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

Timothy W. Guinnane  Philip Hoffman

Abstract
The persistence literature in economics and related disciplines connects recent outcomes to events long ago. Although this influential literature is promising, it raises serious questions about how to distinguish deep causal factors that persist across time from alternative explanations derived from the rapidly changing historical context or misuse of historical sources. We discuss two prominent examples that ground the rise of the Nazi Party in distant historical roots. Several econometric, analytical, and historical errors undermine the papers’ contention that deeply rooted culture and social capital fuelled the Nazi rise. The general lesson for persistence studies is that beyond careful econometrics and serious consideration of underlying mechanisms (including formal theory), they must pay scrupulous attention to the historical context and the limitations of historical data.

Reference Details
CWPE  2271
Published  12 December 2022
Revised   24 June 2023

Key Words  Historical persistence, medieval pogroms, social capital, culture, networks, Nazism, voting behavior, anti-Semitism, political parties, religion, empirical economics, data based estimates, econometrics

JEL Codes C18, D71, D72, D85, D91, L14, N01, N13, N14, Z10, Z12

Website  www.econ.cam.ac.uk/cwpe
Persistence and Historical Evidence: The Example of the Rise of the Nazi Party

Timothy W. Guinnane
Department of Economics, Yale University

Philip Hoffman
Division of Humanities and Social Sciences, Caltech

This version: June 24, 2023

Abstract

The persistence literature in economics and related disciplines connects recent outcomes to events long ago. Although this influential literature is promising, it raises serious questions about how to distinguish deep causal factors that persist across time from alternative explanations derived from the rapidly changing historical context or misuse of historical sources. We discuss two prominent examples that ground the rise of the Nazi Party in distant historical roots. Several econometric, analytical, and historical errors undermine the papers’ contention that deeply rooted culture and social capital fueled the Nazi rise. The general lesson for persistence studies is that beyond careful econometrics and serious consideration of underlying mechanisms (including formal theory), they must pay scrupulous attention to the historical context and the limitations of historical data.

JEL codes: C18, D71, D72, D85, D91, L14, N01, N13, N14, Z10, Z12

Keywords: Historical persistence, medieval pogroms, social capital, culture, networks, Nazism, voting behavior, anti-Semitism, political parties, religion, empirical economics, data based estimates, econometrics

Note: This draft incorporates comments by Robert Allen, Will Damron, Jeremy Edwards, Burkett Evans, Michael Gibilisco, Richard Grossman, Harold James, Noel Johnson, Maggie Jones, Jonathan Katz, Ian Keay, Joshua Lewis, Carolyn Moehling, Tom Nicholas, Sheilagh Ogilvie, Mark Rosenzweig, Juan Carlos Suarez Serrato, Jörg Spenkuch, Mark Spoerer, Jochen Streb, Richard Tilly, Nico Voigtländer, Hans Joachim Voth, Nikolaus Wolf, Noam Yuchtman, Qiyi Zhao, seminar participants at Caltech, and the editor and referees of the Economic History Review. Voigtländer and Voth posted an earlier version of this paper at https://www.anderson.ucla.edu/faculty_pages/nico.v/Research/Response_to_GH.pdf along with their comments. The present draft supersedes that one and our 28 November 2022 CESifo working paper. We have considered all these readers’ feedback in detail. Much of our response to the feedback and important evidence for our argument remain in the appendix at: https://hoffman.sites.caltech.edu/documents/22791/Guinnane_Hoffman_appendix.pdf.
Much influential economic history today aims to demonstrate the persisting influences of long-ago events. Melissa Dell (2010), for instance, ties poverty in Latin America in recent years to institutions established under colonialism. Nathan Nunn (2008) claims to link slow economic growth in late twentieth-century Africa to the devastation of the slave trade. Similar efforts have spread into political science: Avidit Acharya (2016) and his coauthors use tools from economics to connect differences in political attitudes in the United States today to the prevalence of slavery more than 150 years ago.¹

This paper considers an influential example of this genre: studies that invoke earlier historical events to explain the Nazi Party and anti-Semitic behavior in Germany in the 1920s and 1930s.² In “Persecution Perpetuated” (henceforth PP), Nico Voigtländer and Hans Joachim Voth (2012) argue that differences in the local culture of anti-Semitism in the Middle Ages explain cross-sectional patterns in votes for the Nazis and anti-Semitic activities in the early twentieth century. In “Bowling for Fascism” (henceforth BF), Shankar Satyanath, Voigtländer, and Voth (2017) claim that social capital formed in the nineteenth century accounts for cross-sectional differences in Nazi Party membership in the 1920s and 1930s. Both papers argue for the persistent effects of causes in the past and both are widely cited.³

These two articles address one of the central events of the twentieth century, the rise of a regime that triggered a world war and tried to exterminate the entire Jewish people. Current political events make the topic and these two papers’ explanation of great interest.⁴ Careful scrutiny of the two papers, however, shows that both suffer from a number of interrelated weaknesses. First, the econometric results are fragile. Many results depend on outliers or are not robust to reasonable alternative specifications.

¹ For the development of this literature, see Nunn (2009) and Cioni (2022).
² Noteworthy econometric and statistical studies of voting for the Nazi Party and Party membership include Van Riel (1993), King et al. (2008), Spenkuch and Tillmann (2018), and Brustein (1996). For party competition in the Weimar parliaments, see the roll call analysis in Hansen and Debus (2012).
³ PP has 943 Google Scholar citations and BF 352 (as of June 22, 2023).
⁴ Edsall (2022), a thoughtful New York Times opinion piece, cites BF.
Some reflect tendentious specifications. This fragility stems, in part, from flawed use of historical evidence. Second, each article’s argument suffers from the lack of a model, mathematical or verbal, that would clarify the implicit assumptions and suggest possible alternative explanations that would more accurately fit the historical evidence. Third, misinterpretations of the historical context compound the econometric and modelling problems. The two articles largely ignore the influence of political and religious authorities who shaped both the regressors and the outcome variables. Anti-Jewish pogroms in the fourteenth century depended not just on anti-Semitic attitudes, but on the willingness of local authorities to protect Jewish communities. Likewise, some Weimar institutional actors actively discouraged joining the Nazi party. Finally, both articles do injustice to the historical literature and code published data in ways that fail to respect the limitations of the historical sources.

Standard econometric techniques do not overcome such problems. At bottom, the issue is how to address the influence of slowly changing, deep causal factors that persist across time when there are alternative explanations derived from the rapidly changing historical context. Not appreciating historical context (which includes coincidences not taken into account in econometric specifications) can lead to spurious empirical relationships between modern outcomes and deep factors from the past. So can misuse of historical sources. Both PP and BF make such mistakes. Those mistakes may in turn hide sources of true persistence. In any case, we see little firm evidence for enduring social capital or cultural anti-Semitism. Lasting regional differences in politics and religion, however, may provide a more promising alternative explanation for the results in both articles. Germany’s historiography has long stressed the importance of regions.

Related weaknesses afflict other examples of the persistence literature, as research by Kelly (2019) and others suggests. The remedy, from our perspective, is for persistence studies to be serious about models, about historical data, and about doing the necessary historical research. That should be

---

5 See also Austin (2008), Abad and Maurer (2021), Dippel and Leonard (2021).
essential for all persistence studies. Otherwise, researchers risk being snared by explanations that are appealing but ultimately unsupported.

We first examine PP to ask whether the evidence supports its claim that cross-sectional differences in medieval culture explain 1920s and 1930s anti-Semitism. We then turn to BF and ask whether the legacy of nineteenth-century social capital caused variations in Nazi Party membership. To reduce the length of our text, we describe many of our econometric results briefly and then refer readers to the appendix for full documentation.

How Robust is the Evidence in PP?

PP’s authors claim that anti-Semitism in 1920s and 1930s Germany derived from an enduring culture of hostility to Jews that can be traced back to the Middle Ages. This cultural anti-Semitism varied from place to place within Germany, but it persisted in a given place across time.6 Similar claims about persistent behavior and attitudes underlie other econometric studies and can be derived from theoretical models.7 One obvious difficulty is how to measure anti-Semitic attitudes in the Middle Ages. PP uses as a proxy an indicator variable equal to one for Jewish communities that fell victim to pogroms during the Black Death (1347-1351).

PP’s core results (PP Table VI) test the effect of this pogrom proxy variable, POG1349, on six different outcomes: two measures of voting for the Nazis and other extremists in the 1920s; two sets of violent attacks against Jews in the 1920s and 1930s (including the Reichkristallnacht); deportations of Jewish residents; and anti-Semitic letters to the Nazi periodical Der Stürmer. Each regression controls for the locality’s population size and religious composition from the period 1924-33. PP’s authors stress the

6 The spatial variation distinguishes PP from the claim about widespread and uniform anti-Semitism in Goldhagen (1996). That claim (as PP notes) has been controversial. For an overview of the debate, see Deák (1997) and Herbert (1999).

7 For models of cultural persistence, see Bisin and Verdier (2001), Richerson and Boyd (2008), and the works cited in PP. For an application in economic history, see Mokyr (2016).
t-ratio associated with their pogrom indicator, POG1349. If persistent culture causes anti-Semitic behavior, POG1349 should have a positive and statistically significant coefficient in all six regressions.

Our Table 1 reconsiders the results for two of those outcomes, plus a composite of all six. (Appendix A discusses the other four outcomes reported in PP Table VI.) Let us begin with votes for the Nazi Party in the May 1928 federal election. Column (1) replicates the regression reported in PP (Table VI, Column (2)). POG1349 had a significant positive coefficient for the 1928 election, but a partial regression plot from our Column (1) (Figure 1) shows that the result is driven by outliers, many of which are in Bavaria, Germany’s second largest federal state. Column (2) re-estimates Column (1) as a quantile (median) regression, a standard check for outliers. In this specification, POG1349 has little effect on the conditional median. This difference shows that the PP result was driven by outliers, to which OLS is sensitive. Appendix A.5 using more flexible models to show that the pogrom proxy only increased the 1928 Nazi vote share in Bavaria.

There is, in short, no relationship between POG1349 and the 1928 vote outside of Bavaria. To address this regional difference more generally, we add to the PP specification a full set of fixed effects for the German states along with their interactions with POG1349. The result (see Appendix A.5) shows that Bavaria is the only federal state with a significant relationship between the medieval pogrom and the 1928 Nazi vote.9

Our Table 1 reveals analogous problems with PP’s composite measure, the first principal component (p.c.) of all six outcomes used in PP Table VI. The p.c. is supposed to capture “a broader, underlying pattern of attitudes” (PP, p. 1370). Table 1, Column (3), replicates PP Table VII, Column (1),

---

8 In the regression reported in Column (1) of Table 1, there are 16 observations with a “studentized” residual greater than or equal to 2. Fourteen are in Bavaria; the other two are in Baden. 70.6 percent of Bavarian communities experienced a pogrom; overall, this figure is 72.3 percent. In the PP data, Bavaria’s borders are those of the Weimar Republic. This Bavaria therefore included more territory than the medieval Duchy of Bavaria.

9 The issue here is whether the effect of cultural antisemitism is different from zero, and not whether our estimates differ from PP’s (see Appendix Section A.2 for an explanation).
the specification that corresponds to the models presented in their Table VI. The partial regression plot for this specification (Appendix, Figure A4.1) shows that this result, too, is driven by outliers, primarily in Bavaria. The pogrom coefficient is not significant in a quantile regression (Table 1, Column 4), and further analysis of the residuals supports same conclusion. Including fixed effects shows that the pogrom variable has a significant effect only in two tiny states. (Appendix Sections A.4 and A.5)

Table 1 includes one more example, deportations. For this outcome alone, PP adds an additional and redundant control, the log of the Jewish population. This model (like the others in their Table VI) already includes the population and the percentage Jewish. Comparing our Columns (5) and (6) shows that the significant result PP reports depends entirely on adding this superfluous regressor. Appendix Sections A.1-A.5 discuss related problems in the outcomes PP discusses. Although the reported results are consistent with PP’s hypothesis, diagnostic tests (such as examining outliers, considering more general functional forms, or varying the weighting of observations) imply that the pogrom proxy, with one exception, does not have a robust effect on twentieth-century anti-Semitic behavior.

Only one of PP’s regressions survives scrutiny: the Reichkristallnacht attacks (PP Table VI, Column 6). This example illustrates the consequences of ignoring political and religious actors. Medieval pogroms reflected not just cross-sectional variation in anti-Semitism in 1349, as PP assumes, but the actions of political and religious leaders at the time. The same goes for anti-Semitic outrages in the twentieth century. Historians argue that the Kristallnacht attacks were a government and Nazi Party operation that did not mirror the local populace’s anti-Semitism. Here the dependent variable itself ignores the historical context. (See Appendix A.1)

The coding of the pogrom indicator reveals weaknesses in PP’s use of historical sources. PP bases POG1349 on two compendia that summarize the history of Jewish communities across Germany. PP claims that it does not use observations where medieval events were uncertain, but our analysis (Appendix C.1) shows that PP’s dataset includes many observations for which the compendia authors evince doubt. PP highlights the example of Heiligenstadt, calling it a “typical entry.” As Appendix C.1
shows, the entry for Heilgenstadt is in fact highly atypical; in most cases, the historical sources are far less certain of events. There are also examples where PP does not follow its own rules for deciding between conflicting sources.

Placebo tests raise a different and more general question about the pogrom variable. Table 2 uses specifications identical to PP’s vote-share models (Table VI, Columns (2) and (3)) to ask whether POG1349 influences the vote for parties that did not have an anti-Semitic profile. The text table focuses on 1924 votes for two liberal parties, the DDP and the DVP, which formed core parts of the Weimar coalition. The DDP vote constitutes an especially useful placebo. Jewish Germans overwhelmingly supported this party, which had among its leaders a large number of (usually non-observant) Jews. Liepach (1969, pp. 119-120) notes considerable overlap between the leadership of the DDP and that of the major German Jewish organization of the time. Right-wing elements smeared the DDP as the Judenpartei (Frye 1976, pp. 43-44; Niewk 1980, p.28). Yet the pogrom indicator, which supposedly proxies for a long history of anti-Semitic views, shows that a history of medieval violence raised the DDP’s vote share. The DDP results for 1924 thus cast serious doubt on the pogrom proxy’s interpretation. The combined DDP and DVP results for 1924 point to a similar problem. POG1349 fails a basic placebo tests: using PP’s specifications, the proxy for medieval anti-Semitism is significantly correlated with conduct that has nothing to do with anti-Semitism. Appendix A.6 reports similar results for all Weimar coalition parties, as well as the extremist parties, in 1924, 1928, and 1933.

Understanding the Outliers: The Role of Political and Religious Authorities

Bavaria is a problem in many of PP’s models. In our specifications that add state fixed effects and their interaction with POG1349, the medieval pogrom proxy tends to be significant in only a subset of German states (Appendix Sections A.4 and A.5). PP argue that enduring anti-Semitism explains

---

10 There were far too few Jewish Germans voting in 1924 for this result to reflect their effect.
Weimar-era outcomes in Germany. The econometric evidence actually supports the historiographical stress on differences across German regions. To understand why regions play this role, we have to examine the role that political and religious authorities played in both the medieval pogroms and the anti-Semitic behavior in the 1920s and 1930s.

One example illustrates the role those actors played: the 1349 pogrom in Strasbourg, which is mentioned in PP (p.1347), although it is not in the dataset because after World War I, Strasbourg returned to France. In 1349 Strasbourg’s thriving Jewish community was rounded up and burned to death even before the plague reached the city. In 1390 Jews who had returned were expelled. Jews only returned four hundred years later under the influence of the French Revolution. Strasbourg thus seems a clear illustration of the anti-Semitic attitudes at issue in PP, and it is used as an example in S. K. Cohn’s (2007) historical analysis of the European-wide pogroms.11

The story, though, is not just bigotry: persecution always required the cooperation of political and religious authorities. Strasbourg’s 1349 massacre occurred only after three municipal leaders had been deposed and the city’s chief magistrate driven from the city. These authorities were not necessarily philo-Semites; they simply tried to uphold a promise the city had made to protect the Jewish community in return for fiscal benefits. They failed because the city’s influential butchers’ guild, as well as regional nobles and Strasbourg’s bishop, wanted to get rid of the Jews. Had all the local authorities united to oppose violence against the Jews, there would have been no massacre. This is not just speculation: eleven years earlier, the regional nobles and the same bishop joined the city’s leaders to stop a pogrom in the surrounding region. Anti-Semitism may have been a constant in Strasbourg, but authorities could encourage anti-Semitic attacks or (at least sometimes) stop them, even in the Middle Ages. T. Finley and M. Koyama (2018) show that pogroms during the Black Death were in fact more likely where political

11 Our sources for Strasbourg include Mentgen (1995); Ephraïm (1923, 1924); Ginsburger (1908); Haverkamp (1981); and the documents published in Witte and Wolfram (1896).
authority was fragmented, because the rents from taxing the Jews were divided, so any single authority had less incentive to protect the Jewish community.12

Something similar can be said for Strasbourg in the 1920s and 1930s. Despite the deep roots of anti-Semitism in Strasbourg, in the 1920s and 1930s the city did not witness any of the anti-Semitic violence seen in other hotbeds of cultural hostility to Jews. Strasbourg was French again after a period of German control between 1871 and the end of World War I. The French authorities protected the Jewish population, even when the authorities themselves were anti-Semitic (Goodfellow 1993; Caron 1998).

The religious and political authorities at the center of the Strasbourg story play no role in PP’s discussion. Similar authorities mattered elsewhere too, for instance in the Bavarian cities of Nuremberg and Regensburg (Haverkamp 1981, 67-77, 91-92). Power over Nuremberg’s 2000 or so Jews was divided, particularly in 1349, between the Holy Roman Emperor and the city council, which opposed the emperor and wielded more influence locally. As the plague approached, the emperor, fearing a pogrom in Nuremberg, sold his rights to Jewish property there. The city council had pledged to protect the Jews in return for tax revenue, but when a new city council took over in the fall of 1349, the city’s debts convinced them to sell the Jews out. With the emperor’s permission, the council let the pogrom happen (Avneri 1968, 2: 598-613; Haverkamp, 71-73). In Regensburg, by contrast, the city’s mayor, council, and leading citizens joined together in 1349 to carry out their promise to protect the Jewish community in return for tax revenue. They both thwarted residents who had gathered to attack the Jews and defended the Jews against attacks by the Bavarian Duke (Kirmier 2014; Avneri, 2: 679-691).

The Bavarian difference in the 1920s results reported in PP Table VI derived both from Bavaria’s medieval experience and from its role as the home of the Nazi Party. At the time of the plague, Bavaria was different because it was severely fragmented politically, even by the standards of late medieval

12 There were other important causes at work in the 1347-51 Black Death pogrom: the spread of rumors, the severity of the local plague, whether the Jewish community played an important local economic role, and the politics of religious identity. See Cohn (2007), Anderson, Johnson, and Koyama (2017), Johnson and Koyama (2019), and Jedwab, Johnson, and Koyama (2019).
Germany with its many divided lines of authority. For the portion of Bavaria in the medieval Bavarian duchy, political authority splintered after the 1347 death of the Duke (and Holy Roman Emperor) Louis IV, who divided his power and revenue among his sons (Holzapfel 2013; Uhlhorn and Schlesinger 1970, pp. 186-88; Immler 2016). As in Strasbourg or Nuremberg, divided authority made it difficult to protect local Jews. In Ducal Bavaria, the sons in power actually organized pogroms, with a clear goal of financial gain (Kirmeyer 2012, 2014). The rest of Bavaria was politically even more fragmented in the fourteenth century, especially Franconia, the site of many of the outliers in Figure 1.13

Bavaria was different in the twentieth century because it was where Hitler first became known and where his initial followers resided. When the government lifted a ban on the Nazis in 1925, the party held its first rally in Bavaria. Although the party gained support early on in other parts of Germany, in 1928 the Nazis benefitted from having well-organized district offices already at work in Bavaria and from having Hitler able to speak and raise money locally for election propaganda (he was banned from doing so in Prussia). In addition, the Party had a well-known Bavarian general (Franz Ritter von Epp) on their ballot, who helped Hitler raise money and reportedly attracted votes from veterans otherwise reluctant to vote for the Nazis. Their intense electoral propaganda won the party an above-average vote share in Bavaria in the 1928 elections (particularly in what had been Franconia), even though the total Nazi vote there and elsewhere remained small..

13 For the outliers here, see note 8. Of the 14 outliers in modern Bavaria, the history literature suggests that all were fragmented at the time of the Black Death Pogrom (Holzapfl 2013; Immler 2016; Avneri 1968; Flachenecker and Lochbrunner 2021; Hofacker 2015; Laschinger 2011; Musesedes, 2016; Ullmann 2012). If we apply the measures of divided authority used in Finley and Koyama (who rely on somewhat different sources), then at least 9 were fragmented; the other 5 either did not meet their criteria or were not in their data set. If we combine their criteria with our reading of Avneri, then at least 10 were fragmented. Again, the other 4 either did not meet that standard or were not described in sufficient detail in Avneri.

14 Hoser (2007); Ziegler (2019b); Pridham 1973; Selb and Munzert 2018. Selb and Munzert find no direct effect of Hitler’s speeches on Nazi voting, but his talks did raise money for expensive printed propaganda. In the PP replication data, for the election of May 1928, the Nazi Party had a 6.1 percent vote in Bavaria versus 2.5 percent in the rest of Germany. Of our 14 Bavarian outliers, 10 had been in Franconia.
Bavaria was not the only part of Germany where authority was splintered in the medieval period, so it did not have more Black Death pogroms than the rest of Germany. Bavaria was unusual, however, in having both Black Death pogroms and a high Nazi vote share in 1928. That combination produced the Bavarian outliers that stand out in Figure 1. If PP’s regression indeed demonstrates the influence of enduring anti-Semitism in the Weimar Republic, then it is surprising that POG1349 has no effect when Bavaria is excluded from the estimation (see Appendix A.5). The Black Death pogroms struck throughout Germany. An alternative possibility that better fits the data is the historical coincidence of Bavaria’s having both fractured political authority after 1347 and an effective and better funded local Nazi party organization in 1928. That coincidence would also explain why the fixed effects regressions (Appendix A.5) typically show a relationship in Bavaria but nowhere else.

To see this issue more precisely, consider the proxy \( p \) that is used to measure persistent cultural anti-Semitism \( s \) in a town in PP. We cannot observe this latent variable \( s \); we only see the dichotomous proxy \( p \), which equals one if the Jewish community in the town suffered a pogrom in 1348-50. Proxy variables are by definition mismeasured: if \( p = s + u \), then \( u \) is the measurement error, which includes the factors other than anti-Semitism that gave rise to pogroms in 1348-50. If the claim in PP is correct, then the true model for the 1928 vote is \( y = \alpha s + e \), where \( \alpha \) is the effect of enduring anti-Semitism. The error term \( e \) represents the other factors affecting the 1928 Nazi vote.

To use the medieval pogrom as a proxy, the regressions in PP estimate the equation \( y = \beta p + f \). (We will develop this intuition abstracting from other controls, but return to them below.) Because the true model for \( y \) is \( y = \alpha s + e \), the estimate \( \beta \) in PP is:

\[
\frac{\sum yp}{\sum p^2} = \frac{\sum(\alpha s + e)(s + u)}{\sum(s + u)^2}
\]

\(1\)

---

15 In the PP data set, 17.7 percent of Bavarian towns had Black Death pogroms versus 18.3 percent outside Bavaria. However 16.6 percent of Bavarian towns had both a pogrom and an above median Nazi vote share in May of 1928, versus 8.3 percent of towns outside Bavaria.
If the variables in the sums are i.i.d, have finite means and are measured relative to their means, then the
equation to the right of the equal sign in Equation (1) converges to the following as $n$ increases:

$$
\alpha \left( \frac{\text{cov}(s, u) + \text{var}(s)}{\text{var}(s) + 2\text{cov}(s, u) + \text{var}(u)} \right) + \frac{\text{cov}(s, e) + \text{cov}(u, e)}{\text{var}(s) + 2\text{cov}(s, u) + \text{var}(u)}
$$

(2)

where $\text{cov}(s, u)$ is the covariance of $s$ and $u$, $\text{var}(s)$ is the variance of $s$, etc. If all the covariances in
equation (2) are zero, then $\beta$, the estimate for POG1349, will simply be an attenuated estimate of the true
coefficient $\alpha$, a standard result for measurement error in a regressor in a linear model. The covariances
in equation (2) are unlikely to be zero, however, because PP omits a role for political and religious
authorities.

We discuss these issues as they pertain to Bavaria, where the historical evidence is clearest. But
the issues are not limited to Bavaria. First, $\text{cov}(u, e)$ is not zero because the error terms affecting medieval
pogroms and 1928 Nazi vote were correlated in Bavaria when the role of political and religious actors is
omitted. One might assume that they would not be correlated because $u$ pertains to 1348-50 and $e$ to
1928. But in Bavaria historical coincidence connected them. In Bavaria, splintered political authority
would make $u$ large by adding to the factors other than anti-Semitism that gave rise to medieval pogroms.
Historical coincidence inflated $e$ as well in Bavaria, because Hitler got his start there and was able to raise
funds for election propaganda in 1928, which would give the Nazis a higher percentage vote. This
historical coincidence created the Bavarian outliers that biased the coefficient for POG1349. And as

\[16\text{ For textbook treatments, see Greene (2018, pp.102-3 and pp. 281-88) or Wooldridge (2010, pp.78-82).}\]
\[17\text{ We stress that medieval anti-Semitism in Bavaria was not the chief reason Hitler got his start there. His Bavarian success had}
\text{more to do with political events that struck Bavaria in particular: a 1918 revolution that toppled a monarchy and established a}
\text{republic; a failed attempt to establish a Soviet-style regime, which was violently overthrown and caused political sentiment to}
\]
Figure 1 shows, there were outliers outside Bavaria as well. The lesson is that just because a potential causal variable lies in the past does not mean it is exogenous, particularly when political factors can affect observations across time and regions.

Second, the covariance $\text{cov}(s, u)$ between anti-Semitism $s$ and the proxy error term $u$ is also unlikely to be zero because $u$ will include political factors that affect the chances of a pogrom. Strong local anti-Semitism could make it easier for elites with financial motives to find allies for a pogrom that would seize Jewish assets. The expected financial gains would depend on the odds of resistance by local authorities, and hence on unobserved political questions such as how divided local political authority was in 1348-50. If authority was fragmented, as in Bavaria, stopping the pogrom would be less likely. Since greater local anti-Semitism $s$ would make it easier to win support for such a pogrom, $\text{cov}(s,u)$ would not be zero.

PP requires that POG1349 is exogenous to behavior in the 1920s and 1930s. The nonzero covariances we discuss show this is not true. Just because something is in the past does not make it exogenous.

Adding Controls and the Role of Political and Religious Authorities

With the right controls, the covariances we discussed earlier could be driven closer to zero: what appears as part of the error term in PP’s regressions would be partialed out. Such controls would include variables that pin down the changing political and religious context, both for the Middle Ages and the 1920s and 1930s. For the Middle Ages, PP does incorporate characteristics of medieval cities (PP, Tables VIII, A19). But those controls do not account for the sort of fragmented and varying political authority swing to the right; and a 1920 coup, which until 1924 made Bavaria a haven for right wing extremists fleeing a failed right-wing coup in Berlin (Gelberg 2007; Ziegler 2019b; Pridham 1973).
we found in Bavaria. They therefore cannot capture whether local medieval authorities had an incentive to protect the Jews.

Religion is relevant not just in 1348-50 but also in the 1920s and 1930s. J. L. Spenkuch and P. Tillmann (2018) show that until the Vatican reached an accommodation with the Nazis in 1933, many Catholic bishops successfully pressured their faithful not to cast Nazi ballots and to vote instead for the centrist Catholic Zentrum Party and its Bavarian sister party, the Bayerische Volkspartei (BVP). The chief exception was in villages where Catholics were influenced by priests who resisted the bishops and “openly sympathized with the Nazis” (Spenkuch and Tillmann 2018, pp. 20-21, 28-29; Spicer 2008). The influence of local Catholic leaders (and how it varied over time and space) is thus especially important in the 1924 and 1928 elections.

If the actions of religious and political authorities (either in the 1300s or in the 1920s and 1930s) better explain anti-Semitism than does persistent culture, then it would be easier to account for three troublesome patterns in the PP data. First, many towns with a Black Death pogrom were close to places that did not have such a pogrom, sometimes only a few miles away (Figure 2). Such a sharp local difference seems incompatible with the idea of local culture, which would presumably diffuse over neighboring communities as people went to market, sought marriage partners, or looked for work. By contrast, this geographic pattern would fit quite well with the fragmentation of local political authority in the late Middle Ages. Second, Jews soon returned to communities that experienced a pogrom in 1348-50. Among the many examples after the Black Death pogroms, Jews returned to Gießen in 1375, Guben in 1354, Halle in 1368, and Hamm in 1370. The town id numbers are 479, 532, 546, and 559; Avneri’s entries are in Volume I (1968, pp. 278, 307, 320
the actions of local political and religious authorities. The Jews might return when new urban magistrates, bishops, or seigneurial lords offered them credible protection. Third, Spenkuch and Tillmann (2018, Table 6 and p. 31) show a clear role for the Catholic Church in restraining (or not) anti-Semitic conduct in the 1920s and 1930s. Catholic cities witnessed fewer of these later pogroms before 1933, even controlling for having a history of a medieval attack on Jews. Once the Catholic Church reached an accommodation with the Nazis in 1933, however, Catholic cities experienced more attacks on Jews. This rapid change in the historical context is hard to reconcile with a predominant role for deeply rooted cultural anti-Semitism.

*Do the Results in BF hold?*

We now turn to a second paper that links the horror of the Nazi period to deeply-rooted features of the past. BF seeks to explain cross-sectional differences in Nazi Party membership by appealing to another historical cause, variations in the “social capital” embodied in voluntary associations created in nineteenth-century Germany. The idea of social capital spread in the social sciences thanks to the work of Robert Putnam (R. Leonardi, R.Y. Nanetti, R.Y. and R. Putnam1 1992; Putnam 2000). The concept has proven difficult to define. Some have worked to pin down its meaning via the sociology and economic theory of networks (Banerjee et al. 2019; Jackson 2019). The main empirical measure of social capital used by Putnam and others (including BF) is the density of “civil society” organizations. BF addresses an older literature to argue that during the Weimar Republic, the Nazi Party drew on social capital to boost recruits: “… an important strand of the literature on the rise of totalitarianism has argued that the weakness of German civic society facilitated the rise of the Nazis. Our results demonstrate that the opposite is closer to the truth” (BF 2017, p. 482).

354). Finley and Koyama (2018) suggest that the medieval attacks varied in intensity. If so, Jews might return if the dangers was not death but rather the threat of having to flee.
BF constructs a proxy for social capital by counting the number of civil-society associations per capita in a sample of 229 German cities in the mid-1920s.\textsuperscript{20} The authors’ regressions test whether social capital measured in this way explains the percent of the population who joined the NSDAP from each city in this period. The regressions control for city size and religious composition, as in PP, as well as the percentage of the work force that is blue collar. The main results (BF 2017, Table 3) imply that more social capital leads to more Nazis, but additional tests (BF 2017, Table 7) show this was true only in federal states BF considers politically “unstable.” In Prussia and other “stable” states, there is no such relationship. Prussia was the largest single federal state, accounting for 60 percent of Germany’s 1925 population and 52 percent of the BF sample cities. The other states BF labels as stable had about 15 percent of the total German population and about 20 percent of the sample. So BF’s results, taken at face value, imply that social capital only affected Nazi recruiting in one-third of Germany, a remarkable finding that BF’s authors appear not to appreciate. Put differently, this is another case of outcomes driven by regional outliers: BF’s results for “Germany” reflect, if anything, only the “unstable” states. Here we focus on the main results and the stability issue, which is a serious challenge to BF’s results.\textsuperscript{21}

BF devises a stability index as the first principal component of three variables, all defined for the period October/November 1918-May 1932: (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

\textsuperscript{20} BF counts what in German is called a \textit{Verein}. The word can mean “association,” “society,” or “club.” We use these terms interchangeably.

\textsuperscript{21} BF drop from consideration the territories that were allocated to Poland and Russia after World War II. The figures for population in 1925 pertain to the entire country. In the Appendix Section B.1, we show that BF’s main results (as presented in their Table 3) reflect the influence of outliers. Dropping Bavaria does not change the sign of the main result, but the standardized regression coefficient for the social-capital variable declines by 40 percent.
(possibly in different coalitions); and (3) the percentage of that period the state was ruled by the “Weimar coalition.”: the Social Democrats (SPD), the Zentrum, and liberal German Democratic Party (the DDP).22

BF’s authors never explain the logic for the third component. Their argument stresses turnover in state-level leadership, not connections to the federal government. In addition, the party that headed Bavaria’s government for much of the Weimar period (the BVP) had agreements with a Weimar coalition party (the Zentrum) that meant the Zentrum did not stand for office in Bavaria. In BF’s scoring, this means no Weimar coalition was possible in Bavaria and for the third element in the stability index Bavaria thus receives a zero by definition.23 Most important, instead of using the stability index itself in their regressions, BF’s authors convert it into an indicator variable: “we split the non-Prussian part of Germany into a stable and an unstable half (with above- and below-median stability, respectively). (BF (p.508)).” Creating the indicator variable throws away information. Their description of the binary indicator is also inaccurate; the empirical exercises include the median values as part of the unstable group. Many observations bunch around the index’s median, so allocating those median observations to the “above” or “below” groups can, and in this instance does, drive the results.

Our Table 3 re-estimates the regression models reported in BF’s Table 7.24 Column (1) replicates BF’s column (3) for states that are “unstable” by their definition and index. As BF stresses, in unstable states, more social capital means more Nazis. Our Column (2) estimates the model in Column (1) as a median regression. The clubs variable is not significant. Once again, the OLS result reflects outliers. Column (3) drops the “Weimar party” element from the index but retains BF’s binary definition of

---

22 This definition for the third element appears in the notes to BF Table 7 and underlies the values of the three index elements used in their analysis. BF p. 508 defines the third element in a different and conflicting way: “governed by at least one party from the Weimar coalition.”

23 The Zentrum and BVP (the Bayerische Volkspartei) were two Catholic parties. The BVP emerged from the regional wing of the Zentrum during World War I. With only a few exceptions, the two parties cooperated in ways that lead some scholars to call them “sister” parties.

24 Appendix Table B.3.1 allocates the states by alternative definitions of the stability indicator.
stability. The indicator is no longer significant, showing that the BF result requires that third element. Column (4) uses BF’s version of the index, but defines the binary indicator to include the median values among the “stable” states rather than among the unstable ones. This change affects Bavaria alone, which has the median value of the stability index. The estimate for the social-capital variable is, once again, not significant. BF’s results, in short, hinge on that third element, which is unexplained and historically inappropriate, and on assigning the median values to the “stable” group.

The Appendix (Section B.3) reports additional checks for the specifications reported in Table 3 as well as the related robustness checks reported in BF’s Appendix. When we define the stability indicator to assign the median to the “stable” group, none of BF’s results survive. BF’s Table 7 also includes specifications that use the entire sample and interact everything with the BF stability indicator. Our Appendix B.3 shows those regressions are equally sensitive to the problems in the stability index.

Finally, we pose an obvious question: what is the point of taking a continuous index and turning it into a binary indicator? This procedure just throws away information. In Table 3, Column (5), we estimate BF’s model using the interaction of social capital and the continuous stability index. Social capital’s interaction with stability has no effect. In fact, the next effect of social capital as computed from that regression implies that social capital only matters in the stable states, precisely the opposite of the BF argument (Appendix Table B3.6) In short, the entire relationship BF stresses depends on turning the index into a binary indicator.25

---

25 Our Table 3, Column (5) is identical to the regression that underlines BF Appendix Figure A7. They do not report the regression and that figure uses the wrong critical values for the confidence intervals. See our Appendix Section B.3.
BF’s authors construct their measures of Nazi Party joiners from a public-use sample created by earlier researchers. They created the social-capital proxy, however, by counting the number of associations listed in the directories published for most German cities in this period. This proxy raises two distinct issues. First, can BF’s sample capture the idea underlying that proxy? Second, is the distribution of associations in the 1920s exogenous? The answer to both questions has to be yes for the causal argument in BF to work, as BF’s authors recognize. To justify their yes answers, they rely on a persistence argument that is critical for their paper: differences in the density of associations in the 1920s were driven by “deep historical factors that have no direct link with Nazi Party entry” (BF 2017, p. 487). As we shall see, that argument and their claims about their sample and their proxy fail to stand up to scrutiny.

BF lists the cities in their sample but does not state precisely which year’s edition they used, so we cannot examine the actual directories that underlie their data. Our Figure 3 reproduces part of the relevant section from a directory for Worms (one of BF’s cities) from 1925. The directory divides the associations into functional categories; our figure shows the last page of the group that includes charities and cooperatives (Gemeinnützige Vereine und Genossenschaften) and the first page of choral and music societies (Gesang- u. Musikvereine).

The Appendix (Section C.3) discusses possible bias in the selection of cities drives BF’s results. BF’s authors started with the 547 cities that had populations over 10,000 in the 1925 census. They dropped 65 cities now in Poland or Russia, claiming “towns and cities in the formerly German areas of Eastern Europe rarely preserved marginal library holdings such as city directories” (BF, p. 490, footnote 26). 

26 Brustein (1996). BF does not mention that many joiners soon quit; about 40 percent of the 1.4 million people who joined the Nazis prior to 1933 had quit by that date (Brustein 1996, pp. 10-15). If clubs attracted new Nazi Party members who had no serious attachment to the Party, then their joining would add little to Party’s strength as a political force.

27“We use any surviving directory from the 1920s; where several are available, we take the directory nearest in time to 1925” (BF, p. 491).

14). They provide no support for this claim, and, as we show in Appendix C.3, some 70 percent of these places have an extant directory today, usually in a German library. BF’s authors then contacted “libraries and archives” in the remaining cities (BF 2017, pp. 490-91). Some did not reply and others said they had no directories. The final sample includes only 197 places from the original 547 and has some striking omissions, such as Berlin and 11 of the other largest cities.

Our research located directories in German libraries for many of the missing cities (Appendix C.3). To assess the possibility of selection bias, BF Table 1 compares vote shares and socioeconomic statistics for their sample cities and Germany as a whole, but as our Appendix C.3 makes clear, this comparison of observables is not enough. It is easy to imagine city characteristics that correlated with the probability of producing a directory in the 1920s or that would affect the survival of such directories to the present. Those characteristics could in turn be correlated with local political conditions and with social capital, both today and in the 1920s.

There is a second selection problem in the clubs the directories report. To be a useful measure of social capital, the directories have to either include all relevant clubs, or report unbiased samples of such clubs. The historiography says they do not. A study of Tübingen in the late 1920s states that the city’s directory “normally covered nearly two-thirds of all local voluntary associations.” (Koshar 1982, p. 32) The real worry is that the selection of clubs to list in the directories is correlated with their suitability as Nazi recruiting grounds. Directories may systematically undercount the sorts of groups that would be hostile to Nazi recruiting efforts. Workers’ organizations are a clear example, for as BF acknowledges (p. 518 and appendix E.5), they would not be fertile ground for Nazi members. A 1925 directory for Bonn, for instance, includes almost no associations whose members were likely to have been working class.29 Workers’ organizations proliferated in the 1920s. Yet in the BF data set, some large, industrial cities have

29 *Einwohner-Buch der Stadt Bonn*. 1927. Bonn: Druck und Verlag J.F. Carthaus. Professional and business groups account for about one-third of all associations listed in the 1925 directory for Worms (see Appendix Sections C.3 and C.4). BF apparently excludes these groups from the social-capital proxy, although BF does not say that explicitly.
suspiciously few clubs: Essen (population in 1925, 630,000) has 13 clubs total in the BF data. BF does not discuss the issue, but as Appendix C.3 shows, the authors could have checked their club listings against external sources.

We also doubt BF’s use of the history of associations to defend two important assertions. For their study, it is important that the clubs they count are not ideologically akin to the Nazis; in that case, joining the Nazi Party would reflect a political orientation rather than social capital. In addition, the distribution of associations across cities (measurement issues aside) has to be exogenous. To support both assumptions, BF’s authors rely upon a tendentious reading of the history of German associations. In their view, their data for the 1920s reflects a persistent “culture of associational life” created in the nineteenth century (BF 2017, pp. 483-87). “After controlling for city size, the share of Catholics, and the proportion of workers, we believe that differences in the density of associations are reasonably exogenous for the purpose of our study (i.e., driven by deep historical factors that have no direct link with Nazi Party entry).” (BF 2017, p. 487). In other words, associations had nothing to do with the nationalism that drew people into the Nazi Party, and the density of the associations came into being far enough in the past to make it uncorrelated with the error term in the BF regressions.

Nineteenth-century nationalism might have been less xenophobic than the Nazis, but the rich literature on German associational life in the period from 1848 to World War I stresses a rapid growth of civil-society organizations devoted to nationalist goals such as a fleet to challenge Britain and colonies in Africa and elsewhere. Such nationalist concerns continued to permeate associational life in the Weimar

30 The appendix to BF (Section E.5) notes that workers associations “are at best weakly associated with Nazi Party entry,” but does not discuss the possibility that such bodies are undercounted in the directories.
31 For the colonial associations, see Conrad (2011, pp.25-27) and Speitkamp (2014, pp. 19-20). The far larger Navy League (Flottenverein) agitated for a German fleet that could challenge British seagoing supremacy. Several organizations created to honor the memory of the 1870/71 victory over France eventually morphed into right-wing political organizations.
period, as one careful local study (Allen 2014, pp. 16-19) observes. Gardening clubs invited nationalist
speakers. Even choral societies had an ideological bent.

A more specific worry concerns direct Nazi participation in some of the clubs that BF use to
failed, the [Nazi] party dissolved into sports clubs, sharpshooting associations, and hiking organizations.”
Koshar’s claim (see also Anheier 2003) implies something closer to the reverse of what BF argues. The
Nazis later regained the right to recruit members, but the strength of those ersatz-Nazi groups reflected
the Party not through the mechanism stressed in BF, but directly: some were, temporarily, little more than
the Nazi Party in disguise.

What about BF’s argument that the culture of associational life was set in the middle of the
nineteenth century and was therefore exogenous in the 1920s? BF (pp. 481-84) correctly emphasizes that
the pre-March Revolution period (1815-1848) saw both a flowering of liberal and democratic associations
and concerted effort to suppress many of them. Thereafter BF points to the 1848 national Parliament’s
proposal for a constitution that guaranteed freedom of association and maintains (BF, p. 484) that after
1848, “earlier prohibitions never returned with full force.” If in fact German governments had respected
freedom of association after 1848, then the distribution of clubs in the 1920s would have still been shaped
by events and clearly not be exogenous. BF’s argument does not support their conclusion. In any case, the
assertion ignores a continued history of repression, one that distorted the location and types of
associations Germans could form.32 Some of the great political battles of the later nineteenth century
involved suppression of associations. The number of clubs grew dramatically from 1848 to 1918, and
featured organizations for working-class people in particular. The growing number and changing types of
associations undermines BF’s assumption (p. 484) that “the state- and city-level factors driving variation

32 See Brooks and Guinnane (2017). Freedom of association did not have legal guarantees throughout Germany until 1908.
in the repeal of restrictions are plausibly exogenous to NSDAP entry in the 1920s and 1930s.” The historical evidence contradicts this assumption.

To support the exogeneity claim BF’s authors do report a regression that in their view demonstrates the deep historical roots of associational density in the 1920s. For 39 of the 229 cities in their data, they know how many delegates local associations sent to the 1848 Democratic Congress in Berlin (BF 2017, Appendix F). For those 39 cities, this variable explains 13 percent of the number of *Turnverein* (gymnastic club) members in 1863 and 46 percent of their clubs per capita variable for the 1920s. For the early 1860s, they also construct an index using numbers of *Turnverein* members and attendees at a choral festival. The index explains about 20 percent of the variation in the BF Nazi recruitment variable for 1925-1933 for the 150-odd cities for which this information is available.

These statistical results do not reassure. First, they pertain to only part of the BF sample of 229 cities. Second, associations extant in 1848 (so the history demonstrates) had a different social, confessional, and political basis than those in the 1920s. The former survived restrictions imposed by authoritarian governments. BF exacerbates this problem by dropping all religious clubs. Catholic associational life especially took off in the later nineteenth century. BF’s 1848 clubs would include almost no Catholic groups, which (so we will see) were (later) usually hostile to the Nazi Party. In addition, “The number of local voluntary associations grew throughout the 1920s, reaching extremely high levels as measured by both historical and comparative standards.” (Berman 1997, p. 413). Discussing the Weimar period, Koshar (1982, p. 33) notes “the appearance of soccer clubs and shooting clubs with predominantly working-class memberships promoted and reflected deepening class divisions.” Tenfelde (2000, pp. 95-96) also stresses the growth and diversity of clubs devoted to working-class memberships, especially during the Weimar Republic. By then clubs were drawing members from very different social and religious groups and with what were likely be different political leanings as well. It would be very difficult to believe that all these developments are exogenous in the 1920s and 1930s.
BF excludes two sets of clubs from the counts that form their social-capital proxy: the “political” and “religious.” There are two drawbacks to this decision. First, it may not be possible to identify such groups from the information in city directories. In Figure 3, for instance, the first page lists a school association whose leader is a minister (Pfarrer). The organization probably had some type of affiliation with a church. (In contrast, the prior entry says it is a Catholic association.) This directory (like others) has a separate section for religious organizations, but many bodies listed elsewhere, such as leisure-time groups or, in this case, a charity, had the backing of a political or religious body. Dropping clubs as BF does introduces biases the political orientation of sample clubs in unpredictable ways. Second, the wider logic justifying these omissions is unclear. One might want to understand the difference between social capital in a political organization and social capital in a choral society. This question warrants an empirical test, not dropping a large number of clubs from the sample.

BF exclude religious organizations because “we are interested in the ‘bottom-up’ characteristics of grassroots organizations, not in ready-made sociality created by members of the church hierarchy” (BF 2017, p. 486). The discussion points to the Catholic Church in particular. It certainly was hierarchical, but its history in the nineteenth and early twentieth century (and before) is full of local organizations created by the local laity, from traditional confraternities to charitable organizations, women’s associations, and groups linked to professions. More important, it is unclear why “bottom up” associations in general would involve more social capital or have a greater impact on Nazi recruiting. In addition, many of the other associations underlying their proxy were in effect the local branch of a national organization, organized at the behest of canvassers from national organizations. An association

33 The association is the Kinderschulverein, seventh from last on the directory’s page 493.
34 BF tests for the difference between several types of associations, but since what they view as political and religious clubs are not in the data, they could not check to see whether political and religious groups are different.
that is little more than a branch of some larger entity could create relationships among members and help recruit Nazis.

Because of the German Catholic Church’s overt and well-documented hostility to the Nazis, the omission of Catholic associations could seriously bias BF’s estimates. According to H. Mommsen (1988, p. 353), “In Catholic regions, as opposed to their Protestant counterparts, the NSDAP was only rarely able to penetrate the network of middle-class clubs and associations that had played such an important role in its expansion in northern Germany.” Z. Zofka (1979) stresses that many Catholic associations strongly discouraged members from joining the Nazis. He shows that in Bavaria, areas with strong local Catholic bodies had fewer Nazi members.35

BF’s regressions all include a control for the percentage of the city’s population that was Catholic. Unless the city’s religious composition perfectly predicts the number of missing Catholic associations, however, excluding the Catholic clubs from the social capital proxy could easily bias the results in favor of BF’s conclusions. The BF proxy could simply count associations whose members were more likely than average to be favorable to the Nazi Party. In theory, one could test whether excluding religious and political associations affects the results in BF. We cannot do so because BF did not include the relevant counts in their replication data, nor do we know precisely which directories they used.36

What Do BF’s Results Say About Social Capital?

Do the results in BF necessarily imply anything about the role of social capital in Nazi support? BF does not model how the social capital embedded in these associations might have promoted Nazi

35 See in particular Zofka, pp. 168-69. BF (2017, p.489) quotes Zofka as saying the chairmen of local associations “and other opinion leaders increasingly converted to the Nazi creed and induced other members” of associations “to follow.” Zofka, however, stresses that Catholic associations remained hostile to the Nazi party. Brustein makes the same point (1996, pp. 166, 171).

36 While the omission of Catholic organizations is most glaring, given what we know about the history, the omission of other religious organizations, as well as political groups, is also unfortunate.
recruiting, except to say that “associations facilitated Nazi recruitment” by spreading the party’s message (BF 2017, p. 480, 490). 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 Nazi recruiter to join the association and ask other association members to identify the best sources of information in the group (the “gossips” in the association, in the language of an experimental study) (Jackson 2019; Banerjee et al 2019). The recruiter would then pass favorable information about the Nazi Party to these gossips: for instance, telling them about an upcoming Nazi speaker, an effective tactic used by the party (Brustein 1996, p. 163; Allen 2014, pp. 80-82). That would be more efficient than approaching each association member individually or (according to the experiments) going to the group’s leaders, and it would use the association’s social capital, the connections between the members. The result would be 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, here proxied 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 (so we have seen) suggest that was the case in German towns and that information about groups’ membership and their probable political sympathies was often common knowledge, even for groups that were not overtly political. Memberships usually aligned internally along class or religious lines that would make it easy to guess at political leanings. Recruiters could exploit this information and use it for recruiting without ever joining groups or making use of the associations’

37 Allen 2014, pp. 16-19; Brustein 1996, pp. 163-71; Koshar 1982, pp. 31-36. Tenfelde 2000 recounts the history of the Hessian town of Eschau, with two competing sets of clubs. Members of a given club would patronize a given pub, hairdresser, etc. According to Tenfelde, the political associations of the two sets of clubs post-date World War II, but the example serves to show that someone could tell a lot about a person by knowing which associations they belonged to.
social capital, the connections between members. They could, for instance, just give members of a promising group leaflets about Nazi speakers or invite them to a Nazi talk. The tactic would be no different from, say, an American political campaign publicizing a Republican candidate among gun owners or a Democratic candidate advertising on MSNBC. It would involve no social capital, because it did not rely on 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.

Nazi recruiters could exploit this information about memberships even without prospecting among openly political groups. If anything, excluding the Catholic groups might make the remaining ones even more likely to have an above average number of Nazis, and so reinforce the relationship between associations and party membership, all without any involvement of social capital.

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, as we show using a simple model in Appendix B.4. If the Nazi party has some recruiters who use the first method and some who rely on the second, then BF’s regression coefficients would simply add the effect of the two methods of recruiting. If this sum were positive and significant, that would say nothing about social capital, because the whole effect could simply be the other method of recruiting. This problem of interpreting the coefficients’ meaning is serious.

Here one might object that this distinction between recruiters’ knowledge and social capital is interesting but not really a problem for BF’s claims. The number of clubs is a standard proxy for social capital, and it does not really matter what the connections were between members of associations. BF argues that places with more associations had more Nazis and that this evidence says something important about the town and about social capital. We would agree that such a relationship would say something about the town. But it would not necessarily reveal anything about social capital unless it involved the connections between the members of associations. To argue otherwise runs counter to the economic
theory of social capital and to the broader social science research on social capital. Ties between members of groups figure prominently in all that research, and they are essential if we want to pin down what precisely social capital is (Jackson 2019).

The Nazi Party succeeded by crafting nationalistic proposals that attracted a core group of members and then efficiently marketing this program to a broader group of voters (Brustein 1996, pp. 1, 9, 57-60, 118-119, 157-182). In recruiting members, it may have targeted receptive audiences, much as modern political campaigns do, or harnessed connections between individuals. Yet only the second path relied on social capital, and BF’s evidence cannot tell us which path was taken. Only additional historical research about Nazi recruiting would reveal which one it was. Did Nazis draw new party members from associations to which they themselves belonged? Or did the Nazis recruit from groups they themselves had not joined? A careful reading of local historical studies might provide an answer.38

Conclusion

PP and BF muster evidence which they argue shows that deep, slowly changing historical forces played an important role in the extreme anti-Semitism that underlay so much of Weimar political life, including the rise of the Nazi Party. We have established, however, that little of this evidence stands up to scrutiny. Our discussion of the flaws in PP and BF does not rule out a role for persistent social capital or a longstanding culture of anti-Semitism. These factors may well help explain the rise of the Nazi Party in 1920s and 1930s and be important for questions in other times and places. The evidence that PP and BF offer, however, does not demonstrate this was the case in Weimar Germany.

One might argue that we are asking for too much robustness. Our demands (so the argument might go) might raise the evidentiary bar so high that we would mistakenly reject the evidence for the true

38 As BF (4887-489) notes, Anheir (2003) shows that Nazi recruiters relied on social connections to attract new members, at least in places with no party office or district organization. But the associations in question here were far right groups, and many had been covers for the Nazi Party when it had been banned, precisely the sort of connection that undermines BF’s argument.
role played by medieval anti-Semitism and Weimar social capital in the rise of the Nazi Party. Such an argument, however, would not be persuasive. The robustness checks we apply to the important conclusions in PP and BF are simple steps one would expect in any empirical research. Moreover, the weaknesses in the two papers are not just a matter of statistical inference. Neither PP nor BF consider placebo checks that would verify their interpretation of the proxies in question. Similarly, BF’s stability results reflect conceptual flaws as well as inappropriate econometric decisions. And a more powerful econometric test in BF (for instance, a regression with a larger sample) could never tell us whether the Nazi Party relied on social capital or simply on information about social groups when enrolling new members; the regression coefficient measures the sum of the effect of the two methods of recruitment. Only more historical research can solve that problem, just as only more historical research can overcome the limits of BF’s data or explain why Bavaria causes trouble in both papers. And some of that research would have been easy for the authors to do.

We have stressed the role of political actors in both the medieval pogroms and the later Weimar political developments, actors who ranged from medieval political authorities to Weimar-era Catholic bishops to the Nazis who recruited others into the Party. In a changing historical context, they played crucial roles in the outcomes we observe, including the regional differences in Bavaria. Ignoring them makes it impossible to understand the role of anti-Semitism and social capital in the Nazi takeover in 1920s and 1930s Germany. As generations of historians have stressed, medieval pogroms erupted in villages and towns throughout Western Europe. Anti-Semitism was widespread in Europe before World War II, as was complicity with occupying Nazi forces. But there was only one Third Reich.

Weaknesses of this sort are, unfortunately, not unique to PP and BF. They afflict other examples of the persistence literature as well. Analyzing these specific persistence studies, however, does yield some general warnings for avoiding such pitfalls in the future. We see three types of concerns.

First, treatments have to be exogenous. PP assumes that the medieval pogrom can be treated as exogenous because it happened long ago. That misses the correlation arising from the coincidence of two
unrelated things in the federal state of Bavaria. Similarly, BF claims that the distribution of civil-society organizations across German cities was fixed in the middle of the nineteenth century, but does not marshal convincing evidence for that claim.

Second, not having at least a verbal model opens the door to trouble. A simple model of relevant actors would have made the authors of PP consider the political and religious authorities who could encourage or discourage anti-Semitic behavior in the fourteenth century and the 1920s and 1930s. Omitting them (so we show) created the spurious correlations across time. Similarly, BF does not say explicitly how Nazi recruiters used associations to enroll new party members. As our simple model showed, even if there is a correlation between Nazi party membership and associational density, it could be caused by a completely different mechanism that does not involve social capital.

Third, econometric exercises must reflect both an interest in letting the data speak and careful handling of the underlying data. Despite robustness checks in both papers, important empirical results are sensitive to small and reasonable changes in definitions or use of controls, and the conclusions vary considerably with modest alternatives. Neither paper considers the pervasive outlier problem. Both papers rely on proxies but neither one considers placebo checks to verify the proxies. These problems extend to the data. PP’s coding of the medieval information does not respect the many expressions of doubt in the sources. The BF dataset includes fewer than half the relevant cities because its authors did not locate a large number of extant directories, even for the largest cities. BF does not list specific sources for its main proxy, and its results may derive from the omission of many organizations that would naturally be part of the proxy. BF does not mention any tests for the implications of that core decision, which would have been possible with relatively little historical research.

The most important task for persistence studies, however, is getting the history right. Failing to do that can undermine the econometrics, the data, and the modeling. We have noted several instances in which PP and BF ignore historian’s conclusions or over-simplify a debate in the interest of presenting a striking result. “History” is not a series of details; it creates the building blocks for causal connections, the
context for the treatment variables, and the world in which the actors at stake actually lived. History
constrains our use of data but also reveals ways to test our assumptions. We cannot understand whether
history affected later events if we do not have the history right.
REFERENCES


FIGURE 1

Note: The figure shows a partial-regression plot for the 1928 Nazi vote and the POG1349 variable. See Belsey, Kuh, and Welsch (1980, p.30). The x-axis here plots the residuals from a regression of POG1349 on the other regressors (X_i), and the y-axis plots the corresponding residuals from a regression of the Nazi vote variable on the independent variables other than POG1349 (X_0). The solid line plots the implied linear fit, which is (by construction) the regression reported in PP, Table VI, column (2): the 1928 Nazi vote share = .0142*POG1349, standard error = .00567)
FIGURE 2

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.
FIGURE 3
TWO PAGES FROM THE CITY DIRECTORY FOR WORMS 1925
Kirchengefangnereine, siehe unter kirchliche und religiöse Vereine
Arbeits-Sängerbund (Regest XIII Worms). Leitung: Wilhelm Judisch, Peters-
straße 15. » Nr. 2047.
Kreis-Sänger-Mitgliederverband. Vorstand: Karl Schröder, Schmiede-
straße 16. Gesangverein „Atariu“. Vorstand: Wilhelm Beisel, Römer-
straße 22. Gesangverein „Inquisiton“. Worms-Glanztal. Georg Klinger, Melan-
steinstraße 2 MM. Vereinslokal bei Pf. Diegeler, Kreuzstrasse 11 MM.
Gesangverein „Rödertal“. Worms-Glanztal. Johann Kreiser, Kreuzstrasse 12
Gesangverein „Klöppelrin“. Franz Graf, Frankenthaler Straße 29. Vereinslokal:
Goldenes Kreuz, Gernersch. Straße 23.
Gesangverein „Büllesgäu 1“, gest. 1910 Worms. Vorstand: Karl Schneider,
Rehnbauerweg 23. Vorstand: Kapellmeister Albert Weinelt. Vereinslokal:
Zum Nordend, Siegfriedstraße 2.
c) Musikverein der Werke. Leitung: Adam Niehl.
*Rundfunk-Gesellschaft Worms. Vorstand: Karl Rössig, Cornelius-Hen-
Hofbauer, Wallstr. 9.
*Rundfunk-Gesellschaft „Eintrag 1884“. Ludwig Nies, Wielandstraße 1. » Nr. 132.
*Gesangverein „Germania“. Vorstand: Erich Körner, Römerstraße 6.
Gesangverein „Rödertal“. Worms-Glanztal. Johann Kreiser, Kreuzstrasse 12
Gesangverein „Klöppelrin“. Franz Graf, Frankenthaler Straße 29. Vereinslokal:
Goldenes Kreuz, Gernersch. Straße 23.
Gesangverein „Büllesgäu 1“, gest. 1910 Worms. Vorstand: Karl Schneider,
Rehnbauerweg 23. Vorstand: Kapellmeister Albert Weinelt. Vereinslokal:
Zum Nordend, Siegfriedstraße 2.
c) Musikverein der Werke. Leitung: Adam Niehl.
*Rundfunk-Gesellschaft Worms. Vorstand: Karl Rössig, Cornelius-Hen-
Hofbauer, Wallstr. 9.
*Rundfunk-Gesellschaft „Eintrag 1884“. Ludwig Nies, Wielandstraße 1. » Nr. 132.
*Gesangverein „Germania“. Vorstand: Erich Körner, Römerstraße 6.
Gesangverein „Rödertal“. Worms-Glanztal. Johann Kreiser, Kreuzstrasse 12
Gesangverein „Klöppelrin“. Franz Graf, Frankenthaler Straße 29. Vereinslokal:
Table 1: Replication and sensitivity in PP

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) NSDAP28</th>
<th>(2) NSDAP28</th>
<th>(3) PCA_stnd</th>
<th>(4) PCA_stnd</th>
<th>(5) Deported</th>
<th>(6) Deported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pogrom</td>
<td>0.0142**</td>
<td>0.00294</td>
<td>0.290**</td>
<td>0.0588</td>
<td>0.142**</td>
<td>0.135</td>
</tr>
<tr>
<td></td>
<td>(0.00567)</td>
<td>(0.00283)</td>
<td>(0.132)</td>
<td>(0.0670)</td>
<td>(0.0706)</td>
<td>(0.137)</td>
</tr>
<tr>
<td>LogPop</td>
<td>-0.00254</td>
<td>0.00121</td>
<td>-0.0875</td>
<td>-0.0433</td>
<td>0.241***</td>
<td>1.135***</td>
</tr>
<tr>
<td></td>
<td>(0.00219)</td>
<td>(0.000900)</td>
<td>(0.0646)</td>
<td>(0.0296)</td>
<td>(0.0841)</td>
<td>(0.0311)</td>
</tr>
<tr>
<td>Jewish_pc</td>
<td>0.00174</td>
<td>0.000705</td>
<td>0.0215</td>
<td>0.0601</td>
<td>0.0743**</td>
<td>0.384***</td>
</tr>
<tr>
<td></td>
<td>(0.00190)</td>
<td>(0.00131)</td>
<td>(0.0971)</td>
<td>(0.0439)</td>
<td>(0.0348)</td>
<td>(0.0340)</td>
</tr>
<tr>
<td>LogJews</td>
<td>0.815***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0822)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prot_pc</td>
<td>0.000290***</td>
<td>0.000138***</td>
<td>0.284***</td>
<td>0.254***</td>
<td>-0.0039***</td>
<td>-0.00431**</td>
</tr>
<tr>
<td></td>
<td>(8.84e-05)</td>
<td>(4.06e-05)</td>
<td>(0.0757)</td>
<td>(0.0322)</td>
<td>(0.00116)</td>
<td>(0.00178)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0340*</td>
<td>-0.00295</td>
<td>-0.0801</td>
<td>-0.341***</td>
<td>-2.612***</td>
<td>-7.613***</td>
</tr>
<tr>
<td></td>
<td>(0.0195)</td>
<td>(0.00856)</td>
<td>(0.106)</td>
<td>(0.0668)</td>
<td>(0.462)</td>
<td>(0.372)</td>
</tr>
</tbody>
</table>

Observations: 325 325 311 311 278 278
Estimated by: OLS QR OLS QR Poisson Poisson

Source: All models estimated using PP replication data

Notes: Column (1) replicates PP Table VI Column (2). The dependent variable is the Nazi vote share in the 1928 election. Column (2) estimates Column (1) as a quantile (median) regression. Column (3) replicates PP Table VII Column (1). The dependent variable is the first principle component of the six outcome variables in PP Table VI. Column (4) estimates Column (3) as a quantile (median) regression. Column (5) replicates PP Table VI Column (4). The dependent variable is the number of Jews deported from the place. Column (6) estimates the same model but drops the superfluous “LogJews” regressor. Column (6) uses the same sub-sample as Column (5); see text for discussion of coding error that unnecessarily drops observations from PP’s Table VI Column (5). The precise definitions of the controls varies across specifications; this table always uses the definition that underlies the model in PP. In every case, the Pogrom proxy is defined as in the text, and “Prot_pc” is the percentage Protestant in 1925. In Columns (1) and (2), the city population and Jewish percentage are from the 1925 census. In Columns (3) – (6) they are from the 1933 census. In Columns (3) and (4) all variables, including the dependent variable, have been standardized. See the Appendix for additional checks that consider provincial interactions with the pogrom proxy as well as functional-form issues in the poisson models.
Table 2: The liberal parties as placebos

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Pogrom point Estimate</th>
<th>SE</th>
<th>Obs</th>
<th>Adj R-sq</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1924 election</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 DDP24</td>
<td>0.0109**</td>
<td>(0.00544)</td>
<td>325</td>
<td>0.265</td>
<td>OLS</td>
</tr>
<tr>
<td>2 DDP24</td>
<td>0.00682</td>
<td>(0.00523)</td>
<td>325</td>
<td></td>
<td>QR</td>
</tr>
<tr>
<td>3 DVP24</td>
<td>0.00955</td>
<td>(0.00799)</td>
<td>325</td>
<td>0.233</td>
<td>OLS</td>
</tr>
<tr>
<td>4 DVP24</td>
<td>0.0167</td>
<td>(0.0109)</td>
<td>325</td>
<td></td>
<td>QR</td>
</tr>
<tr>
<td>5 DDP_DVP24</td>
<td>0.0205*</td>
<td>(0.0110)</td>
<td>325</td>
<td>0.306</td>
<td>OLS</td>
</tr>
<tr>
<td>6 DDP_DVP24</td>
<td>0.0294**</td>
<td>(0.0116)</td>
<td>325</td>
<td></td>
<td>QR</td>
</tr>
</tbody>
</table>

Note: The table presents placebo checks for models analogous to PP Table VI, Column (3). We report the point-estimate and standard error for the pogrom proxy; every regression includes all the controls in VV’s analogous model. The DDP and DVP grew out of the Wilhelmine-era National Liberal and Progressive parties. DVP_DDP is the sum of the two party’s vote shares. Appendix Tables A6.1-A6.4 for other parties and elections. In 1928, both the DDP and DVP had drifted right. The DVP in particular had shared some electoral lists with a right-wing party that had some ideological overlap with the Nazis (the Volksnationale Reichsvereinigung). The DVP results for 1928 are different from what we show here. Tables A6.1 – A6.4 also indicate that the effect of the pogrom proxy on electoral outcomes varies by region.
### Table 3: Alternative definitions of the stability index

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clubs_all_pc</td>
<td>0.349***</td>
<td>0.263</td>
<td>0.0999</td>
<td>0.198</td>
<td>0.134**</td>
</tr>
<tr>
<td></td>
<td>(0.128)</td>
<td>(0.183)</td>
<td>(0.147)</td>
<td>(0.183)</td>
<td>(0.0524)</td>
</tr>
<tr>
<td>Stability index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.741</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.631)</td>
</tr>
<tr>
<td>Stability index x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.0424</td>
</tr>
<tr>
<td>clubs_all_pc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0329)</td>
</tr>
<tr>
<td>LnPop25</td>
<td>0.192</td>
<td>0.371*</td>
<td>0.0324</td>
<td>-0.0164</td>
<td>0.136**</td>
</tr>
<tr>
<td></td>
<td>(0.134)</td>
<td>(0.218)</td>
<td>(0.125)</td>
<td>(0.168)</td>
<td>(0.0512)</td>
</tr>
<tr>
<td>Cath_pc25</td>
<td>-0.525</td>
<td>0.0644</td>
<td>-0.998**</td>
<td>-1.490**</td>
<td>-0.804***</td>
</tr>
<tr>
<td></td>
<td>(0.388)</td>
<td>(0.554)</td>
<td>(0.442)</td>
<td>(0.597)</td>
<td>(0.115)</td>
</tr>
<tr>
<td>BCollar_pc25</td>
<td>-0.272</td>
<td>1.287</td>
<td>-0.553</td>
<td>-1.511</td>
<td>-1.883***</td>
</tr>
<tr>
<td></td>
<td>(1.929)</td>
<td>(2.128)</td>
<td>(1.427)</td>
<td>(1.497)</td>
<td>(0.385)</td>
</tr>
<tr>
<td>Stability index x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.0224</td>
</tr>
<tr>
<td>LnPop25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0465)</td>
</tr>
<tr>
<td>Cath_pc25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.239**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0924)</td>
</tr>
<tr>
<td>BCollar_pc25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.955***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.282)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.239</td>
<td>-4.869*</td>
<td>-0.0791</td>
<td>0.529</td>
<td>-0.599</td>
</tr>
<tr>
<td></td>
<td>(1.833)</td>
<td>(2.712)</td>
<td>(1.683)</td>
<td>(2.206)</td>
<td>(0.693)</td>
</tr>
<tr>
<td>Observations</td>
<td>58</td>
<td>58</td>
<td>54</td>
<td>35</td>
<td>225</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.108</td>
<td>0.055</td>
<td>0.178</td>
<td>0.217</td>
<td></td>
</tr>
<tr>
<td>Estimator</td>
<td>OLS</td>
<td>QR</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
</tr>
<tr>
<td>Mean (med) dep var</td>
<td>0.463</td>
<td>0.463</td>
<td>0.00923</td>
<td>0.0266</td>
<td>0.0266</td>
</tr>
<tr>
<td>Reg beta</td>
<td>0.440</td>
<td>0.332</td>
<td>0.141</td>
<td>0.265</td>
<td>0.265</td>
</tr>
</tbody>
</table>

Source: Computed from BF replication data

Note: Column (1) replicates BF Table 7, Column (3). The sub-sample includes only “unstable” states as defined by BF. Column (2) estimates the model in Column (1) by quantile (median) regression. Column (3) drops the third element from the stability index, but treats the median state as do BF’s authors, assigning it to the “unstable” category. Column (4) defines the stability index as in BF but considers the median state to be “stable.” Column (5) replicates the regression that underlines BF Appendix Figure A7. (BF does not report the actual regression). The sample for Column (5) is the entire dataset, including Prussia. The model uses the continuous stability index as defined in BF. See our appendix text (section B.3) for additional discussion of this model and BF Figure A7. Our appendix Table B3.6 reports computations for the net effect of stability in selected states, showing that with this specification, social capital only affects Nazi recruitment in stable states.