Persistence and Historical Evidence: The Example of the Rise of the Nazi Party

Online Appendix, February 2024

Contents

This appendix provides further detail on the econometric and data concerns raised in the text. The appendix has three sections. Section A considers robustness and related econometric matters for all six outcomes discussed in PP Table VI, as well as the results for the first principle components of those six outcomes PP reports in its Tab VII, Column (1). Section B discusses the primary econometric results for BF. Section B also includes a formal model that elucidates the ambiguity in BF's claims about "social capital." Section C considers the coding and selection of data for each article.

We offer a general remark about sample sizes. Both articles estimate regressions that rely on different sub-samples. PP, for example, starts with 1428 communities, but most empirical exercises rely on 325 or fewer observations for places the authors think had a Jewish community in 1349. In other places the sample size reflects a definition; BF has a full sample of 229 cities, but for some purposes (for example, their "stable" states) the relevant sub-sample is smaller. Sometimes an estimation sub-sample is reduced because of missing observations for either a regressor or the dependent variable. Our replication efforts display the same variation in sample sizes. When we refer to a descriptive statistic, we are careful to describe the sub-sample to which it pertains. All results here use the sub-sample relevant to the question at hand.

Throughout we use the replication data provided online. Unless otherwise noted, we use the variable definitions and the replication code in the articles.[1] The general notes at the end of this appendix explain some abbreviations. Models labelled "replication" in our tables are precise copies of those found in PP's and BF's text and appendix, except for the cases we note where the original article apparently includes a transcription error.

A.1 PP 1920s Pogroms and the *Reichskristallnacht*

[1] The replication data and code were downloaded from: https://www.anderson.ucla.edu/faculty_pages/nico.v/Research/publications.htm.

The first regression in PP Table VI seeks to explain what the authors call pogroms in the 1920s.[2] Their regression implies that places with a Black Death pogrom were 6 percent more likely to see anti-Semitic violence in the 1920s. PP Table V cross-tabulates this dependent variable (POG20s) with POG1349. The dependent variable, POG20s, equals one in only 20 of the 320 observations in the regression. In only a single place is POG20s one and POG1349 zero.[3]

Our Table A1.1 first replicates the PP Column (1), Table VI, using OLS as in PP, and then estimates the same model as a binary logit and a binary probit (our Columns (2) and (3)). The two limited-dependent variable models imply the same substantive result but estimate the coefficient on POG1349 less precisely. This contradicts the received wisdom that the OLS (linear probability) model is substantively no different from binary logit or probit. The reason is easy to see: the dependent variable pertains to a "tail" event, and these distributions differ most in their tails. If we drop the single observation that had a 1920s pogrom and where POG1349 is coded as zero, the model cannot be estimated as logit or probit because POG1349 perfectly classifies the dependent variable.

Table A1.1 Column (4) demonstrates that the results PP report are not robust to the omission of Bavaria. Below (Section A.5) we consider the role of state fixed effects in these specifications. There we show that the relationship between the pogrom proxy and the outcome variable is different in Bavaria. The single place that had a 1920s pogrom but not a medieval pogrom is in Silesia, so every Bavarian observation that is coded "1" for a 1920s pogrom also had a medieval pogrom. This is not the same as the outcome variable having a different conditional mean in Bavaria; rather, the relationship between the outcome and the pogrom proxy is different in Bavaria. A single dummy for Bavaria could not contend with this problem. This result will be echoed in the other econometric specifications discussed in this

[2] The coding for POG20s relies on Alicke. See PP, p. 1352. This appendix section has seven sub-sections. The first three discuss the results PP present in their Table VI. Section A.4 focuses on the summary measure PP presents in its Table VII, the first principal component of the six outcome variables in PP Table VI. Section A.5 considers the role of provincial-level fixed effects in the PP models, while A.6 discusses placebo tests. The final sub-section discusses PP's matching exercises.

[3] The town is Beuthen, PP town number 234. The sources are especially unclear in this case. Avneri (I, p. 79) states that there was a Black Death pogrom in a place called "Bytom," "mit dem vielleicht unser Ort gemeint ist." Alicke is more certain. The town is not part of the Finley/Koyama dataset, which suggests that they thought the sources were too weak to include it. See Appendix C.1.

appendix; see especially Section A.5. For Column (4), dropping Bavaria reduces the point estimate for the pogrom variable by about one-third and increases the standard error by about the same amount. The regression has 320 observations, 67 of which are in Bavaria. Twelve percent of those Bavaria places experienced a 1920s pogrom, a figure that rises to 17 percent if they also had a Black Death pogrom.

We then show in Columns (5) and (6) that this Bavaria problem does not affect the results for the 1938 *Reichskristallnacht* attacks. Our results can be compared to PP Table VI, Column (6). The difference is instructive. Historians doubt that the "night of broken glass" attacks reflect the local populace's anti-Semitism. National-level government officials set the attacks in motion; the SS and Stormtroopers carried them out and committed most of the violence and pillaging, which ended up being unpopular:

> "Der Novemberpogrom underschied sich von den klassischen antisemitischen Ausbrüchen, die als Pogrom bezeichnet werden, durch den Tatbestand, daß er von Mitgliedern der Regierung ausgelöst und durch den Parteiappparat systematisch durchgeführt wurde, so daß regionale Unterschiede eigentlich nur bezüglich des Zeitpunkts, an dem die Übergriffe einsetzten und an dem sie aufhörten, existierten. Es fehlte jede Volksempörung auch im Ansatz; an deren Stelle trat eine kaltblütige Entfesselung der niedersten Instinkte bei den eingestezten SA- und SS-Leuten (Mommsen 1988, p. 604)."[4]

When Stormtroopers pillaged several local Jewish stores "as part of the nationally organized" *Reichskristallnacht* in the pro-Nazi town of Northeim, the locals' reaction "to this (as was the case all over Germany) was so openly negative that it was the last public anti-Semitic incident in the town" (Allen 2014, pp. 290, 372, note 39).

Dropping Bavaria affects the results for the 1920s attacks, but not for those organized at the national level. The difference between the 1920s pogroms and the *Reichskristallnacht* undermines PP's

[4] "The November pogrom differed from the classic anti-Semitic outbreaks known as pogroms in that it was triggered by members of the government and systematically carried out by the party apparatus. There were regional differences only in the timing of the onset and end of the attacks. There was no popular outrage, even in the beginning; in its place there was a cold-blooded unleashing of the lowest instincts among the SA and SS."

argument. The 1920 attacks were driven by locals, and there, the result is quite fragile. The medieval pogrom proxy supposedly reflects long-standing differences in local attitudes towards the Jews, but only affects the attacks that were organized by the national government and unpopular with the general public.

A.2 PP Deportations

PP uses as one indicator of modern anti-Semitism the number of Jews deported from a community between 1933 and 1944. In PP Table VI, Column (4), the dependent variable is the count of such deportations (deptotal), and the regressors are POG1349 and several controls. POG1349 has a positive and statistically significant effect; the point estimate implies that a community that experienced a Black Death pogrom increased deportations by 30 Jews, compared to a sample mean of 197.

These results do not withstand scrutiny, however. There are three issues. First, as we demonstrated in the text, PP uses an additional and superfluous control for this model only. PP does not justify why this control appears in only one model, and this particular nonlinear transformation of the control, as we show in the text, accounts for the results. Here we present alternative specifications that reinforce this point. Second, the total number of deportees, the dependent variable, has a small number of extremely large values. This handful of observations appears to be responsible for the results PP stresses. This problem parallels the outliers issue for voting that we document in the text. Third, PP rely on the Poisson model for this exercise. This model makes a strong functional-form assumption that the data reject. Reasonable alternative models reject PP's view. (PP calls them "alternate specifications" rather than "robustness checks.") When we construct the corresponding OLS models, we find that the OLS models contradict the Poisson.

Table A2.1 reports descriptive statistics for the deportations variable. PP's authors say they use the Poisson model because the deportations variable is highly skewed (PP p. 1368). Indeed it is; the mean is 213, the skewness is 7.89, and the variance is 760,000. The Poisson distribution assumes that the mean equals the variance, and clearly it does not in this case. The count-data literature calls this phenomenon "overdispersion;" the sample dispersion is greater than the Poisson distribution assumes.

The first column of Table A2.2 replicates PP's Poisson regression (Table VI, Column (4)).[5] As we noted in the text, this model includes a control that is not used in the other regressions reported in PP Table VI: the (log of the) number of Jews present in 1933. PP's authors say that they "add the size of the Jewish population in 1933 to our regular set of controls" (p. 1368), but they do not explain why.[6] The control is puzzling. This specification controls for the log total population, the percent Jewish in 1933, and the percent Protestant in 1925. The "left-out" category is Catholics and those of other religions and the nonreligious, plus changes in the percent Protestant between 1925 and 1933. (They do not explain why they control for Protestants in 1925 rather than 1933; the data are available from the same source as they use for Jews.) The combination of the population and the percentage Jewish control already pins down the number of Jews in the community in 1933. The "log Jews" control just adds a particular nonlinearity in the number of Jews.

Columns (2)-(4) in Table A2.2 explore the implications of that modelling choice. Column (2) drops the additional control. (This is the same result we report in text Table 1, Column (6)). Note that the pogrom variable is no longer significant. The last two columns in Table A2.2 consider alternative ways to control for the number of Jews with an additional variable. Column (3) enters the number of Jews linearly, while (4) adds the square of the percentage Jewish to consider a different kind of nonlinearity from the logarithm PP uses. None of these alternatives yields a statistically significant estimate for POG1349; only models with the additional control in PP's particular logarithmic form yield the result PP stresses.

The PP result also reflects a handful of extremely large values of that "ln Jews" control. The median of the log number of Jews in 1933 is 4.6, and the 99th percentile is 9.6. Table A2.3 reports the

[5] An error in PP's code affects this regression and the related exercises. Their sources lack information on the number of Jewish residents in 1933 for some observations. PP imputes this figure from earlier data, but do so *after* constructing the control called "logjews33" in the code. There are thus 22 observations where the number of Jews in 1933 is missing but the percentage of Jews in 1933 is known. Since these two variables rely on the same information, they should both either be missing or not missing. When we correct the code and re-estimate the PP specification, the results do not change in any meaningful way and are not reported here. In the rest of this discussion, we use the data as it appears in the PP article, to assure apples-to-apples comparisons.

[6] PP mentions a robustness check using the number of Jews in 1939, but this does not address our point.

implications of dropping the observations corresponding to the largest five values of this control. (The PP result is still the first column of Table A2.2, for comparison.) Dropping just two observations (Column (2)) increases the point estimate for the Pogrom variable by about 20 percent, but does not affect the standard error. Dropping three or four observations (Columns (4) or (5)) increases the point estimate by 50 percent, and nearly doubles the standard error. Both the point estimates and the standard errors are sensitive to the inclusion of a small number of observations. A few large values of the "ln Jews" variable drives the result.

A few large values of the dependent variable also drive the results in this model. Much of the skewness in the dependent variable reflects these extremely large values. If we drop just the five largest values of the dependent variable, the skewness falls to 4.6 and the variance to 82,213. Table A2.4 first replicates the PP specification for reference. Our Column (2) re-estimates the model, dropping the 1 percent of the sample corresponding to the largest values of the dependent variable – that is, two observations. Column (3) does the same for the largest 2 percent of the sample (five observations). Dropping just those five observations increases the point estimate by one-third but the estimate is now insignificant.

Table A2.5 considers three additional, related issues. First, the Poisson model can consistently estimate the conditional mean, even with the over-dispersion we document, so long as the "link" function is correctly specified.[7] Second, in that case, however, the maximum-likelihood standard errors will not be correct; the over-dispersion creates a problem analogous to heteroskedasticity in a linear model. PP uses "robust" standard errors. We show that bootstrap standard errors are much larger. The sample appears to be too small for the asymptotic approximation of the "robust" standard errors to work well. Third, over-dispersion implies that distributions other than the Poisson can yield more efficient estimates of the conditional mean (see Cameron and Trivedi 1998, Chapter 3). The Poisson is a special case of those other

[7] In the PP model, the log of the Poisson parameter $\lambda = \ln(X\beta)$, where the Xs are the regressors, β is a vector of parameters to estimate, and λ is the Poisson parameter (both mean and variance). The "link" function is thus the natural log. See Mroz (2012) for a discussion of the sensitivity of the Poisson model in general.

distributions and the nesting makes it straightforward to test against alternatives. We report estimates from the most common generalization, the negative binomial, and find that the data always reject the Poisson special case and imply different results for the pogrom proxy.

Table A2.5, Column (2) reports PP's Poisson specification with bootstrapped standard errors.[8] The point estimates are the same, of course, but the bootstrap standard error for POG1349 is about 50 percent larger than the "robust" standard error, and the Pogrom proxy is not statistically significant. This difference between "robust" and "bootstrap" standard errors reflects the extreme problem the over-dispersion creates in this case. Columns (3)-(6) consider alternative functional forms that nest the Poisson as a special case. Our point is not that the negative binomial is "correct," just that it is more general than PP's Poisson model and implies different results. Each of these NB models estimates an additional parameter that, if zero, collapses to the Poisson. The Poisson model assumes that the mean (λ) is also the variance. Using A. Colin Cameron and Pravin K. Trivedi's nomenclature, negative binomial version NB1 assumes that the variance is $(1+\delta)\lambda$. The additional parameter δ is estimated as part of the model; if $\delta=0$, the NB1 model collapses to Poisson. A different version of the negative binomial, NB2, assumes that the variance is $\lambda+\alpha\lambda^2$. In this case, when $\alpha = 0$, the distribution collapses to the Poisson. To be clear, both NB1 and NB2 make parametric assumptions. They are more flexible than the Poisson, and are a simple diagnostic for the Poisson, but one could do even better. Both the NB1 and the NB2 models reject the Poisson specification.[9] In the NB2 model, the pogrom proxy never has a statistically significant estimate. The Aikake Information Criterion suggests for this case that the NB1 model fits better than NB2, but with bootstrap standard errors, the pogrom proxy in the NB1 specification is not significant, Thus the result that PP reports for deportations reflects the Poisson specifications' inability to model this particular data.[10]

[8] All bootstrap standard errors in this appendix were estimated using 200 replications. Where the original article adjusts for clustering, so do our bootstrap standard errors.
[9] This is, we reject the null hypothesis that the ancillary parameters (δ and α) are zero.
[10] A more general model than the negative binomial, or a non-parametric approach, might imply something different from our Table A2.5. Our point is that PP's result does not survive standard specification checks.

Table A2.6 turns to alternatives to the "counts" approach discussed thus far. Footnote 43 of PP points the reader to an exercise in PP's Appendix. There, Table A.10 reports an OLS model in which the dependent variable is the log of (one plus) the number of deportees, but that specification is not really a robustness check as it does not correspond to the models in their text Table VI. The appendix version lacks two controls that are standard in PP Text Table VI and other exercises: controls for population size and percentage Protestant. Our Table A2.6 reports one-for-one comparisons to the specifications reported in PP Table VI. Column (1) replicates the model reported in the PP appendix. Column (2) uses the same controls as PP use in their Table VI, Column (4). This OLS model explains 78 percent of the variation in the dependent variable, but the pogrom proxy is not significant. Thus a true robustness check that parallels the Poisson specification contradicts their results. Column (3) in Table A2.6 returns to the control for the log of the number of Jews that appears in the PP specification. Dropping that control in this OLS model actually makes the pogrom proxy marginally significant, although that model also has a lower adjusted R^2.

PP Appendix Table A12 reports a different approach: an OLS model in which the dependent variable is the proportion of all Jews deported from a city. This variable is the ratio of their usual dependent variable to the control for the number of Jews in the Poisson models. Our Table A2.7, Column (1) replicates this specification, subject to some transcription errors.[11] PP weights this model by the population of the city in 1933; it does not weight the other models. "An observation where 2,000 out of 10,000 Jews were deported has more informational content than 2 out of 10" (PP Appendix, Note 5). This is a reasonable argument, but applies with equal force to the voting models, which PP does not weight.[12] Using this logic, the correct weights for the deportations model would be the *Jewish* population, however,

[11] PP Appendix Table A12, Column (1) reports the point estimate for POG1349 as 1.09 instead of 10.09, and the table swaps the estimates for the percent Jewish and percent Protestant. Our table correctly reports the regression results.

[12] In the estimate sub-sample for the 1928 Nazi vote (PP Table VI, Column (2), the mean total number of votes cast in a district was 22,453. Twenty-five percent of those districts had 2238 or fewer votes. The mean number of votes cast for Nazis was 778, but in one-quarter of the districts, the Nazis received fewer than 33 votes. If we weight the model reported in PP Table VI, Column (2) by the total number of votes cast, the pogrom proxy remains significant. The point-estimate falls to .011 and the SE also falls slightly, to .005.

not the entire population. Columns (3) and (4) in Table A2.7 weight by the Jewish population. The pogrom proxy remains marginally significant but the point-estimates fall by one-half. Using either set of weights, the “ln Jews” variable is not crucial to the PP result in the way it is in the Poisson models (Columns (2) and (4)). Columns (5) and (6) show that without weights, POG1349 is essentially zero.

Seven of the observations for the proportions deported exceed one; that is, more Jews were deported than lived there in 1933. This is not logically impossible on its face, although one wonders how it came about. The OLS models using the ln(Deportees + 1) as their dependent variable (as in our Table A2.6) are *not* robust to the exclusion of these seven observations; the point estimate for the pogrom proxy is huge (.78) but t-ratio is 1.27 (not reported in tables). The models for which the dependent variable is the proportions deported (as in our Table A2.7) are, however, robust to the exclusion of these seven observations. These seven places were all relatively small (the largest, Kreis Beckum, had a population of 11,500 in 1933; the average population of the seven is 4,900). All had small Jewish populations (Beckum had 86 Jews; the average Jewish population for the seven places is 29).

A.3 PP Letters to *Der Stürmer*

PP also uses a Poisson model for the main results pertaining to anti-Semitic letters that appeared in *Der Stürmer*. Some of the statistical issues discussed in connection with the deportees estimates also affect this model. Some do not. Refer back to Table A2.1, which also reports descriptive statistics for the letters. The ratio of the variance to the mean is much lower for this variable than for the deportees variable, but the variance still greatly exceeds the mean. Thus the Poisson model seems potentially t problematic. In levels, the variable is almost as skewed as the number of deportees.

The letters variable again has a long right-hand tail that the Poisson model does not capture. The results PP reports are sensitive to the exclusion of these observations. Table A3.1, Column (1) reproduces PP Table VI, Column (5). In Columns (2) and (3), we remove the observations corresponding to the largest 1 and 2 percent of values for the dependent variable. Even removing 2 percent of the sample (six observations) reduces the point estimate for the pogrom variable by one-third and leaves it marginally significant. If we drop three percent corresponding to the largest values of letters, the point estimate continues to fall and the standard errors grow. The results in PP are thus driven by as few as six places with unusually large numbers of letters to *Der Stürmer*.

Table A3.2 turns to the standard error and functional form issues. Column (2) re-estimates PP's Poisson model using bootstrap standard errors. In this case, the bootstrap standard errors are slightly larger than the "robust" errors, but the result PP stresses survives. Table A3.2 also reports the two different NB models with both robust and bootstrap standard errors. Both models reject the Poisson specification and the BIC and AIC imply that the NB2 fits better than NB1.The pogrom proxy is not significant in the NB2 models.

Table A3.3 considers the OLS models reported in PP's Appendix Table A.11. The models that PP reports here do not include the full set of controls that appear in their text regressions. Our Table A3.3 uses the full set of controls that appear in PP Table VI. Column (1) is the model PP reports. Column (2) shows that POG1349 is essentially zero in a model that explains 55 percent of the variation in the dependent variable.

Table A3.4 considers an OLS model PP reports in Appendix Table A12, Column (4). The dependent variable here is the number of letters per 10,000 local population. Column (1) replicates the PP result using the total 1933 population as the weight.[13] The pogrom proxy here is indeed significant, but that result depends entirely on the weights. Column (2) estimates the same model without the weights, showing that in the unweighted model, POG1349 has no effect. The weights used in this exercise are most unusual. The 1933 population is the denominator in the dependent variable and in the percentage Jewish and Protestant; it is also, in logged form, a control, and in (1), a weight.[14]

[13] PP Appendix Table 12, Column (4) transposed the point-estimate and standard errors for percentage Jewish and percentage population variables. Our table is correct.

[14] In cases like this, where some values are potentially zero, some researchers now employ an inverse hyperbolic sine transformation instead of the natural log, as the former is defined for zeros. We used this transformation to re-estimate the models in Table A2.6 and A3.3; the results are similar to those for the natural log transformation.

PP use the Poisson model for both the deportees and the *Stürmer* letters investigations. Standard specification checks reject these models in favor of more flexible functional forms. The pogrom proxy does not have a significant effect in some of the alternative forms. The deportees model also reflects the considerable influence of a small number of observations. Further, the OLS models PP reports in its appendix do not correspond to the specifications discussed in the text.

A.4 All six outcome variables

The text and this appendix dig into the details of all six of the models PP reports in Table VI. The only model that survives econometric scrutiny is that for the *Kristallnacht,* which we have argued does not support the article's argument because those attacks reflected government initiative rather than local anti-Semitism. We now consider the summary measure PP constructed from the six outcome variables modelled in Table VI: the first principle component of those six outcomes. PP views this exercise as capturing a "broader, underlying pattern of attitudes" (p. 1370) towards Jews at this time. What does this variable actually capture? That principle component is only correlated with three of the outcome variables in PP Table VI. Its correlation with the Deportations, Letters, and *Kristallnacht* indicators are not significant at any conventional confidence level. On the other hand, it is correlated with the 1920s pogrom variable (.237; p=0) and higly correlated with the 1928 Nazi vote share (.926; p=0) and the 1924 DVFP vote (.923; p=0). The high correlation with the 1928 Nazi vote explains the pattern we showed in the text Figure 1. The "summary measure" looks much like the 1928 Nazi vote share because it is virtually the same as that one outcome.

PP Table VII reports OLS regressions that standardize both the dependent variable and all the regressors. The reason PP gives for doing so is sensible but would apply equally to the voting models reported in PP's Table VI. Our Table A4.1 reconsiders the results reported in PP's Table VII. (Some of these results also appear in our text Table 1). The first column replicates PP Table VII, Column (1), the specification that corresponds most closely to the models presented in PP Table VI. Column (2) re-estimates this model as a median regression; the pogrom indicator is no longer significant. As text Figure

1 shows, the regression for the principle component reflects outliers, just like the regression for the 1928 Nazi vote share PP reports in Table VI. Column (3) re-estimates their model as OLS, while dropping 20 observations for which the "studentized" residuals in Column (1) exceeds 2.0 in absolute value. Columns (2) – (3) show the importance of outliers to their results. The largest outliers are mostly in Bavaria, as text Figure 1 shows. Column (4) drops Bavaria from the OLS model PP reports. The point estimate for the pogrom proxy is cut in half and the variable is only marginally significant.[15]

Here one might argue that our POG1349 point estimates for the principal components variable and for the outcomes in PP Table VI are not significantly different from the corresponding estimates in PP. (One could make a related claim for most of what we show about the econometric weaknesses in PP and BF). This is not, however, the issue. Rather, we want to know whether the true effect of cultural antisemitism is different from zero.[16] Like any replication exercise, we ask whether, starting from scratch, one would come to the same conclusions as PP if we used better econometric approaches. The correct null hypothesis is that the pogrom proxy does not affect Weimar-era outcomes.

A.5: Province-level effects

The outliers in Bavaria raise a more general issue concerning the role of regions. Bavaria accounts for most of the outliers for both the 1928 Nazi vote share model and the principle components model discussed in section A.4. It's natural to ask what happens if we simply drop the Bavaria observations: that is, to ask whether the problem is that the mean values of the dependent variable are different in Bavaria, and, at the same time, whether the slope coefficient (the effect of the pogrom proxy) is different in Bavaria. We performed this exercise (but do not report regression tables). The results differ

[15] The standardized regression coefficient for Columns (1) and (4) are both about .114. The dependent variable in Table A4.1 has the long right-hand tail of the underlying variables, not surprisingly. The mean of this variable is .02. It has a standard deviation of 1.122 and a skewness of 2.44.

[16] If we ignore controls and measure variables as deviations from their means, then PP measure the effect of anti-Semitism on an outcome y by estimating β in $y = \beta * POG1349 + f$ where f is the error term. The PP estimate of β (call it β_1) may not be significantly different from our estimate β_2, a point raised in the discussion of the 1928 election results by PP's authors, who cite Gelman and Stern (2005). But the issue is whether the effect of cultural anti-Semitism is significantly different from zero. The answer is no, as we show in the text and this appendix. That β_1 is sensitive to outliers in the specific case of the 1928 elections pointed us in this direction.

across models. For the "1920s pogrom" model dropping Bavaria is important; the pogrom estimate goes from .061 (SE=.023) to .046 (SE=.029). For the 1928 Nazi vote share model, the estimate goes from .014 (SE=.006) to .007 (SE=.004). The impact for the DVFP's 1924 vote share is reversed; with the full dataset PP uses for its Table VI, the estimate is .0147 (SE=.011). Dropping Bavaria makes the estimate marginally significant (.012, SE=.007). For the combined six outcomes (the principle components model) the point estimate goes from .290 (SE=.132) to .150 (.080), remaining marginally significant.

The effect of dropping the Bavarian observations is not the same for the Deportations model, the Letters model, or the *Kristallnacht* estimates. As noted, we do not think the *Kristallnacht* results tell us much about local anti-Semitism. And the two Poisson models (for Deportations and Letters) have all the weaknesses documented above. Without the Bavarian observations, the bootstrap standard errors for the two Poisson models are even larger than in PP's Table VI, and the two negative binomial models more decisively reject the PP interpretation (tables not reported). Even this simple exercise demonstrates that the PP results are sensitive to taking Germans regions seriously.

We can do better, however. The best way to ask how regions matter is to estimate models that allow the regression intercept to differ across regions, as well as allowing the impact of the pogrom proxy to differ by region. Tables A5.1-A5.3 report such models. In each case, we start with the variables PP use in their Table VI (and other specifications) and add both state-level fixed effects. These specifications allow the dependent variable to have a different mean in each state. We next add to the fixed-effects model the interactions of those fixed-effects with the pogrom proxy. This allows each state to have its own mean for the dependent variable, and, crucially, allows the pogrom proxy to have a different effect in each state. This flexible approach allows the data to tell us how the pogrom proxy's effect differed across regions. Table A5.1 considers the role of province-level fixed effects for two sets of results in PP. In Columns (1) and (2) of Table A4.2, the dependent variable is the 1920s pogroms, which is the dependent variable in PP Table VI, Column (1). In Columns (3) and (4), the dependent variable is the same as in Table A4.1: the first principle component of the six outcomes PP stresses. This is the dependent variable in PP Table VII, Column (1).

Column (1) starts with the model PP reports and adds to it fixed effects for the administrative units PP calls “provinces.”[17] Adding fixed effects in this way adjusts for differences in the mean of the dependent variable across administrative units, but does not address our real concern, which is the fact that the pogrom proxy seems to matter in only a few places. The pogrom effect remains almost unchanged from the PP version. In column (2), we add interactions of all the province fixed effects with the pogrom variable POG1349. This generalizes our concerns about Bavaria; the interactions allow the effect of medieval pogroms to differ from province to province. The regression cannot estimate all of these parameters. In the provinces that are missing interactions, the effect of pogroms is same as the “main” effect for that variable, that is, in Baden.[18] Columns (3) and (4) do the same for the first principle component model discussion in Appendix Section A.4. The F-test shows that the province/pogrom interactions are collectively different from zero for both models That is, PP restricts the model to force the pogrom proxy to have the same effect in every state. The data reject that that restriction

We obtain the pogrom’s total effect on the dependent variable by adding the main effect for Pogrom to the relevant provincial interaction with Pogrom. For the model in column (2), the total Bavarian effect is .018 + .124 = .143 (SE=.053). The only other province for which the pogrom proxy has a non-zero total effect is Braunschweig (.923, SE=.034), which has two observations. For the model in Column (4), Bavaria’s total effect (.931, SE=.384) is the only significant effect of the pogrom indicator. Given this flexible specification , the pogrom proxy has no effect in any part of Germany save Bavaria.

Table A5.2 has the same format as Table A5.1, and presents similar models for the two election variables. In Columns (1) and (2), the dependent variable is the vote for the right wing anti-Semitic DVFP (*Deutschvölkische Freiheitspartei*) party in 1924, which PP (reasonably) treat as a sort of proxy for the Nazis, who were banned at the time In column (3) and (4), the dependent variable is the 1928 Nazi vote. For the 1924 election, Bavaria’s total effect is .05 (SE=.03); the only larger effect is in Württemberg, with

[17] In most of Germany the “provinces” in PP are the German federal states; in Prussia they are that state’s 15 provinces. See PP Appendix, Note 3. The geographical divisions are different in BF. The models we report always use the definitions employed in the relevant article. See below.

[18] These are Bremen (1 observation), Hohenzollern (1), Anhalt (2), Oldenburg (1), and East Prussia (1).

.026 (SE=.01). For the 1928 election Bavaria has the largest effect, at .041 (SE=.016). Württemberg is also significant (.017, SE=.005).

Table A5.3 extends this inquiry to the models for Deportations and letters to *Der Stürmer*. Adding fixed effects to the model for Deportations (Columns (1)) renders the pogrom variable insignificant. The same does not apply to the Letters model, however (Columns (3). For both dependent variables, the interaction terms (which allow the effect of the pogrom variable to be different across administrative units) show that the pogrom results are driven by a handful of sometimes tiny states. As in Table A5.1, the total effect of the pogrom proxy in the regressions with the interactions is the "main" effect (which is for the reference location, Baden) plus the interaction for a particular place. For deportations, the total effect in Bavaria is small. In some places the legacy of a medieval pogrom is to *reduce* the number of deportations. The total effect in Hannover is -.486 (SE=.18) and in Hesse-Nassau, -.65 (.20). [19]

The results for Letters are slightly different. The result PP reports for Letters appears to be driven by enormous effects in two small places. Braunschweig (2 observations) has a combined effect of 14.47 (SE=1.006). Mecklenburg (4 observations) has a combined effect of 12.5 (SE=.73). More generally, both interactions models for Deportations and Letters indicate that constraining the effect of pogroms to be the same in all places, as PP does, is a misspecification.

In the text, we stress that the relationship between Weimar-era outcomes and medieval pogroms that PP stresses is usually driven by Bavaria. Here we have generalized the question to allow the relationship to vary across provinces. For one model (Letters), we have noted Bavaria does not appear to show a different relationship between pogroms and the dependent variable. The general point nonetheless remains: PP's arguments concern Germany, but the results reflect only small parts of the country. Tests of linear restrictions therefore reject the null that the pogrom proxy has the same effect everywhere in

[19] The point-estimates reported for the Poisson model throughout reflect the parameterization that the Poisson parameter (which is both mean and variance) $\lambda=\exp(X\beta)$. When the total effect of a pogrom is less than zero, as in Hannover, this just implies that the pogrom in that province reduced λ, and not that λ is itself negative.

Germany. Clearly, the province-level effects stand in for some traits that are more or less pronounced in particular parts of Germany. This is just a different way of saying that the results PP reports are not robust.

A.6 Placebos: Considering Additional Elections

The entire PP argument rests on the idea that the variable POG1349 reflects enduring anti-Semitism. We have shown that most of the PP results reflect outliers or other econometric issues. PP focuses on outcomes that have a plausible link to anti-Semitism. Equally important in an exercise like this, though, are falsification tests using outcomes that do not seem evidence of anti-Semitism. It is important to know whether the proxy "predicts" outcomes it should not. Does a medieval pogrom appear to "cause" behavior in the 1920s that has nothing to do with anti-Semitism? These placebo checks would assure the reader that the POG1349 can be interpreted as PP asserts.

The multiplicity of political parties and elections in Weimar Germany offers the chance to check many outcomes, but PP's authors do not take full advantage of it.[20] PP relies on a public database of election results in the Weimar Republic, so it is straightforward to consider other outcomes. In this section we treat the vote for each party in the federal elections of 1924, 1928, and 1933 as placebos. This section reports all models that we considered.

PP's main election results, presented in PP Table VI, consists of models for the Nazi Party in 1928 and for the DVFP in 1924. PP Table XI does consider a second right-wing party, the DNVP (*Deutschnationale Volkspartei*), which reflected the merger of several older right-wing parties and was arguably less focused on anti-Semitism than was the DVFP (or the Nazis). In this model, the PP estimate for POG1349 is *negative* and significant. Their Table XI also includes votes for the KDP (the Communist Party) in 1924 and 1928. For those two models, their pogrom proxy is not significant.

[20] Appendix Table A6.5 lists the abbreviations, full names, and political orientation of the parties discussed here.

PP view these models as a check on whether pogrom proxy helps extremist parties in general, and not just anti-Semitic parties like the Nazis. Their interpretation of the DVFP and DNVP results are worth quoting:

> "We have already shown that the DVFP gained more seats in localities with a past of medieval pogroms. If this is a reflection of anti-Semitism – and not more right-wing attitudes generally—then we should expect the closest (but less anti-Semitic) competitor DNVP to register fewer votes in towns and cities with an anti-Semitic past… Votes lost by the DNVP are similar to votes gained by the DVFP in these cities… Because the two parties' programs were similarly right-wing overall, these findings point to anti-Semitism, not extreme political attitudes as the driver of voting behavior in cities with Black Death pogroms (p. 1384)."

This statement embeds a strong assumption about substitution between votes for these two parties and all the other parties competing in the election. If there were only two parties, and if voters saw degrees of anti-Semitism as the only difference between the parties, then PP's authors could, in fact, interpret the result as they do. The first premise is obviously false, however, and the second is far from obvious.

Our text Table 2 asked whether, using the PP specifications, a medieval pogrom increased the vote share for the two main democratic parties in 1924. This is the correct placebo; our prior was that it would not affect the vote share either way. But it did. This failure of a basic placebo test calls into question the entire notion of the pogrom proxy. We continue along these lines, asking whether other parties' vote shares were affected by the pogrom proxy in ways inconsistent with PP's interpretation of the proxy Table A6.1 first looks at what PP considers the extremist parties: the right-wing parties DVFP and DNVP, plus the KPD.[21] For each party, we report regressions estimated for the entire dataset (that is, Germany), for Prussia alone, and for Germany minus Bavaria. We also report median regressions for all of Germany. The logic of the two geographical subsets follows from BF's arguments about political

[21] To make these results readable, for the models reported in this section the associated tables include only the pogrom variable and the associated fit statistics. The replication code generates the full model, which includes the controls PP uses, as can be seen when our specification is identical to theirs. That is, all of our specifications are identical to those reported in PP Table VI, Columns (2) and (3).

stability in Prussia as well as our demonstration that Bavaria is often an outlier that drives PP's results. Rows (1)-(4) parallels the PP regression for the DVFP (PP Table VI, Column (3)), while rows (5)-(8) parallel their DNVP result (PP Table XI, Column (3)). In the DVFP specification, PP report that POG1349 is not significant. Row (4) shows that the same holds for the median regression model, once again. Note that in rows (5) through (7), the vote share for the DNVP is negatively correlated with the pogrom variable, both in Germany as a whole and in Germany without Bavaria or Prussia. This represents a clear violation of PP's interpretation of the pogrom variable. Row (9)-(12) report a model for which the dependent variable is the *difference* between votes for the two right-wing parties.[22] These models apparently support the contention in PP about the difference between the DVFP and the DNVP – except, once again, for the median regression model. But these results do not really address the other part of the claim in PP, which is that anti-Semitic voters preferred both the DVFP or DNVP to other parties. Rows (13) – (16) *combine* the votes for the two right-wing parties. If the pogrom dummy does pick up persistent anti-Semitism, then its coefficient in these regressions should be positive and significant. But it is negative for Germany (although not significant) and negative and significant for Prussia alone.

The Prussian result should be cause for concern: if the variable has the wrong sign in a state that accounted for two-thirds of the population, some questions are in order. The supposed indicator of persistent anti-Semitism *reduces* the vote share for anti-Semitic parties. This important outcome suggests that PP's results for the DNVP alone in 1924, and for the Nazis later, is weak evidence that areas with medieval pogroms were more likely to support anti-Semitic parties. More generally, the PP argument is about all of Germany, not the one-third outside of Prussia.

The last four rows in Table A6.1 report results for the Communists (the KDP). Row (17) replicates a result reported in PP. PP states, correctly, that POG1349 does not increase KPD votes in Germany as a whole and interpret this result as demonstrating that their variable is not correlated with extremism in general. Row (19) shows that areas outside Bavaria with a medieval pogrom were more

[22] Defined as the vote share for the DVFP minus the vote share for the DNVP.

likely to vote for the Communists, however. Both this and the results when we combine the right-wing parties call into question the idea that POG1349 picks up enduring anti-Semitism reflected in Weimar elections.

Table A6.2 expands on results for the 1924 election reported in text Table 2.[23] As noted there, neither the DVP nor the DDP featured anti-Semitic claims as part of their political appeals. PP's argument implies that the pogrom proxy should not augment either's vote share. Our text discusses the implications of the results for the DDP and DVP. Recall that the DDP was viewed as the "Jewish" party: our placebo results show that the medieval pogrom variable raises its vote share. The remainder of our Table A6.2 considers the SPD and the Catholic parties. Neither the SPD nor the Center Party benefited from the pogrom proxy. The Center and BVP worked closely. The results do not suggest any association between POG1349 and voting for these parties, with the exception of the large, negative result for the BVP in all of Germany. That result may not indicate anything important, as the BVP did not really contest elections outside Bavaria. It does offer a useful warning about using a proxy without asking what it does in a variety of circumstances. Combining the two Catholic parties (the variable CathParty24) does not suggest anything surprising, either.

Tables A6.3 and A6.4 report additional checks based on the 1928 and 1933 elections (along with comparisons for the NSDAP, the Nazi Party), adding to the results that text Table 2 presents for 1928. In 1928, the liberal DVP again apparently benefited from the medieval pogrom, judging from the median regression or the OLS models for Prussia and Germany-less-Bavaria. We noted in the text that by 1928, the DVP had drifted to the national right, and the idea of this placebo is not so clear as in 1924. Places with a medieval pogrom outside of Bavaria were also more likely to vote for the Communists.

The 1933 election was fought in an atmosphere of violence and disorder. Results from this election are hard to interpret (as the many scholarly works on the subject attest.) Table A6.4 does not imply the troubling results for the liberal parties we noted in 1928. It does, however, show that outside

[23] Recall that appendix Section A.5 extends the 1924 and 1928 election results that PP presents to consider the role of provincial fixed effects.

Bavaria, the Communists benefited from a medieval pogrom. PP argues that the Nazis had toned down their anti-Semitic rhetoric to make themselves more palatable for the 1933 election. The results (Table A6.4) support that view, if we take it on face value: the legacy of medieval anti-Semitism harmed the Nazis outside of Bavaria.

Much of our discussion of the econometric results in PP and BF deals with standard errors, outliers, and related issues. These placebo tests, on the other hand, go to the heart of the idea underlying PP. The authors view POG1349 as a proxy for medieval anti-Semitism. They interpret the correlation of POG1349 with variables from the Weimar era as evidence that the anti-Semitism persisted into the twentieth century. RGH notes correctly that the more placebos one tries, the more likely one is to find at least some violations. Note that here we have considered and report every sensible voting outcome for the period. We do not just report the violations.

The tests we report here show that only a subset of sensible outcomes are consistent with the results the article stresses. A broader look reveals cases where the history of a medieval pogrom reduced an anti-Semitic party's vote share. More troubling, the POG1349 indicator increases the vote share for parties that had no anti-Semitic profile at all. We also see the regional diversity noted in other contexts: in some parts of Germany, the pogrom proxy affects anti-Semitic voting behavior differently than in others. Our placebo checks relied on the same public-use database of voting results that PP uses for its own exercises. To the extent PP includes any placebo tests, the question is whether the pogrom proxy predicts votes for extremist parties in general, including the KPD. We show that outside Bavaria (that is to say, in most of German) the pogrom proxy does, in fact, predict higher vote shares for the Communists. More importantly, we address the central question: does a medieval pogrom affect an outcome variable with which it should have no correlation?

A.7 PP Matching

PP supplements the main results for each of the models presented in their Table VI with the average treatment effect (ATT) for two different types of matching exercises. We should stress our unease about the entire idea of matching. While it is widespread in some areas of social science, economists tend to be skeptical. The core idea of matching is to mimic an experiment with random assignment into treated and untreated status. History gave us communities with POG1349=0 and POG1349=1, and, unfortunately, those communities differ in other potentially important ways. The idea of matching is to find communities with different values of the treatment variable and very similar values of controls. Then the simple difference in outcome measures supposedly yields a consistent, causal estimate of the treatment effect. This notion approaches a purely experimental setting that assigns treatment and control observations based on observable traits. The problem, of course, is that unobservables likely influence selection into treatment status and thus pollute the results.

PP uses two different matching algorithms. The first (propensity score matching) defines "matches" using a metric based on values of the variables included in the relevant regression. The second (geographic matching) defines the controls by matching treatment communities with the nearest community that did not have a pogrom. By definition, both methods require strong assumptions about the relationship between the observables and the unobservables that may drive selection into treatment status. The simplest assumption, "no unmeasured confounders," requires that no unobservables affect selection into treatment or the outcome for the treated. (Actually, a weaker condition of conditional mean independence is all that is required, but that again makes strong and unverifiable assumptions about the unobservables.)

The core idea of the geographic match also interacts uncomfortably with the basic idea of PP's pogrom proxy. If anti-Semitism reflects persistent local culture, it is difficult to understand why one would find nearby communities with different histories of pogroms. The geographic matching approach relies heavily on this fact: it assumes that two places are quite similar, except one community went crazy and attacked the Jews in the 1340s. We would stress a likely unobservable relevant to both matching exercises: the political issues we have discussed and that PP does not take seriously. The difference between two communities that look similar on observables might well be political arrangements in the 1340s that could be correlated with outcomes in the Weimar era, as we stressed for Bavaria.

The hope for matching estimators is that they may make results less reliant on functional-form assumptions. This is indeed a worthy goal. But it requires strong assumptions about unobservables. Thus it essentially restates a trade-off: we can either rely on functional-form assumptions in a regression or on strong assumptions about unobservables with matching. We stress that most of our discussion of both PP (and BF) have little to do with functional form. Sensitivity to the inclusion of Bavaria and to extreme values of the regressors and dependent variables is not a functional-form issue. And even when we raised the functional-form issue, as with the Poisson model, we have noted other issues that are unrelated to functional form.

PPs view these matching estimators as a robustness check that supports their conclusions. We do not agree. Our Table A7.1 presents an overview of PP's matching results with some checks of our own. As PP says, for the six outcomes presented in Table VI, the ATT is positive and significantly different from zero. This is true for both types of matches. Our table then reports additional specifications motivated by our earlier findings. First, we redo all the matching with a subsample that excludes Bavaria. This changes the results considerably. The ATT for the 1920s pogrom is still significant in both models, but the effect is much smaller in both approaches. Without Bavaria, there is no longer any effect for the 1928 Nazi vote share. The DVFP is no longer significant for the first matching model without Bavaria.

Neither the deportations nor the *Stürmer* ATT changes dramatically when we drop Bavaria. But these models are still sensitive to the problems we noted in the text. The lower panel of Table A7.1 uses the same sample restrictions we tested using the Poisson model that PP employs. If we drop the five largest values of the dependent variable in the deportations model (cf. Table A2.4), the ATT remains statistically significant but the effect is about two-thirds of its size with the full sample. This is true in both matching models. The results for the *Der Stürmer* letters (cf. Table A3.1) are sensitive in a different way: dropping the largest five values of the dependent variable increases the matching ATT's considerably.

For several models, the geographic matching estimator implies larger effects from pogroms than does the OLS estimate. This odd result underscores the core assumption in PP, which is the proxy for anti-Semitism derived from pogroms in the 1349 is a good proxy for anti-Semitism in the 1920s and 1930s. Suppose instead (as we think likely) that any village-level differences in anti-Semitism present in 1349 dissipated over time. In this case, the Weimar-era OLS regressions suffer from measurement error; the variation in the pogroms in 1349 would represent, in the 1920s, largely noise, and the estimate for that variable should be attenuated. The fact that the geographic matching estimator is larger than the OLS calls into question whether these communities are the right level of analysis.

A.8 The proximity of the communities PP studies

PP intends its pogrom proxy to reflect independent degrees of anti-Semitic sentiment in the fourteenth century. The idea is that some places hated Jews more than others, and that the places with the strongest hatreds were more likely to erupt in anti-Semitic violence. We have registered several concerns about the proxy and the econometric conclusions they draw based on it. It's worth noting an important requirement for the regression results in general: the attitudes this variable intends to measure must be independent across communities in 1349. Appendix Figure A8.1 should cast doubt on this idea. The places PP thinks had a Jewish community at risk of a pogrom in 1349 are concentrated in a small number of German regions. Most of these communities are, in fact, close together. People at that distance in the fourteenth century communicated at religious and other festivals and through market relationships. Doubtless some spouses in one community were born in one of the neighbors. The idea that any variable drawn from these places reflects independent attitudes towards the Jews requires more defense than PP musters.

B.1 BF Main Results

Table B1.1 replicates the main results from BF Table 3.[24] We focus on BF Panel A, Column (4), which is the specification the authors stress throughout the article. Column (2) shows that the variable of

[24] This appendix section starts by showing how outliers drive BF's main results, and then turns to the role of spatial standard errors. Section B.3 discusses the BF argument that social capital only mattered in politically unstable states.

interest, associations per capita, has little effect in the median regression model. This result echoes the problem with outliers we documented in PP. The remainder of our table focuses on OLS models. Figure B1.1 displays the dependent variable and OLS residual. As in PP, the problem here is a long right-hand tail; this accounts for the difference between the OLS and quantile regression results. The OLS residuals from BF's specification (our Column (1)) track the dependent variable closely. Our Column (3) drops the two observations that correspond to the largest 1 percent of residuals from Column (1). Our Column (4) increases this to 2 percent. Our Column (5) takes a different approach, dropping 16 observations for which the "studentized" residual in Column (1) has an absolute value exceeding 2.

The OLS results here are sensitive to outliers. Although the effect remains statistically significant, the point estimate and standardized regression coefficients fall by 25 percent if we remove just two (of 229) observations. Our Column (6) drops the 23 Bavarian observations. This cuts the estimate for Clubs per capita nearly in half, and it is only marginally significant. Something is different about Bavaria in these models, as well.

B.2 BF Spatial Standard Errors

In a recent paper, Morgan Kelly (2019) shows that the standard errors reported in PP are underestimated because they do not account correctly for spatial correlation. Many of the PP locations are quite close to one another, as we note, and the intuition that "close things are similar" should have led to some caution.[25] The cities that form the observations in BF are also close to one another. The Ruhr area illustrates the problem (and is admittedly a severe example). Four of BF's cities lie within 10 kilometers of Essen, twelve are within 20 kilometers of that city, and a further 14 percent of the entire sample is within 50 kilometers. Situations like this not only raise the possibility of incorrect standard errors, but also may affect the definitions of both the dependent variable and the social-capital proxy. It is possible,

Section B.4 reports a simple model underlying our argument that BF's results are not necessarily about social capital.

[25] In their Appendix Table A.14, PP reports some regressions with Conley standard errors. Kelly shows that PP used too small a distance parameter; that is, that the correlation in spatial errors occurs at longer distances than PP allows.

for example, that individuals who lived in another place belonged to civil-society organizations in Essen, and that Essen's proximity made it harder to achieve the critical mass needed to form a club in one of its neighbors.

We assess the standard errors question by re-estimating the BF models that we discuss in our text Table 3. In some cases, the standard errors change in meaningful ways. This is especially true for the non-Prussia model reported in BF Table 7, Column (2). But we do not find the consistent pattern that Kelly identifies for PP. To reach this conclusion, we first estimate Conley standard errors allowing observations to be spatially related up to 50 km. The standard errors for the social capital proxy increase by about 7 percent if the influence decays within the cutoff region, but the estimate remains significant. Without the decay, the Conley SE is 50 percent larger than the "robust" standard errors BF reports, and the estimate is not significant. Reducing the distance parameter to 25 km or increasing it to 100 and 200 produces similar results: the SEs are different, but not in a way that suggest the BF results depend on ignoring the spatial issue. We do not report any regression results for this exercise.

B.3 BF Political stability

The text reports investigations of BF's "stability" analysis. BF Table 3 implies that more social capital leads to more Nazis. This is the authors' headline result. Additional tests (BF Table 7), however, show this held true only in federal states they consider politically "unstable." In Prussia and other "stable" states, there is no such relationship. This is a remarkable result. BF's authors argue, in effect, that the results for Germany (such as their Table 3) hinge on the inclusion of cities from a small number of unstable federal states. Without those cities they would not have their headline finding. Put differently, had BF's authors had limited themselves to Prussia, using their own data and methods they would have concluded that social capital did *not* aid Nazi recruitment, and they would thus have had a very different argument. Many economic history studies in fact do limit themselves to Prussia, where some two-thirds of Germans lived. Here we investigate the way BF comes to this conclusion about stability. We show that the results BF reports reflect a series of undefended, unusual definitional and specification decisions, combined with incorrect computations. This is the only place in either PP or BF that considers regional differences, but the paper does not handle the issue correctly.

BF constructs a stability index as first principle component of three features of the coalition governments that led the individual states in the Weimar era. BF's authors provides two different, conflicting definitions of the index. They compute the principle components using the wrong level of analysis. And they do not use the resultant index as a regressor; rather, they use it to construct a binary indicator for "stable" and "unstable" states. (They do use the continuous index in a robustness check discussed below; unfortunately, that check has a serious error.) Since a large number of observations are at or near the median, this splitting of the sample at the median creates important distinctions not really in the index. Finally, BF's regression do not use the binary indicator as defined in the text. The text claims that the authors divide the sample (excluding Prussia) above and below the median value of the index. In fact, Bavaria is right at the median value of the index, and they assign its 23 observations to the unstable group. Bavaria accounts for 22 percent of the 106 observations not in Prussia. Many of their exercises rely on these 106 observations. Bavaria accounts for 40 percent of the observations in the "unstable" subset of the non-Prussian observations. If we instead assign the median values to the stable group, we obtain results that contradict what BF reports.

The stability index comprises the first principal component of three variables defined for the period October/November 1918-May1932. The elements are: (1) The percentage of that period the longest-serving government was in power; (2) the percentage of that period the longest-serving party was in power (possibly in different coalitions); and (3) the percentage of that period the state was ruled by the "Weimar Coalition," meaning the Social Democrats (SPD), the *Zentrum*, and German Democratic Party (the DDP). This, at least, is the definition in the notes to BF Table 7. BF p. 508 defines the third element differently: "… governed by at least one party from the Weimar coalition."[26] The two different definitions of the third element reflect that phrase "one party:" Definition one says the state is only stable if ruled by

[26] An earlier version of our paper and appendix discusses the alternative definition. Here we focus on the definition used the BF regressions.

the Weimar coalition in toto, while definition two says it has to be ruled by at least one Weimar coalition party. BF's empirical work all relies on the first definition, however, so we focus on that.[27]

We are unpersuaded that either version of the index's third variable captures the relevant concept. Political stability, in BF's argument, pertains to continuity in governments that make law and order more likely. We do not see why it matters whether a state's political leadership were in the same parties as those who led the national government. The example of Bavaria shows why: a single party, the BVP, led Bavarian government in the periods September 1921-November 1922 as well as for the six-year period July 1924-August 1930. Surely Bavarians enjoyed a stable government. Yet because the BVP was (technically) not part of the "Weimar coalition," the third element in Bavaria has a value of zero by construction.

BF estimates and uses this index in ways that drive the results. The three variables underlying the index reflect experience at the state level, so the principal components should be estimated with one observation per state. BF instead estimates the components using every city as an observation. This procedure effectively weights the political history by the number of cities in a federal state; Prussia, that is, enters the principal components computation 119 times in a sample of 229 observations.[28] The median value used to create their binary stability indicator excludes Prussia, so that large state contributes more than half of the observations in the index calculation, but BF then excludes it from the definition of the stability indicator. In addition, BF says that the binary indicator divides states at the median, with those above the median labelled as stable, and those below the median labeled as unstable.[29] Given their actual definition, however, Bavaria's 23 observations (23 of the 108 non-Prussian observations) lie precisely at the median of the continuous index. *BF assigns Bavaria to the unstable group*, regardless of what the text definition says. Two other, large states (Baden, with 17 observations, has an index value of -1.28, and

[27] That is, the BF code made it clear that the operating definition is the one in the notes to BF Table 7.
[28] BF cannot define the index in Lübeck (1 observation), Bremen (1), or Saarland (2). Thus the maximum sample size for these stability exercises is 225 instead of 229. See BF Table 7, Columns (5) and (6).
[29] "Next, we split the non-Prussian part of Weimar Germany into a stable and an unstable half (with above- and below-median stability, respectively)…" (BF p. 508).

Saxony with 21 observations, has a value of -1.22) are also very close to the median. These three medium-sized states account for 56 percent of the observations outside Prussia.

Our Table B3.1 reports the stability index as BF computes it, as well as two versions we think better capture a reasonable argument about stability. (BF Appendix Table A.8 reports the values BF uses for these three elements, as well as the computed index value.) We have identified several weaknesses in the BF index do not consider the full set of permutations that would arise from fixing two or more problems at once. Column (1) reproduces the BF version of the index. Note that Bavaria is the median state, with an index value of -1.31. (The median for BF's purposes excludes Prussia.) Depending on how we treat the median, those 23 Bavarian cities, about one-fourth of the non-Prussian subsample, will be either "stable" or "unstable" in the econometric exercises. Baden, with 17 observations, has a similar index value (-1.28), as does Saxony (21 observations, -1.22). BF assigns Bavaria (the median state) to the unstable group but both Baden and Saxony to the stable group. This is one reason not to convert the continuous index to a binary indicator; BF's approach assigns Bavaria to one side of a divide as opposed to two other two states with very similar stability index values. In Column (2) we compute the index at the state level. This shifts Baden's 17 observations from above the median to below, using BF's binary distinction; Baden now counts as "unstable." Column (3) dispenses with the third element of the stability index. The index here is the first principle component of the first two elements in BF's index. As we noted, BF do not defend that third element, and it seems unrelated to the core idea of turnover in state governments. Using the BF binary distinction and this definition of the index, Baden is now unstable while Bavaria and Saxony are stable. In what follows it's important to note that BF assign Bavaria to the unstable group entirely because of the third element in their index.

Text Table 3 shows the sensitivity of BF's results to these changes. Our Appendix Table B3.2 expands on the consequences of these differences by re-estimating BF's regression models using the alternative stability indicators. Column (1) replicates BF Table 7, Column (4) while our Column (2) replicates BF column (3). For the stable states in our Column (1), there is no relationship between social capital and Nazi recruitment; for the unstable states in our Column (2), the relationship is positive and significant. This is the BF argument. The next two columns show that estimating the index at the state level does not produce materially different regression results. For the unstable states, the regression beta is smaller, but the estimate remains significant. Our Columns (5) and (6), however, show that the BF's regression results hinge entirely on the third element of the index. Dropping the third element of the index, as we do, shifts Bavaria (23 observations) from "unstable" to "stable," and Baden (17 observations) from "stable" to "unstable."[30] In these models, social capital has a greater impact on Nazi recruitment in the stable states, although the effect is not significant in either group. The difference between our definition (Columns (5) and (6)) and theirs (Columns (1) and (2)) reflects the way the binary indicator allocates cities in two states.

Table B3.3 has the same format as Table B3.2. In B3.3, we compute the stability index as does BF, but we define the binary stability indicator such that the median state is considered stable (rather than unstable, as in BF). Comparing each column in Table B3.2 to B3.3 shows the role of defining the stability dummy in that particular way. The clubs variable does not have a significant effect on Nazi recruitment in *any* of the specifications reported in Table B3.3. BF's entire "stability" result hinges on assigning the median state, Bavaria, to the unstable group rather than the stable.

We next turn to the last two columns in BF Table 7. BF's Column (5) combines all of the states (stable, unstable, and Prussia) and interacts the stability dummy with the baseline controls. BF adds a dummy for Prussia and its interaction with the stability dummy; Prussia is thus a different kind of stable. Column (1) of our Table B3.4 reproduces BF's Column (5). The important variables are the interaction of the clubs variable with the stability dummy and with Prussia; the BF argument requires that both variables be less than zero. And so they are. If we shift the median state to the "stable" group, however, as in our Column (2), the interactions are no longer significant. Once again, the entire result depends on how one assigns the median observations.

[30] Also shifted to unstable are the two Mecklenburgs, each of which has a single observation.

Our Column (3) replicates BF Table 7, Column (6), which adds state fixed effects to the previous specification. Our Column (4) applies the definition of the stability indicator whereby the median state is "stable." Once again, the result BF stresses disappears: the interaction between clubs and the stability indicator is no longer significant. The same applies for the interaction with Prussia, which in these specifications functions as a second stability indicator. Our Tables B3.3 and B3.4 use the BF stability index as the article constructs it. We simply shift the median state from "unstable" to "stable." The results BF reports disappear.

Why turn a continuous index into a dummy?

We have demonstrated the important consequences of the way BF defines the stability index and the way the article computes and uses it. We have also shown that a very small change to the binary indicator leads to substantively different results. But there is a more general question: why take the continuous variable, the index, and turn it into a binary indicator? This procedure throws away information in general, and in this, case, makes Bavaria look very different from two other large states that have similar values of the stability index.[31]

We first take a close look at the only place BF uses the actual index values. BF's Appendix Figure A7 purports to show the net effect of stability on Nazi recruitment, using the stability index as a continuous regressor. BF does not report the underlying regression; we display it as Column (5) in the text Table 3 and repeat it here as Table B3.5, Column (1). Table B3.5 reports the same model estimated without Prussia (Column (2) and without Bavaria (Column (2)). The crucial interaction between the clubs proxy and the stability index is not significant in any of these specifications. BF's appendix summarizes this regression in a graph of the net effect of the stability indicator on Nazi recruitment. The net effect of social capital is the sum of the main effect of clubs and its interaction with the stability index, evaluated at the value of the stability index for each state. The code for the regression underlying BF's Figure A7

[31] In BF, the regional aggregates are the German states; Prussia is not disaggregated into its provinces, as in PP.

generates correct values for the net effect and its standard errors, but incorrectly constructs the 95 confidence bands the figure plots. To construct the confidence bands, the BF code multiplies the standard error by 1.96, which is the correct critical value only for much larger samples. For the 13 degrees of freedom implied by the clustered standard errors in the regression, the critical value is 2.16. Using the correct error bands, the region of statistical significance apparently shown by BF Appendix Figure A7 disappears.

Table B3.6 reports the net effect of social capital evaluated at the value of the stability index for that state along with the standard errors and the correct confidence bands. The effects are estimated from the regression in Column (1) of our Table B3.5. BF does not report this regression, but it is the regression that underlies BF appendix Figure A7. We include BF's stability index values for reference. The *only* statistically significant impacts of social capital are in stable states: Hesse, Lippe, and Prussia. By contrast, BF argues that social capital only aided Nazi recruitment in politically unstable states. The index implies that Hesse is the most stable state (after Anhalt, which has one observation). Prussia is not far behind. The estimates for some of the states BF considers unstable are larger, but much less precisely estimated. The binary indicator BF relies on puts Baden in the stable and Bavaria in the unstable category but as Table B3.6 shows, the net effect of the index in the two states is almost identical. Thus the stability index itself contradicts the BF claim that social capital only mattered in unstable states.

Table B3.7 takes an entirely different approach that shows that the discussion of stability is a red herring. The regression reported in B3.7 does not use the notion of stability at all; rather, it lets the effect of clubs on Nazi recruitment vary freely across states. The stability argument implies that social capital affected Nazi recruitment differently in different parts of Germany. But BF's test of that argument placed considerable structure on the relationship: the differences had to fall in line along the binary difference implied by the way they constructed the stability index and then turned it into a binary outcome. The regression we report in Table B3.7 has the baseline controls used elsewhere along with a full set of state-level fixed effects and interactions of those fixed effects with the proxy for social capital. Prussia is the reference state, so the main effect for the Clubs variable pertains to Prussia. Nothing in the regression reported in Table B3.7 relies on the stability index or BF's binary indicator. Our approach permits social capital's effect to be different in different states, depending on what the *data* say, without invoking additional assumptions about stability in that state. The BF approach, on the other hand, forces the effects in Baden and Bavaria (for example) to reflect the difference in the stability index in those two states. With the binary stability indicator, the regression forces the effect of social capital to be the same for all states within a "stable" or "unstable" group.

Table B3.7 shows that BF's stability-binary approach creates results that are not actually in the data. Notice first that in Prussia, more clubs leads to more Nazis. Prussia serves as the paradigmatic stable state and accounts for more than half of the observations: this result alone is troubling. Hesse, which the BF index implies is the second-most stable state, does not support the argument, either: the net effect of clubs there is .527 (SE=.106). This result may reflect in part, the small number of observations in that state. Next compare the results for Baden and Bavaria. These two states have similar stability index values, yet in B3.7, their interaction terms have opposite signs. Both effects are precisely estimated. This is why the BF binary stability indicator works as it does: the relationship between social capital and Nazi recruitment in these two states is clearly different, and by placing them on opposite sides of a binary divide, the BF approach makes it look like the difference reflects "stability." Finally, the BF binary index imposes the linear restriction that all "stable" states have the same net effect and that "unstable" states have the effect (which is possibly different from the stable states). The data clearly reject this restriction (F= 57.31, p= 0).

In sum: the underlying index includes a third element for which the interpretation is at best unclear. Including this element drives the results. BF turns this continuous index into a binary indicator for reasons the authors do not explain. Creating a binary from a continuous variable throws away information, and this decision requires defense. The text misstates the binary indicator's definition as actually used in the econometric results. Even using the actual definition of the binary indicator used in BF, if we simply shift the median state (which is a large part of the sample) from "unstable" to "stable," the results stressed in BF disappear. These conclusions hold using every specification reported in the article. We then ask the more general question: why not just use the index itself? The exercise BF's appendix reports has a computation error; when done correctly, regressions using the continuous index itself do not support their contention that social capital only affected Nazi recruiting in political unstable states. Finally, we show that this entire idea of "government stability" places restrictions on the effects that the data rejects.

B.4: What do BF's results say about social capital?

The final section of our paper before the conclusion shows that the lack of model in BF makes it impossible to interpret the meaning of the estimated effect of social capital in BF. The economics literature on social capital and networks suggests that the most effective way to use social capital to recruit people into the Nazi Party would be for a party recruiter to join the organization and then use social ties to its influential members (its "gossips" to use the language of an influential economic study of social capital, who would not necessarily be the same as its leaders) to pass on favorable information about the party to the whole membership. That method of recruiting would use the association's social capital, the connections between the members, and it would result in the relationship highlighted in BF between associations and Nazi Party recruitment.

That is not, however, the only possible interpretation of an empirical relationship between associational density and Nazi recruitment. Social capital is about ties among people proxied in BF by membership in organizations. A different explanation is equally consistent with the findings and has nothing to do with interpersonal ties and thus social capital. It would simply require that Nazi recruiters know something about what sort of person would join which group. Historical studies of the Nazi Party (so we show) suggest that was possible and even highly likely.

It would involve no social capital, because it would have nothing to do with connections among club members. Yet the statistical relationship between Nazi recruitment and the number of clubs would be the same as in BF, because more clubs would give recruiters more chances to find associations whose members would find the Nazi Party appealing.

Either method of recruiting (via social capital or via knowledge about membership) would lead to a positive correlation between party recruitment and the number of associations in a town. To see why, suppose that there are a total N_t of these clubs in the town, and for the sake of simplicity, assume that they are all of the same size. Note that $N_t = N * s$, the town's club density N multiplied by the town's population s. Assume that if the recruiter employs the first method (taking advantage of social capital) and knows who the association's gossips are, then he will enroll k new Nazi Party members from the association. If he does not know the identity of the gossips, he will enroll no one. To avoid wasting his efforts, he will therefore recruit only from clubs where he can identify the gossips (ones where he is a member). If he belongs to a fraction p of the city's associations, he can expect to enroll $kpNs$ recruits. Expected per capita recruits will be kpN. The expected number would be the same if the Nazi Party had several recruiters in the town, only p would now be the fraction of the town's associations with at least one recruiter as a member. If we allow for other correlates z of party membership and presume a linear relationship, then the equation linking per-capita Nazi Party recruiting y and the association density N would be: $y = \alpha kpN + \beta z + e$, where z are other correlates of party membership, e is the error term, and α and β are coefficients.

What if the recruiter relies on the second method and exploits not social capital but information about the associations' members? Assume that he recruits k' new members if the club's associates are politically receptive, but no new members if the club's associates lean in the other direction. To minimize effort, he will canvas only in clubs that are receptive. He enrolls $k'p'Ns$ party members, where p' is now the probability that a club's members are politically receptive. We will end up with a similar linear relationship between per-capita Nazi recruiting y and the association density N: $y = \alpha' k' p' N + \beta' z + e'$, where α' and β' are the new coefficients and e' is the new error term. Here there is no social capital and no identifying the gossips. The Nazi recruiter just has to know the likely political opinions of each club's members. If the Nazi party has some recruiters who use the first method and some who rely on the second, then

$$y = (\alpha' p' k' + \alpha pk)N + (\beta' + \beta)z + e' + e \qquad (3)$$

An estimate of Equation (3) would only tell us whether $\alpha' p' k' + \alpha pk > 0$. It might be the case that social capital does boost Nazi Party membership ($\alpha pk > 0$), as BF claims, but it could also be the case that social capital does nothing ($\alpha pk = 0$) and the regressions reflect the information about groups ($\alpha' p' k' > 0$). This problem of interpreting the coefficients' meaning is a serious one that only further historical research could resolve: for instance, by determining whether successful Nazi Party recruiters actually joined organization that provided new recruits.

C.1 Coding the Pogrom Proxy

PP draws two critical variables from two compendia of German Jewish history, Avneri (1968) and K. Alicke (2008).[32] The first variable is an indicator for whether a Jewish community experienced a Black Death pogrom. The second indicates whether a Jewish community existed in 1349 in a particular place. The former variable is the key regressor in all of their tests, while the latter defines the set of "at-risk" communities and thus the data subset PP studies most intensively.

Each compendium reports a short history of the Jewish community in a particular location; both summarize and comment on other sources. Inspecting the two works conveys different impressions. Avneri is the more scholarly of the two works. After each entry, his compendium briefly lists the sources that underlie its account of that community, and it often contains internal explanations for its reasoning. Alicke contains much less detail; he almost never ties a specific judgement to a specific source. PP (p. 1348) says the authors relied on Avneri, but "supplemented" Avneri with information from Alicke. They do not say how. PP also claims to err on the side of certainty in dealing with these two sources: "Doubtful cases of Jewish settlement or occurring of pogroms in 1349 are not included in the dataset" (p. 1348).

[32] This appendix section discusses inaccuracies in PP coding of the information underlying the pogrom proxy (Section C.1) and the existence of a Jewish community at the time of the Black Death (Section C.2). Sections C.3 and C.4 discuss the way BF located and used the city directories that form the basis for their social-capital proxy.

Anyone who studies these medieval events faces severe versions of the problems fundamental and common to all historical inquiry. The medieval authors of the primary sources sometimes lacked direct and complete information and had their own point of view. The secondary sources available to Avneri and to Alicke suffer from similar constraints. Avneri and Alicke themselves had to put structure on sometimes conflicting information. Finally, PP's authors had to read textual accounts and reduce them to binary distinctions. This process will not lead to universal agreement under the best of circumstances.

We have shown that PP's econometric estimates are fragile in several ways. Appendix section A.1 notes some of the estimates are sensitive to the miscoding of POG1349 in particular. We have no doubt that widespread anti-Semitic violence inflicted horror on many (most?) European Jews at the time of the Black Death. For PP's econometric estimates, however, it is critical to know precisely *which* communities suffered pogroms and which did not. We are concerned that PP coded POG1349 in a way that makes it unreliable as an indicator of this violent experience. Our worries are serious enough to make us doubt whether POG1349 really captures the cross-sectional patterns in ways that make it useful for econometric tests.

There are two, related issues. PP claims that it excluded all doubtful cases. We cannot agree. A fair reading of the two compendia indicates that PP includes *many* doubtful cases. PP (p. 1348) refers to Avneri's entry for Heiligenstadt as "typical." It is not. According to Avneri, "At the time of the Black Death, the Jews of Heiligenstadt were systematically killed. Survivors were recorded in Erfurt in 1365 and in Frankfurt in 1389. Heiligenstadt only admitted Jews again in 1469." We agree with PP's authors that Heiligenstadt should be coded as an example of the 1349 pogrom, as indeed it is. The Heiligenstadt entry is hardly typical, however. That entry is clearer, more certain and more specific about the events than are all but a handful of other entries. Avneri usually indicates his confidence in what his sources can say. Many entries in his compendium describe the period using a phrase such as "the Jews in [this place] fell victim to the general persecution at the time of the Black Death."[33] Something bad happened, the sources show, but just what is not clear, and Avneri does not mislead the reader. PP code this phrase to mean there was a pogrom, and that is indeed one reading of Avneri's intention. But since Avneri is clear and specific in cases like Heiligenstadt, and in other cases refers to Jews as "burned" or murdered, we are less sure that "fell victim to persecution" means – to him – a pogrom. When the evidence is more definite, Avneri says so. For example, in Bonn "the community was destroyed at the time of the Black Death."[34]

Avneri signals his degree of confidence in two different ways. For example, the Jewish community in Villingen, Avneri (II, pp. 854-55) reports, suffered a pogrom at the time of the Black Death. The evidence? First, the territorial ruler later sold a house belonging to a Jewish widow. Second, the *Memorbücher* for neighboring communities refer to Jews murdered in Villingen at the time of the Black Death. Finally, the local synagogue became the property of the hospital (*Spital*). This evidence is all indirect, but for most communities, it is all that would ever be available. And the synagogue especially suggests that a community had existed but did not survive. When Avneri has details that suggest mass murder, he reports that evidence.

In other cases, Avneri uses language that indicates doubt. He uses two approaches. One is to use the term "affected" (*betroffen*), to indicate that something bad happened, although he knows less about the incident than in Heiligenstadt or Bonn. In other examples, he uses terms like *soll*, *vermutlich*, and *angeblich* to refer to the events his sources report. English translations of these terms span the range from "allegedly" to "presumably." He says, for example, that Assenheim was "*anscheinend*" (apparently) affected by the pogroms. We can only assume that as a careful scholar, Avneri wanted his readers to understand the limits of his evidence.

[33] The statement can also be weaker, just saying the community was "affected" (*betroffen*). See, for example, Aldenhoven: "Zur Zeit des Schwarzen Todes wurden die Juden in Aldenhoven von der allgemeinen Verfolgung betroffen."

[34] "Zur Zeit des Schwarzen Todes wurde die Gemeinde vernichten" (I, p. 94). The distinction between a pogrom and persecution might seem too fine, but PP also makes this distinction. The six outcome variables studied in PP Table VI all concern forms of anti-Semitic behavior. The two votes and the anti-Semitic letters to *Der Stürmer* involve no direct physical violence, however. Deportations, on the other hand, were usually the first step towards murder, and we would agree that the 1920s attacks as well as the *Reichskristallnacht* are examples of violence.

To investigate this issue systematically, we drew a 10-percent random sample of PP communities from the replication dataset and checked the entries for each of these places in both Avneri and Alicke. Table C1.1 reports the results. Our Table C1.1 shows that PP did not exclude all doubtful cases. In most of these examples, the language includes the signifiers of doubt mentioned earlier. And the case of Gau-Algesheim (No. 126) should not have been in the dataset if PP excluded all doubtful cases. PP codes the pogrom proxy as one. Avneri, however, does not mention any Black Death pogrom. Alicke suggests there might have been a community and it might have been erased in the Black Death pogroms, but his adverbs make clear he is unsure on either point. Gau-Algesheim is an especially clear problem, but the table shows that many of Avneri's and Alicke's entries indicate considerable doubt.

For comparison, we include the pogrom "intensity" score that Finley and Koyama (2018) report in their paper. These two authors reviewed all of the evidence in Avneri and tried to match it to other sources. One result of their careful research is a score that shows what they think happened, and how strong the evidence is. Finley and Koyama show that pogroms were both more intense and more likely in communities where political authority was fragmented. (A zero in this column means they include no information on this place.[35]) According to their online appendix, Finley and Koyama excluded cases where there was no distinct description of a pogrom that would allow them to measure its intensity, either in Avneri or other sources they consulted. Their addition of additional medieval controls also eliminated some towns.

PP's authors did not respect the care and nuance apparent in Avneri's compendium. Neither the article nor the appendix includes robustness checks that study the consequences of dropping the (many) instances where Avneri indicates doubt. One could also imagine robustness checks that code POG1349 for degrees of ambiguity. A final problem concerns plain mistakes. PP codes the community of Beckum, for example, as experiencing a pogrom. Avneri (I, p. 61) documents the existence of a Jewish community

[35] That is, for communities where our Table C1,1 has a "zero" value, Finley and Koyama think the sources are so weak they did not include the place in their dataset or analysis. Consider the PP sub-sample that corresponds to the two election outcomes in PP Table VI (325 observations). Finley and Koyama could not determine what happened to the Jewish community in 134 of these places (41 percent).

in the early 14th century, but mentions nothing about the Black Death or violence towards the Jewish community: "Andere Nachrichten fehlen."[36]

C.2: The Existence of a Jewish Community at the Time of the Black Death

PP's authors appreciate an important feature of their sources: sometimes, the first mention of a Jewish community is the pogrom of 1349. For their econometric exercises, it is critical to know the full set of Jewish communities at risk of a pogrom. If we first learn about a Jewish community because it was destroyed, however, then there will be bias in the selection of places to study, bias that might affect the way POG1349 affects the 20th-century indicators of anti-Semitism. There are two possible worries. The first is that a pogrom might so thoroughly erase a Jewish community that later scholars cannot know that it ever existed. That possibility seems unlikely, given the material Avneri and Alicke work with. The more serious is the one that PP does address. Their coding implies 325 Jewish communities existed in 1349. Of these, they code 72 percent as experiencing the pogrom of that year. This estimate might be too high if we do not know about communities that are not documented until later because they did not experience a Black Death pogrom

PP's authors address this point in a robustness check. They draw on other variables from the medieval and early-modern period to estimate probit models for the places where the compendia identify a Jewish community. They then use these estimates to predict the probability of an "expected" Jewish community in a town as of 1349. If this constructed probability exceeds 50 percent, they add the expected community to their dataset as a place that existed in 1349 but did not experience the pogrom. Re-estimating the models they report in Table VI from this expanded dataset yields results similar to those in the article. They draw comfort from this approach. We agree that it is a worthwhile exercise, but it offers little reassurance. It leans heavily on the implicit assumption that the relationship between the observables and the "missing" Jewish communities is similar to those they find directly. This might be true. Yet there

[36] "There is no other information."

might also be reasons that some communities have written documentation that precedes the Black Death pogroms and others do not. By relying on the estimated econometric model they make strong assumptions about both what led to a Jewish community's existence and what led it historical knowledge about it.

More important, econometric checks of this sort cannot address a more fundamental problem. PP's authors do not explain how they determined that a Jewish community existed in a place as of 1349. Avneri reports specific information that he interprets as evidence of Jewish presence. To be at risk of a pogrom in 1349, that Jewish presence really must be a community; the murder of a single person is evil but not a pogrom. In some cases, the presence Avneri documents involved evidence of a Jewish quarter (*Judengasse*) or Jewish institutions such as a synagogue or ritual bath. These places definitely should be coded as having a community. In others, Avneri simply reports that a document mentions a Jewish person as coming from a particular location. PP also codes such cases as implying a community. That inference might be correct; perhaps that single individual is all that made it into a written record, but he was just one person in a real community. It is also possible that this was a single person briefly noted as being in a place, such as a travelling merchant. Other cases are knottier. Some Jewish communities were present long before 1349 but were wiped out in an earlier pogrom. Sometimes the source is clear about community's re-establishment at a later date. Often, it is not.

This variable also raises questions about the way PP uses Alicke to "supplement" Avneri. The Appendix (p. 1) outlines the authors' rule: "We first establish the presence of a Jewish community based on its inclusion in GJ [=Avneri (1968) – the authors], volume II, which is for the period 1238-1350. Whenever later work by Alicke (2008) mentions that a Jewish community existed during this period, we use his information instead." They do not defend this procedure or explain why they trusted Alicke over Avneri only in this type of case and in this one type of conflict. This procedure implies (if followed consistently) that Alicke's information, if conflicting with Avneri's on this point, was only used when it implied a Jewish community. The case of Kempten (Bavaria) shows the implications of PP's rule. Avneri (I, p. 479) stresses uncertainty about whether there was a Jewish community there at the time of the Black Death. Alicke is more definite, in the negative: "A first sign of Jewish life in Kempten dates from the year

1373."[37] PP codes Kempten as having a Jewish community in 1349, as the rule implies, but the effect here is to include a community that is clearly a matter of great doubt.

C.3 Locating and Selecting BF's City Directories

BF's authors base their social-capital proxy on counts of civil-society organizations listed in city directories for the 1920s. Because neither the article nor the appendix reports the year of publication (or the publisher) for the directories BF uses, we cannot examine the specific directories that underlie their social-capital proxy. Yet both the way they collected the directories and the source itself raise serious questions of selection bias and use of historical sources. We first discuss the process by which BF's authors located directories to employ. We then discuss the directories' contents.

German city directories were part of a European phenomenon that started in the eighteenth century and continues in other forms. About 1,000 editions were published in the 1920s for German cities overall. The directories played an important role in local administration. While published privately, the directories' publishers had a relationship of "mutual support" with local governments (Shaw and Coles 1997, p. 59). The authorities sometimes provided cash, and often provided and checked some of the directory's information. Most publishers were local entities responsible for only a few directories, but there were exceptions. The enterprise *August Scherl Deutscher Adreßbuch* GmbH grew out of an earlier newspaper business and by the early twentieth century published directories in several major cities, including Berlin, Hamburg, and Frankfurt. Partly because of Scherl, in the 1920s some German cities had more than one directory published by competing firms.

The fixed costs of publishing made directories most profitable in large cities. Yet many relatively small places had directories. One study for 1913 locates a directory in 498 German urban areas. Of these, 116 were for places with fewer than 20,000 people (Shaw and Coles 1997, p. 49). Below we discuss directories from the 1920s for cities with as few as 10 thousand people.

---

[37] Alicke (II, p. 2175): "Ein erster Hinweis auf Juedisches Leben in Kempten stammt aus dem Jahr 1373. "

Most directories provided important information to locals, but they were also useful to people who lived elsewhere. Enterprises elsewhere seeking business with firms in Cologne, for example, would want to consult Cologne's directory. A traveler to Cologne (whether commercial or private) might also want to consult that city's directory before visiting the city. This is why German libraries held collections of directories for other cities. We cannot agree with BF (p.491)'s assertion – for which BF provides no evidence -- that directories were "printed and distributed in a small area" or their claim that "city directories often survived only in the local city library or archive."[38]

BF's sample cities

BF begins with the 547 German cities that had a 1925 population of 10 thousand or more. The actual sample, however, includes only 197 cities of that size.[39] The final sample is missing 64 percent of the universe, BF's authors claim, because directories for many of these cities no longer exist. To assemble their directories data, BF's authors contacted "libraries and archives" in these cities, asking for information on local directories (BF, pp. 490-91).[40] They dropped 65 of the original 547 cities that are no longer in Germany (most are now in Poland). We return to these cities below. BF dropped an additional 170 cities when the inquiry received no reply. Libraries and archives representing a further 115 places responded that they had no directory available. Thus 285 of the original cities within Germany's post-World War II borders with a 1925 population of 10 thousand or more are not in the sample. BF (p. 491,

[38] To take one example: the Bavarian state library (in Munich) today holds an extensive collection of directories from cities outside Bavaria that would fit the criterion used in BF. Some of them perhaps came to the library later, but the point is that the local library is not necessarily the best place to locate a city directory. This collection reflects the practice of collecting directories from other cities. A partial list includes: Koblenz (1929/30), Bonn (1929/30), Fulda (1924), Zwickau (1928), Essen (1927), Weimar (1922), Karlsruhe (1925), Speyer (1924/25), Bochum (1928), Königsberg (1925), Worms (1925), Lübeck (1925), Hamburg (1926), Düsseldorf (1925), Osnabrück (1928), Breslau (1927), Hannover (1924), Bad Godesberg (1925), Aachen (1924/25), Frankfurt (1925), Oppeln (1925), Breslau (1926), Stettin (1927), Wiesebaden (1924/25), Chemnitz (1927), and Jena (1921). These are all relatively large cities, and it might be that the holdings are biased in this way. But we did not search for directories from smaller communities in this library.

[39] The 229 in the PP regression sample includes 32 additional places with a population less than 10 thousand that the authors learned about in locating the others.

[40] They do not provide a list of the institutions they contacted. BF (p.491, Note 15) implies that contact took place by telephone but this is not clear.

note 15) assert, in regard to those 285 cities, that "For towns and cities without coverage, this information was lost, was destroyed during the war, or did not exist in the first place."

BF's approach generates striking omissions. The BF sample does not include either the largest city (Berlin) or the ninth largest city, Frankfurt am Main.[41] In fact, 11 of the 45 cities with a 1925 population of 100 thousand or more do not appear in the BF sample. BF also missed 18 of the 47 cities with populations 50-100 thousand that are currently within Germany's borders.[42] Cities that size almost certainly had a directory, and given the demand, many copies once existed. Appendix Table C3.1 summarizes our effort to locate directories for the cities with 50 thousand or more people that do not appear in the BF sample. We were able to locate a surviving directory for all but one of the cities over 50 thousand that BF claims have no surviving copies. This required looking in library databases and doing some simple web searching.

One might worry that the ability to find directories for larger cities does not tell us what could be done for smaller places. To address this issue we considered the 35 cities of size 10-11 thousand persons in 1925 that remained in Germany after World War II, that is, the smallest cities in the universe of cities BF considers. BF found directories for 11 of these. We located directories for another 20 of these cities, about 83 percent of the places for which BF claims there is no surviving directory (See Table C3.2). We cannot know, of course, precisely how many directories BF could have located for places with more than 11 thousand and fewer than 50 thousand had BF used a different approach. But these results for places still in Germany imply that the sample could have been larger and included far more smaller places. Suppose that BF's authors could have located directories for 95 percent of all cities over 10,000 persons, just as we did for the smallest category. This would imply that their main sample would have been (547 –

[41] BF Appendix Table A.2 lists the cities in the sample. Note that #65 is Frankfurt an der Oder, not the much larger Frankfurt am Main.

[42] The decision to exclude cities no longer in Germany accounts for five more cities with more than 50 thousand people in 1925. Using their 1925 names: Hindenburg o.S. (73 thousand in 1925) Königsberg (280 thousand), Liegnitz (73 thousand), Stettin (254 thousand) and Tilsit (51 thousand).

65) x .94 = 453 cities, even if they had discarded all those on territory lost to Germany after World War II. This would have doubled the sample size.

BF's decision to restrict the sample to Germany's current borders rests on the assertion that "Towns and cities in the formerly German areas of Eastern Europe rarely preserved marginal library holdings such as city directories" (BF p.490, note 14).[43] BF does not explain or defend this assertion. Appendix Table C3.3 shows that it is untrue. At least 70 percent of those cities have an extant directory today. Excluding the now-Polish cities is especially unfortunate because some of those places witnessed especially rapid growth in Nazi Party membership in the late 1920s. We can only speculate on how the relationship between clubs and Nazi recruitment might have differed there.

The concern here is not just sample size, it is the real possibility of selection bias. Surely there are reasons that some cities had directories and some did not. Similarly, there are reasons some directories made it into the repositories BF surveyed and some did not. BF's discussion of selection bias is limited to two brief accounts. BF reports descriptive statistics that compare their sample cities to Germany as a whole (BF Table 1). This exercise, however, can only consider selection on *observables*. BF simply notes that their sample is more urban than the country as a whole and attribute this to the greater likelihood larger cities had directories and archives (BF, p. 492). While perhaps true (although we easily found most of the "missing" directories for the very smallest places), this observation ignores the fact that they did not locate directories for several large cities.

BF also considers the Joseph G. Altonji et al (2005)/Emily Oster (2019) approach to assessing the implications of selection on unobservables. The results suggested by both approaches suggest that selection makes the point estimates for the social-capital proxy smaller; that is, selection biases the result against BF's findings. We would note that although formally correct, the Altonji et al. and Oster tests assume that the selection on observables is informative about the selection on unobservables. This assumption implies restrictive covariances between the observables and the unobservables, as Oster

[43] The appendix says "65 [cities] lay in former German territories in the East (now Poland or Russia), and we cannot obtain city directories for these." (p.521)

emphasizes. Given that the BF dataset includes less than half of the possible universe, the possibilities for selection are considerable. The simplest way to reduce selection-bias concerns here would have been to do the historical research and cover more of the universe.

Directories as a source for associations

BF does not discuss a second, related issue: do the directories provide a good enumeration of the associations in a place? General undercounts of associations would be undesirable, of course, but the more serious issue is systematic undercounts of certain types of associations. The historiography implies that some clubs might be Nazi organizations in disguise, while others, such as worker's clubs, would be actively hostile to the Nazis. We note in the text the suspiciously low number of worker's clubs in the BF dataset. If the directories were more likely to include some clubs than others, and the included clubs were either more or less sympathetic to the Nazis, then a proxy based on these directories is badly flawed. BF never address the question of what kinds of associations were listed in the directories. Many groups were clearly not listed, as we noted. This fact raises the risk of a different kind of selection bias. Comparing directories and registries for the same city would tell them what kinds of associations made it into the directories. This would have allowed them to assess any bias that stemmed from relying on the directories.

A good way to check a source's coverage is to compare it against a systematic source with similar information collected in a different way. BF does not do this. BF could have drawn on at least two sources external to the directories. First, many clubs organized as an *eingetragener Verein* (a registered association, or eV for short, as seen in Figure 3). This legal device gives an association entity status and limited liability at little cost. The public registry of associations (the *Vereinsregister*) would support two different checks. First, the register for the period exists, in principle, for all cities, including those that BF says lacked a surviving directory. Compiling data from the register would support comparisons of association counts for cities in their sample and not in the sample.

BF could also perform a similar check from a different source. Many local clubs belonged to a regional or national association of similar bodies. Those associations often published annual yearbooks that list all local member organizations. (BF uses one such national directory in their Appendix F). By checking whether their sample and attrition cities were equally likely to include the local clubs listed in the regional or national directories, BF's authors could reassure the reader that the attrition cities were not different, and that local decisions did not mean that some associations were listed while others were not.

C.4 Selection of Clubs and Other Issues in BF's Directories

BF does not provide a complete list of the directories the authors use, so we cannot know, in cities that had more than one, which edition the authors employed. We do know even know which year's directory they used for a particular city. They used "any surviving directory from the 1920s; where several are available, we take the directory nearest in time to 1925" (BF, p. 491). This procedure implicitly assumes that a proxy based on the number of clubs in, say, 1923 is as good as a proxy based on, say 1928. This assumption ignores the historical fact noted in the text: associational life exploded in the 1920s. It also assumes that the relationship between Nazi recruiting and social capital did not change during the 1920s, a strong and potentially testable assertion they do not discuss.

BF also provides little detail on how they processed the information contained in the directories. We note in the text that BF dropped political and religious organizations from their dataset. Ordinarily a decision like this would come with a robustness check to assure the reader that eliminating a specific group of clubs did not bias the results in favor of the authors' hypothesis. Once a directory is located, it would presumably be simplest to have the assistants enter the data for all associations, and only then drop from the total count some specific group of associations. Why BF's authors did not pursue this strategy, we cannot say. We also cannot address serious questions about a wide variety of other groups. Table C5.1 provides a statistical summary for the 1925 Worms city directory. We make no claim that this directory is "typical." (While Worms is in the BF dataset, as we say, we actually do not know which year BF used.) Note that about one-third of associations (146/455) are professional or occupational (groups 3-5). BF's authors apparently exclude these bodies from their analysis, although they do not say so. (BF Appendix Table A.3 lists the associations in the sample by type.) Just why is unclear; regular meetings of a professional association could involve as much social capital as a cooperative.

REFERENCES

Adreßbuch Stadt und Kreis Worms 1925. Worms: Buchdruckerei Eugen Kranzbühler Gebr Enyrim.

Alicke, K. *Lexikon der jüdischen Gemeinden in Deutschen Sprachraum*. Gütersloh, 3 vols., 2008

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113, no. 1 (2005): 151-84.

Avneri, Z. *Germania Judaica, Von 1238 bis zur Mitte des 14. Jahrhunderts.* 2 vols., 1968.

Cameron, A. Colin, and Pravin K. Trivedi. *Regression Analysis of Count Data*. New York: Cambridge University Press, 1998.

Finley, T., and M. Koyama. "Plague, Politics, and Pogroms: The Black Death, the Rule of Law, and the Persecution of Jews in the Holy Roman Empire." *The Journal of Law and Economics* 61, no. 2 (2018): 253-77.

Kelly, Morgan. "The Standard Errors of Persistence." CEPR Discussion Paper No. DP13783, 2019.

Mroz, T.A. "A Simple, Flexible Estimator for Count and Other Ordered Discrete Data." *Journal of Applied Econometrics* 27, no. 4) (2012): 646-65.

Oster, Emily. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37, no. 2 (2019): 187-204.

Parente, P.M.D.C., and J.M.C. Santos Silva, "Quantile Regression with Clustered Data." *Journal of Econometric Methods* no. 5 (2016): 1-15.

Shaw, Gareth, and Tim Coles, *A Guide to European Town Directories. Volume I: Germany, Austria, Switzerland, and Scandinavia*. Brookfield VT: Ashgate, 1997.

Voigtländer, Nico and Hans-Joachim Voth, 2022. "Response to Guinnane and Hoffman: Medieval Anti-Semitism, Weimar Social Capital, and the Rise of the Nazi Part: A Reconsideration." CESIfo Working Paper No. 10180

Table A1.1: 1920s pogrom/Kristallnacht

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Pogrom20s	Pogrom20s	Pogrom20s	Pogrom20s	Kristall	Kristall
Pogrom	0.0607***	1.936*	0.896**	0.0461	0.124**	0.135**
	(0.0226)	(1.106)	(0.411)	(0.0291)	(0.0522)	(0.0592)
Logpop25	0.0390**	0.532***	0.281***	0.0232*		
	(0.0152)	(0.172)	(0.0898)	(0.0131)		
Jewish_pc25	0.0135	0.168	0.116	0.00393		
	(0.0114)	(0.201)	(0.0926)	(0.0135)		
Prot_pc25	0.000337	0.00915	0.00433	0.000361	0.000355	0.000175
	(0.000423)	(0.00885)	(0.00407)	(0.000451)	(0.000603)	(0.000666)
Logpop33_dep					0.0498***	0.0411***
					(0.0117)	(0.0123)
Jewish_pc33					0.0262**	0.0157
					(0.0132)	(0.0108)
Constant	-0.393***	-10.49***	-5.517***	-0.235**	0.280*	0.391**
	(0.140)	(1.838)	(0.953)	(0.119)	(0.147)	(0.162)
Observations	320	320	320	253	278	224
Adjusted R-squared	0.054			0.021	0.098	0.089
Sample	Replication	Replication	Replication	-Bavaria	Replication	-Bavaria
Model	OLS	Logit	Probit	OLS	OLS	OLS

Source: Computed from replication file for PP.

Note: Column (1) replicates PP Table VI, Column (1). Columns (2) and (3) estimate the same model as a binary logit and a binary probit. The dependent variable in columns (1)-(4) is coded one for what PP calls pogroms in the 1920s. Column (5) replicates PP Table VI, Column (6). Column (6) estimates the same model, dropping the Bavarian observations. Kristall is coded one for places that experienced the *Reichskristallnacht*

Table A2.1: Descriptive statistics for deportations and letters to *Der Stürmer*

	Deportations		Der Stürmer	
	In levels	In logs	In levels	In logs
Mean	212.61	3.49	3.77	1.19
Variance	760,012.90	2.98	114.84	1.06
Skewness	7.89	0.59	6.34	1
Kurtosis	75.53	3.59	50.18	3.94
Number of observations				
Total	278	263	325	182
Zero	15		143	
Quartiles				
1st	9	2.39	0	.69
2nd	22.5	3.22	1	1.10
3rd	76	4.44	3	1.61
95th percentile	943	6.85	14	3.14
Five largest observations				
1st	10,049	9.2	110	4.7
2nd	6589	8.79	77	4.34
3rd	5523	8.62	74	4.304
4th	3447	8.15	73	4.29
5th	3255	8.09	54	3.99

Source: Computed from the PP replication file.

Notes: “Deportations” is the dependent variable in PP Table VI, Column (4). “Letters” is the dependent variable in PP Table VI, Column (5). The sub-samples here correspond to those used in the PP regressions.

Table A2.2: Role of additional controls: Deportations

	(1)	(2)	(3)	(4)
VARIABLES	Deported	Deported	Deported	Deported
Pogrom	0.142**	0.135	0.146	0.110
	(0.0706)	(0.137)	(0.150)	(0.143)
Logpop33_dep	0.241***	1.135***	1.021***	1.069***
	(0.0841)	(0.0311)	(0.0340)	(0.0213)
Jewish_pc33	0.0743**	0.384***	0.305***	0.616***
	(0.0348)	(0.0340)	(0.0440)	(0.0468)
LogJews33	0.815***			
	(0.0822)			
Prot_pc25	-0.00391***	-0.00431**	-0.00464***	-0.00516***
	(0.00116)	(0.00178)	(0.00163)	(0.00163)
LevelJews33			2.69e-05***	
			(9.08e-06)	
SqJews33				-0.0256***
				(0.00708)
Constant	-2.612***	-7.613***	-6.273***	-7.012***
	(0.462)	(0.372)	(0.401)	(0.243)
Observations	278	278	278	278
Log-lik	-3361	-5818	-5326	-4469
AIC	6735	11646	10664	8951
BIC	6757	11665	10686	8973

Source: Computed from replication file for PP.

Note: Column (1) replicates the model reported in PP Table VI, Column (4). “LevelJews33” is the number of Jews in 1933. “SqJews33” is the square of the percentage Jewish. Columns (2)-(4) demonstrate the reliance of the PP result on the “Log Jews 33” specification. The sample is restricted to the observations used in PP’s regressions. Additional observations are available by fixing a coding error, but using them does not change the results here. See the appendix text for details.

Table A2.3: Role of large values of the log Jews variable

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Deported	Deported	Deported	Deported	Deported
Pogrom	0.131*	0.173**	0.162**	0.195	0.192
	(0.0724)	(0.0719)	(0.0743)	(0.137)	(0.136)
Logpop33_dep	0.182	0.183**	0.197**	0.200**	0.198**
	(0.128)	(0.0803)	(0.0847)	(0.0864)	(0.0882)
Jewish_pc33	0.0118	0.0526	0.0527	0.0524	0.0503
	(0.0824)	(0.0352)	(0.0393)	(0.0391)	(0.0399)
LogJews33	0.864***	0.880***	0.854***	0.848***	0.859***
	(0.123)	(0.0774)	(0.0840)	(0.0899)	(0.0960)
	-	-	-		-
Prot_pc25	0.00416***	0.00346***	0.00297***	-0.00307**	0.00358***
	(0.00121)	(0.00115)	(0.00108)	(0.00125)	(0.00132)
Constant	-2.168***	-2.404***	-2.418***	-2.430***	-2.443***
	(0.728)	(0.465)	(0.484)	(0.492)	(0.493)
Observations	277	276	275	274	273
Sample	Drop 1 obs	Drop 2 obs	Drop 3 obs	Drop 4 obs	Drop 5 obs
Model	Poisson	Poisson	Poisson	Poisson	Poisson
Log-lik	-3272	-3157	-3125	-3118	-3102
AIC	6557	6325	6262	6247	6215
BIC	6578	6347	6283	6269	6237

Source: Computed from the PP replication file.

Note: Each specification corresponds to the model reported in PP Table VI, Column (4). The first column drops the observation with the largest value of the “log Jews” variable. Columns (2)-(5) drop one additional observation each. Compare to Table A2.2, Column (1), which replicates the PP result.

Table A2.4: Role of largest values of the dependent variable: Deportations

VARIABLES	(1) Deported	(2) Deported	(3) Deported	(4) Deported	(5) Deported	(6) Deported
Pogrom	0.142**	0.127*	0.192	0.182	0.182	0.172
	(0.0706)	(0.0728)	(0.136)	(0.134)	(0.134)	(0.136)
Logpop33_dep	0.241***	0.204*	0.198**	0.157*	0.144*	0.188**
	(0.0841)	(0.119)	(0.0882)	(0.0850)	(0.0873)	(0.0827)
Jewish_pc33	0.0743**	0.0236	0.0503	0.0462	0.0413	0.0516
	(0.0348)	(0.0738)	(0.0399)	(0.0365)	(0.0395)	(0.0348)
LogJews33	0.815***	0.830***	0.859***	0.918***	0.945***	0.866***
	(0.0822)	(0.115)	(0.0960)	(0.0898)	(0.0903)	(0.0864)
Prot_pc25	-0.00391***	-0.00320***	-0.00358***	-0.00372***	-0.00387***	-0.00447***
	(0.00116)	(0.00105)	(0.00132)	(0.00135)	(0.00145)	(0.00141)
Constant	-2.612***	-2.256***	-2.443***	-2.344***	-2.356***	-2.338***
	(0.462)	(0.679)	(0.493)	(0.496)	(0.563)	(0.535)
Observations	278	276	273	270	267	265
Sample	Replication	Drop 1 p.c.	Drop 2 p.c.	Drop 3 p.c.	Drop 4 p.c.	Drop 5 p.c.
Model	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
Log-lik	-3361	-3193	-3102	-2965	-2874	-2751
AIC	6735	6397	6215	5943	5760	5513
BIC	6757	6419	6237	5964	5782	5535

Source: Computed from the replication file for PP.

Note: Each specification corresponds to the model reported in PP Table VI, Column (4). Column (1) replicates the model reported in PP Table VI, Column (4). Column (2) drops the one percent of observation with the largest value of the dependent variable. Columns (3)-(6) each drop one percent more.

Table A2.5: Alternative count models for deportations

VARIABLES	(1) Deported	(2) Deported	(3) Deported	(4) Deported	(5) Deported	(6) Deported
Pogrom	0.142**	0.142	0.160**	0.122	0.160	0.122
	(0.0706)	(0.104)	(0.0770)	(0.149)	(0.104)	(0.148)
Logpop33_dep	0.241***	0.241*	0.198**	0.162*	0.198	0.162
	(0.0841)	(0.137)	(0.0944)	(0.0945)	(0.166)	(0.158)
Jewish_pc33	0.0743**	0.0743	0.0649	0.00884	0.0649	0.00884
	(0.0348)	(0.0870)	(0.0419)	(0.0501)	(0.113)	(0.112)
LogJews33	0.815***	0.815***	0.837***	0.844***	0.837***	0.844***
	(0.0822)	(0.133)	(0.0910)	(0.0961)	(0.161)	(0.147)
Prot_pc25	-0.00391***	-0.00391***	-0.00389***	-0.00522**	-0.00389***	-0.00522**
	(0.00116)	(0.00119)	(0.00113)	(0.00208)	(0.00107)	(0.00212)
lndelta			2.968***		2.968***	
			(0.150)		(0.147)	
lnalpha				-0.577**		-0.577**
				(0.227)		(0.243)
Constant	-2.612***	-2.612***	-2.254***	-1.794***	-2.254**	-1.794*
	(0.462)	(0.793)	(0.523)	(0.647)	(0.937)	(1.019)
Observations	278	278	278	278	278	278
Sample	Replication	Replication	Replication	Replication	Replication	Replication
Model	Poisson	Poisson	NB1	NB2	NB1	NB2
SE	Robust	Bootstrap	Robust	Robust	Bootstrap	Bootstrap
Log-lik	-3361	-3361	-1246	-1280	-1246	-1280
AIC	6735	6735	2505	2574	2505	2574
BIC	6757	6757	2531	2600	2531	2600

Source: Estimated from the replication file for PP

Notes: Column (1) replicates the model reported in PP Table VI, Column (4). The bootstrap standard errors were estimated using 200 replications. NB1 and NB2 are two different versions of the negative binomial model; see the Appendix text for the definitions.

Table A2.6: OLS models for Deportations

VARIABLES	(1) LnDeport1	(2) LnDeport1	(3) LnDeport1
Pogrom	0.165	0.138	0.292*
	(0.138)	(0.139)	(0.166)
Logpop33_dep		0.0951	1.022***
		(0.157)	(0.0518)
Jewish_pc33		-0.0160	0.450***
		(0.0869)	(0.0789)
LogJews33	1.047***	0.982***	
	(0.0341)	(0.146)	
Prot_pc25		-0.00350**	-0.00360*
		(0.00174)	(0.00206)
Constant	-1.783***	-2.146**	-6.923***
	(0.190)	(0.890)	(0.489)
Observations	278	278	278
Adjusted R-squared	0.774	0.778	0.691
Sample	PP append	Replication	Replication
Model	OLS	OLS	OLS
SE	Robust	Robust	Robust
Weights	None	None	None

Source: Computed from the PP replication file.

Notes: The dependent variable in each case is ln(Deportations+1), where Deportations is the number of Jews deported from a city. Column (1) replicated PP Appendix Table 10, Column (2). This model is not really a robustness check for the main Poisson result (see PP Table VI, Column (4) because it does not have all the same controls. Our Column (2) here includes all the controls used in the Poisson model. Column (3) drops the “ln Jews” variable.

Table A2.7: OLS models for proportion of Jews deported

VARIABLES	(1) Prop_Deport	(2) Prop_Deport	(3) Prop_Deport	(4) Prop_Deport	(5) Prop_Deport	(6) Prop_Deport
Pogrom	10.09**	9.253***	4.744**	3.999*	4.751	5.664
	(4.067)	(2.717)	(2.102)	(2.198)	(6.382)	(6.721)
Jewish_pc33	-0.640	-3.596	-0.272	2.634*	-3.206	-0.448
	(1.134)	(2.902)	(0.864)	(1.476)	(2.766)	(3.028)
LogJews33		4.607		-6.544**		-5.818
		(4.263)		(3.254)		(4.931)
Logpop33_dep	1.191*	-3.601	1.539***	8.273**	-1.321	4.173
	(0.701)	(4.605)	(0.558)	(3.233)	(2.520)	(4.741)
Prot_pc25	-0.0848*	-0.0846**	-0.126***	-0.135***	-0.185*	-0.186*
	(0.0452)	(0.0413)	(0.0384)	(0.0402)	(0.0984)	(0.0987)
Constant	16.46*	44.60*	18.26**	-15.83	55.46*	27.16
	(9.567)	(26.82)	(8.458)	(16.76)	(30.63)	(33.64)
Observations	278	278	278	278	278	278
Adjusted R-squared	0.066	0.074	0.102	0.117	0.009	0.009
Sample	Replication	Replication	Replication	Replication	Replication	Replication
Model	OLS	OLS	OLS	OLS	OLS	OLS
SE	Robust	Robust	Robust	Robust	Robust	Robust
Weights	Population	Population	Jews	Jews	None	None
Log-lik	-1101	-1101	-1179	-1179	-1101	-1101
AIC	2210	2210	2366	2366	2210	2210
BIC	2225	2225	2381	2381	2225	2225

Source: Computed from the PP replication file.

Notes: The dependent variable is the proportion of Jews who were deported. Column (1) replicates PP Appendix Table A12, Column (1). The PP table misplaced the decimal in Pogrom and switched the point-estimates and standard errors for the Jewish and Protestant percentages; our table correctly reports regression estimates. Columns (3) and (4) weight by the Jewish population, not the total population. Columns (5) and (6) are unweighted. See the text for discussion of the appropriate weights.

Table A3.1: Role of largest values of the letters variable

VARIABLES	(1) Letters	(2) Letters	(3) Letters	(4) Letters	(5) Letters	(6) Letters
Pogrom	0.369**	0.331**	0.245*	0.245	0.192	0.165
	(0.144)	(0.139)	(0.136)	(0.152)	(0.150)	(0.152)
Logpop33_dep	0.848***	0.787***	0.730***	0.690***	0.664***	0.643***
	(0.0419)	(0.0484)	(0.0429)	(0.0437)	(0.0407)	(0.0437)
Jewish_pc33	0.218***	0.210***	0.188***	0.170***	0.161***	0.153***
	(0.0383)	(0.0504)	(0.0523)	(0.0529)	(0.0521)	(0.0520)
Prot_pc25	-0.00532**	-0.00199	-0.00320*	-0.00423**	-0.00253	-0.00211
	(0.00228)	(0.00207)	(0.00194)	(0.00190)	(0.00181)	(0.00185)
Constant	-7.934***	-7.433***	-6.715***	-6.250***	-6.058***	-5.852***
	(0.468)	(0.573)	(0.498)	(0.478)	(0.451)	(0.477)
Observations	325	322	319	316	313	309
Sample	Replication	Drop 1 p.c.	Drop 2 p.c.	Drop 3 p.c.	Drop 4 p.c.	Drop 5 p.c.
Model	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
Log-lik	-690	-659.1	-609.7	-592.1	-563.5	-548
AIC	1390	1328	1229	1194	1137	1106
BIC	1409	1347	1248	1213	1156	1125

Source: Computed from the PP replication file.

Notes: Column (1) replicates the model reported in PP Table VI, Column (5). Columns (2)-(6) drop portions of the sample corresponding to large values of the dependent variable, as noted.

Table A3.2: Alternative models for letters

VARIABLES	(1) Letters	(2) Letters	(3) Letters	(4) Letters	(5) Letters	(6) Letters
Pogrom	0.369**	0.369**	0.291**	0.204	0.291**	0.204
	(0.144)	(0.156)	(0.139)	(0.172)	(0.141)	(0.175)
Logpop33_dep	0.848***	0.848***	0.787***	0.816***	0.787***	0.816***
	(0.0419)	(0.0418)	(0.0460)	(0.0400)	(0.0478)	(0.0436)
Jewish_pc33	0.218***	0.218***	0.204***	0.224***	0.204***	0.224***
	(0.0383)	(0.0606)	(0.0417)	(0.0726)	(0.0702)	(0.0766)
Prot_pc25	-0.00532**	-0.00532*	-0.00396*	-0.00267	-0.00396*	-0.00267
	(0.00228)	(0.00273)	(0.00239)	(0.00189)	(0.00228)	(0.00197)
ln delta			0.914***		0.914***	
			(0.214)		(0.255)	
ln alpha				-0.525***		-0.525***
				(0.201)		(0.199)
Constant	-7.934***	-7.934***	-7.197***	-7.599***	-7.197***	-7.599***
	(0.468)	(0.481)	(0.480)	(0.442)	(0.511)	(0.458)
Observations	325	325	325	325	325	325
Sample	Replication	Replication	Replication	Replication	Replication	Replication
Model	Poisson	Poisson	NB1	NB2	NB1	NB2
SE	Robust	Bootstrap	Robust	Robust	Bootstrap	Bootstrap
Log-lik	-690	-690	-581	-569	-581	-569
AIC	1390	1390	1175	1151	1175	1151
BIC	1409	1409	1197	1173	1197	1173

Notes: Column (1) replicates the model reported in PP Table VI, Column (5). See notes to Table A2.5

Table 3.3: OLS models for the log of Letters

VARIABLES	(1) LnLetters1	(2) LnLetters1
Pogrom	0.103	0.0676
	(0.0841)	(0.0818)
Logpop33_dep	0.451***	0.495***
	(0.0308)	(0.0329)
Jewish_pc33		0.122***
		(0.0423)
Prot_pc25		-0.00158
		(0.00112)
Constant	-3.431***	-3.917***
	(0.283)	(0.317)
Observations	325	325
Adjusted R-squared	0.517	0.547
Sample	Replication	Replication
Model	OLS	OLS
SE	Robust	Robust
Weights	None	None

Source: Computed from the PP replication file.

Note: The dependent variable ln(Letters +1), where Letters is the number of anti-Semitic letters to Der Stürmer. Column (1) replicates PP Appendix Table 11, Column (1). This model is not really a robustness check for the poisson model reported in PP Table VI, Column (5), as it is missing controls used in the poisson model. Column (2) reports a model with the same regressors as PP's Poisson model.

Table A3.4: OLS models for Letters per ten thousand population	(1)	(2)
VARIABLES	Letters10	Letters10
Pogrom	0.254***	-0.0660
	(0.0971)	(0.408)
Logpop33_dep	-0.167***	-0.256*
	(0.0402)	(0.136)
Jewish_pc33	0.283***	0.266
	(0.0694)	(0.178)
Prot_pc25	-0.00433**	0.00121
	(0.00194)	(0.00528)
Constant	2.511***	3.430***
	(0.499)	(1.202)
Observations	325	325
Adjusted R-squared	0.166	0.035
Sample	Replication	Replication
Model	OLS	OLS
SE	Robust	Robust
Weights	Pop33	None

Source: Computed from the PP replication file.

Notes: The dependent variable is the number of letters per ten thousand city population. Column (1) replicates PP Appendix Table 12, Colum (4). Column (2) is the same model but does not weight the regression. See the text for discussion of the weights. The PP Appendix Table reversed the point-estimates and standard errors for the percentage Jewish and percentage Protestant variables. Our table is correct.

Table A4.1: The first p.c. of the six outcome variables

VARIABLES	(1) PCA_stnd	(2) PCA_stnd	(3) PCA_stnd	(4) PCA_stnd
Pogrom	0.290**	0.0588	0.106	0.150*
	(0.132)	(0.0670)	(0.0800)	(0.0799)
Logpop33_stnd	-0.0875	-0.0433	-0.0355	-0.0646*
	(0.0646)	(0.0296)	(0.0396)	(0.0338)
Jewish33_stnd	0.0215	0.0601	0.114	-0.0377
	(0.0971)	(0.0439)	(0.0781)	(0.0654)
Prot25_stnd	0.284***	0.254***	0.204***	0.306***
	(0.0757)	(0.0322)	(0.0399)	(0.0383)
Constant	-0.0801	-0.341***	-0.201***	-0.382***
	(0.106)	(0.0668)	(0.0723)	(0.0657)
Observations	311	311	291	247
R-squared		0.056		
Adjusted R-squared	0.052		0.080	0.228
Sample	Replication	Replication	Outlier-S	-Bavaria
Model	OLS	QR	OLS	OLS

Note: Computed from the replication file for PP.

Note: The dependent variable in each regression is the first principle component of the six outcome variables used in PP Table VI. Column (1) reproduces PP Table VII, Col (1). Column (2) estimates that model as a quantile (median) regression. Column (3) drops observations that in column (1) have a “studentized” residual that exceeds, in absolute value, 2.0. Column (4) drops the Bavarian observations.

Table A5.1: 1920s Pogroms and PCA with province FE

	(1)	(2)	(3)	(4)
VARIABLES	Pogrom20s	Pogrom20s	PCA_stnd	PCA_stnd
Pogrom	0.0655***	0.0184	0.290**	0.158
	(0.0199)	(0.0571)	(0.132)	(0.252)
<u>Pogrom proxy x province</u>				
Bavaria		0.124*		0.774*
		(0.0747)		(0.455)
Brandenburg		-0.0102		-0.330
		(0.0636)		(0.512)
Hannover		0.0826		0.257
		(0.145)		(0.379)
Hesse		0.0301		0.241
		(0.0882)		(0.299)
Hesse-Nassau		-0.0157		-0.168
		(0.0787)		(0.421)
Braunschweig		0.905***		-0.0922
		(0.0643)		(0.250)
Mecklenburg		-0.0422		-0.272
		(0.0568)		(0.248)
Pommerania		-0.0901		0.00769
		(0.0923)		(0.289)
Rhineland		0.0435		-0.0506
		(0.0683)		(0.260)
Saxony		0.0692		-0.256
		(0.102)		(0.327)
Silesia		-0.225		-0.0700
		(0.238)		(0.389)
Westphalia		-0.0679		-0.177
		(0.0647)		(0.273)
Wuerttemberg		0.0421		0.129
		(0.0990)		(0.308)
Constant	-0.378***	-0.334**	-0.0801	-0.132
	(0.145)	(0.138)	(0.106)	(0.212)
Test for the null that all interactions are zero:				
F-stat		236.23		7.7
p-value		0		0
Observations	319	319	311	311
Adjusted R-squared	0.092	0.089	0.052	0.506

Source: Computed from PP replication data.

Notes: All models include but do not report province fixed-effects as well as PP's standard controls for log population and the percentage Jewish and Protestant. In Columns (1) and (3) the "Pogrom" estimate is as in PP. In (2) and (4) it is for the baseline province (Baden). See Appendix A.5 for additional discussion.

Table A5.2 Fixed effects for three PP models

	1		2		3		4	
	1924 election		1924 election		1928 election		1928 election	
Pogrom	0.023*	(0.009)	0.010	(0.016)	0.015**	(0.005)	0.003	(0.017)
Pogrom x province								
Pogrom x Bayern			0.040	(0.038)			0.038	(0.024)
Pogrom x Brandenburg			-0.021	(0.058)			-0.014	(0.019)
Pogrom x Hannover			0.013	(0.024)			0.023	(0.027)
Pogrom x Hessen			0.015	(0.024)			0.022	(0.018)
Pogrom x Hessen-Nassau			0.011	(0.021)			-0.014	(0.026)
Pogrom x Land Braunschweig			0.005	(0.019)			0.013	(0.017)
Pogrom x Mecklenburg			-0.017	(0.018)			0.001	(0.017)
Pogrom x Pommern			0.016	(0.028)			-0.013	(0.017)
Pogrom x Rheinprovinz			-0.005	(0.016)			0.002	(0.017)
Pogrom x Sachsen			-0.003	(0.030)			0.005	(0.019)
Pogrom x Schlesien			0.021	(0.029)			0.004	(0.021)
Pogrom x Westfalen			0.006	(0.021)			-0.005	(0.019)
Pogrom x Wuerttemberg			0.016	(0.022)			0.014	(0.017)
Constant	-0.022	(0.024)	-0.008	(0.028)	0.020	(0.017)	0.028	(0.018)
Observations	325		325		325		325	
Adjusted R-squared	0.526		0.511		0.385		0.384	
F-statistic			1.4				4.2	
p-value			.158				0.0	

Source: Computed from PP replication file.

Notes: The dependent variable in olumns (1) and (2) is the vote share for the DVFP in 1924; in (3) and (4), the vote share for the Nazis in 1928. Compare PP Table VI Columns (2) and (3). All models include a full set of province fixed effects as well as “baseline” controls. The reference province for the interactions is Baden. Given this specification, the effect of the Bavarian interaction in (2), for example. is .158 + .774; that is, the effect of any province is the interaction term plus the main effect for Pogrom. The F-statistic is for the null hypothesis that all interaction terms are jointly zero.

Table A5.3: Fixed effects models with Deportations and Letters

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Deported		Deported		Letters		Letters	
VARIABLES	coef	se	coef	se	coef	se	coef	se
Pogrom	0.129	(0.0968)	0.152	(0.263)	0.386	(0.175)	0.254	(0.609)
Interactions of state FE and the Pogrom proxy:								
Bavaria			-0.0699	(0.301)			-0.201	(0.717)
Brandenburg			-0.498	(0.630)			0.570	(0.902)
Hannover			-0.639	(0.317)			-0.893	(0.729)
Hesse			-0.564	(0.628)			0.135	(1.004)
Hesse-Nassau			0.498	(0.322)			-1.286	(0.768)
Braunschweig			-0.765	(0.274)			14.16	(1.169)
Mecklenburg			0.281	(0.273)			12.25	(0.949)
Pommerania			0.116	(0.363)			-0.0447	(0.779)
Rhine province			0.0880	(0.287)			0.556	(0.823)
Saxony			0.0730	(0.303)			0.525	(0.679)
Silesia			-0.407	(0.302)			0.0438	(0.674)
Westphalia			-0.220	(0.372)			0.0816	(0.737)
Wuerttemberg			-0.581	(0.410)			1.563	(1.257)
Constant	-1.451	(0.543)	-1.304	(0.579)	-8.337	(0.508)	-8.442	(0.689)
Observations	278		278		325		325	
Test interactions = 0			87.41				482.76	
p-value for test			0				0	

Note: The models reported in Columns (1)-(4) are expanded versions of PP Table VI, Column (4). The models in Columns (5)-(8) parallel PP Table VI, Column (5). All models in this table include the controls PP include in their Table VI, along with province fixed effects. The last two lines in Columns (3) and (7) report chi-square tests for the null hypothesis that the state fixed effects interacted with the pogrom proxy are jointly zero. The reference province is Baden. See the note to table A5.1 for the interpretation of these estimates.

Table A6.1: 1924 votes for extremist parties

	Dep variable	Pogrom Estimate	SE	Observations	Adjusted R-squared	Sample	Model
(1)	DVFP24	0.0147	(0.0110)	325	0.080	Replication	OLS
(2)	DVFP24	0.00380	(0.00826)	139	0.464	Prussia	OLS
(3)	DVFP24	0.0126*	(0.00686)	257	0.345	-Bavaria	OLS
(4)	DVFP24	-0.000209	(0.00620)	325		Replication	QR
(5)	DNVP24	-0.0267**	(0.0131)	325	0.372	Replication	OLS
(6)	DNVP24	-0.0435**	(0.0179)	139	0.578	Prussia	OLS
(7)	DNVP24	-0.0378***	(0.0143)	257	0.403	-Bavaria	OLS
(8)	DNVP24	-0.00257	(0.00951)	325		All	QR
(9)	DVFPminDNVP	0.0414**	(0.0167)	325	0.112	Replication	OLS
(10)	DVFPminDNVP	0.0473**	(0.0192)	139	0.329	Prussia	OLS
(11)	DVFPminDNVP	0.0504***	(0.0151)	257	0.161	-Bavaria	OLS
(12)	DVFPminDNVP	0.0119	(0.0132)	325		All	QR
(13)	DVFP_DNVP	-0.0120	(0.0175)	325	0.344	Replication	OLS
(14)	DVFP_DNVP	-0.0397*	(0.0202)	139	0.662	Prussia	OLS
(15)	DVFP_DNVP	-0.0253	(0.0166)	257	0.499	-Bavaria	OLS
(16)	DVFP_DNVP	-0.0186	(0.0167)	325		All	QR
(17)	KPD24	0.00915	(0.00873)	325	0.102	Replication	OLS
(18)	KPD24	0.00185	(0.0130)	139	0.097	Prussia	OLS
(19)	KPD24	0.0190*	(0.0100)	257	0.111	-Bavaria	OLS
(20)	KPD24	0.00609	(0.0110)	325		All	QR

Source: Computed from the PP replication file.

Notes: The Estimate column give the estimated coefficient for POG1349 in regressions for each sample and each dependent variable. Each model corresponds to the specifications reported in PP Table VI, Columns (2) and (3); we use the same regressors, just different political parties as the dependent variable. In some cases we also subset the data, as noted. The models all include a constant term as well as the standard controls used in PP: the percentage Catholic, the percentage Jewish, and the population of the election district. The dependent variable DVFPminDNVP is the vote share for the DVFP minus the vote share for the DNVP. DVFP_DNVP is the sum of those two vote shares. For a summary of the parties see Table A6.5

Table A6.2: 1924 votes for pro-Republic parties

	Dep variable	Pogrom Estimate	SE	Observations	Adjusted R-squared	Sample	Model
1	DVP24	0.00955	(0.00799)	325	0.233	Replication	OLS
2	DVP24	0.0233**	(0.0102)	139	0.276	Prussia	OLS
3	DVP24	0.0171**	(0.00860)	257	0.296	-Bavaria	OLS
4	DVP24	0.0167	(0.0109)	325		Replication	QR
5	DDP24	0.0109**	(0.00544)	325	0.265	Replication	OLS
6	DDP24	0.00915	(0.00750)	139	0.266	Prussia	OLS
7	DDP24	0.0151**	(0.00675)	257	0.249	-Bavaria	OLS
8	DDP24	0.00682	(0.00523)	325		Replication	QR
9	DVP_DDP	0.0205*	(0.0110)	325	0.306	Replication	OLS
10	DVP_DDP	0.0324**	(0.0155)	139	0.324	Prussia	OLS
11	DVP_DDP	0.0322**	(0.0126)	257	0.374	-Bavaria	OLS
12	DVP_DDP	0.0294**	(0.0116)	325		Replication	QR
13	SPD24	0.0109	(0.0110)	325	0.353	Replication	OLS
14	SPD24	0.00194	(0.0126)	139	0.582	Prussia	OLS
15	SPD24	0.00894	(0.0118)	257	0.410	-Bavaria	OLS
16	SPD24	0.00720	(0.0145)	325		Replication	QR
17	Center24	0.0412	(0.0275)	325	0.522	Replication	OLS
18	Center24	0.0112	(0.0151)	139	0.919	Prussia	OLS
19	Center24	-0.0123	(0.0145)	257	0.913	-Bavaria	OLS
20	Center24	0.00371	(0.00982)	325		Replication	QR
21	BVP24	-0.0764***	(0.0258)	325	0.094	Replication	OLS
22	BVP24	-0.00121	(0.00143)	139	0.018	Prussia	OLS
23	BVP24	-0.0278	(0.0194)	257	0.022	-Bavaria	OLS
24	BVP24	-0.000488	(0.00246)	325		Replication	QR
25	CathParty24	-0.0352	(0.0220)	325	0.745	Replication	OLS
26	CathParty24	0.0100	(0.0151)	139	0.919	Prussia	OLS
27	CathParty24	-0.0401	(0.0270)	257	0.766	-Bavaria	OLS
28	CathParty24	-0.000829	(0.00855)	325		Replication	QR

Source: Computed from the PP working file.

Note: See the notes to Table A5.1

Table A6.3: 1928 Votes

	Dep variable	Pogrom Estimate	SE	Obs	Adjusted R-squared	Sample	Model
(1)	NSDAP28	0.0142**	(0.00567)	325	0.043	Replication	OLS
(2)	NSDAP28	0.00107	(0.00454)	139	0.044	Prussia	OLS
(3)	NSDAP28	0.00685	(0.00435)	257	0.057	-Bavaria	OLS
(4)	NSDAP28	0.00294	(0.00283)	325		Replication	QR
(5)	DVP28	0.0128	(0.00774)	325	0.290	Replication	OLS
(6)	DVP28	0.0191*	(0.0112)	139	0.323	Prussia	OLS
(7)	DVP28	0.0158*	(0.00832)	257	0.368	-Bavaria	OLS
(8)	DVP28	0.0193**	(0.00748)	325		Replication	QR
(9)	KPD28	0.0101	(0.00724)	325	0.103	Replication	OLS
(10)	KPD28	0.00276	(0.0116)	139	0.071	Prussia	OLS
(11)	KPD28	0.0174**	(0.00862)	257	0.109	-Bavaria	OLS
(12)	KPD28	0.00518	(0.00724)	325		Replication	QR
(13)	SPD28	0.00463	(0.0134)	325	0.392	Replication	OLS
(14)	SPD28	-0.00866	(0.0143)	139	0.586	Prussia	OLS
(15)	SPD28	0.0105	(0.0148)	257	0.470	-Bavaria	OLS
(16)	SPD28	-0.00290	(0.0158)	325		Replication	QR
(17)	Center28	-0.0128	(0.0128)	325	0.858	Replication	OLS
(18)	Center28	0.000769	(0.0140)	139	0.899	Prussia	OLS
(19)	Center28	-0.00961	(0.0153)	257	0.885	-Bavaria	OLS
(20)	Center28	-0.00229	(0.00696)	325		Replication	QR

Source: Computed from the PP replication file.

Note: See the notes to Table A6.1

Table A6.4: 1933 Votes

	Dep variable	Pogrom Estimate	SE	Obs	Adjusted R-squared	Sample	Model
(1)	NSDAP33	-0.0113	(0.0125)	325	0.426	Replication	OLS
(2)	NSDAP33	-0.0124	(0.0163)	139	0.554	Prussia	OLS
(3)	NSDAP33	-0.0256*	(0.0142)	257	0.459	-Bavaria	OLS
(4)	NSDAP33	-0.0104	(0.0157)	325		Replication	QR
(5)	DVP33	0.00235	(0.00226)	325	0.111	Replication	OLS
(6)	DVP33	0.00174	(0.00463)	139	0.045	Prussia	OLS
(7)	DVP33	0.00236	(0.00294)	257	0.113	-Bavaria	OLS
(8)	DVP33	0.00390***	(0.00105)	325		Replication	QR
(9)	KPD33	0.0125*	(0.00746)	325	0.135	Replication	OLS
(10)	KPD33	0.00491	(0.0114)	139	0.123	Prussia	OLS
(11)	KPD33	0.0213**	(0.00869)	257	0.150	-Bavaria	OLS
(12)	KPD33	0.00692	(0.00873)	325		Replication	QR
(13)	SPD33	0.00116	(0.00966)	325	0.365	Replication	OLS
(14)	SPD33	0.00211	(0.0105)	139	0.575	Prussia	OLS
(15)	SPD33	0.00543	(0.0107)	257	0.441	-Bavaria	OLS
(16)	SPD33	-0.00672	(0.0103)	325		Replication	QR
(17)	Center33	-0.00295	(0.0101)	325	0.865	Replication	OLS
(18)	Center33	0.00482	(0.0134)	139	0.886	Prussia	OLS
(19)	Center33	0.00118	(0.0118)	257	0.884	-Bavaria	OLS
(20)	Center33	6.95e-05	(0.00664)	325		Replication	QR

Source: Computed from the PP replication file.

Note: See the notes to Table A6.1

Table A6.5: An overview of the political parties mentioned in the paper and their abbreviations

NSDAP: *Nationalsozialistische Deutsche Arbeiterpartei*, formal name of the Nazi party

DNVP: *Deutschnationale Volkspartei*. Right-wing.

DFVP: *Deutschvölkische Freiheitspartei*. Right-wing.

DVP: *Deutsche Volkspartei*. Center-right.

DDP: *Deutsche Demokratische Partei*. Liberal.

SPD: *Sozialdemokratischepartei Deutschlands.* Social Democratic

KPD: *Kommunistische Partei Deutschlands*. Communist Party

Center: *Zentrumspartei*. Centrist, Catholic. Operated only outside Bavaria

BVP: *Bayerische Volkspartei*. Center-right, Catholic. Operated only in Bavaria.

The "Weimar coalition" is the SPD, Center, and DDP.

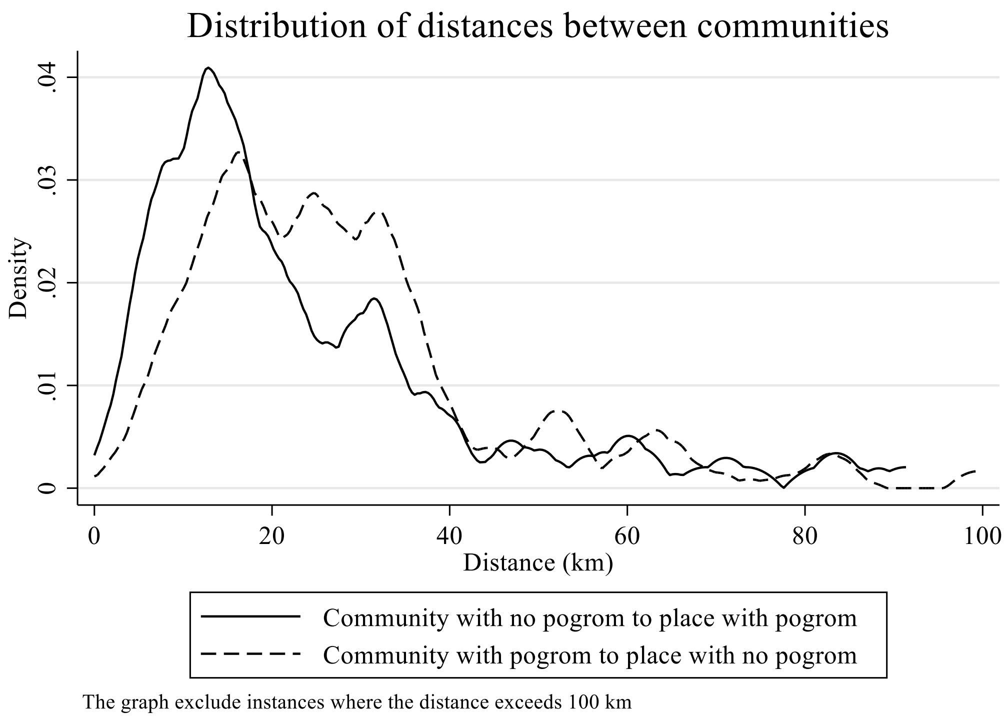
Table A7.1: Summary of PP matching checks

	(1) Pog20s	(2) Nazi 28	(3) DVFP 24	(4) Deports	(5) Sturmer	(6) Pog30s
Replication of VV Table VI						
Matching model						
ATT	0.0744***	0.0133***	0.0203**	161.7***	2.386***	0.103*
	(0.0182)	(0.00486)	(0.0102)	(41.33)	(0.570)	(0.0553)
Observations	320	325	325	278	325	278
Geographic matching						
ATT	0.0819***	0.0116**	0.0238***	195.8***	2.864***	0.152**
	(0.0162)	(0.00456)	(0.00746)	(33.55)	(0.579)	(0.0677)
Observations	320	325	325	278	325	278
Same models dropping Bavaria						
Matching model						
ATT	0.0500***	0.00404	0.00870	173.2***	2.096***	0.124*
	(0.0176)	(0.00406)	(0.00547)	(46.59)	(0.648)	(0.0660)
Observations	253	257	257	223	257	224
Geographic matching						
ATT	0.0595***	0.00131	0.0167***	217.4***	2.508***	0.165**
	(0.0173)	(0.00399)	(0.00403)	(34.62)	(0.666)	(0.0720)
Observations	253	257	257	223	257	224

Matching estimates without 5 largest values of dependent variable

Model/dependent variable	Deportations	Stuermer letters
Match	66.627	0.956
	(23.883)	(0.361)
	275	320
Geographic match	102.378	1.559
	(18.876)	(0.38)
	275	320

Figure A8.1



Note: The dashed line plots the probability density for the distance between a town with a medieval pogrom and the nearest town without one. The solid line plots the density of the distance from a town with no pogrom to the nearest town with one. Sample limited to the 325 observations for which the pogrom proxy is defined. The two distance distributions are not identical because 72 percent of communities are identified as experiencing a pogrom. Often the nearest community to such a place is another community that experienced a pogrom.

Table B1.1: Sensitivity to outliers, main results, BF

VARIABLES	(1) Nazi_pc_stnd	(2) Nazi_pc_stnd	(3) Nazi_pc_stnd	(4) Nazi_pc_stnd	(5) Nazi_pc_stnd	(6) Nazi_pc_stnd
Clubs_all_pc	0.160***	0.0713	0.117***	0.106**	0.0988***	0.0878*
	(0.0538)	(0.0582)	(0.0438)	(0.0425)	(0.0345)	(0.0479)
LnPop25	0.175***	0.167***	0.159***	0.166***	0.184***	0.125**
	(0.0542)	(0.0554)	(0.0528)	(0.0525)	(0.0459)	(0.0555)
Cath_pc25	-0.934***	-0.896***	-0.888***	-0.838***	-0.866***	-1.143***
	(0.164)	(0.175)	(0.161)	(0.155)	(0.125)	(0.159)
BCollar_pc25	-2.774***	-2.272***	-2.671***	-2.592***	-2.438***	-2.435***
	(0.478)	(0.439)	(0.458)	(0.431)	(0.343)	(0.464)
Constant	-0.685	-0.830	-0.497	-0.609	-0.949*	-0.187
	(0.670)	(0.711)	(0.638)	(0.624)	(0.558)	(0.631)
Observations	227	227	225	223	211	204
Adjusted R-squared	0.214		0.200	0.204	0.309	0.252
Model	Replication	Replication	1 % resid	2 % resid	RStud>2	-Bavaria
Estimator	OLS	QR [median]	OLS	OLS	OLS	OLS
Mean (med) dep var	0.00828	-0.221	-0.0261	-0.0554	-0.169	-0.109
Reg beta	0.252		0.193	0.185	0.206	0.148

Source: Computed from the replication file for BF

Note: Column (1) replicates the regression reported in BF Table 4, Panel A, Column (4). Column (2) estimates that model by quantile (median) regression. Columns (3) and (4) drop 1 and 2 percent of the observations corresponding to the largest absolute values of the residuals in Column (1). Column (5) drops observations for which the absolute value of the “studentized” residual from Column (1) exceeds 2. Column (6) drops all observations from Bavaria. “Reg beta” is the standardized regression coefficient for the clubs variable.

Table B3.1: Summary of stability index computations

State	Obs	1	2	3
Anhalt	1	**2.02**	**2.65**	**1.48**
Baden	17	**-1.28**	-0.57	-1.97
Bavaria	23	-1.31	-0.48	**-0.45**
Braunschweig	1	-1.77	-0.93	-1.38
Hamburg	1	-1.91	-1.08	-2.03
Hesse	4	**1.56**	**2.15**	**0.87**
Lippe	2	**1.07**	**1.76**	**1.36**
Mecklenburg-Schwerin	1	-2.48	-1.65	-2.27
Mecklenburg-Strelitz	1	**-0.93**	**-0.15**	-0.78
Oldenburg	2	**-0.92**	**-0.14**	**-0.58**
Prussia	119	**1.29**	**1.95**	**1.11**
Saxony	21	**-1.22**	**-0.45**	**-0.70**
Thuringia	12	-2.23	-1.37	-1.67
Wuerttemberg	20	-2.52	-1.68	-2.60
Medians (excl Prussia)		-1.31	-0.48	-0.78

Note: The table reports the values of the stability index that underlies the binary stability indication used in BF Table 7. Column (1) is the index as computed in BF. (2) Computes the index by state. (3) Drops the third element from the index. The binary indicator as defined in BF is actually all values exceeding the median (where the median excludes Prussia). Bold values indicate states that are “stable” using this version of the binary stability definition. See appendix for discussion.

Table B3.2: Using alternate definitions of the stability index

VARIABLES	(1) Nazi_entry	(2) Nazi_entry	(3) Nazi_entry	(4) Nazi_entry	(5) Nazi_entry	(6) Nazi_entry
Clubs_all_pc	-0.0116	0.349***	0.000914	0.278**	0.147	0.0999
	(0.0619)	(0.128)	(0.0594)	(0.121)	(0.106)	(0.147)
LnPop25	0.114	0.192	0.115	0.209*	0.132	0.0324
	(0.108)	(0.134)	(0.140)	(0.111)	(0.124)	(0.125)
Cath_pc25	-0.704	-0.525	-1.390**	-0.620*	-0.226	-.998**
	(0.427)	(0.388)	(0.635)	(0.369)	(0.447)	(0.442)
BCollar_pc25	0.391	-0.272	-0.320	0.264	-1.038	-0.553
	(1.382)	(1.929)	(1.727)	(1.607)	(2.166)	(1.427)
Constant	-1.005	-2.239	-0.721	-2.434	-0.711	-0.0791
	(1.363)	(1.833)	(1.748)	(1.542)	(1.931)	(1.683)
Observations	48	58	31	75	53	54
Adjusted R-squared	0.033	0.108	-0.017	0.084	-0.017	0.055
Model	Replication	Replication	Replication	Replication	Replication	Replication
Coded as	Stable	Unstable	Stable	Unstable	Stable	Unstable
Estimator	OLS	OLS	OLS	OLS	OLS	OLS
Mean (med) dep var	0.0961	0.463	0.173	0.348	0.578	0.009
Reg beta	-0.0229	0.440	0.00216	0.357	0.244	0.141

Source: Computed from the BF replication file.

Note: This table relies on BF's definition of the binary stability indicator. In each case, the binary stability indicator is coded as "stable" for values that exceed the median. Columns (1) and (2) replicate BF Table 7, Columns (4) and (3). Columns (3) and (4) estimate the same models using the binary indicator as defined in BF, but estimating the stability index at the state level. Columns (5) and (6) drop the third element from the index but retain the BF binary definition.

Table B3.3: Using alternate definition of binary stability indicator

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Nazi_entry	Nazi_entry	Nazi_entry	Nazi_entry	Nazi_entry	Nazi_entry
Clubs_all_pc	0.138	0.198	0.147	0.100	0.147	0.100
	(0.0959)	(0.183)	(0.105)	(0.149)	(0.105)	(0.149)
LnPop25	0.183*	-0.0164	0.131	0.0477	0.131	0.0477
	(0.102)	(0.168)	(0.121)	(0.131)	(0.121)	(0.131)
Cath_pc25	-0.422	-1.490**	-0.229	-0.986**	-0.229	-0.986**
	(0.379)	(0.597)	(0.424)	(0.461)	(0.424)	(0.461)
BCollar_pc25	0.0377	-1.511	-1.050	-0.449	-1.050	-0.449
	(1.600)	(1.497)	(2.042)	(1.442)	(2.042)	(1.442)
Constant	-1.707	0.529	-0.700	-0.274	-0.700	-0.274
	(1.456)	(2.206)	(1.818)	(1.752)	(1.818)	(1.752)
Observations	71	35	54	52	54	52
Adjusted R-squared	0.012	0.178	-0.016	0.051	-0.016	0.051
Model	Replication	Replication	Replication	Replication	Replication	Replication
Coded as	Stable	Unstable	Stable	Unstable	Stable	Unstable
Estimator	OLS	OLS	OLS	OLS	OLS	OLS
Mean (med) dep var	0.430	0.0266	0.579	0.00355	0.579	0.00355
Reg beta	0.218	0.265	0.244	0.140	0.244	0.140

Source: Computed from the BF replication file.

Note: Compare the results in this table to Table B3.3. In each case, the binary stability indicator is coded as “stable” for values that *equal or exceed* the median, in contrast to the BF definition used in Table B3.3. Columns (1) and (2) replicate BF but with the changed definition of stable. Columns (3) and (4) do the same but estimate the index at the state level. Columns (5) and (6) use the changed definition of stable but drop the third element from the index.

Table B3.4: State fixed effects with redefined stability indicator

VARIABLES	(1) Nazi_entry	(2) Nazi_entry	(3) Nazi_entry	(4) Nazi_entry
Clubs_all_pc	0.311***	0.245*	0.219***	0.270**
	(0.0266)	(0.131)	(0.0495)	(0.118)
Stable	-0.643	-1.643		
	(1.367)	(1.125)		
Stab_Clubs	-0.322***	-0.107		
	(0.0422)	(0.166)		
Prussia	0.160	0.615**		
	(0.456)	(0.268)		
Clubs_Prussia	-0.210***	-0.178		
	(0.0532)	(0.148)		
Stable x Clubs_all_pc			-0.217***	-0.199
			(0.0610)	(0.136)
Prussia x Clubs_all_pc			-0.171**	-0.202
			(0.0633)	(0.139)
Constant	-0.362	-0.0648	-1.	-0.274
	(1.185)	(0.370)	(0.673)	(0.810)
Observations	225	225	227	227
Adjusted R-squared	0.255	0.251	0.291	0.378
Stable Def	Above Med	GE median	Above median	GE median
Baseline controls	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes
Stability x Baseline controls	Yes	Yes	Yes	Yes

Note: Columns (1) and (3) replicate BF Table 7, Columns (5) and (6). Columns (2) and (4) change the stability indicator to "median and above." BF's Column (5) reports an incorrect value for the adjusted R-square; the value in the table above is correct.

Table B3.5: Using the BF stability index as a regressor

VARIABLES	(1) pcNSentry_PRS_std	(2) pcNSentry_PRS_std	(3) pcNSentry_PRS_std
clubs_all_pc	0.134**	0.139	0.0693**
	(0.0524)	(0.191)	(0.0313)
govt_stability	0.741	-0.574	0.128
	(0.631)	(1.150)	(0.546)
Clubs x stability	-0.0424	-0.0329	-0.000924
	(0.0329)	(0.0701)	(0.0262)
share_cath25	-0.804***	-0.665	-1.086***
	(0.115)	(0.728)	(0.159)
lnpop25	0.136**	0.269*	0.0780*
	(0.0512)	(0.124)	(0.0424)
bcollar25	-1.883***	0.0963	-2.027***
	(0.385)	(1.277)	(0.411)
Stab x share_cath25	-0.239**	-0.125	-0.0355
	(0.0924)	(0.453)	(0.125)
Stab x lnpop25	-0.0224	0.0699	0.0195
	(0.0465)	(0.117)	(0.0406)
Stab x bcollar25	-0.955***	0.219	-0.733**
	(0.282)	(0.629)	(0.287)
Constant	-0.599	-2.547**	0.171
	(0.693)	(1.155)	(0.601)
Observations	225	106	202
R-squared	0.248	0.096	0.267
Sample	All	Not Prussia	Not Bavaria

Note: Column (1) is the regression that underlies BF Appendix Figure A7. BF does not report the regression.

Table B3.6 The net effect of clubs when political stability is a continuous variable

	Net effect of social capital					
	Index	Estimate	SE	Lower CI	Upper CI	N
Anhalt	2.024	0.048	0.023	-0.002	0.099	1
Baden	-1.278	0.189	0.093	-0.013	0.39	17
Bavaria	-1.31	0.19	0.094	-0.014	0.393	23
Braunschweig	-1.772	0.209	0.109	-0.026	0.445	1
Hamburg	-1.908	0.215	0.114	-0.03	0.461	1
Hessen	1.555	0.068	0.016	0.033	0.104	4
Lippe	1.068	0.089	0.022	0.041	0.136	2
Mecklenburg-Schwerin	-2.477	0.239	0.132	-0.046	0.525	1
Mecklenburg-Strelitz	-0.929	0.174	0.082	-0.003	0.351	1
Oldenburg	-0.919	0.173	0.082	-0.003	0.35	2
Prussia	1.287	0.08	0.018	0.041	0.119	119
Saxony	-1.216	0.186	0.091	-0.011	0.383	21
Thuringia	-2.232	0.229	0.124	-0.039	0.497	12
Wuerttemberg	-2.523	0.241	0.134	-0.047	0.53	20

Source: Computed from replication files.

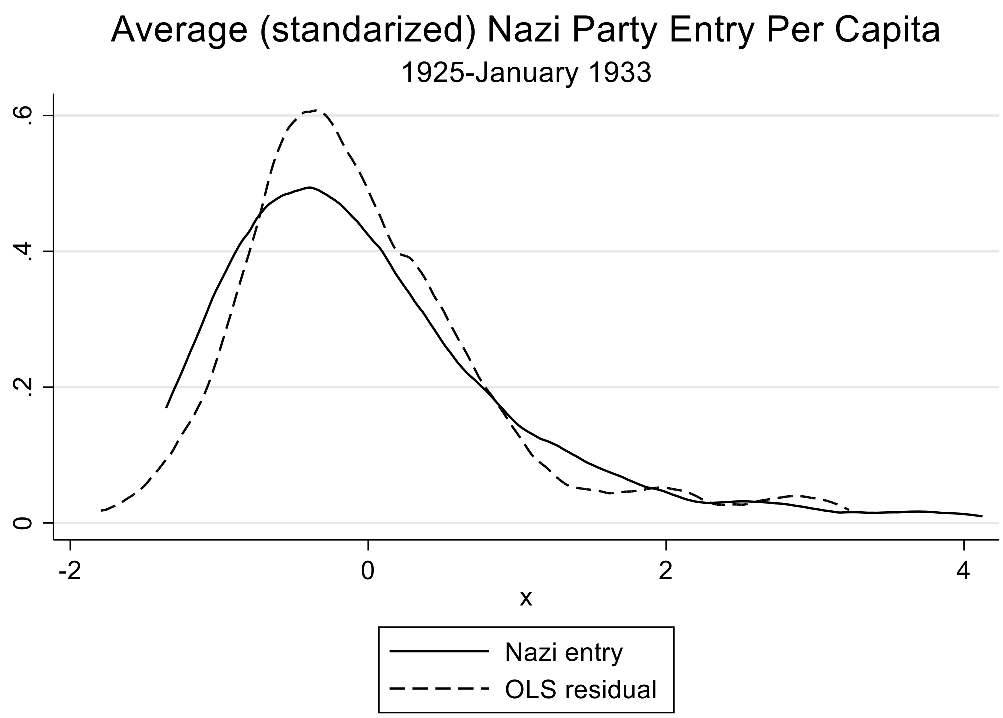
Note: The first column is the state-level mean of the the continuous stability index computed as in BF. “Estimate” is the net effect of clubs per_capita. The lower and upper 95 percent confidence intervals are computed using the correct degrees of freedom from the underlying regression. The calculations are derived from the first regression reported in Table B3.5, which is the regression that underlies BF’s Appendix Figure A7. The estimate is the net effect of social capital in the regression model, evaluated at the index value for each state: that is, it combines the clubs proxy with the product of the clubs/stability interaction and specific stability index value for each state. States in bold have, at the mean value of the index for that state, a net effect of stability for which the confidence interval does not bracket zero. See Appendix Section B.3 for an explanation of why these results differ from BF Figure A7.

Table B3.7: State-level effects of social capital

	(regression)	total clubs effect	Index value	binary indicator	N obs
	(1)				
VARIABLES	Nazi_pc_stnd				
Clubs_all_pc	0.0873***			Stable	119
	(0.0263)				(Prussia)
Clubs interations with:					
Baden	-0.157***	-0.07	-1.28	Stable	17
	(0.00828)	0.022			
Bavaria	0.118***	0.205	-1.31	Unstable	23
	(0.0175)	0.011			
Hesse	0.440***	0.527	1.56	Stable	4
	(0.0854)	0.106			
Lippe	-2.080***	-1.99	1.07	Stable	2
	(0.0696)	0.085			
Oldenburg	-0.998***	-0.911	-0.92	Stable	2
	(0.0989)	0.092			
Saxony	-0.0798***	0.007	-1.22	Stable	21
	(0.0121)	0.015			
Thuringia	0.319***	0.406	-2.23	Unstable	12
	(0.00999)	0.018			
Wuettemberg	0.0309***	0.118	-2.52	Unstable	20
	(0.0101)	0.024			
Constant	-0.0465				
	(0.477)				
Observations	225				
R-squared	0.441				

Note: The regression includes the baseline controls found in BF's regressions. The main effect for the clubs variable corresponds to Prussia, the reference state.

Figure B1.1: The dependent variable and residual in BF



Source: Computed from the BF replication file.

Note: The solid line is the dependent variable in BF Table 3, Panel A, Column (4). Our Appendix Table B1.1, Column (1) replicates the regression. The dashed line in this figure summarizes the OLS residuals from that regression.

Table C1.1 Randomly-selected entries in Avneri and Alicke

PP id	Pogrom	Intensity	town	Random number	Avneri vol/page	Avneri entry	Alicke vol/page	Alicke entry
545	1	0	Schwäbisch Hall	0.001	II/750-53	At the time of the Black Death pogroms most Jews forced to leave or "verbrannt"	III/3734-3738	Pogrom at time of Black Death; many Jews burned and some fled.
644	1	2	Horb/Neckar	0.003	I/370	"Am 20. Dezember 1348, zur Zeit des Schwarzen Todes, erlitten die Juden in Horb den Feuertod."	II/1970-73	„Nach den Pestpogromen von 1348/49 has es jahrhundertenlang – bis auf wenige Ausnahmen – keine Juden in Horb gegeben..."
126	1	0	Gau-Algesheim	0.011	I/269	No mention of the Black Death pogroms	I/1385	"Zum Beginn des 14. Jarhunderts hielten sich vermutlich bereits wenige jüdische Familien im Ort auf; aus diesem Jahrhundert soll auch ein Friedhof ("*Judenkirchove*") stammen. Es gab damals also vermutlich schon eine kleine jüdische Gemeinde in Gau-Algesheim ; dies ging wahrscheinlich in der Zeit der Pestpogrome zugrunde."

18	1	2	Remagen/Rhein	0.012	II/693	"Juden wurden hier vor 1298 und wieder zur Zeit des Schwarzen Todes von Verfolgungen betroffen."	III/3477-3480	Cannot say whether a community as of 1300; "doch urkundlich bestätigt ist die Tatsache, dass einzelne Remagener Juden dem Pestpogrom von 1349 zum Opfer fielen."
795	1	2	Lahr	0.014	I/463	Refers specifically to territorial rulers decisions not to protect the Jews anymore. "Die Juden von Lahr fielen 1349 dem Volke zum Opfer…"	II/2399-23401	No mention of pogroms per se; says first evidence of Jewish community is from just before
560	0	0	Kamen	0.024	I/386	No mention of Black Death pogroms	II/2129-2132	Unclear if a Jewish community
1421	1	5	Rotenburg a.d. Fulda	0.028	II/706	"Zur Zeit des Schwarzen Todes wurden die Rotemburger Juden von der allgemeinen Verfolgung betroffen."	III/3568-3571	No specific mention of Black Death Pogrom; says sometimes tolerated, sometimes persecuted
1302	1	3	Weissenfels/Saale	0.031	II/874-75	"Demnach haben Juden vor dem Schwarzen Tode in Weißenfels gelebt und sind 1350 verbrannt worden."	III/4335-59	reference to Jews being burned by itinerant flagellants
1152	1	0	Schweinfurt	0.037	II/756	(Except for one payment to a Jew, only records concern pogroms, including the Black Death)	III/3756-3762	"betroffen waren"
667	1	3	Jülich	0.042	I/381-2	"Zur Zeit des Schwarzen Todes traf auch die Grafschaft Jülich die große Judenverfolgung. Über das Geschick der jüdischen	II/2105-2109	At the time of the Black Death, Jews were forced out (Vertrieben) and goods confiscated

						Gemeinde in der Stadt Jülich fehlt jede Nachricht.” But evidence that the property of some Jews became the Grafschaft’s property.		
1097	1	5	Reutlingen	0.045	II/694 -696	Emphasizes political conflict before Black Death “Am 8. Dezember 1348 erlagen sie jedoch den Verfolgung zur Seit des Schwarzen Todes.”	III/34 86- 3487	No mention of Black Death pogroms
1253	1	3	Bad Windsheim	0.048	II/909 -10	“Zur Zeit des Schwarzen Todes wurden die Juden zu Windsheim Opfer der allgemeinen Verfolgung.”	II/306 -310	In the pogroms of 1298 and 1348 “fast vollog vernichtet wurde.”
1404	0	0	Bad Homburg	0.052	I/369	“Es ist unsicher, ob während der hier behandelten Zeit Juden in Homberg gelebt haben.”	II/230 -237	Urkunde from 1335 attests to Jewish presence. “Für die folgenden Jahrhunderte liegen keine Daten über das jüdische Leben am Ort vor…”
530	0	0	Greifswald	0.055	I/303- 04	No mention of Black Death pogroms	I/1562 -1566	No mention of Black Death pogroms
1410	1	4	Erfurt	0.058	I/215- 224	Detailed account of Black Death pogrom and the political background to it	I/1135 -1140	Detailed account of pogrom; 100 Jews died in “collective suicide” by setting their houses on fire; the rest left the town.

1114	1	5	Rottweil	0.061	II/720-22	"Die Gemeinde von Rottweil wurde damals vernichtet."	III/3579-3582	"Von der Pestpogromen 1348/49 war die jüdische Gemeinde schwer getroffen."
16	1	0	Lechenich	0.062	I/45	Says specifically there was a pogrom in 1286 or 1287. On Black Death: "Ein weitere Verfolgung traf die Juden zu Lechenich zur Zeit des Scwarzen Todes"	II/2458-2460	Community first mentioned in 13th century. "Ob sich die Juden dauerhaft in Lechenich angesiedelt hatten, ist unbekannt."
784	1	2	Hohebach/Jagst	0.064	I/365-66	"Nach dem Deutzer Memorbuch fielen hier Juden den Verfolgung zur Zeit des Schwarzen Todes zum Opfer"	II/1931-1933	First mention is from 1348, in connection with Black Death pogroms; Jews here victims
493	0	1	Glatz	0.066	I/279	Does not mention of Black Death pogroms	I/1481-83	Not specific about when community first attested; no mention of Black Death pogroms
325	1	3	Erkelenz	0.066	I/225	"Zur des Schwarzen Todes fielen die Juden in Erkelenz der allgemeinen Verfolgung zum Opfer."	I/1141-43	„mit den Pestpogrom wurde die kleine jüdische Gemeinschaft fast vollständig zerstört.“
90	1	3	Bayreuth	0.069	I/60-61	"Für den hier behandelten Zeitraum liegen keine direkten Nachrichten über Juden in Bayreuth vor, wohl aber erfahren wir von Bayreuther Juden, die anderswo, in Straßburg und in Nürnburg ansässig waren." Goes on:	I/370-75	Nothing about the pogroms per se. „Nach der Zeit der mit der Pest einhergehenden Verfolgung Mitte des 14. Jahrhunderts

						Bayreuth was listed as a "Blutort" in a list from the 14th century, but unclear whether this is because of a Black Death pogrom or because of violence against Jews from there in other contexts (he names those contexts). Only evidence for Black Death pogrom in Bayreuth in Avneri makes it clear he has doubts.		übertrug Kaiser Karl IV. das Judenregal – die Steuerhoheit und Gerichtsbarkeit über Juden – dem hiesigen Burggrafen, der „seine" Juden in der Folgezeit schützte."
115	0	0	Beuthen	0.073	I/79	"Unter den Orten der Judenverfolgung zur Zeit des Schwarzen Todes wird im Deutzer Memorbuch ein Bytom gennant, mit dem vielleicht unser Ort gemeint ist."	I/465-467	Does not mention the Black Death pogroms
913	1	2	Miltenberg/Main	0.077	II/540-41	"Die Verfolgung zur Zeit des Schwarzen Todes führte zum Untergang der Gemeinde."	II/2788-2791	"Kurzzeitige Vertreibungen, z.B. in der Pestzeit, überstand die Gemeinde fast schadlos."
891	1	3	Bad Mergentheim	0.078	II/538	Five Jews were killed at the time of the Black Death pogroms	II/256-262	Refers to several pogrom in 14th century, only a few Jews there after 1350
379	1	2	Steinheim/Main	0.079	II/790	"Juden wurden hier von der Verfolgung zur Zeit des Schwarzen Todes betroffen."	III/3933-3935	"wurden vermutlich Opfer der Pestpogrome von 1348/49"
197	1	5	Öhringen	0.081	I/626-27	"Bei den Verfolgungen zur Zeit des Schwarzen Todes	III/3198-3201	Both community and Synagogue

						fanden die Öhringer Juden ihren Untergang.”		destroyed in Black Death pogroms
1037	0	0	Oldenburg	0.081	II-627-28	Does not mention pogroms at the time of the Black Death	III/3203-3208	‘Im Zuge der Pestpogrom Mitte des 14. Jahrhunderts wurden die jüdischen Familien vermutlich von hier vertrieben; denn danach wurden nur noch sehr sporadisch Juden erwähnt”
1427	1	2	Windecken	0.082	II/907-08	“Zur Zeit des Schwarzen Todes wurden sie aus der Stadt vertrieben.”	III/4451-4454	“Während der Pestpogrome von 1348/49 wurden die Juden aus Windecken vertrieben bzw. am Ort erschlagen”
289	1	2	Duderstadt/Eichsfeld	0.089	I/165-77	“Zur Zeit des Schwarzen Todes wurden die Juden verfolgt.”	I/992-96	„Währen der Pestjahre wurden die Juden auch in Duderstadt verfolgt.“
955	1	0	Dormagen	0.090	I/169	“Nach dem Deutzer Memorbuch werden hier Juden von der Verfolgung zur Zeit des Schwarzen Todes betroffen”	I/953-955	„Die Pestverfolgen des Jahres 1349 soll auch die Judenschaft Dormagens zum Opfer gefallen sein.“
351	0	2	Münsterberg	0.094	II/563-64	“Von der Pestverfolgung sind die Münsterger Jüden anscheinend nicht betroffen worden.”	II/2895-2898	No mention of Black Death pogroms.
323	0	0	Emden	0.094	I/209	“Die Tradition der jüdischen Gemeinde in Emden reicht bis in den Anfang des	I/1107-1113	Settlement of Jews in this area is from the second half of the

		14. Jahrhunderts zurück. Urkundliche Nachtrichten aus der hier behandelten Zeit fehlen." Does not mention pogroms at the time of the Black Death	16^{th} century; "die erste Beleg über die Existenz jüdischer Familien in Emden finden sich ab circa 1560."

<u>Source</u>: PP replication files; TF/MK paper

<u>Note</u>: "Pogrom" is the PP pogrom indicator. "Intensity" is the pogrom intensity score from Finley and Koyama (2018). A zero means the place is not in that dataset, because, in the authors' judgement, the sources were too weak to establish what happened to the Jewish community. The other values are: 1, Spared the pogrom; 2, Expelled; 3, Few deaths; 4, Many deaths; 5, Destroyed. For details on the construction of this variable see the appendix to TF/MK. To create the random number, we set the seed to "3222023" (the date we ran the code). The original sample consists of all places PP thinks had a Jewish community in 1349, and for which it defines the pogrom proxy. After sorting by town_id, we used the stata uniform random-number generator (runiform()) to create the numbers reported here. This table includes every place for which the resulting random number had a value less than 0.1.

Table C3.1: Locating directories for large cities

City	Year	Location
Population 100 thousand and more:		
Berlin	1925	
Frankfurt aM	1924	Frankfurt University library
Stuttgart	1925	
Magdeburg	1925	
Barmen	1923, 1927	
Cassel	1925	
Elberfeld	1923	Stadtarchiv Wuppertal
Augsburg	1922	Muenchen Haupstaatsarchiv library
Aachen	1924/25	
Hamborn	1925	
Ludwigshafen	1925	Munich City library 1928
Population 50-100 thousand:		
Bielefeld	1924/25	Bielefeld city library
Brandenburg	1923	City archive
Buer	1925	Wiki.genealogy.net
Coblenz	1925	Dilibri Rheinland Pfalz
Darmstadt	1927	Univesity library
Dessau	1925	City library
Flensburg	1929	Wiki.genealogy.net
Fuerth	1927	Bayerische Staatsbib
Goerlitz	1925	Wiki.genealogy.net
Kaiserslautern	1925/26	Dilibri Rheinland Pfalz
München		
Gladbach	1929	Wiki.genealogy.net
Offenbach	1929	Deutsches Museum Muenchen
Oldenburg	1927	Wiki.genealogy.net
Regensburg	1926	Bayerische Staatsbib
Rostock	1925	Ancestry
Sterkrade		
Wesermuende		
Zwickau	1925	SLUB

Notes: This table lists the cities of 50 thousand or more in 1925 that BF does not include in its dataset. BF uses the directory closest to 1925. The year in the second column corresponds to the edition we were able to find. If we could not locate a directory from the 1920s, the year column is blank. If the location column is blank for a directory we were able to locate, it is listed in the German national library (*Deutsche Nationalbibliothek*).

Table C3.2: Locating directories for small cities

Cities	In BF?	Year	Location
Andernach	no	1928	Dilibri Rheinland-Pfalz
Beckum	yes		
Bensheim	no	1928	DNB
Biberach a Riss	yes		
Bingen	yes		
Borna	no	1922	DNB
Coesfeld	no		
Coswig	no	1925	DNB
Freienwalde a O	no		
Gross Salze	yes		
Grefrath	no	1925	DNB
Haan	no		
Haynau	no	1927	DNB
Heide	yes		
Kirchheim a Teck	yes		
Kitzingen	yes		
Lehrte	yes		
Lingen	no	1925	DNB
Lugau	no		
Neuhaldensleben	yes		
Nienburg a Weser	no	1925	DNB
Norden	no	1926	DNB
Obserstein	no	1925	DNB
Olbernhau	yes	1926	DNB
Oschatz	no	1922	DNB
Perleberg	yes		
Rodewisch	no	1928	DNB
Rottweil	no	1925	DNB
Schmalkalden	no	1925	THULB
Sprottau	no	1926	DNB
Stollberg	no	1928	SLUB
Verden	no	1927	SLUB
Waltrop	no		
Waren	no	1925	DNB
Weida	no	1925	DNB

Notes: For table format, see notes to Table C3.2 The appendix text discusses issues related to identifying these places.

Table C3.3: Locating directories in cities no longer in Germany

City (name as of 1925)	Current name	Year	Other locations
Allenstein	Olsztyn	1925	
Belgard (Persante)	Białogard	1925	
Beuthen	Bytom	1924	
Bobrek	Bytom	1924	
Braunsberg Ostpr	Braniewo		
Breslau	Wrocław	1925	
Brieg	Brzeg	1928	
Bunzlau	Bolesławiec		
Deutsch Eylau	Iława		
Deutsch Krone	Wałcz		
Dittersbach	Dzietrzychowice	1929	https://sbc.org.pl/dlibra/publication/36996/edition/33602
Elbing	Elbląg	1925	
Frankenstein i Schless	Ząbkowice Śląskie	1922	
Glatz	Kłodzko	1920	
Gleiwitz	Gliwice	1924	
Glogau	Głogów	1920	https://www.martin-opitz-bibliothek.de/de/sammlungen/digitale-sammlungen/adressbuecher/e-h/glogau
Gollnow	Goleniów	1921	https://www.digitale-bibliothek-mv.de/viewer/image/PPN1668677024_1921/1/-/
Gottesberg	Boguszów-Gorce		
Gruenberg	Zielona Góra	1924	
Gumbinen	Gusev	1928	
Haynau	Chojnów	1927	
Hindenburg i.OS	Zabrze	1924	
Hirschberg	Jelenia Góra	1927	
Insterburg	Chernyakhovsk	1926	
Jauer	Jawor	1925	
Koeslin	Koszalin	1926	
Kolberg	Kołobrzeg	1924	
Konigsberg	Kalinengrad	1925	

Kreuzburg i OS	Kluczbork		
Landeshut in Schles	Kamienna Góra		
Langenbielau	Bielawa	1924	
Lauban	Lubań	1924	
Lauenburg i.P	Lębork	1926	
Leobschuetz	Głubczyce		
Liegnitz	Legnica	1926	https://wiki.genealogy.net/
Loetzen	Giżycko		
Lyck	Ełk	1926	
Marienburg i.W	Malbork	1926	
Marienwerder	Kwidzyn	1924	
Neisse	Nyse	1927	
Neusalz a Oder	Nowa Sól		
Neustadt i. OS	Prudnik		
Neustettin	Szczecinek		
Oels	Oleśnica	1928	
Ohlau	Oława	1928	
Oppeln	Opole	1925	
Ortelsburg	Szczytno		
Osterode in Westp	Ostróda		
Rastenburg	Kętrzyn	1924	
Ratibor	Racibórz	1926	
Reichenbach Schless	Dzierżoniów		
Rossberg	Rozbark		
Sagan	Żagań	1924	Univ Wroclaw library
Schneidemuehl	Piła	1925	
Schweidnitz	Świdnica	1929	
Sprottau	Szprotawa	1926	
Stargard i. Pom	Stargard	1925	
Stettin	Szczecin	1925	
Stolp	Słupsk	1921	

Strehlen	Strzelin	
Striegau	Strzegom	1929
Swinemuende	Świnoujście	
Tilsit	Sovetsk	1925
Waldenburg Schless	Wałbrzych	1929

Notes: See notes to Table C3.1 for format. Directories without a location entry are in the German National Library. Cities without a year entry could not be located.

Table C5.1: Summary of the Worms city directory for 1925

Section	Type as listed in directory	In English	Number of clubs
1	Gemeinnützige Vereine und Genossenschaften	Charities and cooperatives	32
2	Gesang-und Musikvereine	Choral and music	28
3	Gewerbliche und technische Vereine, Einkaufs-, Verkaufs- und Lieferungsgenossenschaften, Innungen, Handwerkervereinigung	Occupational and technical associations; purchasing and marketing cooperatives; guilds, blue-collar associations	61
4	Gewerkschaften	Unions	77
5	Kaufmännische, Beamten-, Lehrer- u. Lehrerinnenvereine, soweit solche nicht unter Gewerkschaften aufgeführt sind	Professional (white collar)	8
6	Kirchl. u. religiöse Vereine	Religious	45
7	Konsumvereine	Consumer	1
8	Nationale u. politische Vereine	Political	12
9	Spar- und Darlehnsvereine	Savings banks	2
10	Sportvereine	Sports	45
11	Stenographenvereine	Typists clubs	3
12	Vereine für Geselligkeit und Unterhaltung	Associations for sociality	42
13	Vereine für Gesundheitspflege	Health	2
14	Vereine für Landwirtschaft, Obst-und Gartenbau, Tiersucht (Tierschutz) und Fischerei	Agriculture, animal husbandry and fishing	33
15	Vereine mit wissenschaftlichen oder künstl. Bestrebungen	Scientific or artistic	18
16	Wandervereine	Hiking clubs	6
17	Wohltätigkeits-, Hülfs- und Unterstützungs und Anstalten	In service of the poor	39
18	Verschiedene Vereine	Miscellaneous	4
	Total		455

Note: These numbers tabulated from part (*Theil*) V, *Vereine und Körperschaften*, pp. 492-510. The section numbers and groups are as appear in the text.

Source: *Adreßbuch Stadt und Kreis Worms 1925*. Worms: Buch und Druckerei der Buchdruckerei Eugen Kranzbühler

General notes to appendix tables

Samples:

“Replication” corresponds to the sub-sample used in the original article

“Pussia” is only the Prussian observations in the original article

“-Bavaria” is the original sub-sample but excluded observations from Bavaria

Criteria for outliers

“Outlier-S:”drops observations for which the absolute value of the “studentized residual” in the replication model exceeds 2

Information criteria

Log-lik: the log-likelihood at the final values

AIC: Akaike information criterion

BIC: Bayesian information criterion

Note on standard errors:

Results marked “replication” use the standard error/clustering scheme of the original article

QR results use the standard error/clustering scheme of the analogous OLS specification

Bootstrap standard errors respect the clustering (if any) used in the parallel model reported in the original article