. Log(1+new severe cases/million), April 27 to August 30 2020						
East	-.199	-.393	-.364	-.266	-.428*	-.384
	(.345)	(.379)	(.297)	(.320)	(.258)	(.356)
Observations	77	106	138	203	287	varying

Note: The table reports results from a differences in regression discontinuities specification with an interacted local linear RD polynomial and border segment fixed effects. The RD bandwidth is indicated in the column header. The IK "optimal" bandwidth uses the plug-in rule from Imbens and Kalyanaraman (2012). Depending on the outcome, IK bandwidths range from 70.82 km to 140.79 km. Demographics controls are the log of population density, the share of the population older than 45 and younger than 65, and the share of the population older 64. Income refers to log disposable income per capita. Standard errors clustered on the state (*Bundesländer*) level are reported in parentheses.

of the data, and that explains the discontinuously lower COVID-19 prevalence across the border from East to West without any reference to the BCG hypothesis.

An implication of the hypothesis that discontinuities in commuter flows rather than of BCG vaccination explain the jump in COVID-19 intensity is that this discontinuity should weaken over time. If BCG vaccination has a protective effect, this effect should manifest itself in both early and later stages of the pandemic. If mobility matters, once

the virus was introduced everywhere, individuals in East Germany should transmit the virus to others in the East in similar rates to those in the West (especially after the lockdown was relaxed in May and June). In Panels G and H of Table 3 we present estimates of the discontinuity at the former border for log total reported cases and log total reported systematic cases per million between April 27 and August 30. The RKI reported 86,350 new cases over this period. Estimates of the discontinuity in new cases are smaller than the estimates for pre-April 27 cases regardless of the choice of bandwidth. For all bandwidths but the largest, the discontinuity in new cases is statistically insignificant, providing additional evidence that mobility flows rather than differences in BCG vaccination appear to generate the aggregate discontinuity in COVID-19 intensity at the former border. This fact together with the pattern in the discontinuities across cohorts, leads us to conclude that differential BCG vaccine coverage does not play an important role in explaining the geography of the outbreak in Germany.

6. CONCLUSION

We use variation in vaccination policy across the former East and West Germany to test whether the BCG vaccine offers protection against COVID-19. We identify patterns in the data that are inconsistent with the hypothesis that the BCG vaccine limits the spread or the severity of COVID-19. Instead, a more plausible explanation for the stark discontinuity in COVID-19 cases observed at the border is the continued presence of limits to mobility between the former East and West. These limits, coupled with the epidemic beginning in the West, decreased the early COVID-19 exposure in the East.

An important limitation that our paper shares with the nonexperimental literature on the BCG hypothesis is that it looks only at whether or not there is a long-run effect of the BCG vaccine (decades after it was administered). We consider this broad version of the hypothesis to be of first-order importance. However, well-documented protective effects of the BCG vaccine regarding other viral infections, such as yellow fever (Arts et al., 2018), arise from a trained response of the innate immune system which typically occurs within one to twelve months after the vaccine has been administered (Kleinnijenhuis et al., 2015; Covián et al., 2019; Chumakov et al., 2020). Hence, we cannot rule out that the vaccine might have a short-run effect which could offer some protection to risk groups.

Our results may be of interest as decisions are made on allocating resources to various ways of fighting COVID-19. The BCG vaccine is already in low supply (Guallar-Garrido and Julián, 2020) and is an important tool in the fight against tuberculosis—a disease which killed 1.5 million people in 2018 alone. Efforts to combat COVID-19 are already interrupting routine vaccination and detection efforts, which is projected to lead to a steep rise in fatalities from tuberculosis and other infectious diseases (Nature, 2020).

ACKNOWLEDGEMENTS

The authors are grateful to Gerda Asmus, Eli Berman, Victor Chernozhukov, Hunter Clark, Gordon Dahl, Beverly Hirtle, Melanie Krause, Keith Meyers, Linda Oz, Devesh Rustagi and the participants at the UCSD applied micro and omni-methods group seminars for valuable comments. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

REFERENCES

Alesina, A. and N. Fuchs-Schündeln (2007). Goodbye Lenin (or not?): The effect of communism on people's preferences. *American Economic Review 97*(4), 1507–1528.

Almond, D. and J. Currie (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives 25*(3), 153–72.

Aronson, N. E., M. Santosham, G. W. Comstock, R. S. Howard, L. H. Moulton, E. R. Rhoades, and L. H. Harrison (2004, 05). Long-term Efficacy of BCG Vaccine in American Indians and Alaska Natives: A 60-Year Follow-up Study. *JAMA 291*(17), 2086–2091.

Arts, R. J., S. J. Moorlag, B. Novakovic, Y. Li, S.-Y. Wang, M. Oosting, V. Kumar, R. J. Xavier, C. Wijmenga, L. A. Joosten, et al. (2018). BCG vaccination protects against experimental viral infection in humans through the induction of cytokines associated with trained immunity. *Cell Host & Microbe 23*(1), 89–100.

Becker, S. O., L. Mergele, and L. Woessmann (2020). The separation and reunification of Germany: Rethinking a natural experiment interpretation of the enduring effects of communism. *Journal of Economic Perspectives 34*(2), 71–143.

Berg, M. K., Q. Yu, C. E. Salvador, I. Melani, and S. Kitayama (2020). Mandated Bacillus Calmette-Guérin (BCG) vaccination predicts flattened curves for the spread of COVID-19. *Science Advances 6*(32).

Bjørnstad, O. N. and B. T. Grenfell (2008). Hazards, spatial transmission and timing of outbreaks in epidemic metapopulations. *Environmental and Ecological Statistics 15*(3), 265–277.

Campa, P. and M. Serafinelli (2019). Politico-economic regimes and attitudes: Female workers under state socialism. *Review of Economics and Statistics 101*(2), 233–248.

Chinazzi, M., J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. Pastore y Piontti, K. Mu, L. Rossi, K. Sun, C. Viboud, X. Xiong, H. Yu, M. E. Halloran, I. M. Longini, and A. Vespignani (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science 368*(6489), 395–400.

Chumakov, K., C. S. Benn, P. Aaby, S. Kottilil, and R. Gallo (2020). Can existing live vaccines prevent COVID-19? *Science 368*(6496), 1187–1188.

Covián, C., A. Fernández-Fierro, A. Retamal-Díaz, F. E. Díaz, A. E. Vasquez, M. K.-L. Lay, C. A. Riedel, P. A. González, S. M. Bueno, and A. M. Kalergis (2019). BCG-induced cross-protection and development of trained immunity: Implication for vaccine design. *Frontiers in Immunology 10*, 1–14.

Curtis, N., A. Sparrow, T. A. Ghebreyesus, and M. G. Netea (2020). Considering BCG vaccination to reduce the impact of COVID-19. *The Lancet 395*, 1545–1546.

Dehning, J., J. Zierenberg, F. P. Spitzner, M. Wibral, J. P. Neto, M. Wilczek, and V. Priesemann (2020). Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science 369*(6500).

Dell, M. (2010). The persistent effects of Peru's mining Mita. *Econometrica 78*(6), 1863–1903.

Deshpande, M. (2016). Does welfare inhibit success? The long-term effects of removing low-income youth from the disability rolls. *American Economic Review 106*(11), 3300–3330.

Eckardt, M., K. Kappner, and N. Wolf (2020). COVID-19 across European regions: the role of border controls. CEPR Discussion Paper No. DP15178, Centre for Economic Policy Research.

Eder, C. and M. Halla (2018). On the origin and composition of the German east-west population gap. IZA Discussion Papers 12031, Institute of Labor Economics (IZA).

Escobar, L. E., A. Molina-Cruz, and C. Barillas-Mury (2020). BCG vaccine protection from severe coronavirus disease 2019 (COVID-19). *Proceedings of the National Academy of Sciences 117*(30), 17720–17726.

Gelman, A. and G. Imbens (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics 37*(3), 447–456.

Genz, H. (1977). Entwicklung der Säuglingstuberkulose in Deutschland im ersten Jahr nach Aussetzen der ungezielten BCG-Impfung. *Deutsche Medizinische Wochenschrift 102*(36), 1271–1273.

Goldfayn-Frank, O. and J. Wohlfart (2019). Expectation formation in a new environment: Evidence from the German reunification. *Journal of Monetary Economics*. Forthcoming.

Guallar-Garrido, S. and E. Julián (2020). Bacillus Calmette-Guérin (BCG) therapy for bladder cancer: an update. *ImmunoTargets and Therapy 9*, 1.

Gursel, M. and I. Gursel (2020). Is global BCG vaccination-induced trained immunity relevant to the progression of SARS-CoV-2 pandemic? *Allergy 75*, 1815–1819.

Hamiel, U., E. Kozer, and I. Youngster (2020). SARS-CoV-2 Rates in BCG-Vaccinated and Unvaccinated Young Adults. *JAMA 323*(22), 2340–2341.

Hauer, J., U. Fischer, F. Auer, and A. Borkhardt (2020). Regional BCG vaccination policy in former East-and West Germany may impact on both severity of SARS-CoV-2 and incidence of childhood leukemia. *Leukemia 34*, 2217–2219.

Hsiang, S., D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, L. Y. Huang, A. Hultgren, E. Krasovich, P. Lau, J. Lee, E. Rolf, J. Tseng, and T. Wu (2020). The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, 1–9.

Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies 79*(3), 933–959.

Klein, S. (2013). *Zusammenhang zwischen Impfungen und Inzidenz und Mortalität von Infektionskrankheiten: Zeitreihenanalysen mit Meldedaten zu Diphtherie, Pertussis, Poliomyelitis und Tetanus von 1892 bis 2011 in Deutschland*. Ph. D. thesis, Freie Universität Berlin.

Klein, S., I. Schöneberg, and G. Krause (2012). Vom Zwang zur Pockenschutzimpfung zum Nationalen Impfplan. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz 55*(11-12), 1512–1523.

Kleinnijenhuis, J., R. van Crevel, and M. G. Netea (2015). Trained immunity: consequences for the heterologous effects of BCG vaccination. *Transactions of The Royal Society of Tropical Medicine and Hygiene 109*(1), 29–35.

Kraemer, M. U. G., C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, L. du Plessis, N. R. Faria, R. Li, W. P. Hanage, J. S. Brownstein, M. Layan, A. Vespignani, H. Tian, C. Dye, O. G. Pybus, and S. V. Scarpino (2020). The effect of human mobility and control measures on the COVID-19 epidemic in china. *Science 368*(6490), 493–497.

Lalive, R. (2008). How do extended benefits affect unemployment duration? A regression discontinuity approach. *Journal of Econometrics 142*(2), 785–806.

Lippmann, Q., A. Georgieff, and C. Senik (2020, 05). Undoing gender with institutions: Lessons from the German division and reunification. *Economic Journal 130*(629), 1445–1470.

Nature (2020). How to stop COVID-19 fuelling a resurgence of AIDS, malaria and tuberculosis. *Nature 584*(7820), 169. Editorial.

O'Neill, L. A. and M. G. Netea (2020). BCG-induced trained immunity: can it offer protection against COVID-19? *Nature Reviews Immunology 20*(6), 335–337.

Rieckmann, A., M. Villumsen, S. Sørup, L. K. Haugaard, H. Ravn, A. Roth, J. L. Baker, C. S. Benn, and P. Aaby (2016). Vaccinations against smallpox and tuberculosis are associated with better long-term survival: a Danish case-cohort study 1971–2010. *International Journal of Epidemiology 46*(2), 695–705.

Santamaria, M. (2020, February). Reshaping Infrastructure: Evidence from the division of Germany. The Warwick Economics Research Paper Series (TWERPS) 1244, University of Warwick, Department of Economics.

Sharma, A., S. K. Sharma, Y. Shi, E. Bucci, E. Carafoli, G. Melino, A. Bhattacherjee, and G. Das (2020). BCG vaccination policy and preventive chloroquine usage: do they have an impact on COVID-19 pandemic? *Cell Death & Disease 11*(7), 1–10.

The RECOVERY Collaborative Group (2020). Dexamethasone in hospitalized patients with covid-19 – preliminary report. *New England Journal of Medicine*.

Wesolowski, A., E. zu Erbach-Schoenberg, A. J. Tatem, C. Lourenço, C. Viboud, V. Charu, N. Eagle, K. Engø-Monsen, T. Qureshi, C. O. Buckee, and C. J. E. Metcalf (2017). Multinational patterns of seasonal asymmetry in human movement influence infectious disease dynamics. *Nature Communications 8*(1), 1–9.

Wolf, N. (2009). Was Germany ever united? Evidence from intra- and international trade, 1885-1933. *Journal of Economic History 69*(3), 846–881.

Wyrwich, M. (2019). Historical and current spatial differences in female labour force participation: Evidence from Germany. *Papers in Regional Science 98*(1), 211–239.

Manuscript submitted to *The Econometrics Journal*, pp. S1–S14.

The spread of COVID-19 and the BCG vaccine: Online supplement

Richard Bluhm[*,†] and Maxim Pinkovskiy[‡]

[*] *Leibniz University Hannover, Institute of Macroeconomics, 30167 Hannover, Germany*
[†] *University of California San Diego, Department of Political Science, La Jolla, CA 92093*
E-mail: bluhm@mak.uni-hannover.de
[‡] *Federal Reserve Bank of New York, Microeconomic Studies Function, New York, NY 10045*
E-mail: maxim.pinkovskiy@ny.frb.org

S1. HISTORY OF BCG VACCINATION IN GERMANY

Even though tuberculosis was widespread among the war-ravaged population, Germany had a non-vaccination policy until the end of World War II and did not join Red Cross-led vaccination campaigns in the early post-war years. This decision was in part due to the "Lübeck vaccination disaster" in which 251 infants were vaccinated with a BCG vaccine contaminated with live tuberculosis bacteria. Almost all children fell ill with tuberculosis and 72 died, leading the Interior Ministry to reject BCG vaccination as unsafe in 1930 (Loddenkemper and Konietzko, 2018).

BCG policies then diverged quickly when the country was divided. In 1953, the German Democratic Republic (GDR) introduced mandatory vaccination for a variety of diseases, including the BCG vaccine against tuberculosis. Enforcement in the GDR was strict. "From 1954 on, school children who had not yet been vaccinated had to present a letter of exemption not only from their parents but also from a physician" (Harsch, 2012, p. 420). Vaccinations substantially outstripped newborns in the early 1950s, suggesting that young adults born before the GDR existed were also vaccinated ex post (Kreuser, 1967). The policy lasted until the collapse of the GDR in 1990.

The Federal Republic of Germany (FRG) only required mandatory vaccination for smallpox from 1949 until the end of 1975. The BCG vaccine was highly recommended but administered on a voluntary basis. In practice, vaccination of newborns was near universal by the mid-1960s. Due to the decentralized nature of the West German health care system, the initial roll out of the vaccination policy varied strongly by state in the 1950s. However, by 1964, practically all newborns in West Germany were BCG vaccinated shortly after birth (Kreuser, 1967). In 1974, the policy was changed to vaccinate only children in risk groups and, in May 1975, the BCG vaccine was temporarily withdrawn from the market in the West, when a new WHO compatible vaccine was discovered to lead to unintended side-effects. Neo-natal vaccination practically ceased for two years and tuberculosis incidence among newborns doubled (Genz, 1977). Voluntary vaccination of risk groups continued thereafter until 1998 (Robert Koch-Institut, 1976, 1998) but vaccination was no longer universal in West Germany from 1975 onward or in reunified Germany after 1990. Some states continued to recommend universal vaccination until 1998 but with considerably lower compliance and confusion among parents about which recommendations apply (Danner and Qast, 1995). Currently, no BCG vaccine is licensed in Germany. We summarize these changes in Table S.1.

Table S.1. Timeline of vaccination policies in both parts of Germany, 1949 until today

Year	West Germany (FRG)	East Germany (GDR)
1949		First BCG vaccinations
1951-52		Extended program with GDR manufactured BCG vaccine
1953	BCG vaccine is licensed	Mandatory vaccination (with refresher), target rate at least 95%
1955	Recommendation to vaccinate all newborn children	
Mid 1960s	Near universal vaccination of newborns	Near universal vaccination of newborns
1974	National recommendation to only vaccinate children in risk groups, some states continue to recommend universal vaccination of newborns	
1975	BCG vaccination temporarily halted for two years	
1983	Further restriction to only those children that have TB in the family	
1988	Vaccine recommended only for children that tested negative for tuberculin and are risk groups	
1990	Reunification, policies of FRG continue	Reunification, policies of FRG apply
1998	Vaccination no longer recommended, risk groups are no longer vaccinated, last states drop recommendation to vaccinate	

Note: Based on Klein et al. (2012), Klein (2013), and various sources cited in the text.

S2. SIR MODEL WITH COMMUTER FLOWS

We simulate a SIR model with multiple locations and exogenous migration flows between locations (Bjørnstad and Grenfell, 2008; Wesolowski et al., 2017). Let

$$\tilde{I}_{i,t} = I_{i,t} + \frac{N_i \sum_j m_{j,i} \frac{I_{j,t}}{N_j}}{N_i + \sum_j m_{j,i}} \tag{S.1}$$

$$S_{i,t+1} = S_{i,t} - \beta S_{i,t} \frac{\tilde{I}_{i,t}}{N_i} \tag{S.2}$$

$$I_{i,t+1} = I_{i,t} + \beta S_{i,t} \frac{\tilde{I}_{i,t}}{N_i} - \gamma I_{i,t} \tag{S.3}$$

$$R_{i,t+1} = R_{i,t} + \gamma I_{i,t} \tag{S.4}$$

where $m_{j,i}$ is the number of commuters going from location j to location i each period and all other variables are as in the classical SIR model (Kermack and McKendrick, 1927). We take German counties as the locations in our models. We assume $\gamma = 1/7$ (because the incubation period is 7 days on average, and much of the transmission is pre-symptomatic) and $R_0 = \beta/\gamma = 2.5$. We assume the initial counts of infected to correspond to the reported cases by county on February 29 2020. We simulate the model for 60 time periods, assuming that after time period 22, all cross-county commuting flows are shut down to simulate measures taken by the German government.

We have tried other parametrizations of the SIR model and we get similar results provided that the epidemic is not allowed to evolve too close to the long-run equilibrium (which, when migration flows are eventually shut down, is the same for each county and hence, would not generate a discontinuity). The continued growth in cases over the summer of 2020 suggests that assuming that the epidemic did not attain long-run equilibrium is reasonable. The magnitude of the case counts resulting from the epidemic vary widely between parametrizations. We view this exercise not as an attempt to model the COVID-19 epidemic in Germany but to provide an illustration that mobility patterns can generate discontinuities in the spread of an epidemic without there being essential discontinuities in the underlying resistance of the population.

(a) Originating in the West **(b)** Originating in the East

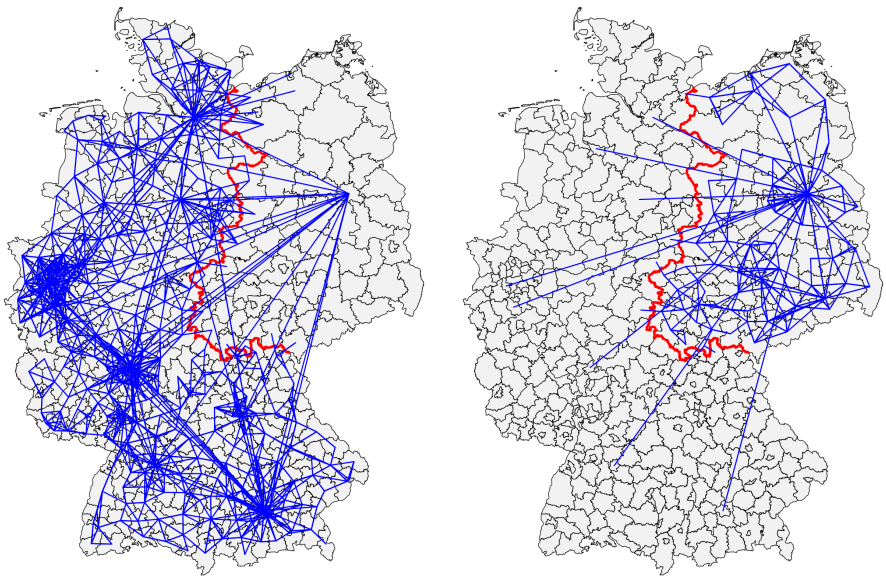

Figure S.1. Major commuting flows (at least 1,000 people) by origin

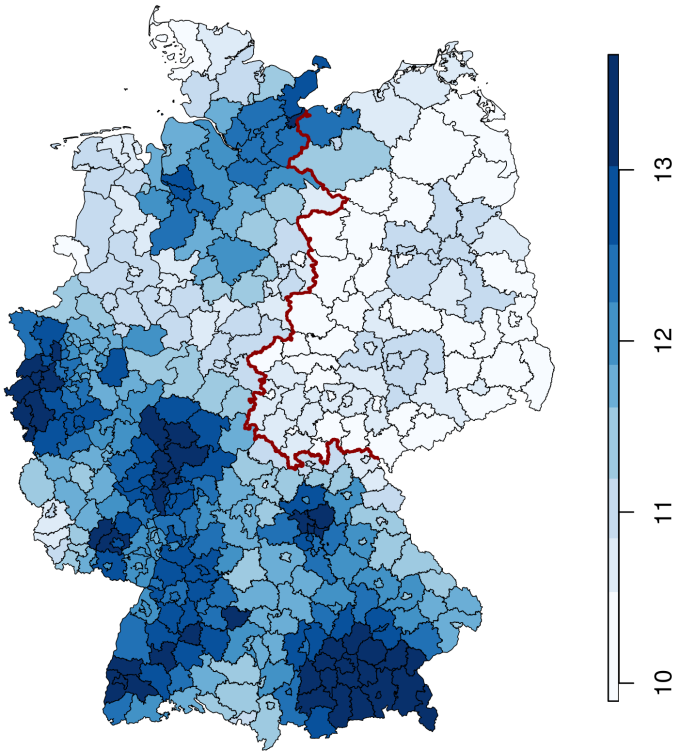

Figure S.2. Spatial distribution of simulated COVID-19 cases in Germany

S3. ADDITIONAL FIGURES

Figure S.3 illustrates the time-series of daily cases in Germany from Jan 27 2020 (week 5) until August 30 2020 (week 35). Week 35 is the last week during which reported cases exceeded 2,000 on at least one day. Coronavirus-related travel restrictions were progressively relaxed starting in May.

Figure S.4 shows the age distribution of cases in the territory of former West and East Germany on April 26, 2020. Cases are all positive COVID-19 tests reported by the RKI. Severe cases display acute respiratory symptoms (including pneumonia) or have died from COVID-19. The distribution is truncated after 79 years, as the RKI reports cases for those aged 80 and above in a single combined category which we omit for display purposes.

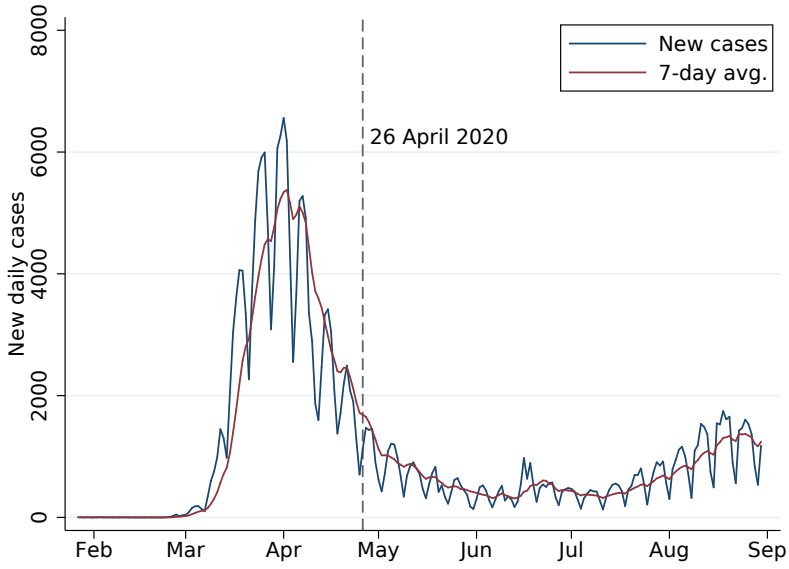

Figure S.3. Time-series of daily cases

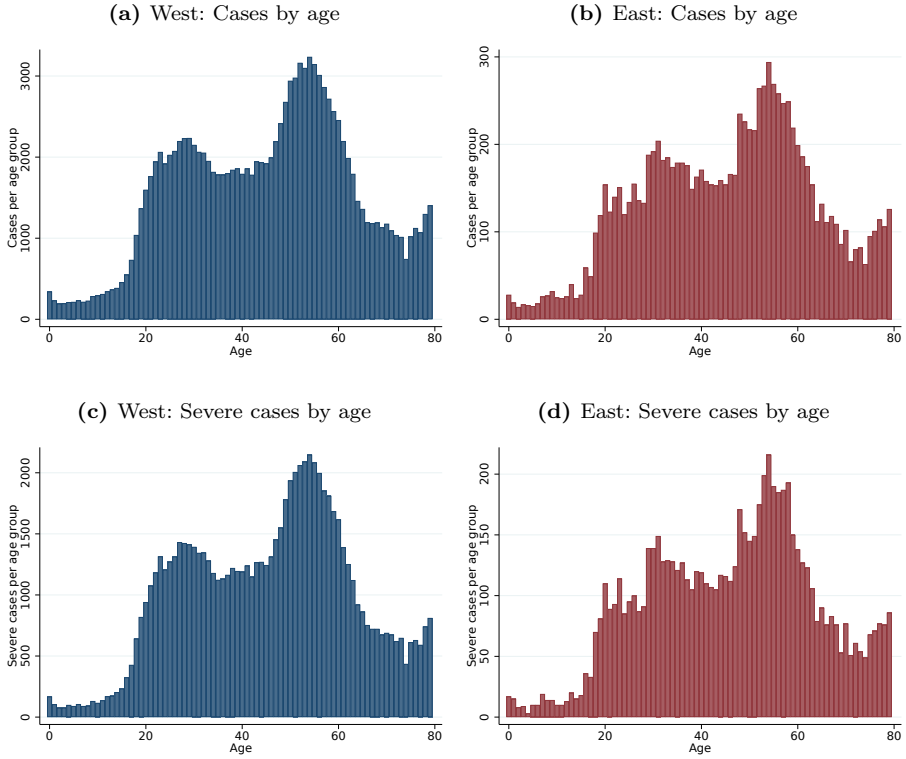

Figure S.4. Distribution of (severe) cases by age on both sides of the former border

S4. ADDITIONAL REGRESSION RESULTS

Table S.2. RD-DD by age window with optimal bandwidths

	Age interval around policy change				
	10 years	8 years	6 years	4 years	2 years
	(1)	(2)	(3)	(4)	(5)
Panel A. 1974 baseline (cases/million)					
East × Treated	-.092	-.196	-.298*	.458	-.037
	(.093)	(.159)	(.161)	(.500)	(.505)
Bandwidth, not treated	116.39	109.36	141.70	125.50	132.67
Bandwidth, treated	112.62	103.56	101.01	131.12	116.71
Observations	313	295	332	357	342
Panel B. 1974 severity measure (symptomatic cases/million)					
East × Treated	.028	-.301	-.501*	.011	-.442
	(.223)	(.211)	(.273)	(.425)	(.661)
Bandwidth, not treated	134.40	108.00	136.35	124.00	131.58
Bandwidth, treated	141.97	130.07	110.55	128.73	139.10
Observations	378	329	340	348	373
Panel C. 1990 East Germany ends mandatory vaccination (cases/million)					
East × Treated	-.531	-.643	-1.00*	-1.37**	-1.30
	(.370)	(.460)	(.541)	(.549)	(.841)
Bandwidth, not treated	126.15	110.23	106.57	85.91	109.74
Bandwidth, treated	117.56	117.12	108.39	116.67	99.16
Observations	333	312	298	281	291
Panel D. 1990 severity measure (symptomatic cases/million)					
East × Treated	-.369	-.547*	-1.41**	-1.15	-1.80
	(.437)	(.323)	(.637)	(.977)	(1.122)
Bandwidth, not treated	166.14	142.09	137.07	188.80	144.20
Bandwidth, treated	131.36	158.66	110.69	110.92	141.31
Observations	411	414	340	422	389

Note: The table reports results from a differences in regression discontinuities specification with an interacted local linear RD polynomial and border segment fixed effects. The RD bandwidth is indicated in each column. The IK "optimal" bandwidth uses the plug-in rule from Imbens and Kalyanaraman (2012) which we compute separately for the treated and untreated cohorts. The number of observations differs per cohort, depending on the bandwidth. Standard errors clustered on the state (*Bundesländer*) level are reported in parentheses.

Table S.3. RD-DD by age window with 50 km bandwidth

	Age interval around policy change				
	10 years	8 years	6 years	4 years	2 years
	(1)	(2)	(3)	(4)	(5)
Panel A. 1974 baseline (cases/million)					
East × Treated	.001	-.211	-.409	1.572***	1.249
	(.125)	(.250)	(.301)	(.464)	(1.321)
Bandwidth, not treated	50	50	50	50	50
Bandwidth, treated	50	50	50	50	50
Observations	154	154	154	154	154
Panel B. 1974 severity measure (symptomatic cases/million)					
East × Treated	-.121	-.479	-.877**	.653	1.191
	(.157)	(.301)	(.361)	(.606)	(1.481)
Bandwidth, not treated	50	50	50	50	50
Bandwidth, treated	50	50	50	50	50
Observations	154	154	154	154	154
Panel C. 1990 East Germany ends mandatory vaccination (cases/million)					
East × Treated	-.298	-.345	-.903*	-.879**	-1.70***
	(.364)	(.393)	(.502)	(.433)	(.603)
Bandwidth, not treated	50	50	50	50	50
Bandwidth, treated	50	50	50	50	50
Observations	154	154	154	154	154
Panel D. 1990 severity measure (symptomatic cases/million)					
East × Treated	.254	.029	-1.02**	-.479	-.183
	(.386)	(.437)	(.448)	(.762)	(1.429)
Bandwidth, not treated	50	50	50	50	50
Bandwidth, treated	50	50	50	50	50
Observations	154	154	154	154	154

Note: The table reports results from a differences in regression discontinuities specification with an interacted local linear RD polynomial and border segment fixed effects. The RD bandwidth is indicated in each column. Standard errors clustered on the state (*Bundesländer*) level are reported in parentheses.

Table S.4. RD-DD by age window with 200 km bandwidth

	Age interval around policy change				
	10 years	8 years	6 years	4 years	2 years
	(1)	(2)	(3)	(4)	(5)
Panel A. 1974 baseline (cases/million)					
East × Treated	-.131	-.202**	-.263**	.304	-.344
	(.082)	(.099)	(.115)	(.348)	(.415)
Bandwidth, not treated	200	200	200	200	200
Bandwidth, treated	200	200	200	200	200
Observations	574	574	574	574	574
Panel B. 1974 severity measure (symptomatic cases/million)					
East × Treated	.042	.031	-.050	.079	-.610
	(.141)	(.191)	(.214)	(.262)	(.452)
Bandwidth, not treated	200	200	200	200	200
Bandwidth, treated	200	200	200	200	200
Observations	574	574	574	574	574
Panel C. 1990 East Germany ends mandatory vaccination (cases/million)					
East × Treated	-.273	-.217	-.469	-.449*	-.754
	(.205)	(.249)	(.299)	(.247)	(.523)
Bandwidth, not treated	200	200	200	200	200
Bandwidth, treated	200	200	200	200	200
Observations	574	574	574	574	574
Panel D. 1990 severity measure (symptomatic cases/million)					
East × Treated	-.061	-.314	-.617	-.713	-1.39
	(.237)	(.246)	(.435)	(.642)	(.875)
Bandwidth, not treated	200	200	200	200	200
Bandwidth, treated	200	200	200	200	200
Observations	574	574	574	574	574

Note: The table reports results from a differences in regression discontinuities specification with an interacted local linear RD polynomial and border segment fixed effects. The RD bandwidth is indicated in each column. Standard errors clustered on the state (*Bundesländer*) level are reported in parentheses.

Table S.5. Discontinuities in other variables

	The dependent variable varies by panel					
	The bandwidth is					
	50 km	75 km	100 km	150 km	200 km	IK optimal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Disposable income per capita						
East	-.084***	-.100***	-.104***	-.128***	-.134***	-.109***
	(.026)	(.014)	(.005)	(.015)	(.014)	(.008)
Panel B. Population density						
East	-.692**	-.584***	-.659***	-.474***	-.102	-.675***
	(.279)	(.097)	(.230)	(.154)	(.193)	(.245)
Panel C. Percent older than 64						
East	2.585***	2.175***	2.424***	3.132***	3.109***	3.232***
	(.889)	(.830)	(.658)	(.616)	(.613)	(.742)
Panel D. Percent older than 45 and younger than 65						
East	2.251***	1.653***	2.105***	1.714***	.992**	2.049***
	(.778)	(.567)	(.546)	(.478)	(.445)	(.468)
Panel E. Days since first case						
East	-1.57	-1.23	-2.16	-3.45***	-3.09***	-2.35*
	(1.928)	(1.541)	(1.460)	(1.248)	(.829)	(1.257)
Panel F. Age-adjusted overall death rate per million						
East	.044***	.040***	.044***	.057***	.042***	.040***
	(.016)	(.010)	(.012)	(.013)	(.015)	(.009)
Panel G. Age-adjusted infectious diseases death rate per million						
East	2.048**	1.723*	2.190**	2.630**	2.688**	2.634**
	(.841)	(.998)	(1.087)	(1.146)	(1.119)	(1.149)
Panel H. Age-adjusted respiratory diseases death rate per million						
East	2.613**	2.191*	2.785**	3.216**	3.310**	3.330**
	(1.109)	(1.258)	(1.372)	(1.457)	(1.437)	(1.454)
Panel I. Age-adjusted overall hospitalization rate per million						
East	.084***	.059*	.071**	.101***	.076**	.071**
	(.019)	(.031)	(.034)	(.034)	(.033)	(.034)
Panel J. Age-adjusted infectious diseases hospitalization rate per million						
East	.109	.121	.112	.151**	.126**	.106
	(.070)	(.091)	(.081)	(.073)	(.063)	(.090)
Panel K. Age-adjusted respiratory diseases hospitalization rate per million						
East	.062	.032	.035	.092*	.063	.038
	(.039)	(.059)	(.047)	(.052)	(.046)	(.054)
Observations	77	106	138	203	287	varying

Note: The table reports results from a regression discontinuity specification with an interacted local linear RD polynomial and border segment fixed effects. The RD bandwidth is indicated in the column header. The IK "optimal" bandwidth uses the plug-in rule from Imbens and Kalyanaraman (2012). Disposable income per capita, population density and all age-adjusted rates are measured in logs. Standard errors clustered on the state (*Bundesländer*) level are reported in parentheses.

Table S.6. Regression discontinuity results with controls

	The bandwidth is					
	50 km	75 km	100 km	150 km	200 km	IK optimal
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. No Controls						
East	-.788***	-.708***	-.828***	-.995***	-.899***	-1.03***
	(.230)	(.150)	(.154)	(.196)	(.146)	(.223)
Panel B. Population density						
East	-.747***	-.669***	-.802***	-.984***	-.894***	-.994***
	(.214)	(.157)	(.184)	(.203)	(.146)	(.237)
Panel C. Disposable income p.c.						
East	-.673***	-.605***	-.662***	-.826***	-.699***	-.863***
	(.246)	(.189)	(.170)	(.199)	(.180)	(.202)
Panel D. Disposable income p.c. and population density						
East	-.599**	-.534**	-.598**	-.789***	-.669***	-.796***
	(.237)	(.234)	(.242)	(.236)	(.195)	(.249)
Panel E. Percent aged 45-64 and percent older than 64						
East	-.588***	-.531***	-.716***	-.811***	-.705***	-.831***
	(.153)	(.109)	(.157)	(.174)	(.171)	(.187)
Panel F. Controls from panels D and E						
East	-.594***	-.531***	-.718***	-.814***	-.711***	-.835***
	(.163)	(.110)	(.154)	(.165)	(.155)	(.185)
Panel G. Days since first case						
East	-.780***	-.698***	-.812***	-.994***	-.871***	-1.02***
	(.230)	(.152)	(.155)	(.203)	(.153)	(.234)
Panel H. Controls from panels F and G						
East	-.594***	-.530***	-.715***	-.823***	-.701***	-.845***
	(.165)	(.113)	(.156)	(.175)	(.161)	(.193)
Observations	77	106	138	203	287	varying

Note: The table reports results from a differences in regression discontinuities specification with an interacted local linear RD polynomial and border segment fixed effects. The RD bandwidth is indicated in the column header. The IK "optimal" bandwidth uses the plug-in rule from Imbens and Kalyanaraman (2012). Disposable income per capita and population density are measured in logs. Standard errors clustered on the state (*Bundesländer*) level are reported in parentheses.

Table S.7. Bivariate OLS regressions of log(1+cases/million) on control variables

	West (1)	East (2)	All (3)
Disposable income p.c.	2.355***	3.502*	3.374***
	(.737)	(2.014)	(.610)
Population density	-.030	.154***	.085
	(.082)	(.023)	(.074)
Percent older than 64	-.089***	-.017	-.122***
	(.031)	(.061)	(.020)
Percent older than 45 and younger than 65	-.012	-.046***	-.063***
	(.018)	(.006)	(.021)
Age-adj. overall death rate	-3.17**	-3.50***	-4.01***
	(1.305)	(.967)	(1.020)
Age-adj. infectious diseases death rate	-.120***	-.078	-.150***
	(.031)	(.247)	(.032)
Age-adj. respiratory diseases death rate	-.103***	-.210	-.127***
	(.025)	(.412)	(.022)
Age-adj. hospitalization rate	-.617	-1.98**	-1.48**
	(.641)	(.810)	(.716)
Age-adj. infectious diseases hospitalization rate	-.014	-1.32***	-.052
	(.034)	(.298)	(.038)
Age-adj. respiratory diseases hospitalization rate	-.034	-1.36***	-.062**
	(.024)	(.133)	(.030)
Observations	324	76	400

Note: The table reports results from bivariate ordinary least squares regressions for the samples indicated in the column headers. Standard errors clustered on the state (*Bundesländer*) level are reported in parentheses.

REFERENCES

Bjørnstad, O. N. and B. T. Grenfell (2008). Hazards, spatial transmission and timing of outbreaks in epidemic metapopulations. *Environmental and Ecological Statistics 15*(3), 265–277.

Danner, K. and U. Qast (1995). Neuere Impfempfehlungen — Impfschemata. In H. Schneemann, G. Wurm, R. Batty, A. Berg, J. Cope, K. Danner, S. Dhillon, P. Elias, W. Feldheim, R. Großklaus, R. Grüttner, G. Gündermann, H. Haindl, H.-J. Hapke, H. Hehenberger, P. E. Heide, G. Heil, J. Keul, R. Kilian, U. Kirschner, A. Klaus, F. Klingauf, A. Kostrewski, I. Krämer, A. Liersch, N. P. Lüpke, G. Mould, H. Müller, A. Obermayer, D. Paar, U. Quast, A. Quilling, A. Rabitz, F. v. Rheinbaben, R. S. Roß, J. E. Schmitz, H. Schütz, E. Telser, E. J. Verspohl, J. Wachsmuth, U. Wahrburg, C. Ward, E. Wisker, and G. Wurm (Eds.), *Hagers Handbuch der Pharmazeutischen Praxis: Waren und Dienste Folgeband 1*, pp. 515–529. Berlin, Heidelberg: Springer.

Genz, H. (1977). Entwicklung der Säuglingstuberkulose in Deutschland im ersten Jahr nach Aussetzen der ungezielten BCG-Impfung. *Deutsche Medizinische Wochenschrift 102*(36), 1271–1273.

Harsch, D. (2012). Medicalized social hygiene? Tuberculosis policy in the German Democratic Republic. *Bulletin of the History of Medicine 86*(3), 394–423.

Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies 79*(3), 933–959.

Kermack, W. O. and A. G. McKendrick (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A 115*(772), 700–721.

Klein, S. (2013). *Zusammenhang zwischen Impfungen und Inzidenz und Mortalität von Infektionskrankheiten: Zeitreihenanalysen mit Meldedaten zu Diphtherie, Pertussis, Poliomyelitis und Tetanus von 1892 bis 2011 in Deutschland.* Ph. D. thesis, Freie Universität Berlin.

Klein, S., I. Schöneberg, and G. Krause (2012). Vom Zwang zur Pockenschutzimpfung zum Nationalen Impfplan. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz 55*(11-12), 1512–1523.

Kreuser, F. (1967). Stand der Tuberkulose-Bekämpfung im Bundesgebiet, in West-Berlin und in Mitteldeutschland. In F. Kreuser (Ed.), *Tuberkulose-Jahrbuch 1964/65 — Band 14*, pp. 33–147. Berlin, Heidelberg: Springer.

Loddenkemper, R. and N. Konietzko (2018). Tuberculosis in Germany before, during and after World War II. In *Tuberculosis and War*, Volume 43, pp. 64–85. Karger Publishers.

Robert Koch-Institut (1976). Impfempfehlungen der Ständigen Impfkommission. *Bundesgesundheitsblatt 19*, 270–273.

Robert Koch-Institut (1998). Impfempfehlungen der Ständigen Impfkommission. *Epidemiologisches Bulletin 15*, 101–114.

Wesolowski, A., E. zu Erbach-Schoenberg, A. J. Tatem, C. Lourenço, C. Viboud, V. Charu, N. Eagle, K. Engø-Monsen, T. Qureshi, C. O. Buckee, and C. J. E. Metcalf (2017). Multinational patterns of seasonal asymmetry in human movement influence infectious disease dynamics. *Nature Communications 8*(1), 1–9.