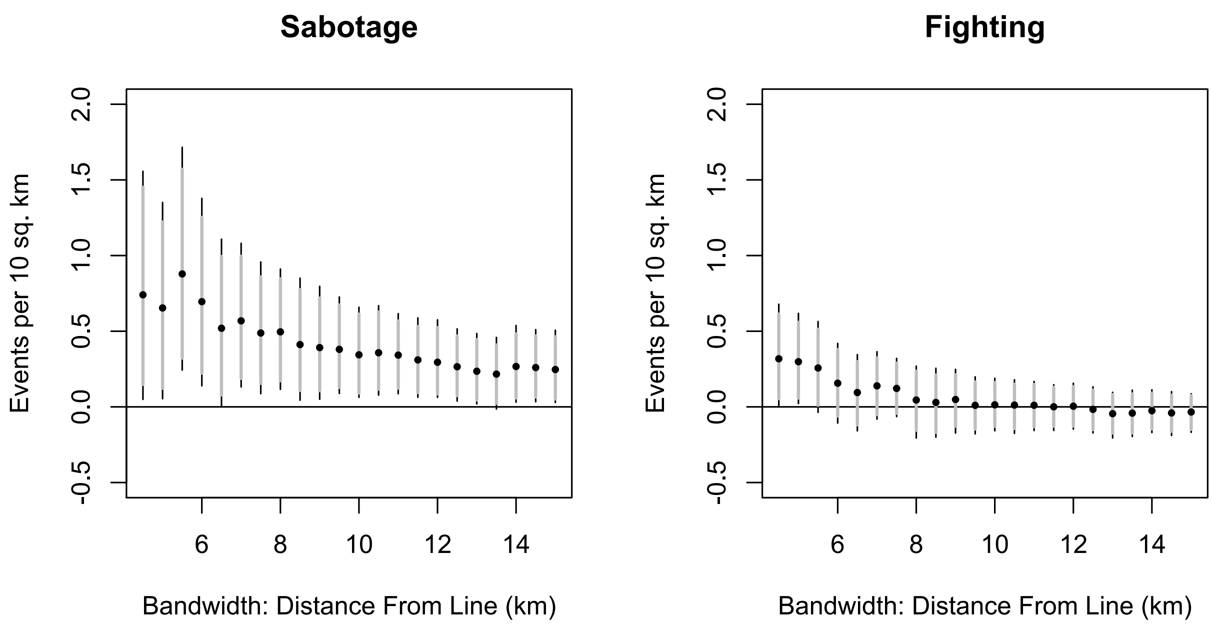
# Political Devolution and Resistance to Foreign Rule: A Natural Experiment

Appendix

## Contents

## 1. Supplementary Tables/Figures

Figure A1: Difference in Resistance Rates, with Bootstrapped SEs.




Supplement to Table 3. Displays differences in resistance rates and bootstrapped standard errors (n=1000) across several bandwidths. Thick lines represent 90% confidence intervals while thin lines represent 95% confidence intervals. The estimate refers to the difference in the number of events per 10 square kilometers: ($\sum Y_{D=1}$ / $\sum L_{D=1}$) - ($\sum Y_{D=0}$ / $\sum L_{D=0}$) for each bandwidth, where D is 1 when the commune is in the German Zone and L is the land area of the commune in 10 $km^2$.

Table A1: Nearest Neighbor with Exact Matching on Department

As a robustness check for the entropy balancing results (Table 4), we use nearest neighbor matching.

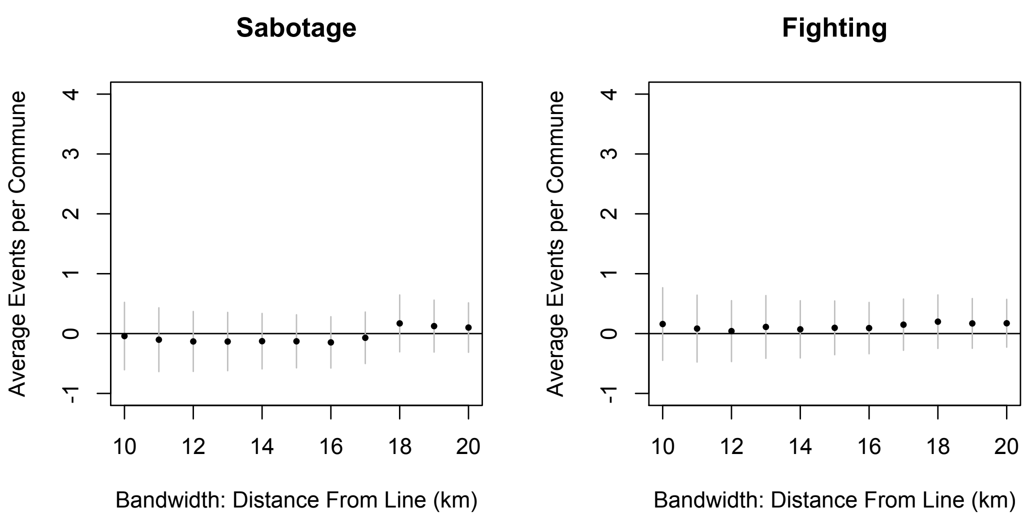
| Bandwidth | **30km** | | **20km** | | **10km** | |
|---|---|---|---|---|---|---|
| | Est. | t | Est. | t | Est. | t |
| **OLS** | | | | | | |
| Sabotage | 0.36 | 2.31 | 0.38 | 1.87 | 0.88 | 2.26 |
| Fighting | -0.11 | -0.97 | -0.08 | -0.60 | 0.05 | 0.19 |
| **Neg. Binomial Model** | | | | | | |
| Sabotage | 0.69 | 2.77 | 0.67 | 2.40 | 1.08 | 2.52 |
| Fighting | 0.32 | 1.50 | 0.37 | 1.63 | 0.30 | 1.03 |
| Post-Match Balance | Diff | t | | | | |
| Train Distance | -0.43 | -1.70 | -0.16 | -0.45 | -0.74 | 1.64 |
| Communications | 0.07 | 0.56 | 0.08 | 0.61 | 0.25 | 1.30 |
| Farmed Area | 3.91 | 1.43 | 2.96 | 0.76 | 4.63 | 0.90 |
| Log Population | 0.09 | 1.18 | 0.13 | 1.55 | 0.20 | 1.52 |
| % Right 1936 | 0.01 | 0.51 | 0.01 | 0.25 | 0.01 | 0.37 |
| % Left 1936 | -0.02 | -1.70 | -0.02 | -1.04 | -0.02 | -0.63 |
| Ruggedness | 0.02 | 0.38 | -0.02 | -0.33 | 0.05 | 0.56 |
| Log Mean Elev | -0.06 | -1.82 | -0.07 | -1.85 | -0.05 | 1.04 |

Number of matches=1, with replacement. Treatment = Vichy Zone (signs inverted for consistency). Matched on all covariates in the bottom half of the table, with the addition of distance to the line in kilometers and department (exact). Models fit on the matched data adjust for remaining post-match imbalance. The negative binomial model includes a logarithmic offset for land area.

Figure A2: Additional Placebo Tests

The placebo test introduces a false demarcation line within each zone, and tests whether a discontinuity in the dependent variable can be detected using several different bandwidths (as in the main model, set at a distance of 10km to 20km from the false line). The selection of a placebo point is arbitrary, but bounded by the following criteria: the bandwidths should not overlap the (true) demarcation line. In the following set of plots, we show an additional placebo test with a false demarcation line set at the 75% distance percentile within each zone. As with the 50% percentile, no statistically significant effect is visible.

North Placebo




South Placebo

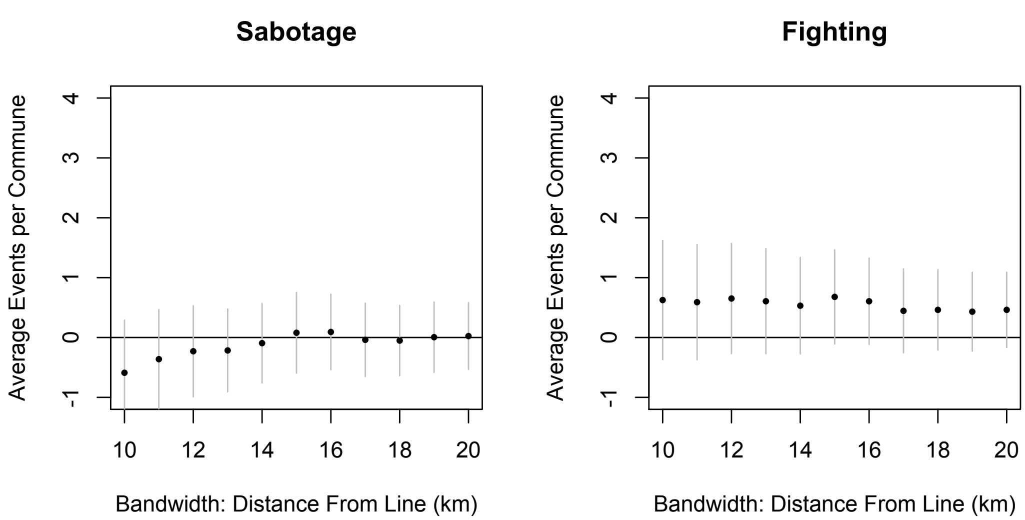


Table A2: Robustness Check - Border Effects

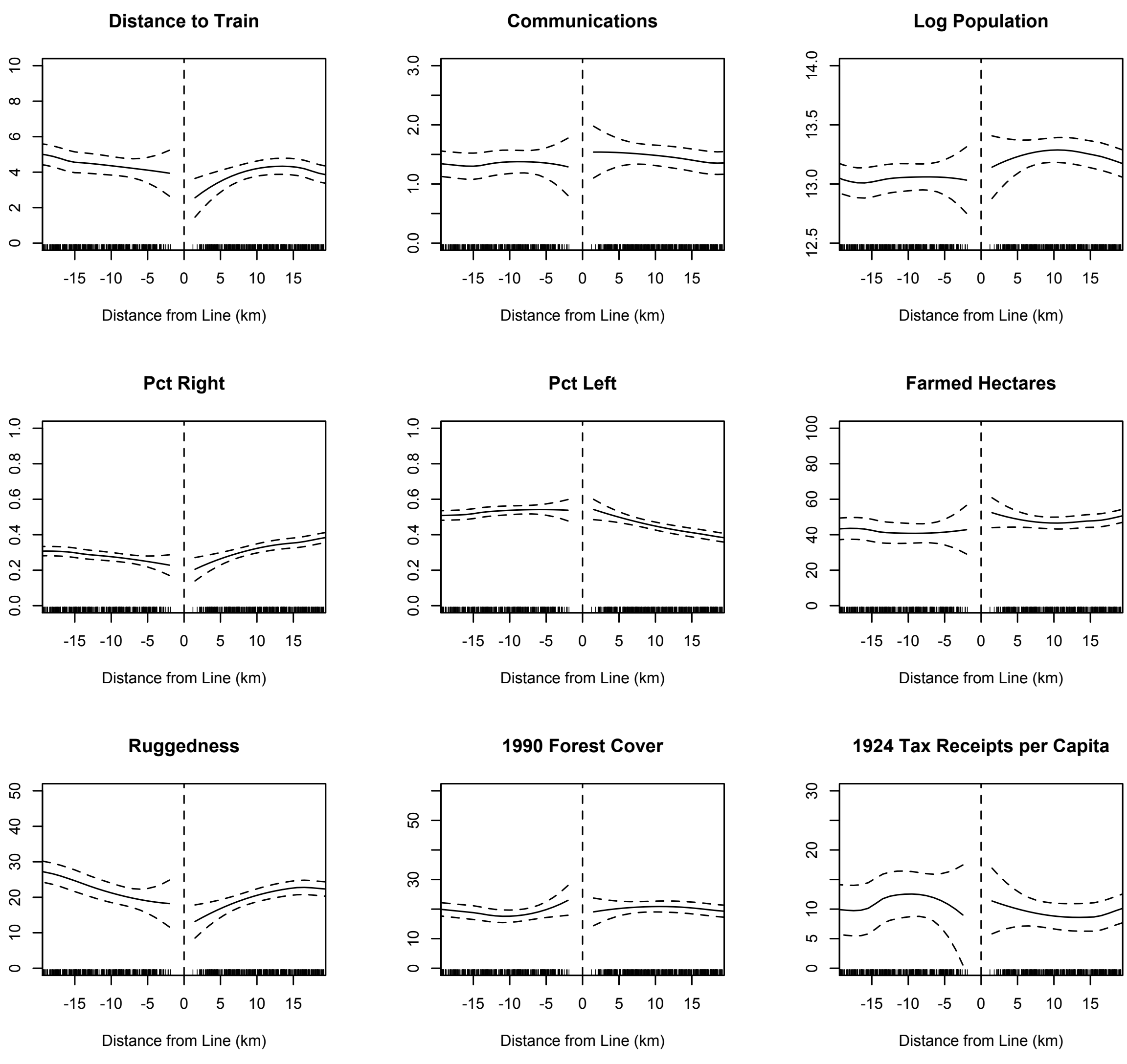
To assess whether the results are driven purely by a border effect, in which the border itself motivates resistance, we drop all communes within 5km of the line. Since this leads to poor balance, we use matching (entropy balancing) to reduce imbalance in observed covariates, and estimate a treatment effect for the German Zone. Although attenuation is visible, the point estimates for sabotage remain positive and statistically significant, suggesting that the effect is not driven solely by communes abutting the border.

| Bandwidth | **Full Sample** | | **30km** | | **20km** | |
|---|---|---|---|---|---|---|
| | Est. | t | Est. | t | Est. | t |
| **Difference in Means** | | | | | | |
| Sabotage | 0.31 | 2.33 | 0.43 | 2.43 | 0.34 | 1.71 |
| Fighting | 0.06 | 0.74 | 0.08 | 0.83 | 0.10 | 0.87 |
| **Neg. Binomial Model** | | | | | | |
| Sabotage | 0.52 | 2.13 | 0.70 | 2.70 | 0.69 | 2.27 |
| Fighting | 0.06 | 0.38 | 0.00 | 0.00 | 0.15 | 0.64 |
| Max Weight | 4.10 | | 3.07 | | 3.11 | |
| Mean Weight | 0.70 | | 0.84 | | 0.91 | |
| SD Weight | 1.03 | | 0.96 | | 0.88 | |
| Max t | 1e-4 | | 5e-4 | | 3e-5 | |

Matching parameters similar to Table 4, in text.

## 2. Additional Balance Tests

Figure A3 : Covariate Density Plots
Although Table 3 demonstrates that covariates are balanced close to the demarcation line, in the following series of plot we provide a visible depiction of the distribution of each covariate using a LOESS smoother. Vichy communes are presented with negative distances. We also include two additional covariates not included in the main table: % Forest Cover per Commune in 1990[1], and a measure of wealth: Tax Receipts per Capita in 1924. [2]




[1] 1990 is the earliest year for which GIS forest cover data is available for all of France. As a result, 1936 farming data, which we opted to use in-text, is likely a better approximation of forest cover in this case.

[2] More proximate data on commune-level finances is unavailable. After 1924, only records for communes with a population greater than 5,000 were published.

Table A3: Regression Balance Tests

As a final balance check, we analyze whether point estimates for $\beta_1$ differ significantly between the following two models at several bandwidths, using OLS:

$$(1) \qquad y = \beta_0 + \beta_1 D$$

$$(2) \quad y = \beta_0 + \beta_1 D + X$$

Where y is sabotage, D is 1 when the commune is in the German zone, and X includes all covariates listed in Table 2. [3]

| Bandwidth | **Full Sample** | **20km** | **15km** | **10km** | **5km** |
|---|---|---|---|---|---|
| Model 1 | 0.102 | 0.395 | 0.456 | 0.633 | 1.077 |
| Model 2 | 0.244 | 0.370 | 0.415 | 0.607 | 1.077 |

The results of this test are in agreement with Table 2 and suggest that while the sample is unbalanced at larger bandwidths, balance is achieved close to the line. In these narrower bandwidths, the inclusion of covariates results in a very small (and substantively insignificant) shift in the point estimates.

[3]Population, Farmed hectares, and Ruggedness are logged due to skewed distributions. Pctright is omitted due to the highly multicollinear relationship with Pctleft. Omitting Pctleft instead of Pctright leads to similar results.

## 3. Alternate Specifications

Table A4: Geographic Matching and LLR

In this approach, we use matching to balance the sample on a set of geographic covariates before we estimate the main model. Although this approach discards observations, it ensures that each treated unit is paired with a control unit that is geographically proximate (with respect to both latitude and longitude).

To implement this design, we trim the sample to the set of communes within 30km of the demarcation line, and then use nearest neighbor matching without replacement. Units were matched on department (exact), geographic distance ($\sqrt{latitude^2 + longitude^2}$), with political vote share in 1936 as a tiebreaker (given that vote shares are often geographically distributed). This resulted in a matched sample of 343 Vichy and 343 German communes.

The upper table displays balance statistics for several bandwidths using the revised sample, while the lower figure displays estimates from local linear regression using a rectangular and triangular kernel, respectively.

| | **<= 20km** | | **<= 15km** | | **<= 10km** | |
|---|---|---|---|---|---|---|
| | Diff | t | Diff | t | Diff | t |
| Train Distance | -0.21 | -0.62 | -0.22 | - 0.57 | -0.68 | -1.50 |
| Communications | 0.09 | 0.72 | 0.06 | 0.38 | 0.05 | 0.26 |
| Farmed Area | -4.24 | -1.13 | -3.63 | -1.01 | -1.22 | -0.34 |
| Log Population | 0.16 | 1.92 | 0.19 | 1.89 | 0.13 | 1.07 |
| % Right | 0.05 | 2.79 | 0.04 | 1.80 | 0.02 | 0.54 |
| % Left | -0.02 | -1.34 | -0.02 | -0.99 | -0.00 | -0.10 |
| Ruggedness | 0.08 | 1.21 | 0.06 | 0.85 | 0.01 | 0.06 |

| | **Sabotage** | | | | **Fighting** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Rectangle** | | **Triangle** | | **Rectangle** | | **Triangle** | |
| Bandwidth | Est. | t | Est. | t | Est. | t | Est. | t |
| 10km | 2.69 | 2.57 | 2.26 | 2.11 | 0.72 | 1.42 | 0.71 | 1.23 |
| 11km | 2.32 | 2.37 | 2.32 | 2.25 | 0.53 | 1.18 | 0.66 | 1.27 |
| 12km | 2.30 | 2.55 | 2.33 | 2.36 | 0.35 | 0.82 | 0.62 | 1.29 |
| 13km | 2.48 | 2.75 | 2.38 | 2.42 | 0.75 | 1.55 | 0.62 | 1.39 |
| 14km | 1.85 | 1.91 | 2.38 | 2.62 | 0.34 | 0.79 | 0.61 | 1.45 |
| 15km | 1.85 | 2.15 | 2.26 | 2.56 | 0.35 | 0.87 | 0.55 | 1.37 |

Table A5: Local Linear Specifications

To demonstrate that the substantive results hold across multiple specifications, we display local linear regressions using six alternate models, with *Sabotage* as the dependent variable. The first two specifications, referenced in text, are variants of the main model (1) without covariates and (2) with a population covariate. Note that the population covariate slightly attenuates the estimate: this is due to several large outliers rather than a lack of balance (see previous section).[4] The third and fourth model substitute a triangular kernel for a rectangular kernel. Although the triangular kernel is optimal for boundary estimation, due to the lack of observations within 2km of the line (communes intersected by the line were dropped), the triangular kernel places undue weight on regions without substantial support, and is thus not ideal for this application. As a potential correction, for the remaining example we buffer the running variable (distance to the demarcation line) by 2km; the resulting estimates can be interpreted as the estimated effect for communes that abut a 'demarcation zone' corresponding to the group of intersected communes. [5]

| | **1** | | **2** | | **3** | | **4** | | **5** | |
|---|---|---|---|---|---|---|---|---|---|---|
| Bandwidth | Est. | t | Est. | t | Est. | t | Est. | t | Est. | t |
| 10km | 1.94 | 2.03 | 1.60 | 2.04 | 1.80 | 1.77 | 1.73 | 1.95 | 1.43 | 2.08 |
| 11km | 1.64 | 1.88 | 1.45 | 1.97 | 1.78 | 1.86 | 1.65 | 1.99 | 1.45 | 1.98 |
| 12km | 1.59 | 2.00 | 1.35 | 2.05 | 1.74 | 1.92 | 1.57 | 2.02 | 1.43 | 2.08 |
| 13km | 1.69 | 2.22 | 1.41 | 2.23 | 1.75 | 2.03 | 1.52 | 2.10 | 1.43 | 2.17 |
| 14km | 1.29 | 1.71 | 0.97 | 1.50 | 1.72 | 2.11 | 1.47 | 2.16 | 1.41 | 2.22 |
| 15km | 1.25 | 1.86 | 0.93 | 1.63 | 1.62 | 2.08 | 1.35 | 2.09 | 1.34 | 2.19 |
| 16km | 1.18 | 1.92 | 1.08 | 1.95 | 1.54 | 2.06 | 1.27 | 2.05 | 1.29 | 2.17 |
| 17km | 1.02 | 1.75 | 0.98 | 1.89 | 1.45 | 2.03 | 1.22 | 2.05 | 1.23 | 2.15 |
| 18km | 0.90 | 1.65 | 0.79 | 1.67 | 1.35 | 1.98 | 1.15 | 1.99 | 1.15 | 2.10 |
| 19km | 0.86 | 1.69 | 0.72 | 1.61 | 1.26 | 1.94 | 1.07 | 1.94 | 1.09 | 2.07 |
| 20km | 0.98 | 1.97 | 0.78 | 1.85 | 1.20 | 1.92 | 1.01 | 1.91 | 1.05 | 2.06 |
| Kernel | Rectangle | | Rectangle | | Triangle | | Triangle | | Triangle | |
| Running Var. | Raw | | Raw | | Raw | | Raw | | 2km Buffer | |
| Area Covariate | | | | | Y | | | | Y | |
| Pop Covariate | | | Y | | | | Y | | | |

[4]Holding area or population constant when evaluating resistance activity is primarily a theoretical choice. Given that resistance groups may be partially mobile, we prefer an estimate that controls for the land area of a commune – the spatial plane on which events can occur. The observed attenuation is not present for narrower bandwidths; a bandwidth of 9km returns a point estimate of 1.92 for Model 2.

[5]Note that for comparability with the previous four models, the bandwidths have been preserved: 10km thus includes all communes within 10km of the demarcation line across both approaches. The only element that differs is the treatment of the running variable.

Table A6: Polynomial Specifications

In contrast to the non-parametric approach, which is based on relatively simple density estimates in close proximity to the threshold, the parametric approach typically relies on upper-order polynomials to model the distribution of the data. Although the flexibility of these models allow the user to capture the causal effect, they require the user to make assumptions about the underlying distribution and may use more degrees of freedom than a non-parametric local linear approach when multiple higher-order terms are included (Imbens et al 2008).

As a robustness check, in the table below we present polynomial estimates using an:

1. OLS model with 2nd order terms and interactions.
2. OLS model with 3rd order terms and interactions.
3. OLS model with 4th order terms and interactions.
4. Negative Binomial with 2rd order terms, interactions and a log offset for land area. [6]

We display results for the *Sabotage* variable for five different bandwidths. These bandwidths, which are larger than those used for the local linear estimates, represent a compromise between the sample size and the introduction of higher order terms. Although Table 2 indicates that covariate balance is not fully achieved at these wider bandwidths, it should be noted that this is not a requirement for a parametric RDD design. While it is crucial that covariates converge as the threshold is approached (as seen in Table 2), the full sample used for estimation does not have to be balanced provided that the model is sufficiently flexible to measure the actual distribution of the dependent variable in the neighborhood of the threshold (see Lee 2008).

| | **Model 1** | | **Model 2** | | **Model 3** | | **Model 4** | |
|---|---|---|---|---|---|---|---|---|
| Bandwidth | Est. | t | Est. | t | Est. | t | Est. | t |
| 25km | 1.51 | 2.65 | 2.85 | 1.96 | 2.86 | 3.07 | 1.93 | 2.14 |
| 30km | 1.60 | 3.35 | 3.33 | 2.89 | 2.04 | 2.69 | 2.20 | 2.79 |
| 35km | 1.34 | 3.25 | 2.54 | 2.65 | 2.09 | 3.24 | 1.65 | 2.26 |
| 40km | 1.28 | 3.38 | 2.53 | 3.32 | 1.84 | 3.21 | 1.49 | 2.18 |
| Full Sample | 1.01 | 3.20 | 1.26 | 2.70 | 1.55 | 2.27 | 1.52 | 2.43 |

The treatment is set to 1 when a commune is located in the German Zone.

Given that our sample size is relatively small and the dependent variable consists of non-normal, overdispersed, count data, we prefer the local linear approach. Nevertheless, the results above confirm the presence of a strong discontinuity in the *Sabotage* variable.

[6]The negative binomial models exhibit poor convergence with higher order terms and multiple interactions. Estimates for 3rd and 4th order models were statistically significant, but the point estimates were inconsistent across bandwidths.

Table A7: Spatial Coarsening

Since events occur on a spatial plane, it may be more appropriate to view the data in terms of geoegraphic bins than as commune-level observations. To do so, we pool communes into bins of width 0.5km or 1km (based on the commune centroid), and calculate the dependent variable as the observed rate of resistance activity per 10 $km^2$ within each bin. Although this approach discards detail, it a) allows us to obtain an unbiased rate and b) avoids potential bias that may result from an unequal number of communes on either side of the line (due to arbitrary commune borders and variation in commune size). Using this coarsened data, we then obtain point estimates using Local Linear Regression with rectangular and triangular kernels:

Sabotage

| | **Bin=0.5km** | | | | **Bin=1.0km** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Rectangle** | | **Triangle** | | **Rectangle** | | **Triangle** | |
| Bandwidth | Est. | t | Est. | t | Est. | t | Est. | t |
| 10km | 1.24 | 3.52 | 1.33 | 3.65 | 1.16 | 3.21 | 1.19 | 3.05 |
| 11km | 1.06 | 3.22 | 1.26 | 3.73 | 1.02 | 3.11 | 1.15 | 3.21 |
| 12km | 0.98 | 3.22 | 1.19 | 3.73 | 0.97 | 3.23 | 1.10 | 3.29 |
| 13km | 1.03 | 3.39 | 1.12 | 3.71 | 1.03 | 3.63 | 1.05 | 3.35 |
| 14km | 0.81 | 2.53 | 1.08 | 3.72 | 0.83 | 2.77 | 1.00 | 3.42 |
| 15km | 0.77 | 2.70 | 1.01 | 3.61 | 0.80 | 3.03 | 0.96 | 3.46 |

Fighting

| | **Bin=0.5km** | | | | **Bin=1.0km** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Rectangle** | | **Triangle** | | **Rectangle** | | **Triangle** | |
| Bandwidth | Est. | t | Est. | t | Est. | t | Est. | t |
| 10km | 0.66 | 2.20 | 0.77 | 1.94 | 0.57 | 4.06 | 0.66 | 5.07 |
| 11km | 0.57 | 2.08 | 0.73 | 2.08 | 0.48 | 3.11 | 0.63 | 5.07 |
| 12km | 0.47 | 1.88 | 0.67 | 2.13 | 0.46 | 3.10 | 0.59 | 4.73 |
| 13km | 0.59 | 2.48 | 0.62 | 2.15 | 0.55 | 4.07 | 0.56 | 4.50 |
| 14km | 0.44 | 1.99 | 0.59 | 2.23 | 0.39 | 2.41 | 0.54 | 4.59 |
| 15km | 0.41 | 2.03 | 0.56 | 2.25 | 0.35 | 2.60 | 0.50 | 4.32 |

## 4. Miscellaneous

Table A8 : Population Flows

The population measure we use for each commune is drawn from pre-war data. However, it is possible that migration during the 1940-1944 period influenced the population of each commune; moreover, this migration may not have been exogenous with respect to the demarcation line. Although no intra-war census or population survey exists, we provide an indirect control for this possibility by examining relative population shifts experienced between the 1936 and 1946 censuses. Specifically, we implement a difference-in-differences (DID) design.

| Bandwidth | **20km** | **15km** | **10km** |
|---|---|---|---|
| Δ 1946 - 1936 (German Zone) | 39.17 | -2.12 | 10.17 |
| Δ 1946 - 1936 (Vichy Zone) | -22.22 | -20.42 | -23.31 |
| DID | 61.39 | 18.3 | 33.48 |
| t | 2.14 | 0.67 | 1.32 |

Change in mean persons per commune between 1936 and 1946.

The results suggest that no statistically or substantively significant difference exists across the two zones in the rate of population change between 1936 and 1946 within narrower bandwidths.

Figure A4 : Distribution of the Dependent Variable

Given the large degree of variance within the dependent variable, we display the distribution of sabotage and fighting activity across both zones. All histograms exclude communes with 0 events. Of the 17 communes with $> 10$ resistance events, 9 are in the German Zone.

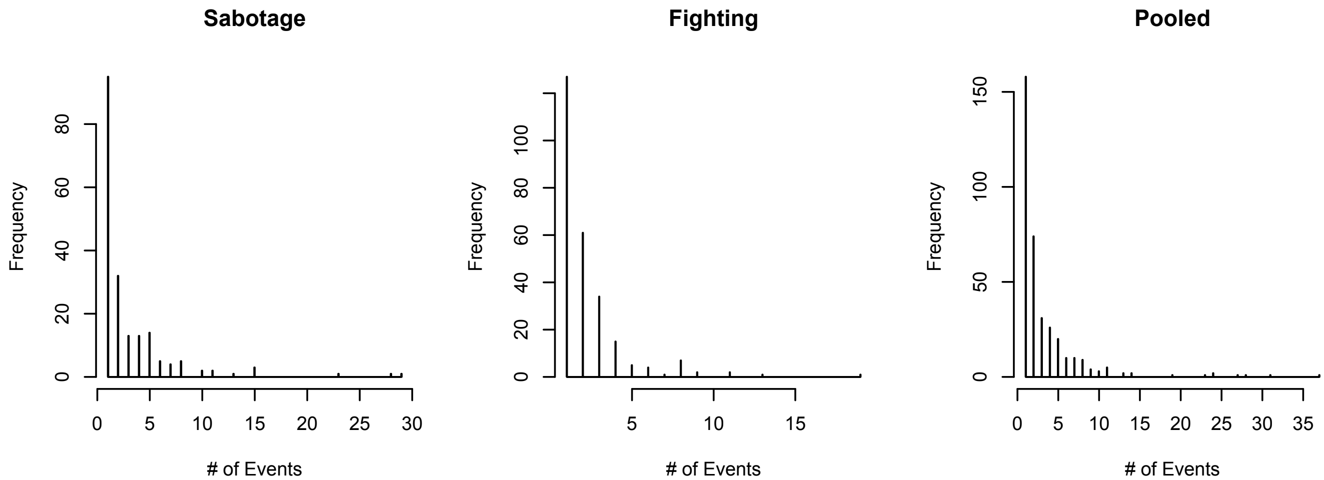
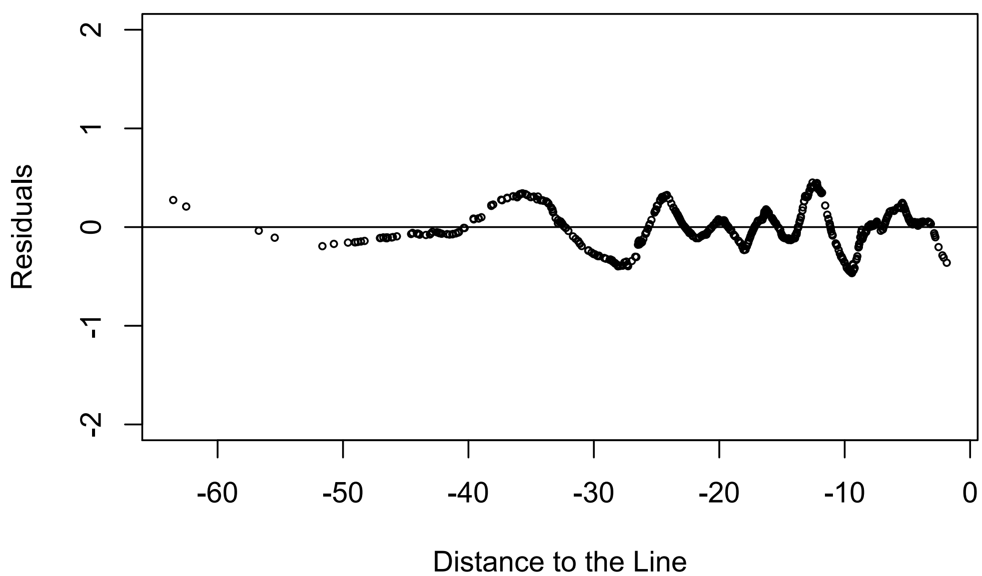


Figure A5: Robustness Check - Spillover

In order to assess the degree of spillover, we use a loess plot to display the smoothed residuals for Vichy communes from Model 3 in Table A6 (4th order OLS) as a function of distance from the demarcation line. If significant spillover is present, the model should overpredict Vichy resistance activity close to the line (due to displacement of these events to the German side). However, no such pattern is visible.




LOESS Bandwidth: 0.1km

Table A10: Proxies for Pre-Nov 1942 Repression: German Zone.

The next two tables address the possibility that heightened repression in the pre Nov-1942 period accounts for subsequent levels of resistance within the German zone. The results suggest that this repression cannot explain the findings.

| | **Intern. Camp** | | **Reported Arrests** | | **Res. Deaths** | | **Combined** | |
|---|---|---|---|---|---|---|---|---|
| | Direct | Bordering | Direct | Bordering | Direct | Bordering | Direct | Bordering |
| **Controls** | | | | | | | | |
| None | 1.59 | 0.62 | 1.79 | 0.44 | 1.30 | 0.58 | 1.48 | 0.32 |
| | (2.44) | (1.42) | (2.11) | (0.71) | (1.75) | (0.61) | (2.58) | (0.78) |
| Non-Spatial | 0.41 | -0.04 | 0.41 | -0.55 | -0.46 | -0.06 | 0.42 | -0.39 |
| | (0.54) | (-0.10) | (0.66) | (-1.33) | (-0.39) | (-0.08) | (0.64) | (-0.90) |
| All | 0.41 | - 0.11 | 0.55 | -0.48 | -0.50 | -0.36 | 0.50 | -0.42 |
| | (0.77) | (-0.35) | (1.19) | (-1.44) | (-0.66) | (-0.57) | (1.11) | (-1.52 ) |

The dependent variable is the count of resistance events occurring after November 1942. We subset the sample to cover the German-occupied zone only. The coefficients are from a negative binomial model with a logarithmic offset for land area and varying controls: 'None' indicates no covariates, 'Non-Spatial' indicates all covariates in Table 2, and 'All' includes cubic terms for the distance to the line and departmental dummies. The key independent variables are 'Direct', which is set to '1' when an event was reported in the commune, and 'Bordering' which is set to '1' if an event occurred in the commune or in a bordering commune. Results are similar when examining sabotage only (not shown).

Source for Internment Camp and Arrests Data:
AJPN: Anonymes, Justes et Persécutés durant la période Nazie dans les communes de France.

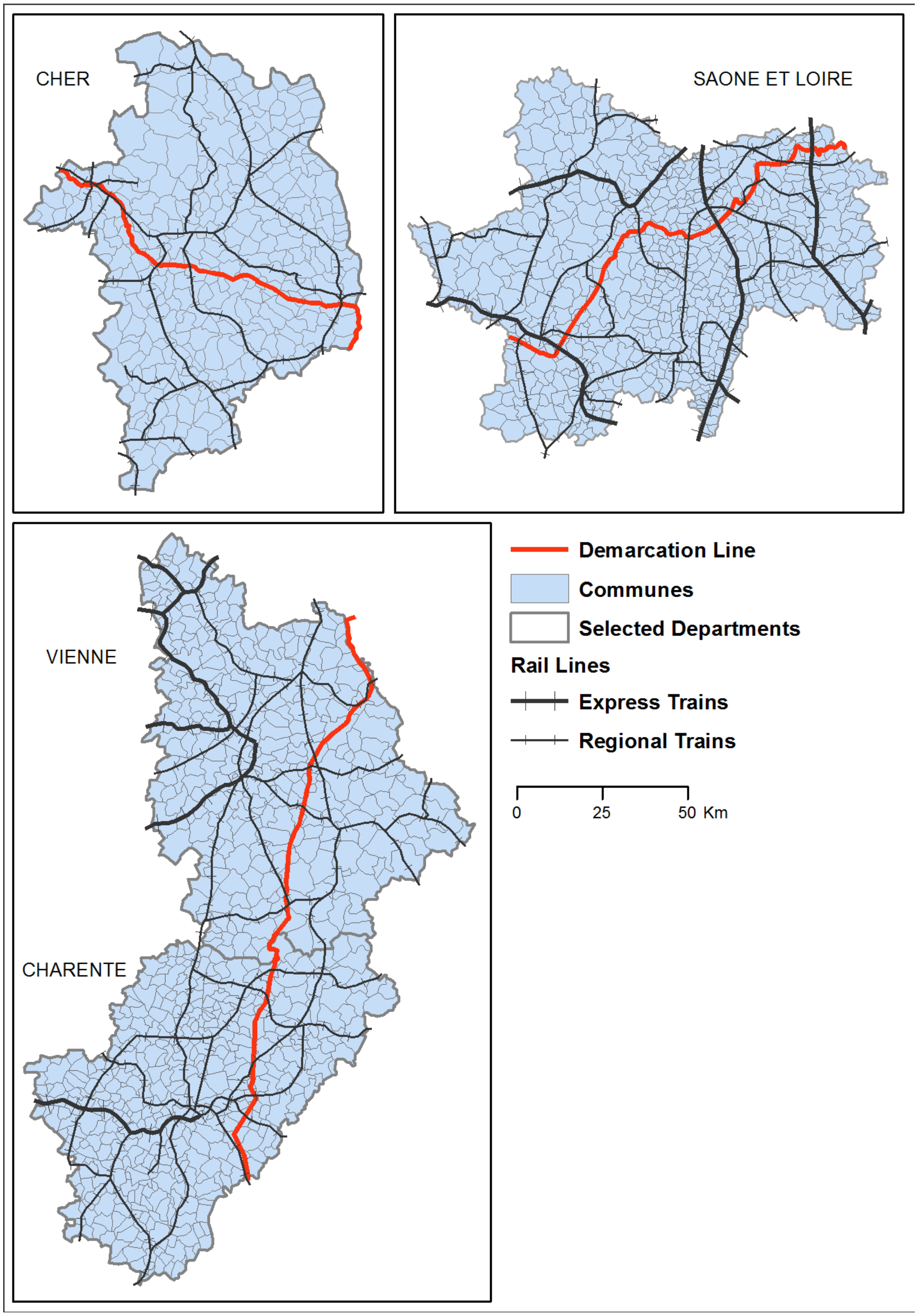
Source for Deaths Data: French Department of Defense

Table A11: Robustness Check with Repression - Matching (Entropy Balancing)

| Bandwidth | **Full Sample** | | **30km** | | **20km** | | **10km** | |
|---|---|---|---|---|---|---|---|---|
| | Est. | t | Est. | t | Est. | t | Est. | t |
| **Difference in Means** | | | | | | | | |
| Sabotage | 0.38 | 2.83 | 0.48 | 2.988 | 0.38 | 2.17 | 0.61 | 2.12 |
| Fighting | 0.01 | 0.06 | 0.01 | 0.09 | 0.02 | 0.14 | 0.04 | 0.18 |
| Max Weight | 3.24 | | 2.89 | | 2.71 | | 2.14 | |
| Mean Weight | 0.71 | | 0.85 | | 0.92 | | 0.92 | |
| SD Weight | 0.92 | | 0.85 | | 0.82 | | 0.61 | |
| Max t | 1e-5 | | 5e-3 | | 1e-4 | | 7e-4 | |

Identical to Table 4 in the text, with the addition of a covariate measuring whether an internment camp or a reported resistance death occurred in the commune prior to Nov 1942.

Figure A6: Train Lines




Source: "Carte du Resau des Chemins de Fer de L'Etat 1926" and "Carte du Resau des Chemins de Fer de Paris a Lyon 1927." L'Institute Geographique de Paris.