

# The origins of common identity: Evidence from Alsace-Lorraine\*

Sirus H. Dehdari<sup>†</sup>  
IIES, Stockholm University

Kai Gehring<sup>‡</sup>  
University of Zurich

August 5, 2018

## Abstract:

The quasi-exogenous division of the French regions Alsace and Lorraine after the Franco-Prussian War allows us to provide novel evidence about group identity formation within historically homogeneous regions. Using several measures of stated and revealed preferences spanning over half a century, we show that being exposed to occupation and repression for many decades caused a persistently stronger regional identity. The geographical RDD results are robust across all specifications. We document two mechanisms using data on regional newspapers and regionalist parties. The differences are strongest for the first two age cohorts after WWII and associated with preferences for more regional decision-making.

*Keywords: Group identity, Regional Identity, Identity formation, Persistence of preferences, Homogenization policies, Assimilation, Alsace-Lorraine*

*JEL: D91, H70, H80, N40, Z19*

---

\*We thank Alberto Alesina, Olivia Bertelli, Carles Boix, Konrad Burchardi, Davide Cantoni, Eric Chaney, Melissa Dell, Axel Dreher, Vasilika Fouka, Boris Gershman, Nathaniel Lane, Horacio Larreguy, Raphael Frank, Andreas Madestam, Shom Mazumder, Pierre-Guillaume Méon, Stelios Michalopoulos, Nathan Nunn, Per Petterson-Lidbom, Torsten Persson, James Robinson, David Strömberg, Jakob Svensson, Hans-Joachim Voth, Fabian Wahl, and conference participants at the NBER Summer Institute 2017, the CesIFO workshop in Venice 2017, the History Tea (Harvard University), Solvay Business School Brussels 2017, the BBQ 2017 conference at Lago di Garda, the EEA/ESEM 2016 in Geneva, the Silvaplana Workshop in Political Economy 2016, the SSES in Lugano 2016, AEL in Heidelberg 2016, ASWEDE in Gothenburg 2016, the Politics and History Network at Princeton University 2016, the Royal Economic Society Junior Symposium in Brighton 2016, Stockholm University and the University of Freiburg for feedback and great suggestions. The same holds for seminar comments at Barcelona, ETH Zurich, Hannover University, Harvard University, Heidelberg University, and Mannheim University. Excellent research assistance was provided by Janine Albiez, Philipp Kerler, Max Huppertz, Eva Marti, Norah Schmidlin, Alfred Sutter and Lukas Willi. We thank Valerie Baldinger and Michele McArdle for proof-reading. Kai Gehring acknowledges funding from the Swiss National Science Foundation Ambizione project PZ00P1174049.

This is a new version of CESifo Working Paper Series No. 6556.

<sup>†</sup>sirus.dehdari@iies.su.se

<sup>‡</sup>mail@kai-gehring.net

# Online Appendix to “The origins of common identity: Division, homogenization policies and identity formation in Alsace-Lorraine”

Sirus H. Dehdari\*  
IIES, Stockholm University

Kai Gehring†  
University of Zurich

August 5, 2018

## Contents

<b>A Theoretical framework</b>	<b>4</b>
<b>B Referendum 1969</b>	<b>8</b>
<b>C Alternative explanations</b>	<b>9</b>
C.1 Support driven by urban agglomerations . . . . .	9
C.2 Religiosity and EU support, relevant for 1992 and 2005 referenda . . . . .	9
C.3 Differences in benefits from trade . . . . .	10
C.4 The relative importance of homogenization policies . . . . .	10
C.5 The influence of Germanization . . . . .	10
<b>D Twitter</b>	<b>12</b>
<b>E Links to examples for other regions mentioned in paper</b>	<b>13</b>
<b>F Tables</b>	<b>14</b>
<b>G Figures</b>	<b>40</b>

---

\*sirus.dehdari@iies.su.se

†mail@kai-gehring.net

## List of Tables

A1	Detailed Overview of Repressive Policies in Alsace and Lorraine . . . . .	14
A2	Variable description and sources . . . . .	18
A3	Survey questions (i.) . . . . .	19
A4	Survey questions (ii.) . . . . .	20
A5	Survey questions (iii.) . . . . .	21
A6	Survey questions (iv.) . . . . .	22
A7	Survey results, Alsace and Lorraine . . . . .	23
A8	Overlap strength of regional and European identity in treated and control areas (A+L)	23
A9	OIP Survey results, 1999 and 2001: European and regional attachments . . . . .	23
A10	Descriptive statistics for outcome variables and treatment . . . . .	24
A11	Descriptive statistics for RDD control and pre-treatment variables . . . . .	24
A12	Pre-treatment variables balance test, within Lorraine . . . . .	25
A13	Pre-treatment variables balance test, Alsace and Lorraine . . . . .	25
A14	OLS estimates using all municipalities in Alsace and Lorraine . . . . .	26
A15	Discontinuities in turnout, within Lorraine . . . . .	27
A16	RD results: Nationalism (Le Pen) and turnout, Alsace and Lorraine . . . . .	28
A17	Covariate balance test: 4 categories . . . . .	29
A18	Covariate balancing . . . . .	30
A19	Discontinuities in referenda results, within Lorraine, excluding Metz . . . . .	31
A20	Share of Yes votes and religion . . . . .	32
A21	Effects at the language border . . . . .	33
A22	Differences in population changes between 1916-1946 (within Lorraine) . . . . .	34
A23	Discontinuities in referenda results, within Lorraine, controlling for historical migration	34
A24	Placebo test: Border between Alsace and Lorraine, and the rest of France . . . . .	35
A25	Placebo test: RD estimates at the pre-1870 border between historical Moselle and Meurthe, within current Moselle . . . . .	36
A26	Newspaper subscription shares: excluding Metz, and discontinuity at language border	37
A27	RD results: Subscription shares of regional newspaper, controlling for the number of sales points . . . . .	37
A28	List of Twitter Keywords . . . . .	38
A29	Twitter data, within Lorraine . . . . .	39
A30	Survey results: policy preferences . . . . .	70

## List of Figures

A1	Threshold costs for teaching regional tradition . . . . .	6
A2	Sample from <i>L'Est Republicain</i> showing voting results . . . . .	8
A3	Map of Lotharingia around 1000 A.D. . . . .	40
A4	Map of Lotharingia around 1000 A.D., zoomed in with 1870 border . . . . .	41
A5	Map of Lorraine in the 1378 century . . . . .	42
A6	Map of Lorraine in the 17th century . . . . .	43
A7	Timeline of outcomes . . . . .	44
A8	Agreement referendum about establishing regions as political entity, 1969 . . . . .	45
A9	Election and referendum results, 1968 and 1969 . . . . .	46
A10	The average and standard deviation of critical parameters . . . . .	47
A11	RD plots, within Lorraine (50 kilometers, 20km in paper), 1st degree polynomial . . . . .	48
A13	Estimation plots for 1969 referendum, within Lorraine . . . . .	52
A14	Estimation plots, 1969 referendum, controlling for border segments, and distance to language border . . . . .	53
A15	Estimation plots, 1969 referendum, controlling for longitude, latitude and their interaction . . . . .	54
A16	Estimation plots for 1992 and 2005 referenda, within Lorraine . . . . .	55
A17	Estimation plots for 1992 and 2005 referenda, whole border . . . . .	56
A18	Estimation plots for 1992 and 2005 referenda, no controls . . . . .	57
A19	Estimation plots, controlling for longitude and latitude . . . . .	58
A20	Estimation plots, controlling for longitude, latitude and their interaction . . . . .	59
A21	Estimation plots for 1992 and 2005 referenda, controlling for border segments . . . . .	60
A22	Estimation plots for 1992 and 2005 referenda, controlling for border segments (within Lorraine) . . . . .	61
A23	Estimation plots, controlling for distance to language border . . . . .	62
A24	Estimation plots, newspaper subscription shares, 1st and 2nd degree polynomial . . . . .	63
A25	Estimation plots, newspaper subscription shares, controlling for longitude, latitude and their interaction . . . . .	64
A26	Estimation plots, newspaper subscription shares, controlling for distance to language border . . . . .	65
A27	RD plot, share of households with subscription of “Le Republicain Lorraine”, 2nd degree polynomial . . . . .	66
A28	Newspaper subscription shares . . . . .	66
A29	Subscription page Le Republicain Lorraine (1) . . . . .	67
A30	Subscription page Le Republicain Lorraine (2) . . . . .	67
A31	Homepage (main) Le Republicain Lorraine . . . . .	68
A32	Homepage (regional) Le Republicain Lorraine . . . . .	68
A33	Vote shares of regionalist parties . . . . .	69
A34	Identity differences by age cohort, relative to National identity . . . . .	70



## A Theoretical framework

The common identity of an individual  $i$  and a group  $j \in \{R, N\} = \mathbf{J}$ , with  $R$  and  $N$  corresponding to *Region* and *Nation*, depends on the perceived distance to the average group member:

$$h^{i,j} = 1 - \left( \sum_{k \in K} \omega_k (p_k^i - p_k^j)^2 \right)^{1/2},$$

where  $p_k^i$  represents the preferences (or traditions, values and norms) of individual  $i$  regarding an attribute indexed  $k$ ,  $p_k^j$  represents the preferences of the average member of the region or the nation, and  $K$  is the set of all attributes. An important part of this heterogeneity function are the  $\omega_k$ , which can be understood as attention weights. Higher weights indicate that the tradition, value or norm  $k$  has a larger influence on the strength of common identity.

Assume for simplicity that the attributes in  $K$  can be categorized in a number of subsets:  $K_R$ ,  $K_N$ , and  $K_o$ .  $K_R$  are those attributes that the individual has in common with the other people in his region, for instance speaking the local dialect or in Alsace cooking the local specialty “tarte flambée”. The vector  $\omega_R$  comprises of the weights for all attributes belonging to  $K_R$ . For these attributes, we assume  $p^i - p^R = 0$ , meaning that individuals within a region share the attributes.<sup>1</sup> We use the scalar  $\omega_R = \sum_{k \in K_R} \omega_k$  as the sum of all weights put on common regional culture.

$K_N$  are the attributes that the individual has in common with the rest of the nation. In France, consider common history or traditions that are widely shared, for instance celebrating the 14th of July, the French language or French cuisine. As with regional attributes, the scalar  $\omega_N = \sum_{k \in K_N} \omega_k$  is the sum of all weights put on national culture. The remaining attributes are represented by  $K_o$  and are neither clearly aligned with the region nor the nation, for example preferences about social or economic questions that show a lot of variation both within regions and nations. Other identities relating to, for instance, their municipality can also be thought of as based on attributes contained in  $K_o$ , but we focus on regional and national identity as the main distinction between treated and control area. All weights sum up so that  $\omega_o + \sum_{j \in \mathbf{J}} \omega_j = 1$ , where  $\omega_o$  is the sum of the weights put on the remaining attributes.<sup>2</sup>

When deciding how to invest in the education of their children, parents maximize the expected utility their children derive from a joint regional and national identity. We choose a specific form for the sake of easier exposition and drop the  $i$  subscript for individuals, as we focus on differences between people in the treated and untreated area, equivalent to using one representative citizen for each area. Hence, we can write the utility of a representative parent based on the weights of their child as

$$U = \omega_R^\alpha + \omega_N^\alpha - C,$$

<sup>1</sup> This is a simplifying assumption that makes the following comparisons much clearer. One could instead define the set of common regional or national attributes as those with a distance lower than some positive threshold value.

<sup>2</sup> We assume the  $p$ 's to be fixed, and only  $\omega$  to vary. In other words, we assume that perceived distance to other group members rests on underlying differences which an individual herself cannot influence. Of course, there are exceptions in reality but it is also true that many attributes that are crucial for common identities rest on such factors like place of birth, joint mother tongue or skin color. What varies is whether these differences are relevant when individuals assess their degree of common identity with a particular group. Take for instance the controversial case of Crimea in Ukraine: Before the tensions between Russia and the Ukraine there was no strong separatist movement in the region. Russia's claim to the region is based on the existence of a Russian speaking minority and a common history, and an important policy aim was to increase the salience of these attributes among people in the region.

with  $0 < \alpha < \frac{1}{2}$ . This means parents assign positive utility to their children sharing their regional identity ( $\omega_R$ ), but they also take into account the potential benefits the children will have from alignment with the rest of the nation ( $\omega_N$ ), as argued above. We assume  $\alpha$  to be the same for both identities but this could easily be adapted. Accordingly, both identities are to some degree substitutes, but the optimal choice will usually be to possess some regional and some national identity as  $\alpha < \frac{1}{2}$ . As we describe below in detail, it is costly for parents to actively be involved in influencing their children's identities. This cost is given by  $C$ .

The transmission of weights ( $\omega_R$  and  $\omega_N$ ) is influenced by parental investment and public schooling. Hence, the  $\omega_j$  of a child is a function of the traditions the parents chose to transmit and the traditions transmitted via public schooling. Just like parents, public schooling can spend time on teaching both regional and national culture, as well as on other subjects unrelated to identity. The weights of the child when growing up are then formed as  $\omega_j = \frac{t_j^P + t_j^S}{2}$  for  $j = \{R, N\}$ , with  $t_j^P$  and  $t_j^S$  denoting the time invested by parents and public schooling. Let  $t_R^S + t_N^S \leq 1$ , but in most situations it is more realistic to think of it as smaller than one as schooling also spends time on teachings subjects like math or sciences. For parents, we assume  $t_R^P + t_N^P = 1$  for simplicity if the benefits from teaching regional or national culture exceeds the costs, as discussed below. The total amount of teaching decides the magnitude of the sum of the weights  $\omega_R$  and  $\omega_N$ , which translates into the weights children will put on these sets of attributes and the strength of their identities.<sup>3</sup>

When parents choose  $t_R^P$  and  $t_N^P$ , they weight the benefits of transmitting regional or national culture against a (fixed) costs  $C_j^P \tau_j \geq 0$ . Take for instance the ability to teach regional music or dances to children. Parents need to learn the text or moves and how to convey this information or skill, which is an important fixed cost. One central, but according to us, plausible assumption is that children who repeatedly experienced a tradition within their own family inherit the ability to teach it to their own children. Accordingly,  $\tau_j = 0$  if parents were themselves exposed to  $t_j^P > 0$ .<sup>4</sup>

The (fixed) cost of teaching for parents is then given by the following cost function:

$$C = C(t_R^P, 1 - t_R^P) = \begin{cases} C_R^P \tau_R & \text{if } t_R^P = 1 \\ C_N^P \tau_N & \text{if } t_N^P = 1 \\ C_R^P \tau_R + C_N^P \tau_N & \text{if } 0 < t_R^P < 1 \\ 0 & \text{if } t_R^P = t_N^P = 0 \end{cases}$$

If time is the limiting factor, teaching one culture also creates opportunity costs reflecting less time spent on transmitting other traditions. With the public schooling parameter exogenously given, plugging in the expressions for the weights into the utility function maximized by the parents gives

$$\begin{aligned} U(t_R^P, 1 - t_R^P) &= \left( \frac{t_R^P + t_R^S}{2} \right)^\alpha + \left( \frac{(1 - t_R^P) + t_N^S}{2} \right)^\alpha - C(t_R^P, 1 - t_R^P) \\ &= B(t_R^P, 1 - t_R^P) - C(t_R^P, 1 - t_R^P), \end{aligned}$$

where  $B(t_R^P, 1 - t_R^P)$  is the benefit from teaching. The optimal choice of parents is a function of the degree to which regional and national culture is taught by the public schooling system, the

<sup>3</sup> This means that all attributes belonging to  $\omega_j$  (for  $j \in \{R, N\}$ ), receives equal weights of  $\omega_j/|K_j|$ . The weight put on the remaining attributes is given by  $\omega_o = 1 - \omega_R - \omega_N$ .

<sup>4</sup> The complete notation including the subscript  $i$  for individuals is  $\tau_j = \mathbf{1}[i \in T], \forall i \in I$  and  $T \subset I$ .  $I$  is the set of all individuals, and  $T$  is the subset of individuals that did not inherit the ability to teach  $j$  culture. We assume that engaging in a joint tradition as a family has a different effect than being told about a tradition in school. Observing parents and copying behavior arguably has a large influence on education style, notwithstanding exceptions where children deliberately deviate from their parents behavior.

utility they derive from both identities and the costs associated with transmission. This leads to an optimal parental investment of  $t_R^{P*} = \left(\frac{1+t_N^S-t_R^S}{2}\right)$ , conditional on being incentive-compatible, i.e. if the utility from teaching the optimal level exceeds the utility from not teaching at all. Let  $\tilde{B}(t_R^P, 1-t_R^P) = B(t_R^P, 1-t_R^P) - B(0, 0)$  denote this excess utility. The first number in the parentheses here and in the following refers to regional traditions, and the second number to national traditions. Consider four different cases:

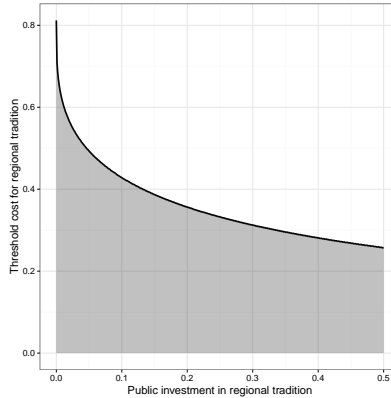
**Case 1** If  $\tilde{B}(t_R^{P*}, 1-t_R^{P*}) \geq C(t_R^{P*}, 1-t_R^{P*})$  for  $0 < t_R^{P*} < 1$ , then  $t_R^P = t_R^{P*} = \left(\frac{1+t_N^S-t_R^S}{2}\right)$  and  $t_N^P = t_N^{P*} = 1 - \left(\frac{1+t_N^S-t_R^S}{2}\right)$ . This means the parents will invest time in learning how to teach and transmit both **regional** and **national** traditions.

**Case 2** If  $\tilde{B}(t_R^{P*}, 1-t_R^{P*}) < C(t_R^{P*}, 1-t_R^{P*})$  and  $U(1, 0) > U(0, 1)$ , and  $\tilde{B}(1, 0) \geq C(1, 0)$ , then  $t_R^P = 1$  and  $t_N^P = 0$ . This means the parents will only invest time in learning how to teach and then transmit **regional** traditions.

**Case 3** If  $\tilde{B}(t_R^{P*}, 1-t_R^{P*}) < C(t_R^{P*}, 1-t_R^{P*})$  and  $U(1, 0) < U(0, 1)$ , and  $\tilde{B}(0, 1) \geq C(0, 1)$ , then  $t_R^P = 0$  and  $t_N^P = 1$ . This means the parents will only invest time in learning how to teach and then transmit **national** traditions.

**Case 4** If  $U(0, 0) = \max U(t_R^P, 1-t_R^P)$ , then  $t_R^P = t_N^P = 0$ . This means the parents will not invest time in learning how to teach and then transmit any traditions.

Figure A1: Threshold costs for teaching regional tradition



*Notes: The solid black line indicates the threshold costs  $\bar{C}_R^P$  for investments in learning how to teach regional culture and traditions. The gray area represents those parameter constellations where the costs are lower than the threshold costs, so that parents will invest in learning regional traditions. The less regional traditions are taught in public schools, the higher the costs parents are willing to pay to maintain regional culture and traditions.*

Figure A1 shows the distribution of costs for which it is optimal for parents to invest time in teaching regional traditions and culture. A decrease in  $t_R^S$  makes teaching regional traditions the best choice for parents along a larger range of parameter values. We can now use this framework to analyze the natural experiment, which can best be described in the three stages introduced above.

### Stage 1

In the first stage, public schooling policy is identical in both areas. Parents decide to teach either regional or national traditions, both traditions, or none of them. The optimal choice of teaching

depends on *i*) the public investment in teaching regional and national traditions, and *ii*) the cost of learning to teach regional and national traditions. For public investments  $t_{R,stage1}^S, t_{N,stage1}^S$ , there exist costs  $C_R^P > \bar{C}_{R,stage1}^P$  and  $C_N^P > \bar{C}_{N,stage1}^P$  such that parents decide not to invest in teaching any traditions, where  $\bar{C}_R^P$  and  $\bar{C}_N^P$  are the maximum allowed (threshold) costs for parents to invest time in regional and national traditions, respectively. Parents invest time if the costs of doing so are lower than the threshold cost  $\bar{C}_{R,stage1}^P$  and  $\bar{C}_{N,stage1}^P$  for the respective traditions. We assume that in the first stage, the costs are above the threshold in the treated and control area so that parents decide not to learn and teach privately.

## Stage 2

After occupation and reflecting the intrusive policies, public schooling in the treated area does not teach regional traditions any more, so that  $t_{R,stage2}^S = 0$  in the treated region. This increases the threshold cost and it is now optimal for parents to invest in teaching regional traditions for a larger range of costs  $C_R^P$ . As national traditions are still taught to a high degree by the state, parents decide to spend all their time teaching regional traditions and  $t_{R,stage2}^P = 1$ . In the control area there was no comparable shock, and public and private investments remain unchanged.

## Stage 3

In the third stage, the temporary shock is over and  $t_{R,stage3}^S$  reverts to the same level in both the treated and the untreated area. If public investment in regional traditions becomes high enough, for instance comparable to stage 1, parents in the untreated area are not willing to bear the cost of learning the regional traditions as  $C_R^P > \bar{C}_{R,stage3}^P$ . However, if regional traditions were taught and transmitted in the treated area during stage 2, parents in the area do not have to bear the fixed costs ( $\tau_j = 0$ ) and they choose  $t_R^P = t_R^{P*} > 0$ . Accordingly, a higher level of teaching regional culture can persist after the shock is over. This difference persists for the first generation; its long term persistence depends on whether  $t_j^P > 0$ , i.e. parents put enough value and time on regional culture so that their children acquire and imitate this behavior.

## B Referendum 1969

In 1968 Charles de Gaulle observed widespread dissatisfaction with the political system and a growing demand for institutional change. In an attempt to satisfy this demand, he announced a constitutional referendum to be held in 1969. The main policy change proposed in the referendum was increasing the political power of regional governments. De Gaulle was convinced that increasing regions' autonomy to settle local affairs locally would restore political balance.<sup>5</sup> Moreover, he believed that the provinces were still close to the heart of the french people.<sup>6</sup> Figure A2 shows a sample of the newspaper we use the primary data source for the referendum outcome. Figure A2b shows samples of voting results disaggregated on the municipality level.

Figure A2: Sample from *L'Est Republicain* showing voting results



(a) *L'Est Republicain* title page

(b) Voting results on municipality level

<sup>5</sup> "Rien n'est plus important pour l'équilibre moral et social de la France que l'organisation, une organisation nouvelle, des contacts et de la coopération, entre ceux qui dirigent et ceux qui sont dirigés." (De Gaulle, 1969)

<sup>6</sup> "Et cependant, bien que les régions fussent officiellement ignorées depuis, les régions, je veux dire, les provinces, fussent officiellement ignorées depuis 179 ans. Elles n'ont jamais cessé d'exister dans l'esprit et dans le coeur des français" (De Gaulle, 1969)



## C Alternative explanations

### C.1 Support driven by urban agglomerations

Another potential concern is whether the effect is driven by outliers. More specifically, it might be driven by urban agglomerations for two potential reasons. Historically, cities enjoyed greater autonomy and might have developed a stronger local identity relative to national identity. Moreover, cities today attract people from a diverse set of places, who could on average be more likely to support the EU. A visual inspection of the maps in Figures 4a and 4c in the main text suggests that the area surrounding Metz does in these cases feature high shares of yes votes. We test whether this is a problem by excluding municipalities belonging to the metropolitan area as defined by INSEE (Table A19 also uses 10 or 15 kilometers from Metz as an alternative cut-off). Depending on bandwidth length, this means that between 30 and 38 municipalities are excluded. Table A26 presents the results for the analysis of newspaper subscriptions within Lorraine excluding Metz (Panel A). All point estimates are very similar and still statistically significant.

### C.2 Religiosity and EU support, relevant for 1992 and 2005 referenda

One distinct feature in which the local laws strongly differ from the rest of France is with regard to religion. Historically, the church played a larger role in the average citizens life in the treated area until after WWI, and still does to some degree until today. In contrast to the rest of France, pupils in the area are still subjected to compulsory religious classes at school (usually two hours per week). This is not uncommon in other European countries, for instance, many of the southern German states feature a similar policy. Usually these classes are not dogmatic, but transmit information about religions in general, of course still with an emphasis on Christianity. If religion or religious denomination is related to a more favorable attitude towards the EU, part of the effect we measure and attribute to differences in exposure to intrusive policies might be driven by differences in religious identity.

However, the available literature indicates no direct relationship between religious attachments and European integration and “even indirect effects of religion on Euroscepticism are small or appear to cancel each other out” (Boomgaarden and Freire, 2009, p.1). To the opposite, albeit minimally, it is argued that “actors such as religious parties and the churches have strayed from the integrationist path and contributed to Euroscepticism” (Minkenberg, 2009, p.1190).

To make sure this is really no concern, we examine the purported relationship in a more systematic way as well. In the specific French context, there are no municipal level measures on religious affiliation and the share of people who consider themselves secular, due to the specific secular constitution and approach in France. Nonetheless, we can use outcomes aggregated at the département level for all of France to assess the relationship between religion and voting in the EU referendum. Table A20 shows results for two variables that measure the intensity of religiousness and religious denomination. *Attendance* measures how often subjects attend religious services, both as a continuous variable and coded as a set of dummies with *never attending* as the reference category. Denomination relates to the share of people who perceive themselves as *Roman Catholic*, *Protestant*, *Christian Orthodox*, *Jewish*, *Moslem* or *other faiths*, with *no religious affiliation* as the reference category.

The results show no difference for *Attendance* in both 1992 and 2005. With *Attendance* coded as individual dummies, there is also no stable relationship. Only very enthusiastic churchgoers have a marginally significant positive correlation compared to those who never attend in 2005, but not in 1992. The pattern is similar for denomination. The only positive correlation which is significant at the 10 percent level is with *Protestant* in 1992, but it also disappears in 2005. Overall, this supports

the existing literature that religion does not play a major role for attitudes towards the EU. Thus, the concern that religious differences would contaminate our main results appears unfounded.

### C.3 Differences in benefits from trade

One of the main benefits of more integration that is usually mentioned is increased gains from trade stemming from lower trade costs (Alesina and Spolaore, 1997). Accordingly, we need to assume that these benefits are comparable close to the border. Clearly, distance to the respective neighboring states correlates with trade costs; municipalities that are closer to the country borders could benefit more from increased trade and thus exhibit higher agreement to more EU integration. At the same time, relying less on trade with the rest of France and more on exports could also foster a stronger regional relative to national identity. There are two ways to evaluate whether this is problematic in our cases.

Firstly, our smallest bandwidth is 10 kilometers only, so that it seems implausible that the relatively small additional distance between treated and control municipalities affects trade costs sufficiently to explain the results. Moreover, our estimates are robust to controlling for distance to the German as well as to other borders. Secondly, the point estimates of the treatment effect barely change when we increase the bandwidths and include more municipalities (Figure A16). Thirdly, if distance to the border has a significant effect, we would expect to see a significant, or at least positive difference between former Lorraine and the rest of France as well. As the differences in Table A24 are neither always positive, nor significant, differences in trade benefits do not seem to be problematic.

### C.4 The relative importance of homogenization policies

By design of the experiment we exploit, it is impossible to exactly distinguish the effect of homogenization policies from the effect of occupation and repression in general. It seems plausible that repression itself provokes a backlash, but the historical literature specifically emphasizes the crucial role of homogenization policies steered at suppressing regional identity (e.g. De La Valette, 1925; Goodfellow, 1993; Harp, 1998; Harvey, 1999; Heffernan, 2001; Zanoun, 2009). Even more than German policies, French policies after World War I clearly aimed at eliminating all signs of regional particularities that were deemed dangerous. Many of these policies plausibly affected the Alemanic-dialect speaking areas more severely, for instance repeated prohibitions of specific newspapers and parties associated with the usage of the German language.

As intrusive French homogenization policies comprised the second and more recent part of the treatment period, there could be a stronger treatment effect on the German-dialect speaking part of Lorraine. Panel B and C in Table A21 indeed shows a significantly higher share of yes votes on the German speaking side in both 1992 and 2005. Of course, this heterogeneous treatment effect could partly be driven by other unobserved differences due to language. Accordingly, while keeping the caveats in mind, this is suggestive evidence supporting the important role of homogenization policies in creating the backlash.

### C.5 The influence of Germanization

Although feeling more German would not directly explain a stronger regional identity, being exposed to German ideas, newspapers and institutions for nearly fifty years could affect preferences. In our model, however, there is no reason to expect a persistently stronger German identity after the occupation ended. Although identities based on different levels (regional, national) need not to be substitutes, national identities probably are to some degree. Accordingly we would expect that a

stronger German identity is related to a weaker French identity. Although we find no such difference in the survey results, we also code a variable based on tweets issued using Twitter about the French and German national football team during the World Cup in 2014 as a robustness test. When using this as an alternative measure of German and French national identity at the local level within Lorraine, we find no significant difference at the 10 kilometers and at the optimal IK bandwidth (see Online Appendix Section D). The analysis rests on relatively few tweets, but the results are in line with the survey evidence and suggest no difference in German or French national identity.<sup>7</sup>

---

<sup>7</sup> The historical and sociological literature also argues that although citizens accepted their legal belonging to Germany, they did so “without feeling German themselves” (Höpel, 2012, p.37). De La Valette (1925) refers to a disillusioned German journalist saying “Alsace does not want us; the Alsations are lost to us”. Carrol (2010, p.66) cites a government official stating that “Prussian methods had failed to instil alien national sentiments into the minds of a people who were proud of their history”. It also seems to be partly misleading to frame the regionalist parties in the 1920s and 30s as pro-German. The “Landespartei” is described as “referring in its manifesto to the right of peoples to self-determination and looked forward to the day when a ‘free Alsace- Lorraine’ would be the mediator between France and Germany in a United States of Europe” (Anderson, 1972). Similarly, the UPR called for “administrative decentralization, a regional elected council and the recognition of bilingualism” rather than for a return to Germany.



## D Twitter

### Georeferencing

There are two ways in which Twitter users indicate their geographic location:

1. **User-provided georeferencing:** User can tag a location in their tweet directly. This type of tweet is unreliable for research, because the location tagged doesn't necessarily coincide with the location of the person tweeting.
2. **GPS-provided georeferencing:** The GPS function in mobile phones allows Twitter messages sent via the phone to contain the coordinates of the user's location. Due to the optionality of the GPS function, only 2- 3 percent of all Twitter users can be georeferenced this way. Due to the abundance of tweets, this method still generates a large number of possible observations.

### Availability of data

It is possible to collect a random selection of tweets at any given point in time via Twitter's API (Application Programming Interface).

### Twitter data Lorraine

The relevant tweets were identified and analyzed in a three-step process.

1. Over the period of the Football World Cup 2014 a random sample of tweets was obtained via Twitter's API. This method resulted in 18'278 observations.
2. Because Twitter only allows for data selection in geographic areas of rectangular shape, ArcGIS was used to identify the tweets specifically located in Lorraine.
3. The content of the selected tweets were then analyzed based on a selection of keywords about the German and French national football teams. The lists of keywords are displayed in Table [A28](#).

## E Links to examples for other regions mentioned in paper

- Scania, Sweden

<https://sverigesradio.se/sida/artikel.aspx?programid=83&artikel=1915851>

- Silesia, Bohemia, Kaliningrad and Danzig, Poland and the Czech Republic

<http://homepage.univie.ac.at/philipp.ther/breslau/html/Entdeutschung%20und%20Polonisierung.%20Die%20Umwandlung%20Breslaus%20in%20eine%20polnische%20Stadt.html>

<https://web.archive.org/web/20071208130441>

<http://www.haus.pl/de/opis/arb4.html>

<https://www.britannica.com/place/Silesia>

<https://www.economist.com/eastern-approaches/2013/05/07/the-expulsion-of-sudeten-germans-is-still-raw>

<https://www.britannica.com/place/Bohemia>

<https://www.britannica.com/place/Gdansk>

<http://dfk-danzig.eu/de/deutsche-in-danzig/deutsche-in-danzig>

- Chechnya, Russia

<https://www.britannica.com/place/Chechnya>

[http://linguistics.berkeley.edu/ingush/ingush\\_people.html](http://linguistics.berkeley.edu/ingush/ingush_people.html)

<https://www.economist.com/news/2003/03/25/putins-proposition>

- Kurds in Turkey, Iran, Iraq and Syria

<https://www.nytimes.com/2008/02/17/magazine/17turkey-t.html?ex=1361854800&en=df64cf85326e2103&ei=5124&partner=permalink&exprod=permalink>

<https://www.bbc.com/news/world-middle-east-29702440>

- South Tyrol, Italy

<https://www.tandfonline.com/doi/full/10.1080/1369183X.2017.1352465>

- Xinjiang, China

<https://www.economist.com/china/2015/06/27/tongue-tied>

<https://www.economist.com/briefing/2018/05/31/china-has-turned-xinjiang-into-a-police-state-like-no-other>

- Tibet, currently occupied by China

<https://www.economist.com/china/2016/09/17/the-plateau-unpacified>

<https://www.nytimes.com/2010/07/25/world/asia/25tibet.html>

- Crimea, formerly in Ukrain, annexed by Russia 2014

<https://www.economist.com/news/2015/06/11/bad-memory>

<https://www.economist.com/eastern-approaches/2014/03/17/ukraines-amputation>

## F Tables

Table A1: Detailed Overview of Repressive Policies in Alsace and Lorraine

Time Period	Ruled By	Policy	Policy Category	Source
1871-1902	Germany	Reactivation of the 1849 “dictatorship paragraph”: permitted house searches, the expulsion of agitators and prohibiting political organizations.	Social, political, military freedom, equality	Carrol (2010); Grasser (1998)
Beginning 1871/72	Germany	Bismarcks <i>Kulturkampf</i> : government seriously restricted Catholic education as well as the Catholic press. Moreover, some religious orders were expelled from the Reichsland.	Regional institutions and administrative personnel	Silverman (1966)
May 1872	Germany	Strasbourg University is reopened as “Kaiser-Willhelm-Universitaet”.	Language	Höpel (2012)
Oct. 1872	Germany	Introduction of obligatory military service.	Social, political, military freedom, equality	Grasser (1998)
1873	Germany	French is prohibited to be taught in schools.	Language	Grasser (1998)
1878	Germany	Legislation to restrict the political participation of the people.	Social, political, military freedom, equality	Carrol (2010)
1882	Germany	The use of French is prohibited in the Delegation.	Language	Grasser (1998)
1887	Germany	Choral and gymnastic societies are banned as they are seen as opportunities for the coming-together of pro-French minded people.	Social, political, military freedom, equality	Carrol (2010)
1890 onward	Germany	Unwelcome legislation (e.g. German trade regulations) is introduced in Alsace-Lorraine.	Regional institutions and Administrative Personnel	Höpel (2012)
1890 onward	Germany	German becomes the only official language and district and county councils become obliged to embrace German as their only language.	Language	Grasser (1998)
Until 1898	Germany	Restrictions are imposed on the press.	Media	Silverman (1966)
1914	Germany	Citizens sympathizing with the French are taken in “protective detention” without trial.	Separation and segregation; Social, political, military freedom, equality	Harvey (1999)

1917/18	France	Approximately 100 000 Germans are deported.	Separation and segregation	Carrol and Zanoun (2011), Callender (1927)
1918	France	Establishment of French Currency.	Regional institutions and administrative personnel	Callender (1927)
Dec. 1918	France	An identity-card system is implemented: Locals are classified and receive a specific civil status according to the origin of their parents. Lower classification is often associated with discrimination.	Separation and segregation	Harvey (1999)
Dec. 1918 to Oct. 1919	France	“Commissions de Triage” are established: Designed to assert the Frenchness of the population in re-annexed areas, individuals suspected of faulty loyalties are investigated and either exonerated, placed under surveillance, taken into custody or expelled from France. In this context, some pro-German Alsatiens are forcefully emigrated.	Separation and segregation; Social, political, military freedom, equality	Carrol and Zanoun (2011); Harvey (1999)
1920	France	French becomes the only language to be taught in schools. The so-called ”direct method”, where students are immersed in the French language with no reference to German, leads to considerable difficulties for a majority of French-speaking Alsatiends.	Language	Grasser (1998); Goodfellow (1993)
1920s	France	French becomes the official legal language. Due to this, many bureaucrats, who had previously built their career under the German system, are in danger of losing their jobs or being denied promotions as the French government now regards them as incompetent or politically problematic.	Language	Goodfellow (1993)

June 1924	France	The Ministerial Declaration by Premier Edouard Herriot introduces a centralised French administration as well as all French laws and institutions into the recovered territories. The Declaration also introduces the separation of church, secular education and a number of anti-clerical laws.	Regional institutions and administrative personnel	Carrol and Zanoun (2011); Goodfellow (1993)
1925	France	The post of Commissioner General is abolished and the regional government returned to the Government of Paris	Regional institutions and administrative personnel	Callender (1927)
1927/28	France	Three autonomist journals become banned as they are seen to have had a central role in a campaign against the French: The "Volksstimme" ("voice of the people"), the "Wahrheit" ("truth") and the "Zukunft" ("future").	Media	Goodfellow (1993)
1927/28	France	Colmar trials: 15 prominent autonomists are arrested and tried with the reason given that they had participated in a plot to separate Alsace from France. 4 of the 15 are sentenced to 1 year in prison, while 5 are sentenced to be exiled.	Social, political, military freedom, equality	Goodfellow (1993)
1939	France	15 autonomists are arrested for relations with the enemy. One autonomist leader is later executed by a fire squad in 1940 in Champigneulles.	Social, political, military freedom, equality	Goodfellow (1993)
1940	Germany	The French language is prohibited from use and street signs must be renamed in German. French names must be replaced by German equivalents.	Language	www.nithart.com; Encyclopédie
1940	Germany	Germans prohibit the Alsatian dialect as it is regarded as a means of protest against the Nazi-government.	Language	Encyclopédie
1940	Germany	Germans prohibit typically Alsatian gatherings and celebrations as they are seen as expressions of specifically regional culture and therefore against the Germanisation efforts of the Nazi regime.	Social, political, military freedom, equality	Encyclopédie

1940	Germany	German is made the official language of the administration.	Language	Grasser (1998)
1945-1952	France	Teaching of German is de jure prohibited in schools, de facto this is applied in about half of the schools.	Language	<a href="http://www.alsace-lorraine.org">www.alsace-lorraine.org</a> ; Anderson (1972)
1953	France	Bordeaux trials: 13 Alsatian <i>malgré-nous</i> are sentenced to death due to their involvement in the massacre of Oradour-sur-Glane.	Social, political, military freedom, equality	Boswell (2008) Collins (2007)

Notes: Encyclopédie refers to [www.encyclopedie.bseditions.fr](http://www.encyclopedie.bseditions.fr).

Table A2: Variable description and sources

Variable	Definition	Source
<i>Dependent Variables</i>		
Share Yes 1969	Share of Yes votes in the 1969 constitutional referendum	<i>L'Est Republicain</i>
Share Yes 1992	Share of Yes votes in the 1992 referendum (Maastricht Treaty)	Centre de données socio-politiques (CDSP)
Share Yes 2005	Share of Yes votes in the 2005 referendum (European Constitution Treaty)	Centre de données socio-politiques (CDSP)
Share of Le Pen votes, 1992	Share of votes for Jean-Marine Le Pen in the 2007 presidential election (first round)	Centre de données socio-politiques (CDSP)
Turnout, 1969	Voter turnout in the 1969 constitutional referendum	<i>L'Est Republicain</i>
Turnout, 1992	Voter turnout in the 1992 referendum (Maastricht Treaty)	Centre de données socio-politiques (CDSP)
Turnout, 2005	Voter turnout in the 2005 referendum (European Constitution Treaty)	Centre de données socio-politiques (CDSP)
Turnout, 2007	Voter turnout in the 2007 presidential election (first round)	Centre de données socio-politiques (CDSP)
Subscription regional newspaper	Subscriptions to "Le Republicain Lorraine"/No.households in 2014	<i>Le Republicain Lorraine</i>
Share Tweets Germany	Number of tweets about Germany during the 2014 World Cup	Twitter
Share Tweets France	Number of tweets about France during the 2014 World Cup	Twitter
<i>Pre-treatment variables</i>		
Ruggedness	Index of variance of elevation in each municipality	Global elevation data set
Elevation	Raw elevation data	NASA SRTM data set
Potato	Soil suitability for production of potatoes (medium input intensity and irrigation)	IIASA/FAO, 2012
Wheat	Soil suitability for production of wheat (medium input intensity and irrigation)	IIASA/FAO, 2012
Barley	Soil suitability for production of barley (medium input intensity and irrigation)	IIASA/FAO, 2012
<i>Covariates</i>		
Median income	Median income in 2008	INSEE
Mean age	Mean age in 2006	INSEE
Education	Share of people with a high school degree	INSEE
Occupation	Share of blue-collar workers	INSEE
Workers, 2006	Share of workers in 2006	INSEE
Farmers, 2006	Share of farmers in 2006	INSEE
Artisans, 2006	Share of artisans in 2006	INSEE
Executives, 2006	Share of executives in 2006	INSEE
Intermediate prof., 2006	Intermediate professionals in 2006	INSEE
Companies, 2011	Number of companies per capita in 2011	INSEE
Commercial est., 2011	Number of commercial establishments per capita in 2011	INSEE
Industrial est., 2011	Number of industrial establishments per capita in 2011	INSEE
Building est., 2011	Number of building establishments per capita in 2011	INSEE
Public est., 2011	Number of public establishments per capita in 2011	INSEE
Theatre rooms, 2013	Number of theatre rooms per capita in 2013	INSEE
Athletic centers, 2013	Number of athletic centers per capita in 2013	INSEE
Multisport fac., 2013	Number of multisport facilities per capita in 2013	INSEE
Swimming fac., 2013	Number of swimming facilities per capita in 2013	INSEE
Psychiatric est., 2013	Number of psychiatric establishments per capita in 2013	INSEE
Service houses, 2013	Number of service houses per capita in 2013	INSEE
Health care, 2013 (short)	–	INSEE
Health care, 2013 (medium)	–	INSEE
Health care, 2013 (long)	–	INSEE
Post offices, 2013	Number of post offices per capita in 2013	INSEE
Elementary schools, 2013	Number of elementary schools per capita in 2013	INSEE
High schools, 2013	Number of high schools per capita in 2013	INSEE
Vocational training, 2013	Number of secondary schools with vocational training per capita in 2013	INSEE
Tech. vocational training, 2013	Number of secondary schools with technical vocational training per capita in 2013	INSEE

Notes: Variable description and source for all variables used in the paper and this Online Appendix.

Table A3: Survey questions (i.)

Variable	Question	Categories/Scale	Source
Regional identity	"Could you tell me whether you feel very attached, rather attached, not very attached or not attached at all to [name of region]?"	4 = very attached; 3 = rather attached; 2 = not very attached; 1 = not attached at all	OIP 99/2001 Q2a3
National identity	"Could you tell me whether you feel very attached, rather attached, not very attached or not attached at all to France?"	4 = very attached; 3 = rather attached; 2 = not very attached; 1 = not attached at all	OIP 99/2001 Q2a2
European identity	"Could you tell me whether you feel very attached, rather attached, not very attached or not attached at all to Europe?"	4 = very attached; 3 = rather attached; 2 = not very attached; 1 = not attached at all	OIP 99/2001 Q2a1
Regional relative to National identity (standardized)		Relation of two identities, standardized with standard deviation 1 and mean 0	OIP 99/2001
European relative to national identity (standardized)		Relation of two identities, standardized with standard deviation 1 and mean 0	OIP 99/2001
Democracy works well within France	"Personally, do you reckon the democracy in France to function very well, fairly well, not very well or not well at all?"	4 = very well; 3 = fairly well; 2 = not very well; 1 = not well at all	OIP 99/2001 Q4
I feel well informed about regional policies	"You personally, do you think that you are well or badly informed about the actions of the regional council of [name of region]?"	4 = very well; 3 = rather well; 2 = rather badly; 1 = very badly	OIP 99/2001 Q14
Democracy works well within the region	"And in [name of region], do you reckon the democracy to function very well, fairly well, not very well or not well at all?"	4 = very well; 3 = fairly well; 2 = not very well; 1 = not well at all	OIP 99/2001 Q5
I am concerned regional administration would increase interregional inequality	"If the region takes action in all those domains instead of the state, are you concerned about the development of interregional inequality?"	4 = Yes, very much so; 3 = Yes, somewhat; 2 = No, not very much; 1 = No, not at all	OIP 2003 Q11a2

*Notes:* Description of survey questions from the Observatoire Interrégional du Politique (OIP) 1999 and 2001. The values of the categories are reversed compared to the original question categories. Questions were originally in French and have been translated.



Table A4: Survey questions (ii.)

Variable	Question	Categories/Scale	Source
Power_Transfer_Region	<b>"Are you in favor of the transfer of all the power and means of the state to the regions?" (Average across 10 policy dimensions)</b>	Value between 1 and 4. 4 = "Strongly in favor" and 1 = "Strongly against"	
1	"Are you in favor of the transfer of all the power and means of the state to the regions regarding the choice in setting up high schools?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2001 Q36a1
2	"Are you in favor of the transfer of all the power and means of the state to the regions regarding the management of high school teachers?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2001 Q36a2
3	"Are you in favor of the transfer of all the power and means of the state to the regions regarding the management of administrative personnel in high schools?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2001 Q36a3
4	"Are you in favor of the transfer of all the power and means of the state to the regions regarding the definition of school programmes and certificates?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2001 Q36a4
5	"Are you in favor of the transfer of all the power and means of the state to the regions regarding the choice in setting up university centers in the region?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2001 Q36a5
6	"Are you in favor of the transfer of all the power and means of the state to the regions regarding the choice of high school creation?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2001 Q36a6
7	"Are you in favor of the transfer of all the power and means of the state to the regions regarding environmental policies like water policy?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2001 Q36a7
8	"Are you in favor of the transfer of all the power and means of the state to the regions regarding cultural policies like heritage conservation?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2001 Q36a8
9	"Are you in favor of the transfer of all the power and means of the state to the regions regarding sport policies?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2001 Q36a9
10	"Are you in favor of the transfer of all the power and means of the state to the regions regarding the support of social housing?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2001 Q36a10

*Notes:* Description of survey questions from the Observatoire Interrégional du Politique (OIP) 2001. The values of the categories are reversed compared to the original question categories. Questions were originally in French and have been translated.

Table A5: Survey questions (iii.)

Variable	Question	Categories/Scale	Source
Autonomy_Region	<b>"Could you tell me whether reforms empowering the regional councils are a very good thing, a rather good thing, a rather bad thing or a very bad thing for the years to come?"</b> (Average across 5 areas)	Value between 1 and 4. 1 = "It's a very bad thing." and 4 = "It's very good thing."	
1	"Here are a certain number of reforms that are under way or under discussion. Could you tell me, for each one of these, whether it is a very good thing, a rather good thing, a rather bad thing or a very bad thing for the years to come? - Authorizing the regional councils to adapt the national laws and regulations in their respective regions, under the control of the Parliament."	4 = A very good thing; 3 = A rather good thing; 2 = A rather bad thing; 1 = A very bad thing	OIP2001 Q35a1
2	"Here are a certain number of reforms that are under way or under discussion. Could you tell me, for each one of these, whether it is a very good thing, a rather good thing, a rather bad thing or a very bad thing for the years to come? - Authorizing the regional councils to negotiate and manage the European funding without state involvement."	4 = A very good thing; 3 = A rather good thing; 2 = A rather bad thing; 1 = A very bad thing	OIP2001 Q35a2
3	" Here are a certain number of reforms that are under way or under discussion. Could you tell me, for each one of these, whether it is a very good thing, a rather good thing, a rather bad thing or a very bad thing for the years to come? - Giving the regional councils more freedom in deciding over their financial resources without depending on the state."	4 = A very good thing; 3 = A rather good thing; 2 = A rather bad thing; 1 = A very bad thing	OIP2001 Q35a3
4	"Here are a certain number of reforms that are under way or under discussion. Could you tell me, for each one of these, whether it is a very good thing, a rather good thing, a rather bad thing or a very bad thing for the years to come? - Developing the study of regional languages at school."	4 = A very good thing; 3 = A rather good thing; 2 = A rather bad thing; 1 = A very bad thing	OIP2001 Q35a4
5	"Here are a certain number of reforms that are under way or under discussion. Could you tell me, for each one of these, whether it is a very good thing, a rather good thing, a rather bad thing or a very bad thing for the years to come? - Assigning new fields of competence to the regional councils."	4 = A very good thing; 3 = A rather good thing; 2 = A rather bad thing; 1 = A very bad thing	OIP2001 Q35a5

*Notes:* Description of survey questions from the Observatoire Interrégional du Politique (OIP) 2001. The values of the categories are reversed compared to the original question categories. Questions were originally in French and have been translated.

Table A6: Survey questions (iv.)

Variable	Question	Categories/Scale	Source
Education_Region	<b>"Are you in favor of the transfer of all the power and means of the state to the regions regarding education policy and standards?" (Average across 5 questions)</b>	Value between 1 and 4. 1 = "Strongly against" and 4 = "Strongly in favor"	
1	"Are you in favor of the transfer of all the power and means of the state to the regions in the following field: - The choice in setting up high schools?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2003 Q12a1
2	"Are you in favor of the transfer of all the power and means of the state to the regions in the following field: - The management of high school teachers?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2003 Q12a2
3	"Are you in favor of the transfer of all the power and means of the state to the regions in the following field: - The management of administrative personnel in high schools?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2003 Q12a3
4	"Are you in favor of the transfer of all the power and means of the state to the regions in the following field: - The definition of school programmes and certificates?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2003 Q12a4
5	"Are you in favor of the transfer of all the power and means of the state to the regions in the following field: - The choice in setting up university centers in the region?"	4 = Strongly in favor; 3 = Somewhat in favor; 2 = Somewhat against; 1 = Strongly against	OIP2003 Q12a5
Proud of French history	"How proud are you of the History of France?"	1 = Very proud to 4 = Not proud at all	ISSP 2003 – National Identity II
Proud of French sport achievements	"How proud are you of France's achievements in sports?"	1 = Very proud to 4 = Not proud at all	ISSP 2003 – National Identity II
Proud of French science/technology	"How proud are you of France's scientific and technological achievements?"	1 = Very proud to 4 = Not proud at all	ISSP 2003 – National Identity II
More power to UN	"Thinking about the United Nations, which comes closest to your view?"	1 = The UN has too much power to 3 = The UN has too little power	ISSP 2004, Citizenship
Intervention of the UN	"Which of these two statements comes closer to your view?"	1 = If a country seriously violates human rights, the UN should intervene, 2 = Even if human rights are seriously violated, the country's sovereignty must be respected, and the UN should not intervene	ISSP 2004, Citizenship

*Notes:* Description of survey questions from International Social Survey Programme (ISSP) 2003, National Identity (II), and ISSP 2004, Citizenship, and the Observatoire Interrégional du Politique (OIP) 2003. The values of the categories are reversed compared to the original question categories. Questions were originally in French and have been translated.

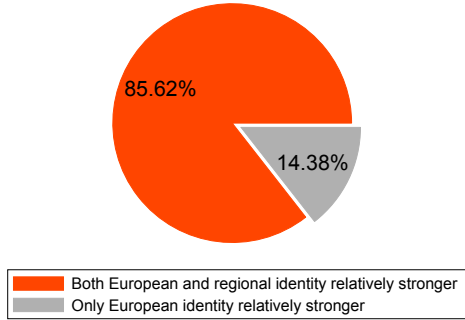
Table A7: Survey results, Alsace and Lorraine

Survey question	Mean, control	$\Delta$	P-value	No. obs.
Feel close to region (Regional identity)	3.362	0.209	<0.001	2617
Feel close to nation (National identity)	3.635	-0.003	0.906	2617
Regional identity/National identity (standardized)	-0.138	0.226	<0.001	2614

*Notes:* Sources are the Observatoire Interrégional du Politique (OIP) 1999, 2001, and 2003, using respondents in all of Alsace and Lorraine, on département level. Identity is measured on a 4-point Likert-scale. The parameter  $\Delta$  comes from the equation:  $y_i = \pi + \Delta Treatment_i + \Gamma'_i \lambda + \eta_i$ , where  $Treatment_i = \mathbf{1}$ [individual in treated region] and  $\Gamma$  comprises of controls for (reported) age, employment status and sex. A positive  $\Delta$  indicates that people in the treated region agree more with the statement.

Table A8: Overlap strength of regional and European identity in treated and control areas (A+L)

Identity differences treated compared to control area  
(conditional on stating stronger EU identity)



*Notes:* Higher (lower) means that an individual in the treated area exhibited a higher (lower) ratio of Regional to National or European to National identity compared to the mean ratios in the untreated area. Higher is mathematically defined as larger or equal. Very few observations are exactly equal to the mean. We are mostly interested in the overlap of the two, but also the overall sum. The overlap is also visualized in the pie chart on the right. The red area indicates the share of persons which answered with both higher or equal European identity and Regional identity. Data is from the OIP 1999, 2001, and 2003, using respondents in all of Alsace and Lorraine.

Table A9: OIP Survey results, 1999 and 2001: European and regional attachments

Dep. Var: Attachment: Europe Variable	Within Lorraine		All of France	
	(1)	(2)	(3)	(4)
Attachement: Region	0.186*** (0.030)	0.185*** (0.031)	0.097*** (0.007)	0.097*** (0.007)
Obs.	1388	1388	25602	25602
Controls	No	Yes	No	Yes

*Notes:* Observatoire Interrégional du Politique (OIP) survey results from 1999 and 2001, asking question on how strong respondents attachment is to Europe, and respondent's Region. Attachment is based on a 1-4 scale, with 1 corresponds to *Disagree strongly*, and 4 corresponds to *Strongly agree*. Controls are age, sex, employment status, and survey year. \*\*\*, \*\* and \* indicate statistical significance at 1%, 5% and 10% levels, based on heteroscedasticity-consistent standard errors.

Table A10: Descriptive statistics for outcome variables and treatment

Variable	Mean	Std. dev.	Min.	Max.
Treatment	0.52	0.50	0.00	1.00
Yes 69	59.69	14.28	5.65	94.74
Yes 92	53.91	11.39	0.00	86.25
Yes 05	45.51	9.96	6.67	81.01
Le Pen 07	15.98	5.36	0.00	55.56
Newspaper subscriptions	14.62	7.63	0.00	32.90
Turnout 69	84.59	7.56	7.41	100.00
Turnout 92	74.40	6.04	52.44	100.00
Turnout 05	73.28	6.40	50.79	100.00
Turnout 07	86.29	4.16	63.38	100.00

*Notes:* Descriptive statistics for the binary treatment variable, *Share Yes 1992* and *Share Yes 2005*, in the respective referenda, and *Share Le Pen 2007* is the share of voters voting in favour of Jean-Marie Le Pen in the 2007 presidential election (first round), whereas *Turnout 1992*, *2005*, and *2007*, refers to turnout in the respective year.

Table A11: Descriptive statistics for RDD control and pre-treatment variables

Variable	Mean	Std. dev.	Min.	Max.
Distance to Metz	83.47	44.39	1.60	203.16
Distance to Strasbourg	107.53	50.32	0.02	223.02
Distance to Nancy	73.97	34.89	0.06	164.98
Distance to Germany	50.87	35.48	0.33	141.55
Elevation	300.51	119.71	110.12	1045.90
Ruggedness	0.73	0.68	0.01	5.18
Potato	7091.57	474.12	3665.80	7848.00
Wheat	6104.37	326.52	3873.60	6687.00
Barley	6099.83	323.85	3873.6	6687
Median income 2008	31.56	6.00	17.69	53.55
Mean age 2006	39.60	3.01	28.26	63.07
Education 1999	0.20	0.07	0.00	0.58
Occupation 2006	0.19	0.07	0.00	0.50

*Notes:* Descriptive statistics for variables used as covariates (for variables used in the main paper) and pre-treatment variables. Distances are in kilometers. Potato and wheat refer to the suitability of the soil to grow the respective crop, based on FAO data. Other variables were chosen with the aim to have the date date closest to our main outcome variables.

Table A12: Pre-treatment variables balance test, within Lorraine

Dep. Variable:	(1) Barley	(2) Wheat	(3) Potato	(4) Elevation	(5) Ruggedness
<b>Bandwidth 10 km</b>					
Treatment	10.799 (224.053) [0.962]	76.157 (234.694) [0.746]	66.839 (143.681) [0.455]	6.491 (7.660) [0.603]	10.914 (8.497) [0.244]
Obs.	403	403	403	408	408
<b>Bandwidth 20 km</b>					
Treatment	-270.251 (191.811) [0.159]	-124.417 (204.683) [0.543]	-159.527 (119.636) [0.183]	-6.532 (7.990) [0.414]	5.023 (6.274) [0.424]
Obs.	756	756	756	765	765

*Notes:* Tests for discontinuities in pre-treatment variables for the whole border. *Ruggedness* is the mean index of the variation in elevation, while *Elevation* is the mean elevation. *Potato*, *Wheat*, and *Barley* refer to the soil suitability for potato and wheat production, respectively. Details and sources are provided in the Online Appendix. Controls included are: distance to Germany (border), distance to Metz, distance to Strasbourg, and distance to Nancy. Standard errors in parentheses and *p*-values in brackets.

Table A13: Pre-treatment variables balance test, Alsace and Lorraine

Dep. Variable:	(1) Barley	(2) Wheat	(3) Potato	(4) Elevation	(5) Ruggedness
<b>Bandwidth 10 km</b>					
Treatment	278.059 (242.550) [0.252]	348.124 (244.659) [0.389]	129.626 (140.356) [0.155]	-25.229 (19.798) [0.203]	-3.949 (11.726) [0.356]
Obs.	614	614	614	619	619
<b>Bandwidth 20 km</b>					
Treatment	-190.426 (175.961) [0.279]	-103.692 (179.980) [0.565]	-202.730** (102.171) [0.047]	-6.090 (14.113) [0.666]	5.911 (8.133) [0.467]
Obs.	1164	1164	1164	1173	1173

*Notes:* Tests for discontinuities in pre-treatment variables for the whole border. *Ruggedness* is the mean index of the variation in elevation, while *Elevation* is the mean elevation. *Potato*, *Wheat*, and *Barley* refer to the soil suitability for potato and wheat production, respectively. Details and sources are provided in the Online Appendix. Controls included are: distance to Germany (border), distance to Metz, distance to Strasbourg, and distance to Nancy. Standard errors in parentheses and *p*-values in brackets.

Table A14: OLS estimates using all municipalities in Alsace and Lorraine

Variable	<b>A: Share Le Pen 2007</b>		<b>B: Turnout 2007</b>	
	(1)	(2)	(3)	(4)
Treatment	-0.691	-0.969	-1.412	0.223
	(0.236)	(0.450)	(0.175)	(0.313)
	[0.003]	[0.031]	[<0.001]	[0.477]
Obs.	3142	3142	3142	3142
Controls	No	No	No	No
Variable	<b>C: Share Yes 1992</b>		<b>D: Turnout 1992</b>	
	(1)	(2)	(3)	(4)
Treatment	11.941	4.865	-0.652	2.081
	(0.473)	(0.789)	(0.262)	(0.470)
	[<0.001]	[<0.001]	[0.013]	[<0.001]
Obs.	3137	3137	3137	3137
Controls	No	No	No	No
Variable	<b>E: Share Yes 2005</b>		<b>F: Turnout 2005</b>	
	(1)	(2)	(3)	(4)
Treatment	6.990	6.185	-3.115	-0.023
	(0.434)	(0.855)	(0.276)	(0.470)
	[<0.001]	[<0.001]	[<0.001]	[0.960]
Obs.	3141	3141	3141	3141
Controls	No	No	No	No

*Notes:* OLS estimates using whole sample of municipalities in all départements in Alsace and Lorraine. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, distance to Nancy. Conley standard errors in parentheses and *p*-values in brackets.

For *Share Le Pen 2007*, *Share Yes 1992* and *Share Yes 2005*, the coefficients indicate both lower levels of national identity and higher regional identity in the treated region. Although the interpretation of the regression coefficient for the treatment variable is the average difference in percentage points between treated and untreated municipalities, it is important to relate them to the average vote share of the whole region. The small differences in turnout in 2005 and 2007 become insignificant when we add controls (Panel B and F). The coefficient for *Turnout 1992* changes signs when controls are added, and becomes insignificant in the RDD at the border.

Table A15: Discontinuities in turnout, within Lorraine

Dep. Variable:	Turnout 2007				Turnout 1969			
	(1)	(2)	(3)	(4) <sup>a</sup>	(5)	(6)	(7)	(8) <sup>a</sup>
Treatment	0.387 (0.862) [0.654]	-0.173 (0.763) [0.821]	-0.552 (0.694) [0.427]	-0.269 (0.535) [0.616]	2.089 (1.086) [0.055]	0.554 (1.006) [0.582]	-0.545 (0.963) [0.571]	1.762 (0.914) [0.054]
Obs.	394	583	744	1325	375	550	693	1037
Dist	10 km	15 km	20 km	42.28 km	10 km	15 km	20 km	33.29 km
Dep. Variable:	Turnout 1992				Turnout 2005			
	(1)	(2)	(3)	(4) <sup>a</sup>	(5)	(6)	(7)	(8) <sup>a</sup>
Treatment	-0.861 (1.229) [0.484]	-1.145 (1.056) [0.278]	-1.646 (0.967) [0.089]	-0.899 (0.908) [0.322]	0.804 (1.222) [0.511]	-0.650 (1.124) [0.563]	-2.413 (1.092) [0.027]	-1.774 (0.898) [0.048]
Obs.	394	583	744	873	394	583	744	1153
Dist	10 km	15 km	20 km	24.21 km	10 km	15 km	20 km	34.58 km

*Notes:* Discontinuity at the treatment border using municipalities in Lorraine. Outcomes are turnout in the 2007 presidential election (first round), turnout in the 1969 referendum, turnout in the 1992 referendum, and turnout in the 2005 referendum. Conley standard errors in parentheses and  $p$ -values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.



Table A16: RD results: Nationalism (Le Pen) and turnout, Alsace and Lorraine

Variable	Panel <b>A</b> : Share Le Pen 2007				Panel <b>B</b> : Turnout 2007.			
	(1)	(2)	(3)	(4) <sup>a</sup>	(5)	(6)	(7)	(8) <sup>a</sup>
Treatment	-0.236 (0.852) [0.782]	-0.232 (0.692) [0.737]	-0.288 (0.644) [0.655]	-1.099 (0.548) [0.045]	0.446 (0.701) [0.525]	0.089 (0.611) [0.884]	0.232 (0.544) [0.670]	0.833 (0.351) [0.018]
Obs.	603	886	1149	1707	603	886	1149	2727
Dist	10 km	15 km	20 km	30.37 km	10 km	15 km	20 km	58.33 km
Variable	Panel <b>C</b> : Turnout 1992				Panel <b>D</b> : Turnout 2005			
	(1)	(2)	(3)	(4) <sup>a</sup>	(5)	(6)	(7)	(8) <sup>a</sup>
Treatment	-0.529 (1.077) [0.623]	-0.288 (0.889) [0.746]	-0.458 (0.793) [0.564]	0.368 (0.733) [0.616]	0.219 (0.994) [0.826]	-0.573 (0.874) [0.512]	-1.238 (0.801) [0.122]	0.500 (0.569) [0.380]
Obs.	604	887	1150	1365	603	886	1149	2443
Dist	10 km	15 km	20 km	24.25 km	10 km	15 km	20 km	48.66 km

*Notes:* RD estimates using bandwidths of 10, 15, and 20 kilometers from the former French-German border. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, and distance to Nancy. Conley standard errors in parentheses and *p*-values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

Table A17: Covariate balance test: 4 categories

Panel <b>A</b> : Alsace and Lorraine								
Variable	Median income 2008		Mean age 2006		Education 1999		Occupation 2006	
	(1)	(2) <sup>a</sup>	(3)	(4) <sup>a</sup>	(5)	(6) <sup>a</sup>	(7)	(8) <sup>a</sup>
Treatment	1.138	1.864	-0.353	-0.645	0.002	0.006	0.009	-0.010
	(0.947)	(0.731)	(0.541)	(0.270)	(0.005)	(0.003)	(0.014)	(0.008)
	[0.230]	[0.011]	[0.515]	[0.017]	[0.621]	[0.023]	[0.515]	[0.218]
Obs.	507	1445	604	2393	604	2368	604	1808
Dist	10 km	29.92 km	10 km	47.14 km	10 km	46.33 km	10 km	32.54 km
Panel <b>B</b> : Alsace vs. Vosges								
Variable	Median income 2008		Mean age 2006		Education 1999		Occupation 2006	
	(1)	(2) <sup>a</sup>	(3)	(4) <sup>a</sup>	(5)	(6) <sup>a</sup>	(7)	(8) <sup>a</sup>
Treatment	4.627	3.543	-1.414	-1.056	0.010	0.019	0.016	-0.011
	(1.135)	(0.803)	(0.841)	(0.409)	(0.008)	(0.004)	(0.026)	(0.014)
	[<0.001]	[0.094]	[0.010]	[0.257]	[<0.001]	[0.526]	[0.455]	
Obs.	196	813	210	1022	210	1284	210	727
Dist	10 km	38.61 km	10 km	49.54 km	10 km	72.07 km	10 km	33.19 km
Panel <b>C</b> : Within Lorraine								
Variable	Median income 2008		Mean age 2006		Education 1999		Occupation 2006	
	(1)	(2) <sup>a</sup>	(3)	(4) <sup>a</sup>	(5)	(6) <sup>a</sup>	(7)	(8) <sup>a</sup>
Treatment	0.236	0.815	0.059	-0.239	0.002	0.002	0.009	-0.013
	(1.015)	(0.868)	(0.641)	(0.382)	(0.006)	(0.003)	(0.016)	(0.010)
	[0.816]	[0.348]	[0.927]	[0.532]	[0.696]	[0.523]	[0.589]	[0.193]
Obs.	311	719	394	1284	394	1617	394	1031
Dist	10 km	25.13 km	10 km	40.45 km	10 km	60.08 km	10 km	29.60 km

*Notes:* Panel **A** tests for discontinuities in covariates using all départements in Alsace and Lorraine, Panel **B** uses only municipalities in Bas-Rhin, Haut-Rhin, and Vosges, while Panel **C** uses municipalities within Lorraine. Education refers to the share of people above 18 with a high school degree and occupation to the share of blue-collar workers in the total population. Controls: distance to Germany (border), distance to Metz, distance to Strasbourg, and distance to Nancy. Conley standard errors in parentheses and  $p$ -values in brackets. Strong differences would indicate problems in the exogenous nature of our treatment assignment, or the comparability of our treatment and control group. There are no clear or significant differences in these main variables.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

Table A18: Covariate balancing

Variable	$\hat{\beta}_{10km}$	$\hat{\beta}_{IK}^a$	Dep. var: Yes 92	Dep. var: Yes 05
<b>Occupation</b>				
Workers, 2006	0.009 [0.589]	-0.013 [0.193]	-10.519 [< 0.001]	-9.359 [< 0.001]
Farmers, 2006	0.003 [0.724]	-0.007 [0.262]	-24.457 [< 0.001]	30.485 [< 0.001]
Artisans, 2006	-0.002 [0.650]	-0.005 [0.060]	-4.197 [0.046]	2.824 [0.046]
Executives, 2006	-0.007 [0.355]	0.009 [0.100]	29.686 [< 0.001]	58.089 [< 0.001]
Intermediate prof., 2006	-0.006 [0.541]	0.002 [0.763]	9.230 [< 0.001]	11.015 [< 0.001]
<b>Economic activity</b>				
Companies, 2011	-3.729 [0.316]	1.320 [0.575]	0.020 [0.014]	0.041 [0.014]
Commercial est., 2011	-0.855 [0.770]	2.292 [0.236]	-0.008 [0.224]	0.020 [0.224]
Industrial est., 2011	-3.344 [0.007]	-0.977 [0.213]	0.037 [< 0.001]	0.012 [< 0.001]
Building est., 2011	1.028 [0.523]	0.386 [0.689]	-0.053 [< 0.001]	-0.100 [< 0.001]
Public est., 2011	-0.699 [0.358]	0.931 [0.058]	0.043 [0.001]	0.003 [0.001]
<b>Public goods</b>				
Theatre rooms	-0.003 [0.299]	-0.001 [0.592]	-0.334 [0.305]	-0.116 [0.305]
Athletic centers	-0.025 [0.617]	0.038 [0.370]	0.129 [0.367]	0.025 [0.367]
Multisport fac.	-0.615 [0.141]	-0.749 [0.008]	0.467 [< 0.001]	0.196 [< 0.001]
Swimming fac.	-0.007 [0.633]	-0.005 [0.861]	-0.010 [0.901]	-0.137 [0.901]
Psychiatric est.	0.003 [0.810]	0.008 [0.253]	1.433 [0.075]	0.968 [0.075]
Service houses	-0.017 [0.137]	-0.014 [0.040]	-0.271 [0.260]	0.052 [0.260]
Healthcare (short)	-0.002 [0.645]	0.001 [0.856]	0.433 [0.708]	0.122 [0.708]
Healthcare (medium)	-0.007 [0.733]	-0.002 [0.942]	0.684 [0.008]	1.004 [0.008]
Healthcare (long)	-0.002 [0.911]	-0.005 [0.653]	2.227 [0.045]	1.669 [0.045]
Post offices	-0.074 [0.186]	0.030 [0.412]	0.504 [< 0.001]	-0.919 [< 0.001]
Elementary schols	-0.205 [0.311]	0.006 [0.950]	0.842 [< 0.001]	0.381 [< 0.001]
Highschools	-0.002 [0.729]	0.009 [0.135]	2.351 [0.006]	1.496 [0.006]
Vocational training	0.001 [0.870]	0.000 [0.963]	2.141 [< 0.001]	0.485 [< 0.001]
Tech. vocational training	0.002 [0.356]	0.002 [0.427]	0.265 [0.213]	0.942 [0.213]
<b>Demographics</b>				
Population density	-77.246 [0.287]	91.480 [0.058]	0.001 [< 0.001]	0.000 [< 0.001]

*Notes:* This table demonstrates the balancing in our respective samples using all départements in Alsace and Lorraine, for different bandwidths. The time period chosen are partly determined by data availability. The different public goods and population density are all measured in the year 2011. All estimations include the same distance controls as our main specification. *p*-values in brackets. There are on average no systematical differences. In the cases where we find a difference in some specifications, it would bias us against our main result as the third and fourth column show.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

Table A19: Discontinuities in referenda results, within Lorraine, excluding Metz

Panel A: Excluding Metz Agglomeration						
Dep. Variable:	Share Yes 1969		Share Yes 1992		Share Yes 2005	
	(1)	(2) <sup>a</sup>	(3)	(4) <sup>a</sup>	(5)	(6) <sup>a</sup>
Treatment	13.194	12.029	4.082	3.599	4.283	5.791
	(2.621)	(1.949)	(1.940)	(1.269)	(2.087)	(1.452)
	[<0.001]	[<0.001]	[0.036]	[0.005]	[0.041]	[<0.001]
Obs.	337	817	355	1152	355	779
Dist	10 km	34.09 km	10 km	53.78 km	10 km	26.22 km
Panel B: Excluding within 10 kilometers from Metz						
Dep. Variable:	Share Yes 1969		Share Yes 1992		Share Yes 2005	
	(1)	(2) <sup>a</sup>	(3)	(4) <sup>a</sup>	(5)	(6) <sup>a</sup>
Treatment	12.670	10.678	3.822	5.867	4.000	7.341
	(2.633)	(1.795)	(1.850)	(1.137)	(2.082)	(1.340)
	[<0.001]	[<0.001]	[0.040]	[<0.001]	[0.055]	[<0.001]
Obs.	372	1145	392	1436	392	1171
Dist	10 km	38.47 km	10 km	49.06 km	10 km	35.71 km
Panel C: Excluding within 15 kilometers from Metz						
Dep. Variable:	Share Yes 1969		Share Yes 1992		Share Yes 2005	
	(1)	(2) <sup>a</sup>	(3)	(4) <sup>a</sup>	(5)	(6) <sup>a</sup>
Treatment	12.367	10.585	3.940	5.483	4.450	7.082
	(2.668)	(1.891)	(1.889)	(1.151)	(2.033)	(1.225)
	[<0.001]	[<0.001]	[0.038]	[<0.001]	[0.029]	[<0.001]
Obs.	353	1017	372	1316	372	1339
Dist	10 km	34.56 km	10 km	44.81 km	10 km	45.98 km

*Notes:* Discontinuity at the treatment border using municipalities in Lorraine, excluding Metz. Panel **A** excludes all municipalities in Metz Agglomeration, Panel **B** excludes all municipalities within 10 kilometers from Metz, and Panel **C** excludes all municipalities within 15 kilometers from Metz. Outcomes are share of Yes votes in the 1969 referendum, share of Yes votes in the 1992 referendum, and share of Yes votes in the 2005 referendum. Conley standard errors in parentheses and *p*-values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

Table A20: Share of Yes votes and religion

	Dep. Variable: Share of Yes votes 1992			Dep. Variable: Share of Yes votes 2005		
	(1)	(2)	(3)	(4)	(5)	(6)
Attendance [mean]	-1.839			-1.774		
	[0.167]			[0.113]		
Attendance: Weekly		0.114			0.099	
		[0.167]			[0.135]	
Attendance: 2-3 times a month		0.002			0.025	
		[0.983]			[0.788]	
Attendance: Once a month		-0.052			-0.097	
		[0.625]			[0.164]	
Attendance: Sev. times a year		0.057			0.054	
		[0.114]			[0.144]	
Attendance: Less freq.		0.036			-0.001	
		[0.391]			[0.988]	
Roman Catholic			0.029			0.004
			[0.291]			[0.902]
Protestant			0.353			0.146
			[0.054]			[0.321]
Christian Ortodox			0.115			0.267
			[0.846]			[0.585]
Jewish			0.847			1.095
			[0.116]			[0.278]
Moslem			-0.092			0.008
			[0.437]			[0.955]
Other Religions			-0.155			0.010
			[0.495]			[0.971]
Obs.	94	94	94	94	94	94

*Notes:* This table tests whether there is a clear relationship between religious affiliation and voting in the two referenda 1992 and 2005. The OLS estimates use aggregate survey results at the département-level. *Attendance* refers to how often the respondents attend religious services. *Never attending* is the omitted reference category for attendance, *no religious denomination* is the omitted reference category for religion. Controls: Sex, Age, Years of schooling, Urban vs Rural, Union membership, Degree, Income, and Household size. *p*-values in brackets. There is no systematic effect of religion, which is reassuring as the areas in former Alsace-Lorraine has a slightly different history with regard to schooling. Accordingly, these differences and schooling should not explain our results.

Table A21: Effects at the language border

Panel A: Share Yes 1969				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	-3.094	-0.595	1.486	0.169
	(2.544)	(2.182)	(1.999)	(2.100)
	[0.225]	[0.785]	[0.458]	[0.936]
Obs.	285	386	479	408
Dist	10 km	15 km	20 km	15.81 km
Panel B: Share Yes 1992				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	2.033	3.454	4.738	4.557
	(1.399)	(1.247)	(1.101)	(0.949)
	[0.147]	[0.006]	[<0.001]	[<0.001]
Obs.	534	733	954	1265
Dist	10 km	15 km	20 km	31.20 km
Panel C: Share Yes 2005				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	2.622	4.360	4.552	4.654
	(1.075)	(0.976)	(0.902)	(0.963)
	[0.015]	[<0.001]	[<0.001]	[<0.001]
Obs.	535	734	955	778
Dist	10 km	15 km	20 km	15.89 km

*Notes:* RD estimates using bandwidths of 10, 15, and 20 kilometers from the language border within Moselle. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, and distance to Nancy. Conley standard errors in parentheses and  $p$ -values in brackets.

*Short interpretation:* The differences in 1992 and 2005 could signal that individuals who spoke German would profit more from European integration, e.g. through more exchange with Germany, or were exposed to the EU friendly German media to a higher extent. We exclude those municipalities as a robustness test.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

Table A22: Differences in population changes between 1916-1946 (within Lorraine)

Panel A: Population difference 1916 to 1926				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	-5.980 (6.117) [0.329]	-2.309 (6.747) [0.732]	-0.494 (6.582) [0.940]	-4.909 (5.374) [0.361]
Obs.	394	581	740	1402
Dist	10 km	15 km	20 km	47.13 km
Panel B: Population difference 1936 to 1946				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	-4.525 (3.632) [0.213]	-4.018 (2.894) [0.166]	-3.866 (2.571) [0.133]	-3.944 (2.105) [0.061]
Obs.	393	581	741	1153
Dist	10 km	15 km	20 km	34.80 km
Panel C: Population difference 1916 to 1946				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	-13.061 (10.206) [0.201]	-6.966 (11.342) [0.539]	-2.662 (11.130) [0.811]	-10.720 (9.039) [0.236]
Obs.	393	580	739	1433
Dist	10 km	15 km	20 km	48.95 km

*Notes:* All estimates include population differences for municipalities only within Lorraine. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, distance to Nancy. Conley standard errors in parentheses and  $p$ -values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

Table A23: Discontinuities in referenda results, within Lorraine, controlling for historical migration

Dep. Variable:	Share Yes 1969		Share Yes 1992		Share Yes 2005	
	(1)	(2) <sup>a</sup>	(3)	(4) <sup>a</sup>	(5)	(6) <sup>a</sup>
Treatment	11.937 (2.515) [<0.001]	10.809 (1.590) [<0.001]	3.637 (1.813) [0.046]	5.777 (1.124) [<0.001]	3.547 (2.064) [0.087]	7.227 (1.366) [<0.001]
Obs.	373	1260	393	1508	393	1102
Dist	10 km	44.43 km	10 km	53.22 km	10 km	32.86 km

*Notes:* Discontinuity at the treatment border using municipalities in Lorraine, controlling for migration between 1916 and 1946 (changes in population between 1916 and 1926, between 1936 and 1946, and between 1916 and 1946). Outcomes are share of Yes votes in the 1969 referendum, share of Yes votes in the 1992 referendum, and share of Yes votes in the 2005 referendum. Conley standard errors in parentheses and  $p$ -values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

Table A24: Placebo test: Border between Alsace and Lorraine, and the rest of France

Panel <b>A</b> : Share Yes 1992				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	-3.168	-0.649	0.058	3.170
	(2.040)	(1.728)	(1.465)	(0.769)
	[0.121]	[0.707]	[0.968]	[<0.001]
Obs.	404	606	814	11416
Dist	10 km	15 km	20 km	218.68 km
Panel <b>B</b> : Share Yes 2005				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	0.208	1.045	1.496	0.135
	(2.006)	(1.666)	(1.453)	(0.735)
	[0.917]	[0.531]	[0.303]	[0.854]
Obs.	405	608	816	10899
Dist	10 km	15 km	20 km	209.71 km

*Notes:* RD estimates using bandwidths of 10, 15, and 20 kilometers from the border between Alsace and Lorraine, and the rest of France. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, and distance to Nancy. Conley standard errors in parentheses and *p*-values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.



Table A25: Placebo test: RD estimates at the pre-1870 border between historical Moselle and Meurthe, within current Moselle

Panel A: Share Yes 69								
Variable	Within current Moselle			Within current Meurthe-et-Moselle				
	(1)	(2)	(3)	(4) <sup>a</sup>	(5)	(6)	(7)	(8) <sup>a</sup>
Treatment	1.185	-2.383	-3.833	-3.923	-1.512	7.934	6.108	-6.839
	(2.558)	(2.226)	(2.081)	(2.002)	(6.570)	(6.059)	(5.789)	(4.120)
	[0.644]	[0.285]	[0.066]	[0.051]	[0.819]	[0.195]	[0.294]	[0.098]
Obs.	188	270	361	424	47	75	108	525
Dist	10 km	15 km	20 km	23.86 km	10 km	15 km	20 km	70.74 km

Panel B: Within current Moselle								
Variable	Share Yes 92			Share Yes 05				
	(1)	(2)	(3)	(4) <sup>a</sup>	(5)	(6)	(7)	(8) <sup>a</sup>
Treatment	-1.020	-0.892	1.196	0.665	-1.832	-2.947	-1.295	0.866
	(2.228)	(2.047)	(1.909)	(1.933)	(2.271)	(1.853)	(1.659)	(1.664)
	[0.648]	[0.664]	[0.532]	[0.731]	[0.421]	[0.113]	[0.436]	[0.603]
Obs.	186	270	361	340	189	273	364	462
Dist	10 km	15 km	20 km	18.74 km	10 km	15 km	20 km	25.65 km

Panel C: Within current Meurthe-et-Moselle								
Variable	Share Yes 92			Share Yes 05				
	(1)	(2)	(3)	(4) <sup>a</sup>	(5)	(6)	(7)	(8) <sup>a</sup>
Treatment	-7.063	-10.198	-8.504	-2.539	-9.075	2.516	7.873	6.737
	(8.412)	(6.030)	(4.937)	(2.479)	(5.273)	(6.121)	(5.652)	(3.249)
	[0.406]	[0.095]	[0.088]	[0.306]	[0.093]	[0.682]	[0.167]	[0.039]
Obs.	50	83	116	578	50	83	116	239
Dist	10 km	15 km	20 km	76.16 km	10 km	15 km	20 km	33.66 km

Notes: RD estimates at pre-1871 border between the départements Moselle and Meurthe. Panel A uses municipalities within modern Moselle while Panel B uses municipalities within modern Meurthe-et-Moselle. Controls added. Conley standard errors in parentheses and *p*-values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

Interpretation: The historical border within current Moselle provides a good placebo test, as it does mostly not follow the current borders. Note that the estimates within current Meurthe-et-Moselle have different signs and switch signs for the *Share Yes 05* estimations.

Table A26: Newspaper subscription shares: excluding Metz, and discontinuity at language border

Panel A: Excluding Metz				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	7.980	7.667	6.927	6.891
	(1.527)	(1.361)	(1.315)	(1.317)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Obs.	259	365	455	450
Dist	10 km	15 km	20 km	19.71 km
Panel B: Effect at the language border				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	-0.763	0.088	0.110	0.245
	(0.954)	(0.804)	(0.801)	(0.809)
	[0.424]	[0.913]	[0.891]	[0.762]
Obs.	291	394	490	452
Dist	10 km	15 km	20 km	17.71 km
Panel C: Excluding German-speaking municipalities				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	10.000	9.815	9.777	10.247
	(1.421)	(1.247)	(1.149)	(1.092)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Obs.	385	553	684	937
Dist	10 km	15 km	20 km	34.32 km

*Notes:* Discontinuity in newspaper subscription shares at the treatment border using municipalities in Lorraine (Moselle, Meurthe et Moselle, and Meuse), and at the language border using municipalities in Moselle. Panel **A** excludes all municipalities in the Metz agglomeration, panel **B** tests for discontinuities at the language border, and panel **C** excludes all German-speaking municipalities. Conley standard errors in parentheses and  $p$ -values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

Table A27: RD results: Subscription shares of regional newspaper, controlling for the number of sales points

Share households with subscription of “Le Republicain Lorraine”				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	9.880	9.958	9.979	11.122
	(1.376)	(1.218)	(1.112)	(0.950)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Obs.	394	583	744	1392
Dist	10 km	15 km	20 km	46.23 km

*Notes:* RD estimates using bandwidths of 10, 15, and 20 kilometers from the border between Alsace and Lorraine, and the rest of France. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, distance to Nancy, and number of sales points where the newspaper can be bought locally. Conley standard errors in parentheses and  $p$ -values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

Table A28: List of Twitter Keywords

List France		List Germany	
les Bleus	#BLEUS	mannschaft	allemagne
#SPAFRA	#FRA	DFB_Team	#GER
#ESPFRA	#UKRFRA	#GERPOL	#FRAGER
#SWEFRA	#SUIFRA	#FRADEU	#FRAALL
#SWIFRA	#ECUFRA	#TeamGermany	#DEU
Landreau	Lloris	#ALL	#HOLDEU
Ruffier	Debuchy	#NEDGER	#NEDALL
Digne	Evra	#DENDEU	#DANDEU
Koscielny	Mangala	#DANGER	#DENGER
Sagna	Sakho	#DANALL	#DENALL
Varane	Cabaye	#USAGER	#USAALL
Matuidi	Mavuba	#USADEU	#BRADEU
Pogba	Schneiderlin	#BRAALL	#BRAGER
Sissoko	Valbuena	Neuer	Wiese
Benzema	Cabella	Zieler	Badstuber
Giroud	Griezmann	Boateng	Höwedes
Rémy	Deschamps	Hummels	Lahm
Carrasso	Mandanda	Mertesacker	Schmelzer
Clichy	Mexès	Bender	Götze
Rami	Réveillère	Gündogan	Khedira
Arfa	Diarra	Kroos	Özil
M'Vila	Malouda	Reus	Schweinsteiger
Marvin Martin	Nasri	Gomez	Klose
Ribéry	Valbuena	Müller	Podolski
Ménez	Blanc	Schürrle	Löw
Boghossian	Gasset	Flick	Köpke
Raviot		Weidenfeller	Durm
		Großkreutz	Mustafi
		Draxler	Ginter
		Kramer	

Table A29: Twitter data, within Lorraine

Dep. Variable:	Share Tweets Germany		Share Tweets France	
Variable	(1)	(2) <sup>a</sup>	(3)	(4) <sup>a</sup>
Treatment	2.955 (2.103) [0.162]	-1.990 (1.564) [0.204]	0.092 (0.570) [0.872]	-1.020 (1.037) [0.326]
Obs.	169	430	169	481
Dist	10 km	31.27 km	10 km	38.24 km

*Notes:* Testing for discontinuities in the share of tweets about the German and French national football team using municipalities in Moselle, Meurthe et Moselle, and Meuse. The dependent variable is coded as the number of tweets about Germany during World Cup 2014 in Brazil, divided by the total number of tweets in each municipality. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, and distance to Nancy. Conley standard errors in parentheses and  $p$ -values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.



## G Figures

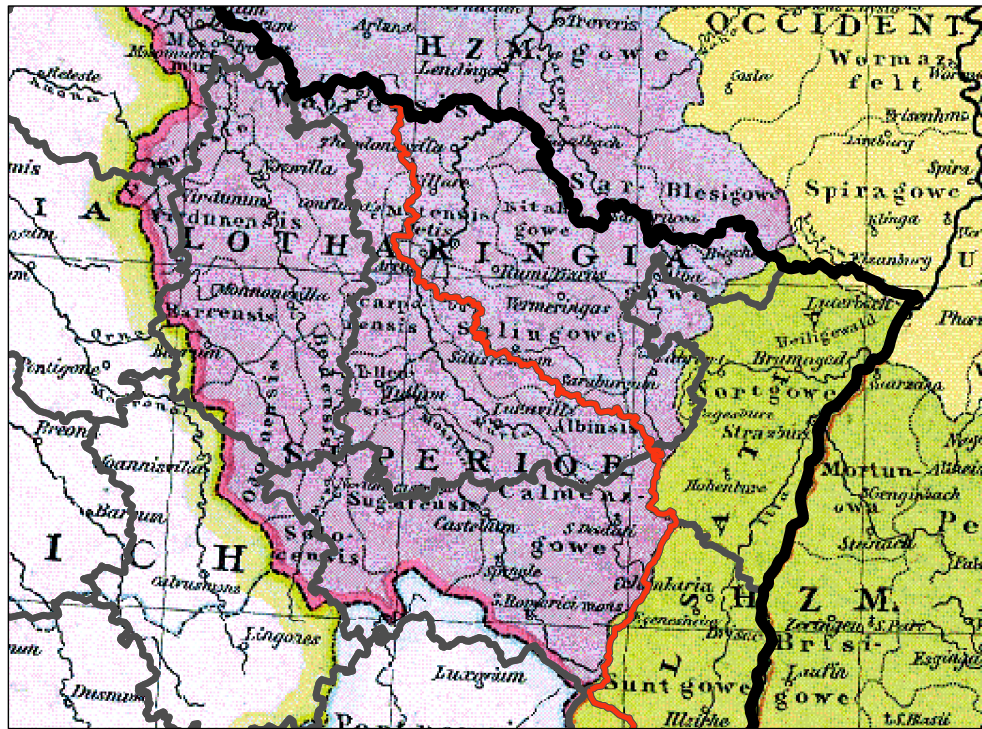
Figure A3: Map of Lotharingia around 1000 A.D.



Notes: Map depicting the former Duchy of Lotharingia, around 1000: Pink= Lower Lorraine, Purple = Upper Lorraine, Orange = Frisia (effectively detached from Lotharingia). This map is used in the *Allgemeiner historischer Handatlas* by Gustav Droysen in 1886. Alsace was a part of the duchy of Swabia at that time.



Figure A4: Map of Lotharingia around 1000 A.D., zoomed in with 1870 border



**Legend**

French National Border
 — Border Alsace-Lorraine
  French Department Border

Notes: Map depicting the former Duchy of Lotharingia, around 1000: Pink= Lower Lorraine, Purple = Upper Lorraine, Orange = Frisia (effectively detached from Lotharingia). This map is used in the *Allgemeiner historischer Handatlas* by Gustav Droysen in 1886. Alsace was a part of the duchy of Swabia at that time.

Figure A5: Map of Lorraine in the 1378 century



Notes: Map of Lorraine in the 14th century. This is a modified extract from the map *Deutschland beim Tode Karl IV.* by Karl Wolf in Meyers Lexikon 6. Auflage. The red line shows the border from the Franco-Prussian war, clearly not following the pre-existing borders and cutting through historical entities. Created from authors' own version of the map.



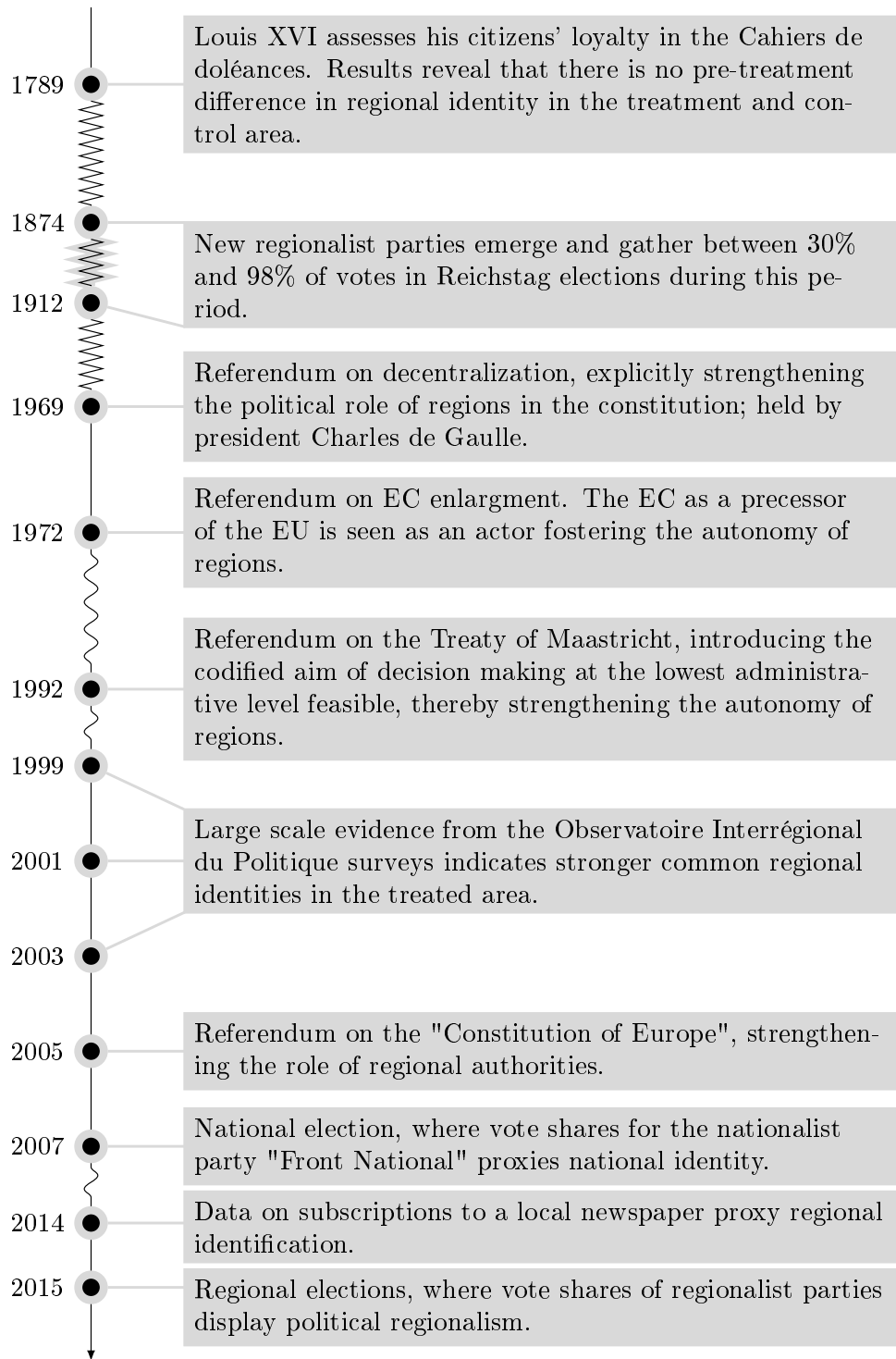
Figure A6: Map of Lorraine in the 17th century



Notes: Map of Lorraine in 1790. The map is an extract from *Carte de la Lorraine, du Barrois et des Trois Evêchés de Metz, Toul et Verdun. Divisée par Baillages, Dans laquelle se trouve Comprise la Généralité de Metz* created by Robert de Vaugondy, Didier (1723-1786) Dezauche, Jean-Claude (1745-1824) in 1756. The original is in the *Bibliothèque nationale de France, département des Cartes et plans, GE C-9972*. A scanned online version is accessible at <http://gallica.bnf.fr/ark:/12148/btv1b7710337x>. It shows the duchy of Lorraine as well as the area of the partly independent enclaves Metz, Verdun and Toul. Although it is admittedly hard to distinguish which area us belongs to which (another version is available at <http://gallica.bnf.fr/ark:/12148/btv1b53099747j/f1.item.zoom>), it is apparent that the borders do not coincide with the border drawn after the Franco-Prussian war. It is also apparent that partly independent enclaves existed on both sides of the border which we use to distinguish in a treatment and control area.

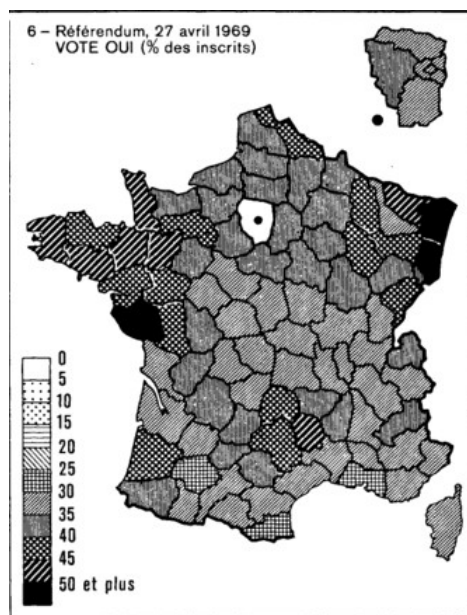


Figure A7: Timeline of outcomes



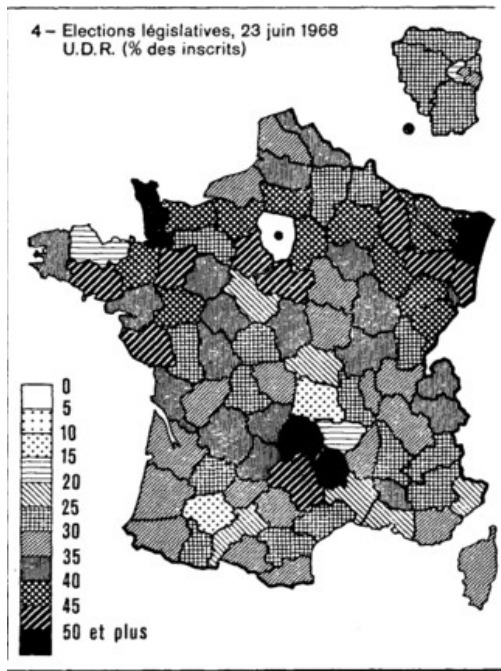
Notes: Distances on the straight parts of the timeline are proportional to years. The curled line is proportional to five years and the zigzag line is proportional to 25 years.

Figure A8: Agreement referendum about establishing regions as political entity, 1969

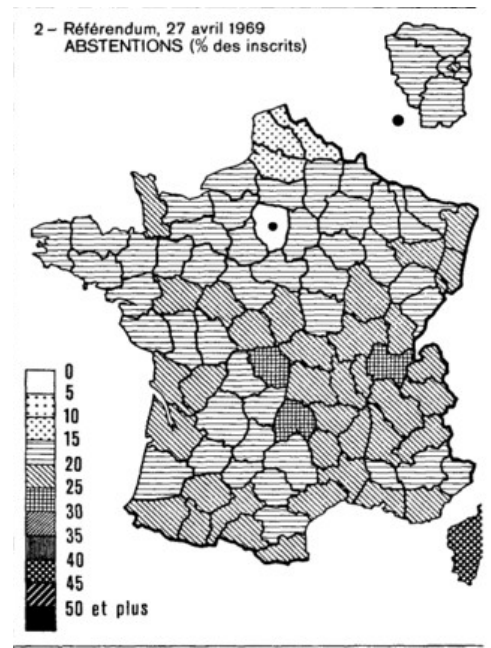


*Notes:* Referendum on creating regions as political entity (1969). Vote shares out of all eligible voters, i.e. out of yes votes, no votes, blanks and abstentions. There is no comparable map showing only the yes share out of valid votes, but this Online Appendix shows maps of abstentions that do not differ between départements. *Source:* [Lancelot and Lancelot \(1970\)](#).

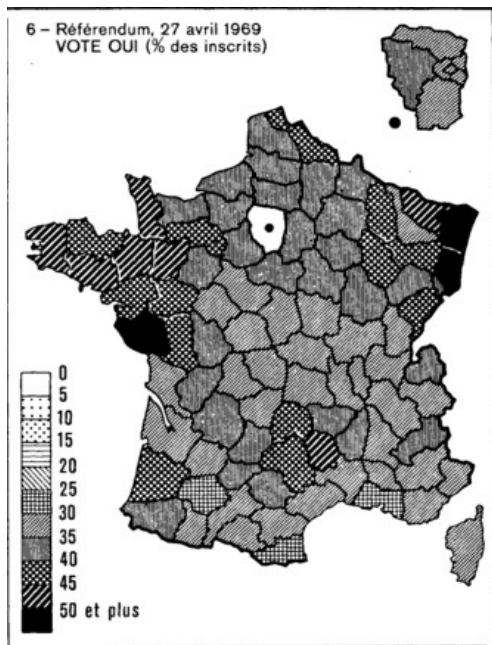
Figure A9: Election and referendum results, 1968 and 1969



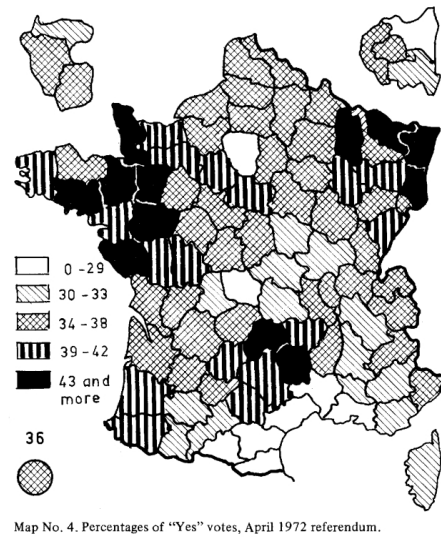
(a) Legislative election 1968



(b) Abstention 1969 referendum



(c) Share of yes votes in 1969 referendum

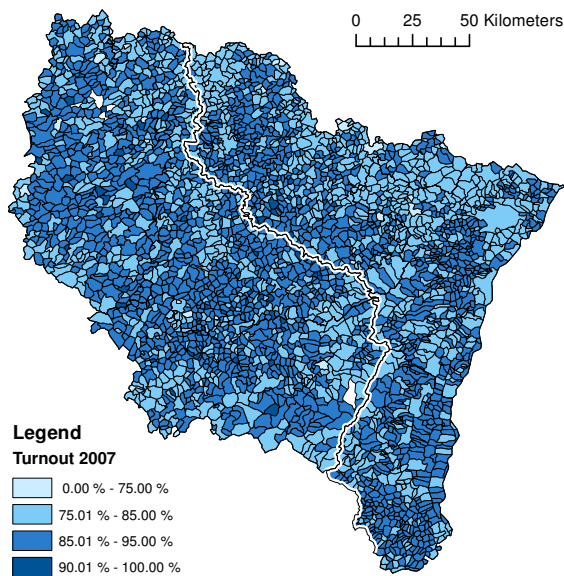


(d) Share of yes votes in 1972 referendum

Notes: Figure a) shows vote shares for the Gaullist right-wing party *Union for the Defense of the Republic* (U.D.R.) in the legislative elections of 1968. Figures b) and c) shows the share of absentees and share of yes votes (among all votes, including invalid/blank votes), respectively, in the 1969 constitutional referendum about decentralization and establishing the regions as an important political unit in the Constitution. Figure d) presents results for the 1972 referendum, which was about "The Treaty of Accession" the question was about whether Denmark, Ireland, Norway and the United Kingdom should be allowed to become members of the "European Communities", a predecessor of the European Union. There were no differences in vote shares for U.D.R or share of absentees between the Moselle (treated) and Meurthe-et-Moselle (non-treated), while the share of yes votes in both the 1969 and the 1972 referenda was higher in Moselle.

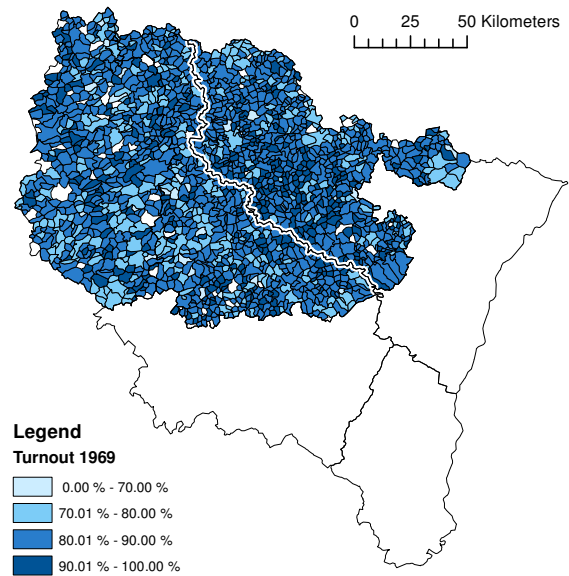
Source: Figures a), b) and c) are from Lancelot and Lancelot (1970). Figure d) is from Leleu (1976).

Figure A10: Maps of municipal level turnout in referenda in 1969, 1992 and 2005; and the presidential election of 2007



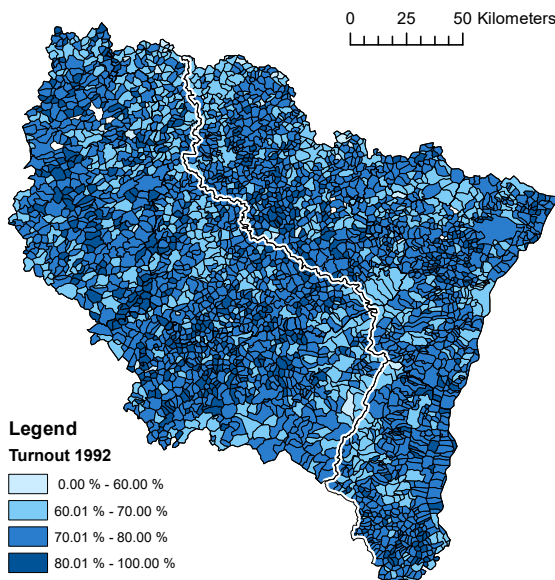
*Notes (a):* Turnout for the presidential election 2007. The treatment border formerly dividing the area is highlighted in white. Darker shades reflect higher values.

(a) Turnout 2007



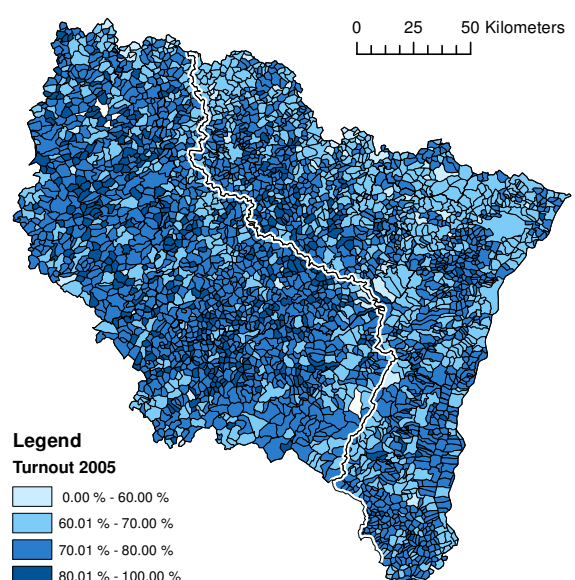
*Notes (b):* Turnout in the constitutional referendum in 1969. Areas where data is not available are left blank. Data is available for the départements of Meuse, Meurthe-et-Moselle and Moselle. The treatment border formerly dividing the area is highlighted in white. Darker shades reflect higher values.

(b) Turnout 1969



*Notes (c):* Turnout in the referendum in 1992. The treatment border formerly dividing the area is highlighted in white. Darker shades reflect higher values.

(c) Turnout 1992

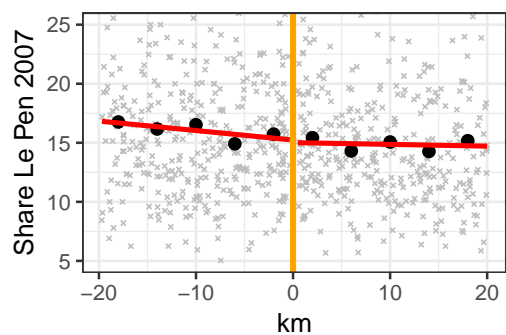


*Notes (d):* Turnout in the referendum in 2005. The treatment border formerly dividing the area is highlighted in white. Darker shades reflect higher values.

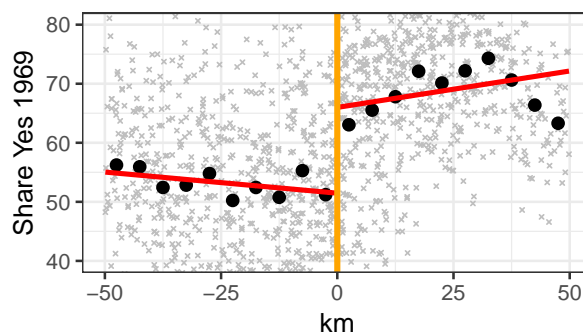
(d) Turnout 2005



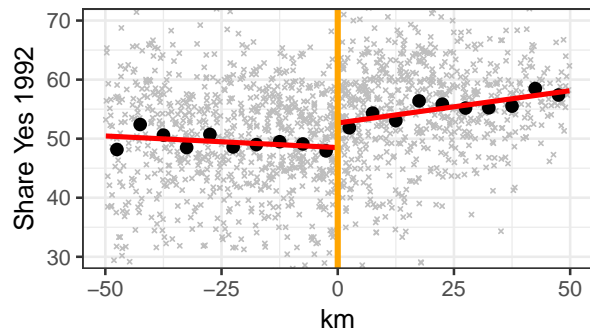
Figure A11: RD plots, within Lorraine (50 kilometers, 20km in paper), 1st degree polynomial



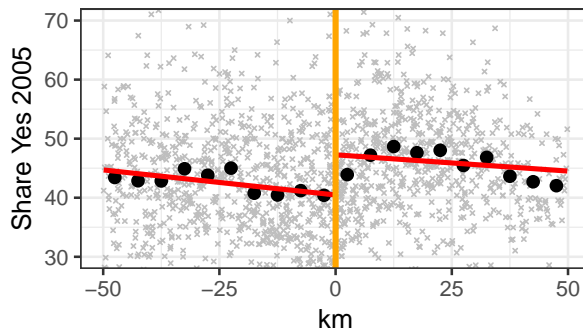
(a) Share Le Pen 2007



(b) Share Yes 1969



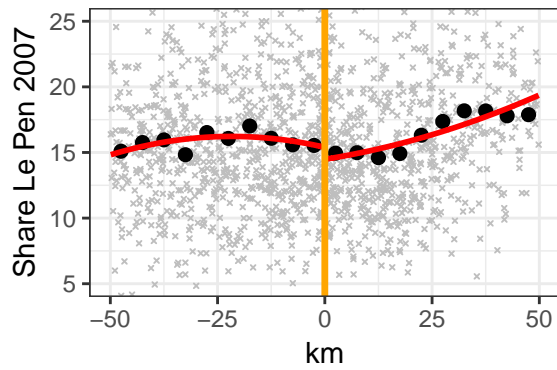
(c) Share Yes 1992



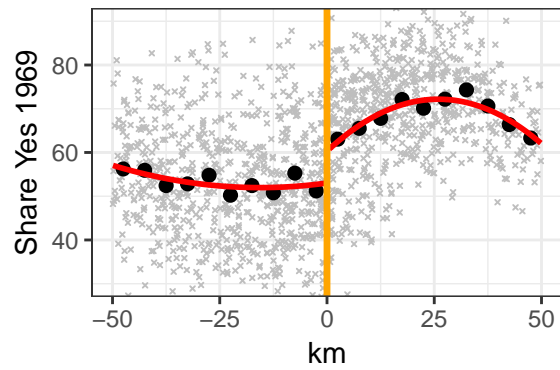
(d) Share Yes 2005

Notes: RD plots, using municipalities in Lorraine. Fitted line based on first degree polynomial. Black dots represent means using 5km bins.

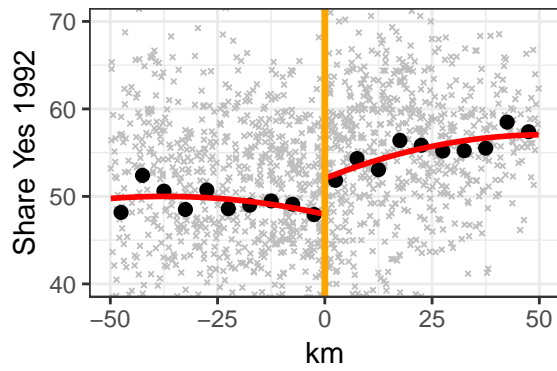
Figure A12: RD plots, within Lorraine (50 kilometers), 2nd degree polynomial



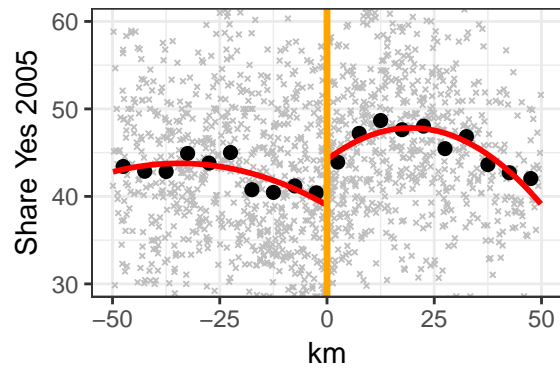
(a) Share Le Pen 2007



(b) Share Yes 1969



(c) Share Yes 1992



(d) Share Yes 2005

Notes: RD plots, using municipalities in Lorraine. Fitted line based on 2nd degree polynomial. Black dots represent means using 5km bins.

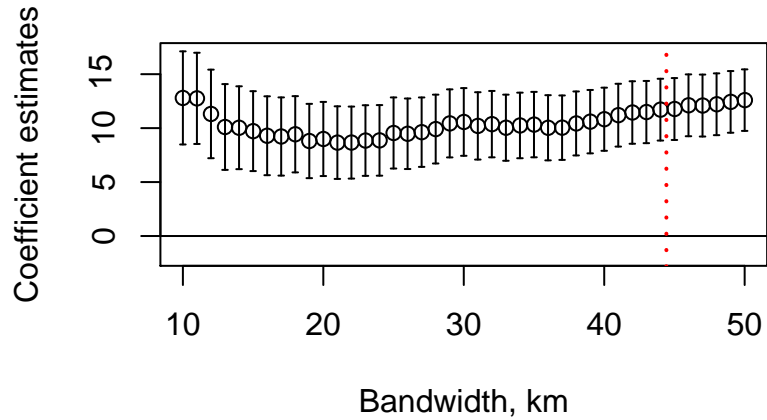
## **RDD: Alternative specifications**

This page provides an overview over the subsequently shown bandwidth plots in which we show the robustness of our results against variations in control variables and geographic areas.

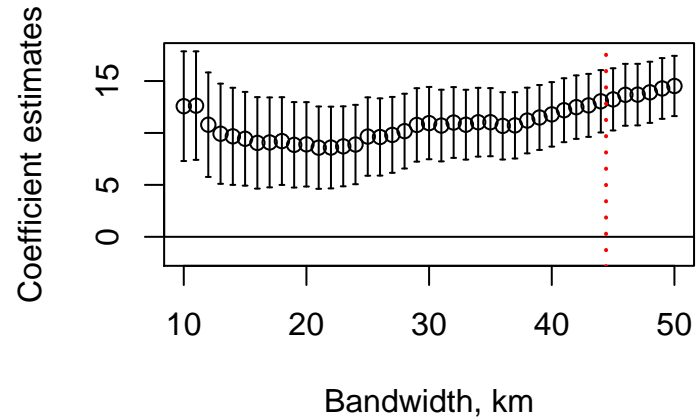
- A13a Varying bandwidths (10km - 50km) for the treatment effect on the 1969 referendum including control variables
- A13b Varying bandwidths (10km - 50km) for the treatment effect on the 1969 referendum without control variables
- A14a Varying bandwidth (10km - 50km) for the treatment effect on the 1969 referendum controlling for border segments
- A14b Varying bandwidth (10km - 50km) for the treatment effect on the 1969 referendum controlling for the distance to the language border
- A15a Varying bandwidths (10km - 50km) for the treatment effect on the 1969 referendum controlling for longitude and latitude
- A15b Varying bandwidths (10km - 50km) for the treatment effect on the 1969 referendum controlling for longitude, latitude and their interaction
- A16a Varying bandwidths (10km - 50km) for the treatment effect on the 1992 referendum within Lorraine
- A16b Varying bandwidths (10km - 50km) for the treatment effect on the 2005 referendum within Lorraine
- A17a Varying bandwidths (10km - 50km) for the treatment effect on the 1992 referendum in Alsace and Lorraine
- A17b Varying bandwidths (10km - 50km) for the treatment effect on the 2005 referendum in Alsace and Lorraine
- A18a Varying bandwidths (10km - 50km) for the treatment effect on the 1992 referendum within Lorraine using no controls
- A18b Varying bandwidths (10km - 50km) for the treatment effect on the 2005 referendum within Lorraine using no controls
- A19a Varying bandwidths (10km - 50km) for the treatment effect on the 1992 referendum within Lorraine, controlling for longitude and latitude
- A19b Varying bandwidths (10km - 50km) for the treatment effect on the 2005 referendum within Lorraine, controlling for longitude and latitude
- A20a Varying bandwidths (10km - 50km) for the treatment effect on the 1992 referendum within Lorraine, controlling for longitude, latitude and their interaction
- A20b Varying bandwidths (10km - 50km) for the treatment effect on the 2005 referendum within Lorraine, controlling for longitude, latitude and their interaction
- A21a Varying bandwidths (10km - 50km) for the treatment effect on the 1992 referendum controlling for border segments
- A21b Varying bandwidths (10km - 50km) for the treatment effect on the 2005 referendum controlling for border segments
- A22a Varying bandwidths (10km - 50km) for the treatment effect on the 1992 referendum within Lorraine controlling for border segments
- A22b Varying bandwidths (10km - 50km) for the treatment effect on the 2005 referendum within Lorraine controlling for border segments
- A23a Varying bandwidths (10km - 50km) for the treatment effect on the 1992 referendum within Lorraine controlling for distance to the language border
- A23b Varying bandwidths (10km - 50km) for the treatment effect on the 2005 referendum within Lorraine controlling for distance to the language border
- A24 Varying bandwidths (10km - 50km) for the treatment effect on regional newspaper subscriptions, first and second degree polynomial
- A25a Varying bandwidths (10km - 50km) for the treatment effect on regional newspaper subscriptions controlling for latitude and longitude
- A25b Varying bandwidths (10km - 50km) for the treatment effect on regional newspaper subscriptions controlling for latitude, longitude and their interaction
- A26b Varying bandwidths (10km - 50km) for the treatment effect on regional newspaper subscriptions using no controls
- A26a Varying bandwidths (10km - 50km) for the treatment effect on regional newspaper subscriptions controlling for distance to the language border



Figure A13: Estimation plots for 1969 referendum, within Lorraine



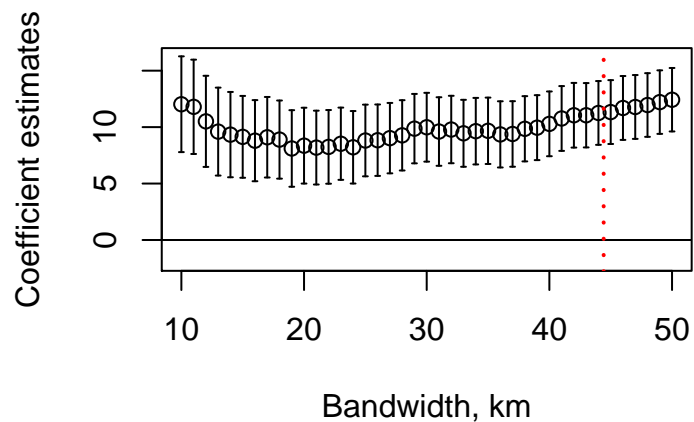
(a) Share Yes 1969



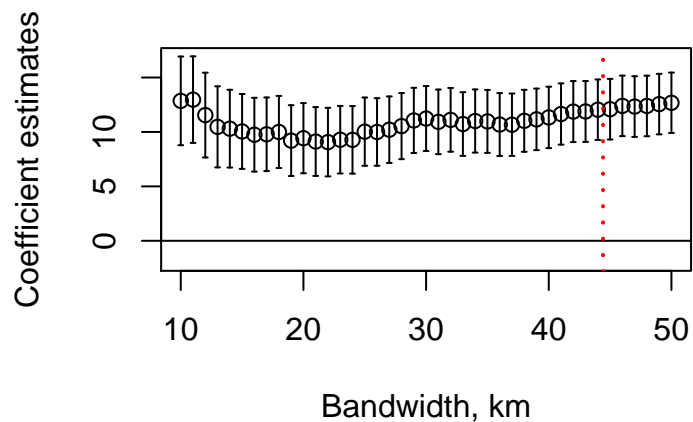
(b) Share Yes 1969, no controls

Notes: Estimates of treatment effect, bandwidths varying between 10 to 50 kilometers, within Lorraine. 1st degree polynomial. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors). (a) shows the results with controls, (b) without controls.

Figure A14: Estimation plots, 1969 referendum, controlling for border segments, and distance to language border



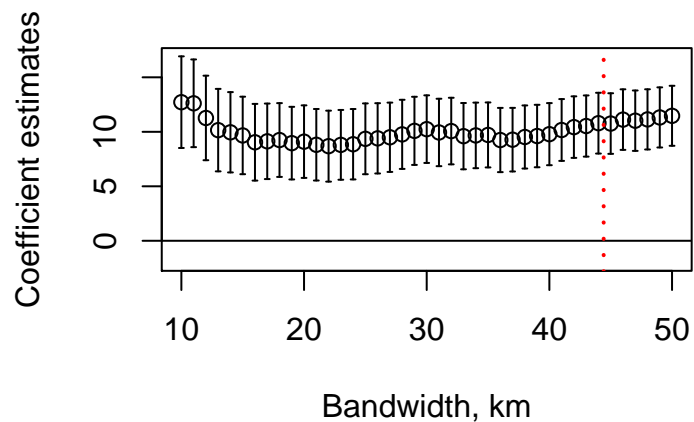
(a) Control for border segments



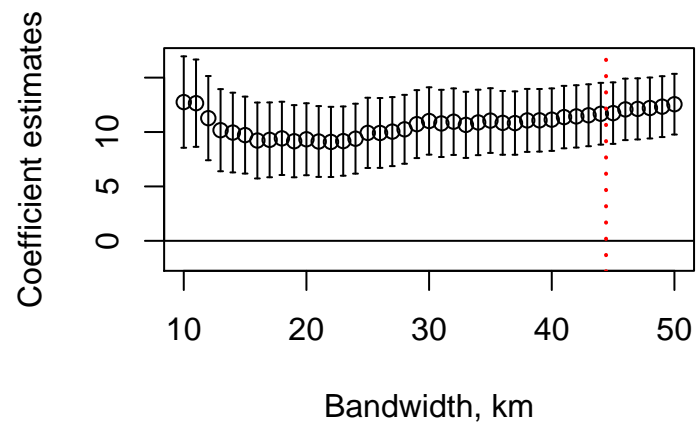
(b) Control for distance to language border

*Notes:* Estimates of treatment effect, bandwidth of 10 to 50 kilometres, within Lorraine. (a) adds controls for border segments and (b) controls for distance to language border. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors).

Figure A15: Estimation plots, 1969 referendum, controlling for longitude, latitude and their interaction



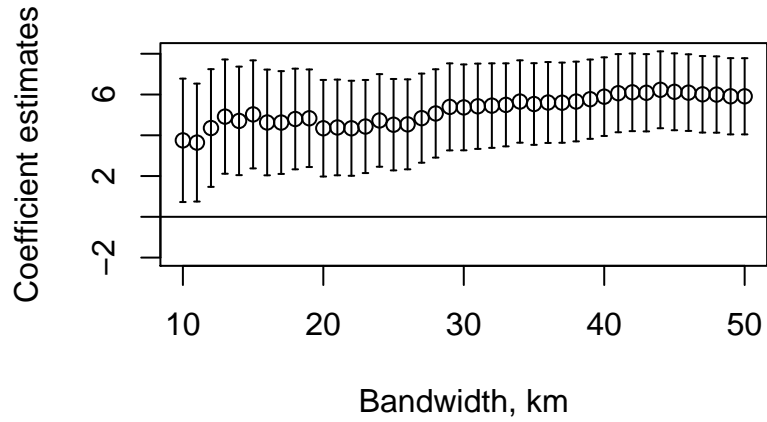
(a) Controlling for longitude and latitude



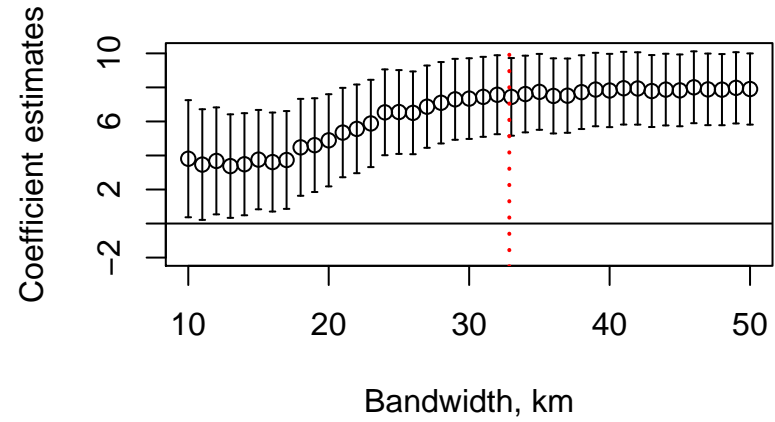
(b) Controlling for longitude and latitude, and interaction

*Notes:* Estimates of treatment effect, bandwidth of 10 to 50 kilometres, within Lorraine, controlling for longitude, latitude and their interaction. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors). It is unclear whether controls should be included in these kind of regressions, but as the graphs show this does not affect our results.

Figure A16: Estimation plots for 1992 and 2005 referenda, within Lorraine



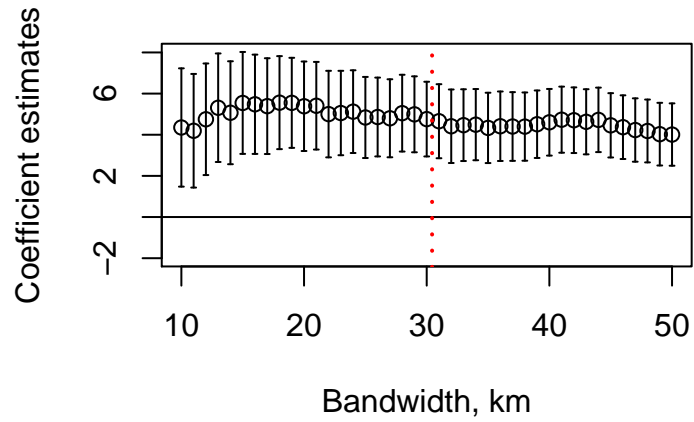
(a) Share Yes 1992



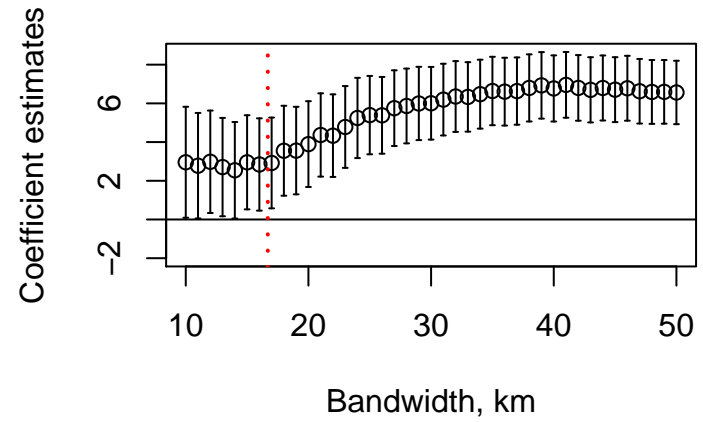
(b) Share Yes 2005

Notes: Estimates of treatment effect, bandwidths varying between 10 to 50 kilometers, within Lorraine. 1st degree polynomial. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors).

Figure A17: Estimation plots for 1992 and 2005 referenda, whole border



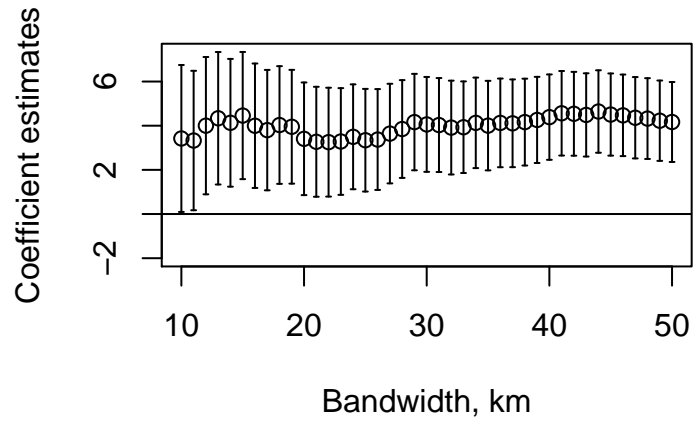
(a) Referendum 1992



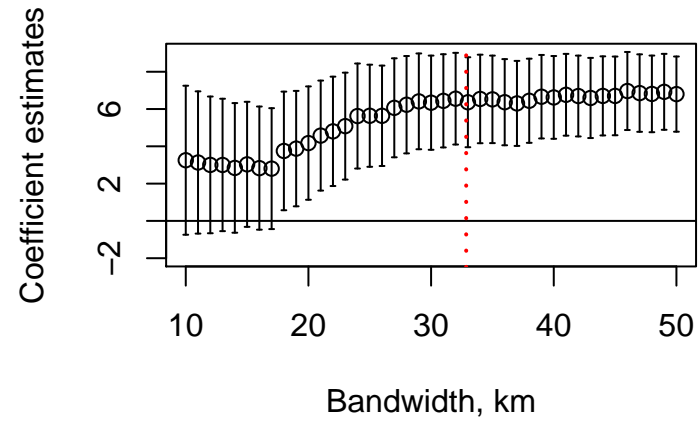
(b) Referendum 2005

Notes: Estimates of treatment effect, bandwidths varying between 10 to 50 kilometres, for the whole border. Local linear regressions, i.e. using a 1st degree polynomial. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors).

Figure A18: Estimation plots for 1992 and 2005 referenda, no controls



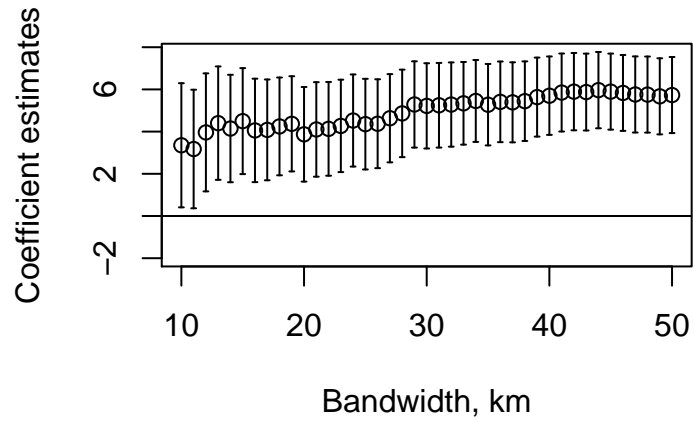
(a) Referendum 1992



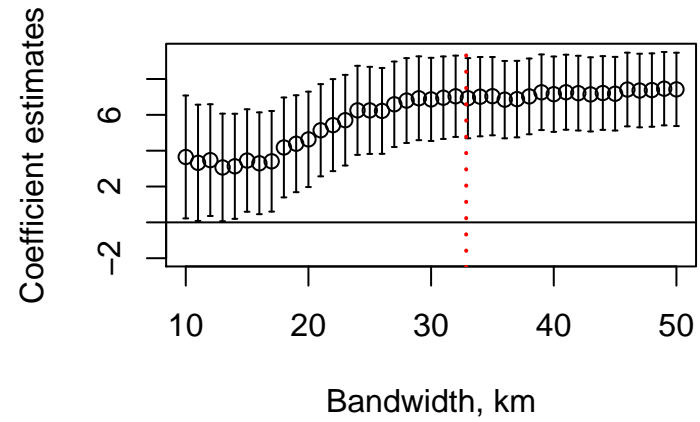
(b) Referendum 2005

*Notes:* Estimates of treatment effect, bandwidth of 10 to 50 kilometres, within Lorraine. Local linear regressions, i.e. using a 1st degree polynomial. This specification is including no controls to show that these are not driving our main result. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors).

Figure A19: Estimation plots, controlling for longitude and latitude



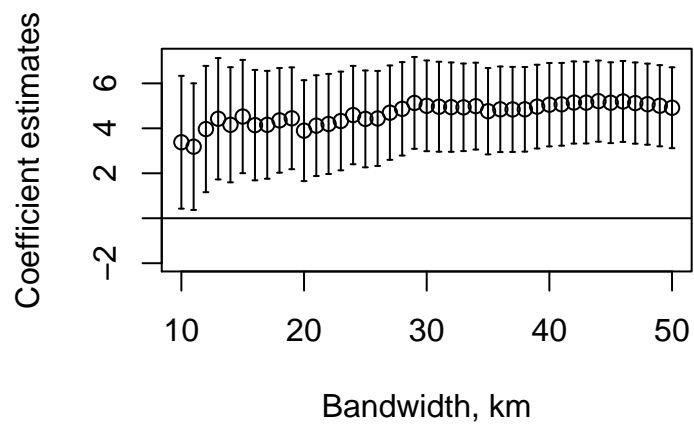
(a) Referendum 1992



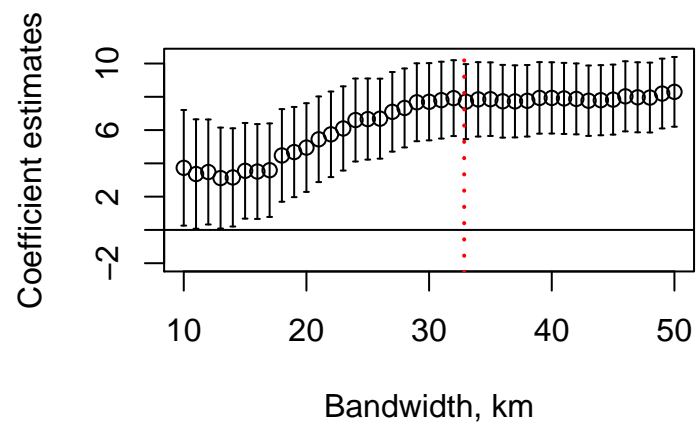
(b) Referendum 2005

Notes: Estimates of treatment effect, bandwidth of 10 to 50 kilometres, within Lorraine. These specifications are in addition controlling for longitude and latitude. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors). As the graphs clearly show that the results are not substantially altered by the inclusion.

Figure A20: Estimation plots, controlling for longitude, latitude and their interaction



(a) Referendum 1992

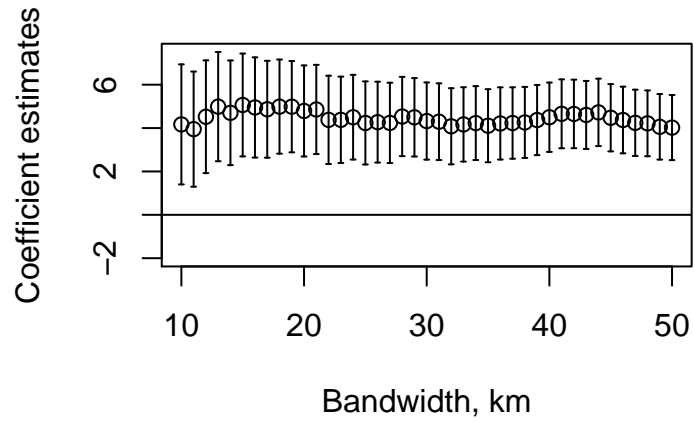


(b) Referendum 2005

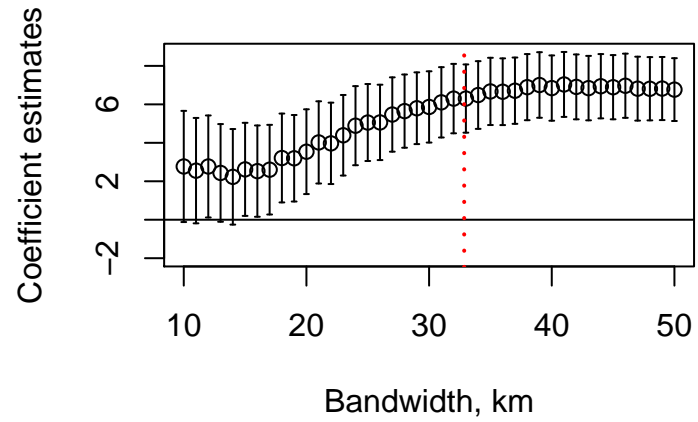
Notes: Estimates of treatment effect, bandwidth of 10 to 50 kilometres, within Lorraine, controlling for longitude, latitude and their interaction. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors). It is debated whether these controls should be included in these kind of regressions, but as the graphs clearly show our results are not depending on it.



Figure A21: Estimation plots for 1992 and 2005 referenda, controlling for border segments



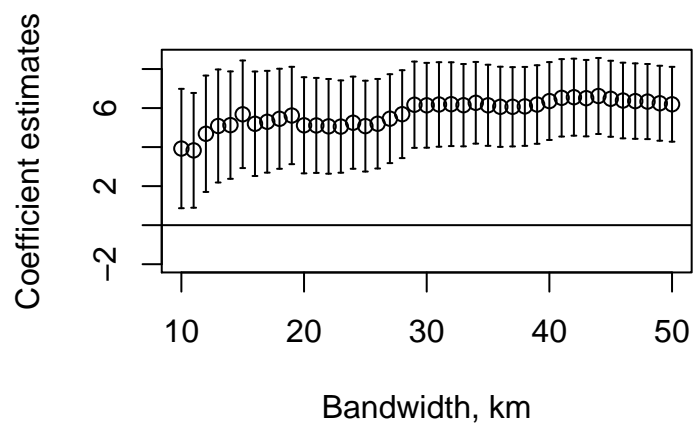
(a) Referendum 1992



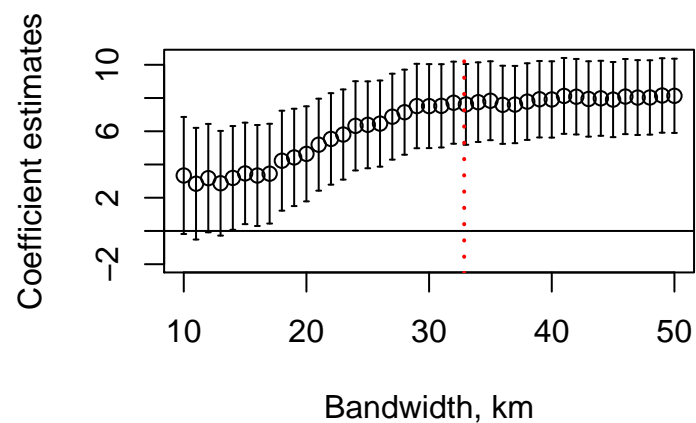
(b) Referendum 2005

Notes: Estimates of treatment effect, bandwidth of 10 to 50 kilometres, all of Alsace and Lorraine, controlling for north, mid, and south border segments. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors). It is an ongoing debate whether this type of control should be included in this type of regression, but as the graphs clearly show our results are not affected by this.

Figure A22: Estimation plots for 1992 and 2005 referenda, controlling for border segments (within Lorraine)



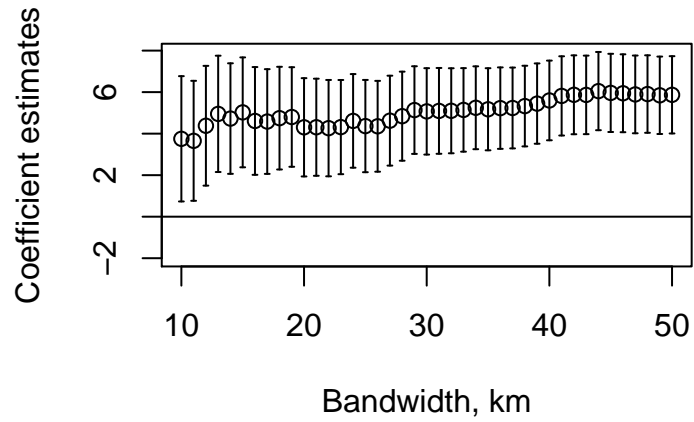
(a) Referendum 1992



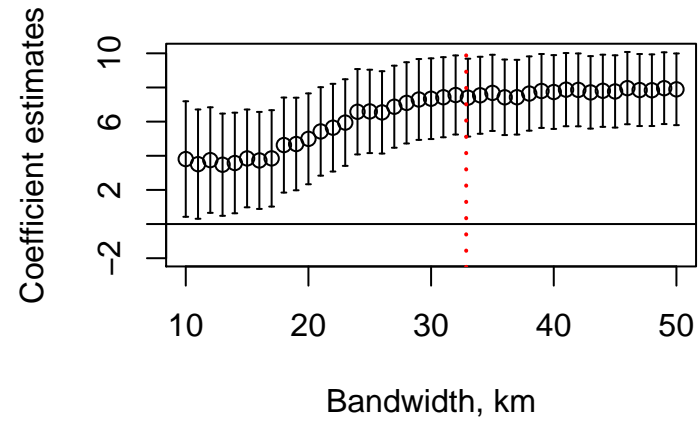
(b) Referendum 2005

*Notes:* Estimates of treatment effect, bandwidth of 10 to 50 kilometres, within Lorraine, controlling for north, mid, and south border segments. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors). It is debated whether these controls should be included in these kind of regressions, but as the graphs clearly show our results are not depending on it.

Figure A23: Estimation plots, controlling for distance to language border



(a) Referendum 1992

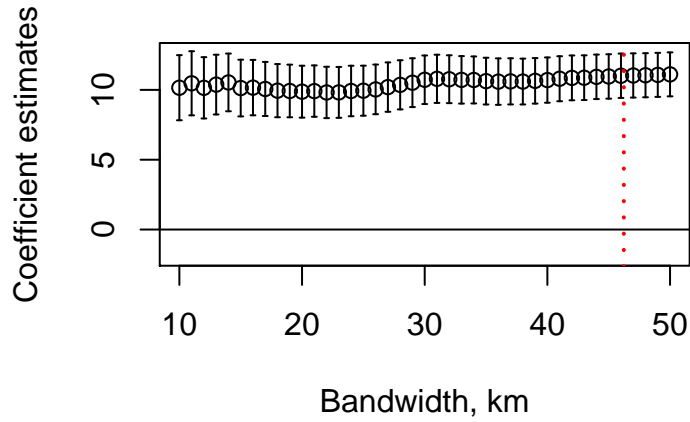


(b) Referendum 2005

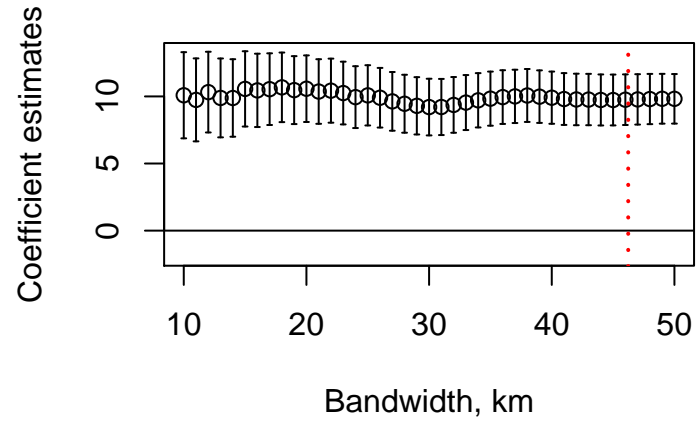
*Notes:* Estimates of treatment effect, bandwidth of 10 to 50 kilometres, within Lorraine, controlling for distance to the historical language border. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors). In addition to omitting municipalities that were formerly German-speaking, this is an additional test that our results are not driven by linguistic differences. It is also an indication that the border within Lorraine was truly exogenous to our outcome (and not endogenous to pre-existing linguistic differences) as the coefficients are barely affected by including the distance.

### Estimation Plots: regional newspaper subscription shares

Figure A24: Estimation plots, newspaper subscription shares, 1st and 2nd degree polynomial



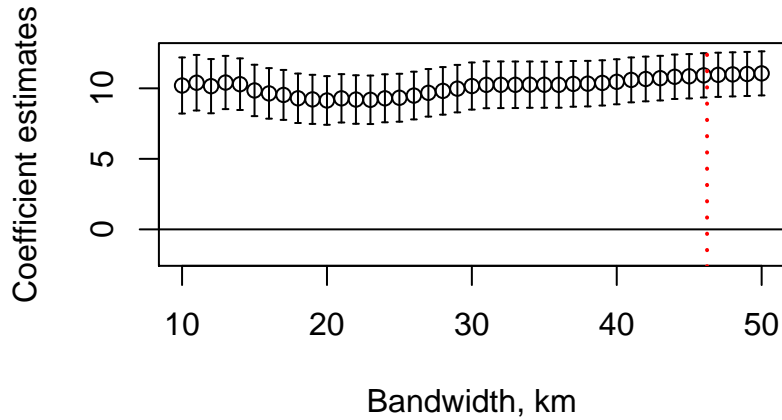
(a) 1st degree polynomial



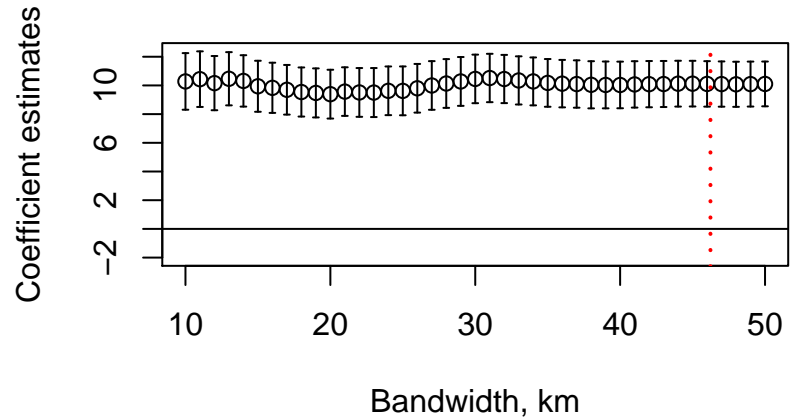
(b) 2nd degree polynomial

*Notes:* Estimates of treatment effect, bandwidth of 10 to 50 kilometres, within Lorraine. These regressions are based on a 1st degree polynomial (a) and 2nd degree polynomial (b). Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors). Our preferred specification chooses a very small bandwidth, and the local linear regression design. These graphs show that for larger bandwidths we get comparable results using higher order polynomials. The coefficient estimates are similar and results become significant with larger bandwidths at conventional levels.

Figure A25: Estimation plots, newspaper subscription shares, controlling for longitude, latitude and their interaction



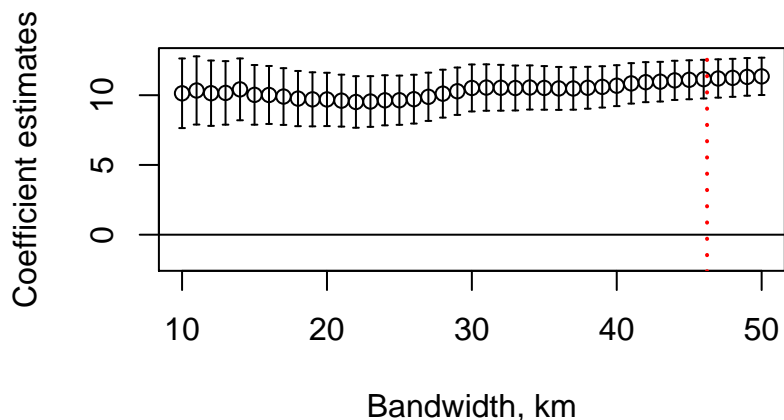
(a) Controlling for longitude and latitude



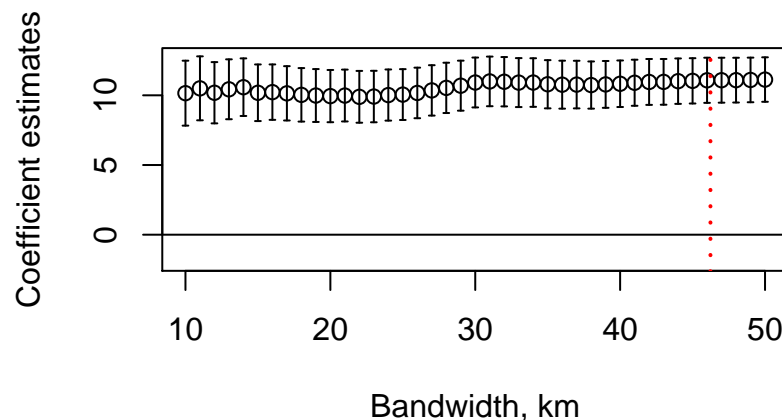
(b) Controlling for longitude and latitude, and interaction

*Notes:* Estimates of treatment effect, bandwidth of 10 to 50 kilometres, within Lorraine, controlling for longitude, latitude and their interaction. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors). It is debated whether these controls should be included in these kind of regressions, but as the graphs clearly show our results are not depending on it.

Figure A26: Estimation plots, newspaper subscription shares, controlling for distance to language border



(a) No controls

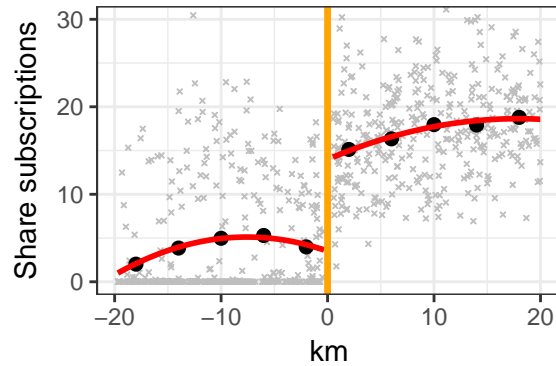


(b) Controlling for distance to language border

*Notes:* Estimates of treatment effect, bandwidth of 10 to 50 kilometres, within Lorraine, controlling for distance to the former/historical language border. Dashed vertical line at the optimal IK bandwidth. Solid vertical lines represent 90% confidence intervals (based on Conley standard errors). In addition to omitting municipalities that were formerly German-speaking, this is an additional test that our results are not driven by linguistic differences. It is also an indication that the border within Lorraine was truly exogenous to our outcome (and not endogenous to pre-existing linguistic differences) as the coefficients are barely affected by including the distance.

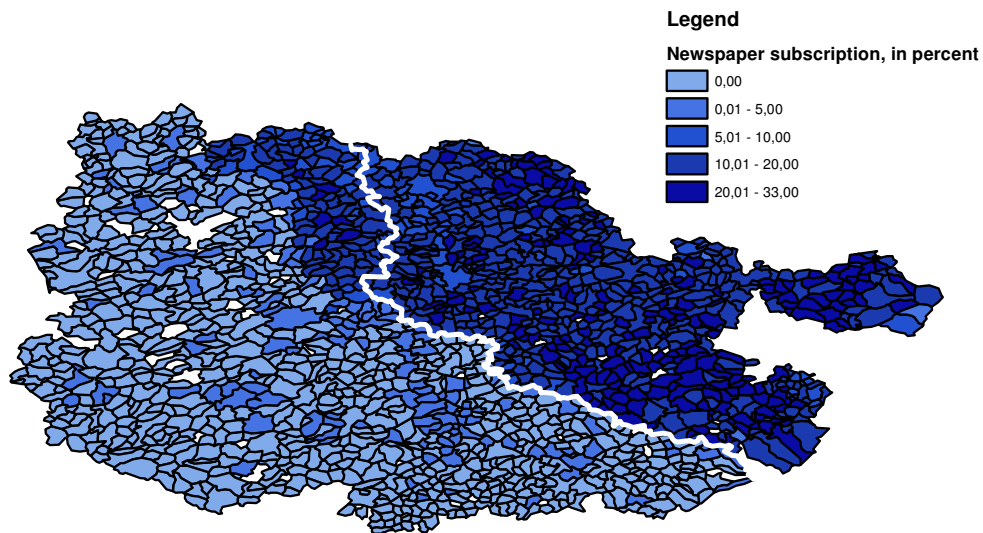
## RD Plots and maps for regional newspaper subscription

Figure A27: RD plot, share of households with subscription of “Le Republicain Lorraine”, 2nd degree polynomial



Notes: RD plots using only municipalities within Lorraine. Fitted line based on 2st degree polynomial.

Figure A28: Newspaper subscription shares

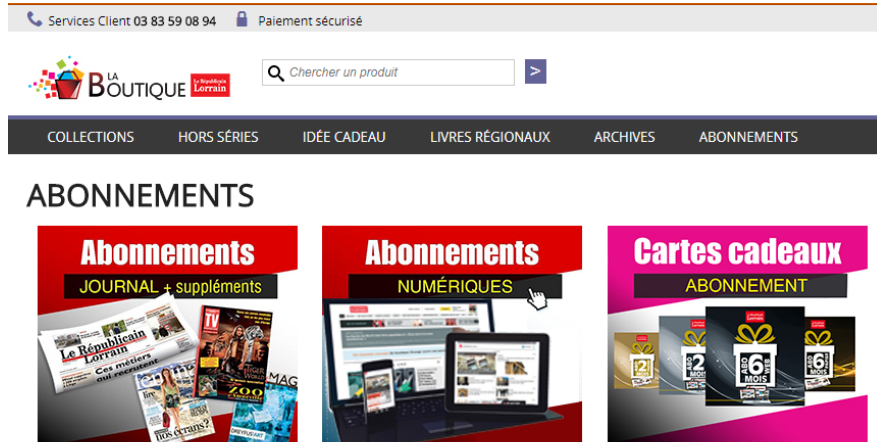


Notes: Municipal level averages share of newspapers subscribers to *Le Republicain Lorraine* within Lorraine. The white solid line indicates the treatment border that divided the region. The treated area is on the right hand side of the white line. White municipality polygons indicate missing data. Darker colors reflect higher shares, and indicate a higher regional identity.

# Le Republicain Lorraine



Figure A29: Subscription page Le Republicain Lorraine (1)



Notes: This is from the subscription page of the newspaper. We use the number of all subscriptions, but our source suggested that almost all subscriptions were still print subscriptions in 2014.

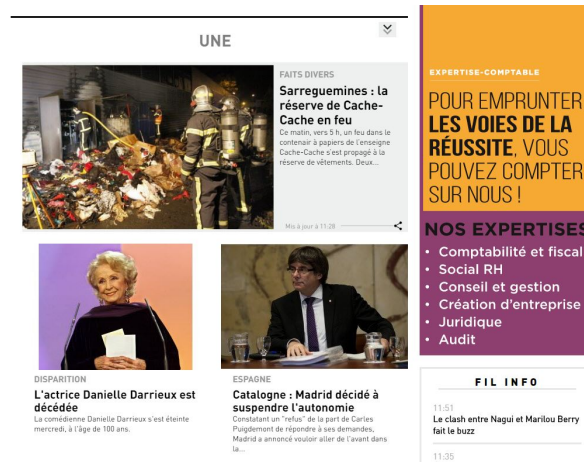
Figure A30: Subscription page Le Republicain Lorraine (2)



Notes: This is from the subscription page of the newspaper. We use the number of all subscriptions, but our source suggested that almost all subscriptions were still print subscriptions in 2014.

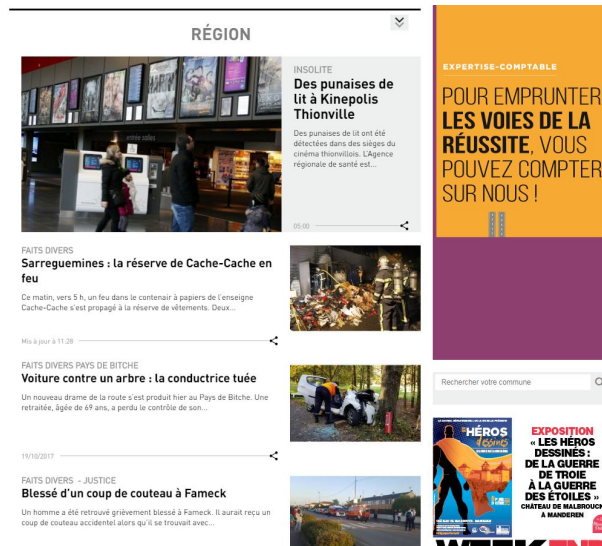


Figure A31: Homepage (main) Le Republicain Lorraine



Notes: This screenshot shows a random example of the main news contained in the newspaper (Date: 2017.19.10).

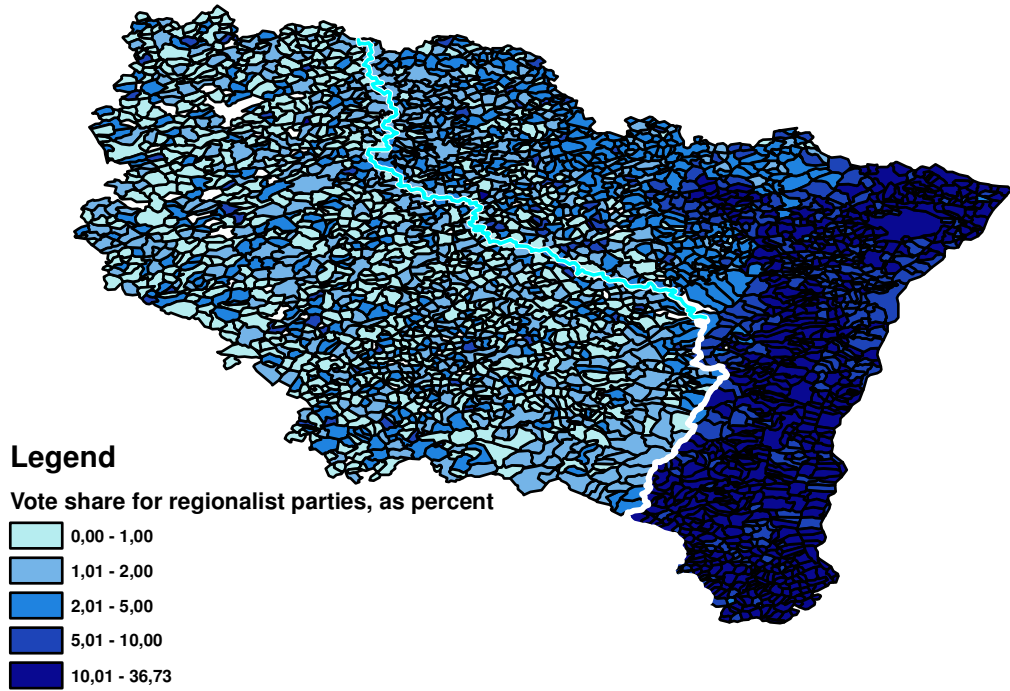
Figure A32: Homepage (regional) Le Republicain Lorraine



Notes: This screenshot shows an example of the regional news contained in the newspaper (Date: 19.10.2017).

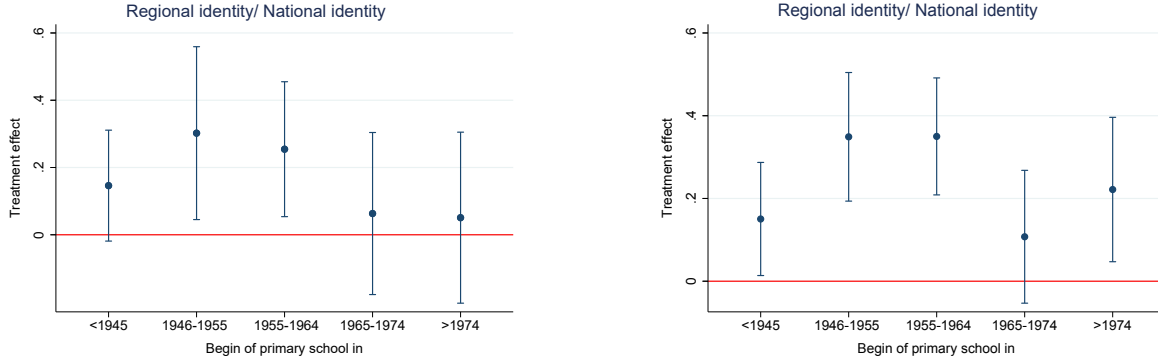
## Regionalist parties

Figure A33: Vote shares of regionalist parties



*Notes:* Municipal level vote shares for the list “Non à l’ACAL, Oui à nos régions!” in the 2015 regional elections. The list comprised of the parties “Unser Land”, “Parti des Mosellans”, and “Parti Lorrain”. The white solid line represents the treatment border formerly dividing the region. Darker colors reflect higher shares, and indicate a higher regional identity.

Figure A34: Identity differences by age cohort, relative to National identity



(a) Treatment effect in Lorraine

(b) Treatment effect in Alsace and Lorraine

Notes: The treatment effects refer to the parameter  $\Delta$  which is part of the equation:

$y_{ig} = \pi + \sum_g \Delta_g \times Age_g \times Treatment_{ig} + \Gamma'_i \lambda + \eta_{ig}$ , where  $Treatment_{ig} = \mathbf{1}$ [individual in treated region] and  $\Gamma$  comprises controls for (reported) age, employment status and sex.  $g$  indicates to which age cohort an individual belongs, the group of untreated participants act as the baseline category. Age cohorts are selected such that the second group started schooling after the end of the treatment and the end of WWII. A positive  $\Delta$  indicates that people in the treated region exhibit a higher value compared to the control area. Sources are the Observatoire Interrégional du Politique (OIP) 1999 and 2001.

Table A30: Survey results: policy preferences

Survey question	Mean, control	$\Delta$	P-value	No. obs.
Democracy works well in France	2.536	-0.023	0.616	1316
Democracy works well within region	2.630	0.111	0.008	1290
Well informed about regional policies	2.704	0.089	0.021	1308
In favor: transfer policy competence to region (avg. 10)	3.031	0.092	0.005	605
In favor: allow more autonomy at reg. level (avg. 5)	2.134	0.108	0.025	1315
Educ. policy should be set at reg. level (avg. 5)	2.855	0.112	0.024	574
Concerned reg. admin. would increase interreg. inequality	3.208	-0.172	0.037	574

Notes: Sources are the Observatoire Interrégional du Politique (OIP) 1999, 2001, and 2003, using respondents within Lorraine, on département level. The paper shows similar results for Alsace and Lorraine. The parameter  $\Delta$  comes from the equation:  $y_i = \pi + \Delta Treatment_i + \Gamma'_i \lambda + \eta_i$ , where  $Treatment_i = \mathbf{1}$ [individual in treated region] and  $\Gamma$  comprises of controls for (reported) age, employment status and sex. A positive  $\Delta$  indicates that people in the treated region agree more with the statement. Avg. "x" indicates that the factor is composed of "x" underlying survey items.

## Additional References

- ALESINA, A. AND E. SPOLAORE (1997): "On the Number and Size of Nations," *The Quarterly Journal of Economics*, 112, 1027–1056.
- ANDERSON, M. (1972): "Regional Identity and Political Change: The Case of Alsace from the Third to the Fifth Republic," *Political Studies*, 20, 17–30.
- BOOMGAARDEN, H. G. AND A. FREIRE (2009): "Religion and Euroscepticism: Direct, Indirect or no Effects?" *West European Politics*, 32, 1240–1265.
- BOSWELL, L. (2008): "Should France be ashamed of its history? Coming to terms with the past in France and its eastern Borderlands," *Totalitarian Movements and Political Religions*, 9, 237–251.
- CALLENDER, H. (1927): "Alsace-Lorraine Since the War," *Foreign Affairs*, 5, 427–437.
- CARROL, A. (2010): "Socialism and National Identity in Alsace from Reichsland to République, 1890-1921," *European History Quarterly*, 40, 57–78.
- CARROL, A. AND L. ZANOUN (2011): "The View from the Border: A Comparative Study of Autonomism in Alsace and the Moselle, 1918–29," *European Review of History*, 18, 465–486.
- COLLINS, J. (2007): "Beyond history: Memory and justice at oradour-sur-glane," *Contemporary French Civilization*, 31, 203–230.
- DE GAULLE, C. (1969): "Entretien avec le journaliste Michel Droit, avant le référendum sur la création des régions et la réforme du Sénat fixé au 27 avril," <https://fresques.ina.fr/de-gaulle/fiche-media/Gaulle00282/entretien-avec-michel-droit.html>.
- DE LA VALETTE, J. (1925): "The Absorption of Alsace and Lorraine," *Journal of the British Institute of International Affairs*, 4, 30–46.
- GOODFELLOW, S. (1993): "From Germany to France? Interwar Alsatian National Identity," *French History*, 7, 450–471.
- GRASSER, J.-P. (1998): *Une Histoire de l'Alsace*, vol. 1, Éditions Jean-Paul Gisserot.
- HARP, S. L. (1998): *Learning to be Loyal: Primary Schooling as Nation Building in Alsace and Lorraine, 1850-1940.*, De Kalb: Northern Illinois University Press.
- HARVEY, D. A. (1999): "Lost Children or Enemy Aliens? Classifying the Population of Alsace After the First World War," *Journal of Contemporary History*, 34, 537–554.
- HEFFERNAN, M. (2001): "History, Geography and the French National Space: The Question of Alsace-Lorraine, 1914-18," *Space and Polity*, 5, 27–48.
- HÖPEL, T. (2012): "The French-German Borderlands: Borderlands and Nation-building in the 19th and 20th Centuries," *European History Online (EGO)*.
- LANCELOT, M.-T. AND A. LANCELOT (1970): "Atlas des élections françaises de 1968 et 1969," *Revue française de science politique*, 312–328.
- LELEU, C. (1976): "The French Referendum of April 23, 1972," *European Journal of Political Research*, 4, 25–46.

- MINKENBERG, M. (2009): "Religion and Euroscepticism: Cleavages, Religious Parties and Churches in EU member States," *West European Politics*, 32, 1190–1211.
- SILVERMAN, D. P. (1966): "Political Catholicism and Social Democracy in Alsace-Lorraine, 1871-1914," *The Catholic Historical Review*, 52, 39–65.
- ZANOUN, L. (2009): "Interwar Politics in a French Border Region: the Moselle in the Period of the Popular Front, 1934-1938," Ph.D. thesis, The London School of Economics and Political Science (LSE).

# 1 Introduction

The formation of a common group identity at the regional, ethnic or country level is a highly important, yet poorly understood aspect of human behavior. One reason for the difficulty to disentangle the factors influencing the identity formation process is that laboratory experiments can only study groups of limited size for a short time period and have to rely on artificial manipulations. For instance, it is hard to emulate violent repression, even though occupations and changes in nation status occurred frequently in history, and are often associated with the suppression of existing local identities. Observational studies, in contrast, can exploit historical events, but almost always struggle with distinguishing the effect of a shock or a set of policies from other aspects that are specific to a certain region.

Historical examples include, for instance, South Tyrol, which was Austrian but was annexed by Italy after WW1, followed by initial bans on the German language from public service and teaching, and the censoring of regional newspapers. Other European administrative regions with a strong regional identity and tensions with national policies include Catalonia and the Basque country in Spain. More violent examples of homogenization policies and repressive policies today are Chechnya and Crimea with their mixed populations in Russia, the Kurdish region in Turkey, as well as Tibet and the Uighurs in China.<sup>1</sup> All those regions were or are exposed to a shock composed of repression and the suppression of group identities, often along with attempts to impose a national identity. The econometric challenge is that in most cases there is no suitable counter-factual to assess the impact of such a shock. This paper provides causal evidence by using a natural experiment that divided historically homogeneous regions in a quasi-exogenous way. The setting allows us to compare two parts of a region that differ only in the prior degree of exposure to repressive policies, but can be observed in the same institutional environment today.

Group identities matter both economically and politically (see [Kranton \(2016\)](#) for an overview). The insufficient alignment of identities fuels separatism in regions like Catalonia, Flanders and Scotland. Arbitrarily determined national borders are associated with strong ethnic identities and weak common national identities in Africa, often related to conflict, violent struggles for autonomy and inferior development (e.g., [Besley and Reynal-Querol, 2014](#); [Michalopoulos and Papaioannou, 2014, 2016](#); [Rohner et al., 2013](#)). At the same time, there are culturally seemingly heterogeneous countries like Switzerland or the United States, which exhibit a strong sense of common identity. Secessionism and separatist conflict can be driven by economic factors ([Gehring and Schneider, 2016](#)) and cultural differences (modeled as preference heterogeneity in [Alesina and Spolaore, 1997](#)). We aim to better understand the sources of existing cultural differences, more specifically the perceived alignment of preferences exemplified in a common regional or national identity.

Although it is widely believed that historical shocks and state policies are crucial in explaining identity formation, conclusive causal evidence is lacking. [Tilly \(1975\)](#), for instance, emphasized war

---

<sup>1</sup> The Polish regions of Silesia and Bohemia, as well as Kaliningrad and Danzig originally featured a strong influence of German culture, which the central government tried to eliminate after WW2. Scania in Sweden was once Danish, and still has a distinct regional identity. Selected Sources can be found in [E](#).

Figure 1: Geographical location of the treated and untreated area



*Notes:* The map shows the division of Alsace and Lorraine after 1871. The treated area is shaded in light grey, and the untreated control area in dark grey. Both areas are again French today. Current national borders in bold black.

and conflict as a source of state formation, but focused on the cooperation between leaders and capital to form a state. Nonetheless, even if group identity and its strength is influenced by shared history and experiences (Depetris-Chauvin et al., 2018) and shared ethnic or social traits, this is neither necessary nor sufficient to form a stable group identity. Recent evidence shows that heterogeneity within groups is on average much larger than heterogeneity between groups (Desmet et al., 2017), but we still observe strong existing group identities. Social psychology stresses subjective perceptions; the importance of group members sharing the collective perception of belonging to a joint group (Turner, 1982), created by emphasizing factors that are common to the group. To incorporate these insights, we build on Shayo (2009) and model common identity as the degree to which an individual perceives her preferences, values and norms to be aligned with her region, nation, or other groups.

We then study how the division of the initially homogeneous border regions Alsace and Lorraine between France and Germany following the Franco-Prussian War in 1870-71 affected the formation of identity in the occupied (treated) area compared to the untreated area over the medium and long term. Figure 1 shows the treated and untreated area. People in the treated area experienced a change in nation status from French to German, and were over more than half a decade exposed to repression and the suppression of their regional identity. Historians postulate that the strong regional identity in the treated area is a plausible reaction to intrusive homogenization policies enforced by the German and the French government between 1870 and the 1950s (Carrol and Zanoun, 2011; Höpel, 2012; Rothenberger, 1975).

We gather and process data from a large variety of sources to show the evolution of identity over a sufficiently long time period. First, we make use of the Cahiers de doléances from 1789 to verify that there are no pre-treatment differences in regional and national identity. The absence of discontinuities in geographical data supports our arguments about the exogeneity of the border. We then use a first referendum about higher regional autonomy in 1969 as a proxy revealing differences in regional identity. To document the persistence of the effect, we continue by using referenda in 1992 and 2005, which would also have given regions more political decision-making power. We augment

these measures of revealed preferences with survey evidence from 1999-2003 stating the strength of regional and national identity. All results show that regional identity has become persistently stronger in the treated area.

At the same time, we find no differences in national identity in the survey or with regard to nationalism in the RDD. We provide a theoretical model of identity transmission (Online Appendix A) to explain how a temporary historical shock can lead to persistent differences in regional, but not necessarily in national identity. We rule out alternative explanations by considering, for instance, post-treatment differences in socio-economic characteristics at the treatment border, accounting for the language border between Germanic and French dialect, as well as using placebo tests, controlling for distances to neighboring countries and removing large urban agglomerations.

Turning to mechanisms, an important question is whether the measured difference is solely due to a psychological shock that changed preferences and influences the transmission of regional identity to new generations through unobservable aspects of raising children or communication in the society. While this is entirely plausible, it would be reassuring to have measurable evidence of potential transmission mechanisms. We provide evidence for two of those channels. Already during the treatment period, regionalist organizations and newspapers emerged, and regionalist parties were politically dominant. Regionalist parties suffered a huge blow in Lorraine after being associated with Nazi Germany in World War II, but are still significantly stronger in the treated area, in particular in Alsace.

Moreover, we got access to internal company data about subscription rates to a regional newspaper. Subscribing to a regional newspaper is an informative proxy variable, as it is both a signal of regional attachment as well as a source of information about regional culture for the subscribers and their children. We find that subscription rates are significantly higher in the treated area within Lorraine. Furthermore, distinguishing the prior survey results by age cohort shows that the differences are strongest for the 2 age cohorts beginning primary school after World War II. They remain stronger when incorporating Alsace in the analysis, which could be related to the stronger current presence of regionalist parties. Finally, we show that those differences in regional identity have important policy implications. People in the treated area show a significantly stronger preference for more regional decision-making in many dimensions including schooling.

Our research adds and relates to different strands of literature. First, the literature on identity economics (e.g, Akerlof and Kranton, 2000; Bordalo et al., 2016; Kranton, 2016) and on the persistence and transmission of culture, identities and values (e.g, Bisin and Verdier, 2000, 2010; Gennaioli and Rainer, 2007; Guiso et al., 2016; Giuliano and Nunn, 2016; Nunn and Wantchekon, 2011; Voigtländer and Voth, 2012 and Tabellini et al., 2008). Most existing models consider the case of two groups, a minority and majority group, and the choice whether to transmit certain values to the next generation via parental investment. The minority group in our setting is the treated area as opposed to first the German and then French majority, who both try to assimilate them by force. Bisin et al. (2011) explicitly model a mechanism that can explain how oppositional identities can persist and Fouka (2016) provides a model how both vertical (parental investment) and horizontal



(schooling) socialization influence the strength and transmission of a group identity. Our results can be interpreted as reflecting both mechanisms, as we also document how a “discriminated” group strengthens their identity as a response.

In addition, there is a large literature on identity in different disciplines of social science, ranging from political science to sociology and social psychology. It is widely accepted that a common identity needs not to be based on objectively aligned preferences, but that the collective perception of social unity can be sufficient to form a group (Turner, 1982). This is also the base of the identity definition in Shayo (2009), which we adapt. It can account for strong group identities despite large preference heterogeneity within groups (Desmet et al., 2017). In social psychology, the social identity model (Tajfel et al., 1971, p.16 & 27) argues that group identity “has primarily a perceptual or cognitive basis” and that “awareness of a common category membership” is a necessary and sufficient condition for individuals to act as a group. It seems plausible that the intrusive assimilation policies strengthened the awareness of Alsations and Lorrainians of their cultural distinctiveness and led to an “alienation” of the affected citizens (Goodfellow, 1993, p.454).

Leed (1981) argues that fighting together against a common enemy in a conflict induces people to form a common identity, by increasing the perceived importance of connecting experiences and traits. In the case of AL, a plausible explanation is that the exposure to intrusive and discriminating policies creates an incentive for parents to invest in teaching regional culture to their children, which persistently increases the salience of attributes common to the inhabitants of the region. The idea that feeling rejected or suppressed by a majority increases group identification also relates to the rejection-identification hypothesis in social psychology (Branscombe et al., 1999). It argues that the perceived common identity between an individual and a group, can be changed not only by changing actual norms or preferences, but also by adapting the importance that an individual assigns to different attributes. For instance, Depetris-Chauvin et al. (2018) show that the success of a common national team can increase national identity in Africa in the short term, arguably without changing actual between-group differences.

We also relate to an emerging literature in economics examining the use and effect of different policies on identity formation and nation building. Alesina and Reich (2018) model when and which assimilation policies are used to instill a common identity, creating the distinction between benevolent and intrusive (“odious”) policies. Our results are in line with some existing evidence of how intrusive policies can backfire and increase the affected group’s identity. Dell and Querubin (2017) use exogenous variation in US bombing patterns in Vietnam, and document that more bombing increased communist military activities, lowered civic engagement and worsened attitudes towards the central government and the US. Carvalho (2013) suggests that banning veils on Muslim women can actually lead to higher religiosity, hence a stronger religious identity.

The long run persistence of the treatment effect in our setting - about half a century - is not unusual and in line with other papers documenting persistence in culture over periods stretching more than a century. These differences are, for instance, associated with outcomes like stated preferences regarding trust (Becker et al., 2015) and different proxies of civic capital (Guiso et al., 2016), but

also with revealed preferences like cheating in a trust game (Lowes et al., 2017), following traditional practices (Giuliano and Nunn, 2016), and differences in homicide rates among Scottish-Irish settlers in the US South (Grosjean, 2014).

There is also a related strand of literature studying schooling as a specific mechanism through which the state can influence identity formation (e.g., Bandiera et al., 2017; Lott, 1999; Ortega and Tangerås, 2008). Carvalho and Koyama (2016) provide a model of how an education system that marginalizes a certain identity can cause cultural resistance on part of the marginalized group. Regarding empirical papers, studies of compulsory language laws in schools are closely related in many respects. Aspachs-Bracons et al. (2008) and Clots-Figueras and Masella (2013) find that within Catalonia, the forced imposition of Catalan is related to an increase in Catalan identity measured by various proxies.

The study that is closest and complementary to ours is Fouka (2016), who provides evidence on the forced imposition of the English language on German pupils in some US states after WWI. As in our study, this shock leads to an increase in German identity, measured via name choices, inter-group marriages and volunteering rates for the army. Her paper is able to carve out the short to medium-term effects of a particular set of policies on migrants, a selected subgroup of the population. Our study, in contrast, is able to measure the medium and long-term effects of occupation and repression on the group identity of the suppressed group in their home region.<sup>2</sup>

The paper is structured as follows. Section 2 explains the historical background of Alsace and Lorraine, and describes our theoretical framework. Section 3 discusses the data and identification strategy, whilst Section 4 presents the main results. Section 5 discusses mechanisms, persistence and policy implications, and Section 6 discusses potential threats to identification. Section 7 concludes the paper.

## 2 Historical background and theoretical framework

### 2.1 History of Alsace and Lorraine: Division, borders and homogenization policies

John Stuart Mill stated that a certain degree of homogeneity is necessary as “unassimilated democratic states will tend to dissolve into as many democracies as there are nations within them” (cited by Conversi, 2004, p.35). Gellner and Breuilly (2008) argue that in an industrial society, different ethnicities, cultures, and in particular languages, act as barriers that reduce efficiency by increasing the costs of communication. France, a country featuring historically diverse regions with distinct languages or dialects, is a well-suited place to study attempts to overcome such barriers. It is nearly

---

<sup>2</sup> A plausible explanation why our two studies find different effects than Clots-Figueras and Masella (2013) on imposing Catalan is that this was not perceived as oppositional to the identity of migrants in Catalonia, whereas in the US case and in our setting, policies were clearly perceived as discriminatory. This is in line with explaining the below-average school performance of African Americans in the US with the perception of investments in education as acting “white” and opposed to black group identity (Fryer Jr. and Torelli, 2010), while for Asian Americans no such effects are observed. Our research design compared to these papers exploits the exogenous border creation within a region, which allows us to compare people who formerly possessed the same identity as a counter-factual.

universally recognized as the birthplace of nation building (Conversi, 2008), with Napoleon probably being the first to systematically attempt enforcing a national identity (see, e.g., Hobsbawm, 1994; Conversi, 2004). France thus serves as a prime example of the formation of a group identity.

Homogenization policies to build a common identity are still a highly relevant issue in many modern states today, as Lott (1999) shows in a cross-country and the specific South African context. Johan Gottfried Herder (1724-1804) was among the first to emphasize the importance of a common language to create a common national identity, which was later also highlighted by Gellner and Breuilly (2008). Generally, homogenization policies can include more benevolent measures like lowering the costs of travel and exchange through institutions and improved infrastructure, but also the imposition of a state religion, the prohibition of regional cultures and in extreme cases genocide and the extermination of certain groups (e.g., Tilly, 1975). Conversi (2008, p.1289) describes the nation building process as a “top-down process entailing assimilation and the forced erosion of cultural differences”, which can lead to existing ethnic and regional identities being perceived as oppositional to national identity.

To put our natural experiment into perspective, it is shortly helpful to discuss some important aspects of the history of Alsace and Lorraine. Both regions have been autonomous political entities as far back as the 7th century. After the Treaty of Verdun, Lorraine became a part of Middle Francia and Alsace of East Francia. Under Charles the Bald, all of modern Lorraine became a part of the Duchy of Lotharingia. Over the centuries, both regions developed strong regional identities with specific traditions and norms. After the Thirty Years’ War (1618-1648) all of Alsace and the Lorrainian cities of Metz, Verdun and Toul were ceded to France in the Treaty of Westphalia. The rest of Lorraine was given to the French Crown through the Treaty of Vienna (1738) and effectively became French in 1767. At the time of the Franco-Prussian War in 1870/71, Alsace and Lorraine had thus been a part of France for more than a century and were exposed to the same nation building policies by Napoleon and other central French leaders.

The peace treaty ending the Franco-Prussian War (July 19, 1870 to May 10, 1871) then stipulated that large parts of Alsace and the eastern part of Lorraine were ceded to the newly created German state. Our identification exploits that disagreements between German chancellor Bismarck and his military leaders as well as emperor Wilhelm I., and the complex negotiation process with France resulted in a quasi-random final border demarcation that was exogenous to our outcome (see Figure 2b).<sup>3</sup> The cautious statesman Bismarck wanted to restrain territorial expansion to the alemannic-dialect speaking parts of Alsace and Lorraine (Lipgens, 1964), while the military, led by the charismatic General von Moltke, wanted to extend the German territory as far as possible.<sup>4</sup>

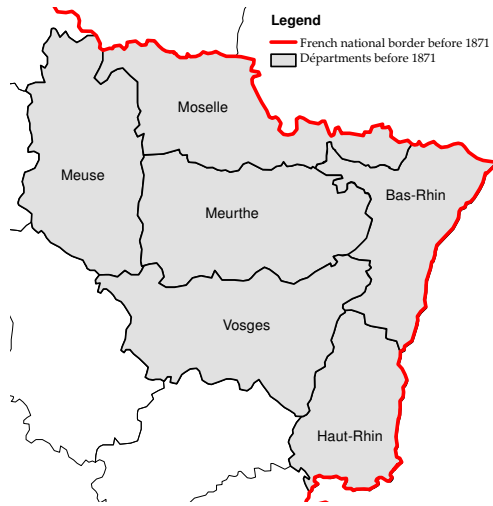
These conflicting interests on the German side and the intense negotiations with the French

---

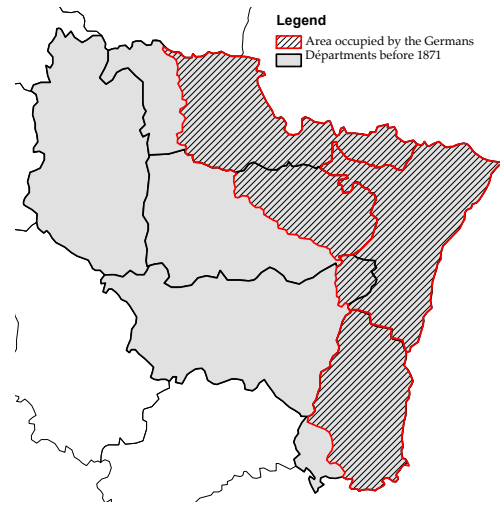
<sup>3</sup> There were strategic considerations involved, mostly regarding certain fortresses or positions like Belfort. The strategic importance of locations could be related to geological features, but, as we show, there are no indications that they are linked to prior differences the relative strength of regional compared to national identity and no discontinuities in ruggedness or elevation.

<sup>4</sup> The literature indicates that von Moltke had from the onset of the war on planned to march as far into France as possible (Förster, 1990), with the aim of territorial gains and as defensive measures to weaken the arch-enemy in anticipation of the plausible next conflict. Bismarck on the other hand feared that excessive annexations might increase the risk of a new conflict.

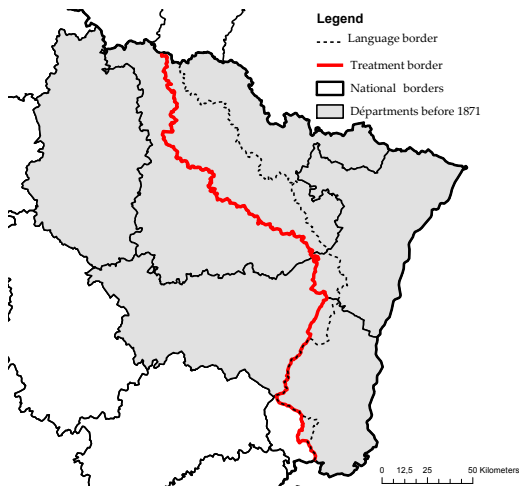
Figure 2: Historical maps: before, during and after German occupation



(a) France before German occupation (1870)



(b) France after German occupation (1871)



(c) France after German occupation (1871) and former language border.



(d) France after the occupied region was returned (1918 - today)

*Notes:* Moselle is the treated part of Lorraine, Meuse and Meurte-et-Moselle are the untreated parts. Bas-Rhin and Haut-Rhin compose Alsace, and Vosges serves as their counterfactual.

leader Adolphe Thiers resulted in the compromise splitting Alsace, and in particular Lorraine, rather arbitrarily (Förster, 1990; Lipgens, 1964; Messerschmidt, 1975). As an example of the complex nature of these negotiations, Bismarck was willing to “save Metz for France”, and considered keeping the French part of Lorraine altogether a “folly of the first order” (Wawro, 2005 p.206). Moltke and Wilhelm I., however, refused to return it, as the military considered taking Metz one of their great achievements and a return a “national humiliation”. Moreover, Thiers succeeded in stretching the border a little further towards Germany by offering the German military to conduct a victory

parade through the Champs Elysees in Paris.<sup>5</sup> Maps from previous centuries show that there was no apparent overlap between any historical borders and the border we use (see Figure 2 and Figures A3 - A6 for older historical borders). We also use historical evidence from 1789 to show that there were no apparent pre-treatment differences in regional identity.

The annexed area was incorporated into the German Empire as the *Reichsland Elsass-Lothringen*. In Alsace, the départements already in place during French rule were converted into the German districts of *Oberelsass* and *Unterelsass*, corresponding to the former (and existing) départements Haut-Rhin and Bas-Rhin, respectively. In Lorraine, the district *Lothringen* was created from parts of the former départements *Moselle* and *Meurthe*, and corresponds to today's département *Moselle* (see Figure 2). Due to the suspicion about the loyalty of the new citizens, the treated area was never recognized as an integrated part of the German Empire – instead it was an imperial territory under the direct authority of Kaiser Wilhelm I. (Carrol and Zanoun, 2011). France regained control of the lost provinces after the Treaty of Versailles (1919), which it kept with the short exemption of WWII, when both areas together with other parts of France were occupied by Germany. The homogenization policies aimed at realigning the preferences and values of the lost citizens into France are sometimes described as even more repressive than the German ones (Anderson, 1972; Harvey, 1999).

Hence, the "treatment" we examine is the exposure to repressive homogenization policies and the suppression of group identity associated with occupation. This does not allow us to trace the treatment effects back to any particular policy, or distinguish exactly what share was caused by French and German policies. Nonetheless, we think it is informative to study such a composite treatment. In almost all comparable historical, current and potential future cases, some of which like South Tyrole or the Kurds in Turkey we describe above, measures against the regional population were not restricted to one particular policy and go hand in hand with tensions against a central state. Moreover, historians emphasize homogenization policies as the crucial aspect of the treatment period in Alsace-Lorraine (Goodfellow, 1993).

To enable the reader to better understand the full range of those measures, Table 1 presents examples of the repressive homogenization policies divided into five categories. *Language* policies, aiming to oust local languages and foster the use of the national language; *Media* policies, restricting the freedom of press; *Social, political, military freedom, and equality* policies, aiming to restrict political rights, participation, socio-regional gatherings and the choice to serve in the military; *Separation and segregation* policies, aiming to separate or segregate locals according to origin or nationality; and *Regional institutions and administrative personnel*, aiming at replacing regional

---

<sup>5</sup> After elections in both French and German-occupied parts of France lead to the anti-war conservative party winning 500 out of 676 seats, their leader Adolphe Thiers negotiated with Bismarck for 5 days. The result was in its details unpredictable and the planned border changed frequently during the negotiation process. For example, Bismarck was willing to "save Metz for France", and considered keeping the French part of Lorraine altogether a "folly of the first order" (Wawro, 2005 p.206). Moltke and the Kaiser Wilhelm I refused to return it however, as the military considered taking Metz one of their great achievements and a return a "national humiliation" (Wawro, 2005 p.206). The final result was a compromise between both positions and it is documented that, at least partly, "Bismarck, [...], quite uncharacteristically wilted under the pressure" (Wawro, 2005 p.305). The northern border thus rather arbitrarily divides the former duchy of Lorraine in two parts.

institutions and administration. Table A1 shows a comprehensive list of both German and French policies, ranging from early 1870s until the early 1950s.

Table 1: Overview of policy categories and examples (see details in Table A1)

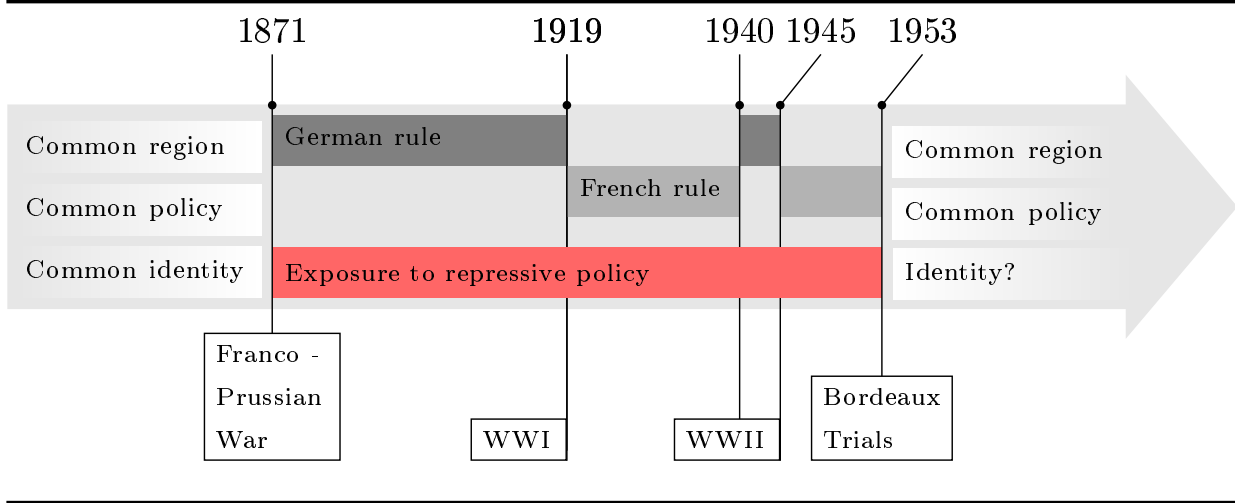
Policy category	Example
Language	1920: French becomes the only language taught in school (Grasser, 1998).
Media	1927/ 28: Banning of three autonomist journals, the "Volksstimme", the "Zukunft" and the "Wahrheit" (Goodfellow, 1993).
Social, political, military freedom, equality	1927/28: Colmar trials: 15 prominent autonomists are arrested and tried for participation in a plot to separate Alsace from France (Goodfellow, 1993).
Separation and segregation	1918: Locals are classified according to an identity-card system. Lower classification leads to, e.g. travel bans (Harvey, 1999).
Regional institutions and administrative personnel	1924: Ministerial Declaration by Premier Edouard Herriot imposes a centralized administration, French laws and intuitions (Carrol and Zanoun, 2011).

Examples of political restrictions under German rule include that the region initially no representatives in the *Bundesrat* or the *Reichstag* (Vajta, 2013). As part of the “Kulturkampf”, government regulations restricted particular types of education (Silverman, 1966) and restrictions on the press were not lifted until 1898. The government also kept the French dictatorship paragraph of 1849 in force, which allowed house searches, expelling agitators and prohibiting political organizations (Carrol, 2010). Strasbourg University was reopened as “Kaiser-Wilhelm-Universität” with the specific aim to replace regional traditions and to homogenize the annexed region (Höpel, 2012).

France regained control of the "lost provinces" after the Treaty of Versailles (1919). Among others, it was prohibited to teach in the Germanic dialect, which was the mother tongue of a majority of the population, in school, and German was removed as an official language (German as a second language was not taught in schools until the early 1950s). The families of the about 200,000 Germans who had settled in the region after 1871 were deported in order to “remove any trace of German influence” (Carrol and Zanoun, 2011, p.469).

Moreover, a special commission, called *Commissions de Triage*, was formed to ascertain the *Frenchness* of the population in the re-annexed area (Carrol and Zanoun, 2011). Municipal names, street names and family names were almost all changed to French. Between 1926 and 1930, several newspapers promoting the regional cause were forbidden, and members of regionalist parties were put into jail. France consequently replaced bureaucrats and local teachers with external bureaucrats who were not familiar with the local circumstances and traditions. Historical evidence suggest that the repression and the suppression of their regional culture contributed to the formation of a stronger regional identity (Harvey, 1999).

Figure 3: Timeline of events



## 2.2 Theoretical framework

This section introduces our definition of group identity and describes a simple model of cultural transmission with multiple identities and its predictions (Online Appendix A presents the formal model). Most existing models describe a setting where people have to choose between different, potentially oppositional, identities, but cannot hold more than one identity. Our setting requires a model where everyone possesses multiple identities, for instance, as a citizen of her municipality, region or country. An important feature of these multiple identities is that they are not necessarily substitutes, at least not perfect substitutes.

Our approach relates to the literature on the size of nations, which models common identity or the lack of it as preference heterogeneity, as well as to the literature on identity formation (Akerlof and Kranton, 2000) and oppositional identities (Bisin et al., 2011). We want to emphasize a definition of a common identity that builds on Shayo (2009), and relies on the *perceived* heterogeneity or distance to other members of a group. Hence, the common identity of an individual  $i$  and a group  $j \in \{R, N\} = \mathbf{J}$ , with  $R$  and  $N$  corresponding to *Region* and *Nation*, depends on the perceived distance to the average group member:

$$h^{i,j} = 1 - \left( \sum_{k \in K} \omega_k (p_k^i - p_k^j)^2 \right)^{1/2},$$

where  $p_k^i$  represents the preferences (or traditions, values and norms) of individual  $i$  regarding an attribute indexed  $k$ ,  $p_k^j$  represents the preferences of the average member of the region or the nation, and  $K$  is the set of all attributes. An important part of this heterogeneity function are the  $\omega_k$ , which can be understood as attention weights. Higher weights indicate that the tradition, value or norm  $k$  has a larger influence on the strength of common identity.

These weights are an important distinction compared to standard models. [Desmet et al. \(2017\)](#) use the World Value Surveys to show that within-group variation in values and preferences is larger than between-group differences. Accordingly, the fact that strong group identities (e.g., regional or ethnic) nevertheless exist is only feasible when recognizing that it is the perception of heterogeneity that matters. The intuition of this approach is easy to understand. People from a region differ in their shared history, in the spoken dialect, local cuisine or music from other regions in the country. The degree to which this affects common national identity, however, depends on how much people emphasize these differences compared to other regions.

Individuals benefit from a strong common regional identity, as it helps them to feel socially compatible with fellow group members in their region of residence. A higher perceived distance to the average group member of the region lowers individuals' regional identity and can make them feel isolated. The cost associated with isolation is not only psychological: a lack of social compatibility can also hurt business and/or employment opportunities. The same holds for a common national identity. For instance, if someone does not know how to comply with national traditions, it will be more difficult to find a job in the regional public administration (if that is controlled by the central state) or to trade with other regions.

Identity and its transmission to a child can of course be affected by a wide range of factors, including peers, regionalist associations, newspaper and parties, or public events. To illustrate how a temporary historical shock can lead to persistent differences in regional, but not necessarily in national identity, our model focus on public schooling and parents as two main factors. It relates to [Bisin et al. \(2011\)](#) to the extent that children's identity is influenced by both parents and other outside factors (in their case, peer effects, in our case, public schooling). For tractability reasons, we focus on schooling as one plausible government-led identity transmission mechanism (cf. [Lott, 1999](#)) and parenting as exemplary for private investments in identity formation. As in [Bénabou and Tirole \(2011\)](#), people care about identity and respond to threats by adjusting their identity investments.

The model helps to explain the way in which an exogenous shock on how the central state, e.g. through public schooling, treats regional identity can lead to persistent differences in identities. Every individual is a member of two groups, region and nation. People gain utility from feeling closer to their region, which is their closest reference group, but also from a common national identity with the other regions, e.g. through lower transaction costs. Identity formation is affected by public schooling, which is modeled as an exogenous decision imposed by the nation state, and by parental investment. As in [Doepke and Zilibotti \(2017\)](#), parents combine Beckerian altruism about the future economic well-being of their children with a paternalistic value assigned to their own (regional) identity. They maximize the expected utility their children derive when determining parental investment, weighting the benefits of regional and national identities against the costs of teaching the respective traditions and norms. Our model could be extended to include variable costs or the time spent on teaching, but for simplicity assume these costs as a one-time fixed cost.

The game then unfolds in three stages, resembling the historical events illustrated in [Figure 3](#). In *Stage 1* (until 1870/71), both areas are exposed to the same public schooling policy. Because



they belong to homogeneous regions, there is no reason to expect differences in parental decisions on how much to invest in transmitting regional or national traditions and norms to their children. In *Stage 2* (1871 – ~1953), people in the treated area are exposed to intrusive policies and repression, exemplified by a public schooling policy that does not teach regional culture sufficiently. If their utility from regional identity is high enough, parents choose to pay the fixed costs of learning how to transmit regional culture to their children themselves. Finally, in *Stage 3* (after ~1953), the temporary shock is over and public schooling returns to teaching regional and national culture at similar levels in both areas. However, the optimal level of transmitting regional culture through direct parental investment remains higher in the treated area if parents chose to invest the fixed costs in Stage 2 during the treatment period. In contrast, as long as parents did not have an incentive to invest in the ability to teach national traditions, the model predicts that national identity converges back to the same level after the treatment is over.

The model requires us to make one, we think plausible, assumption. If parents were taught and experienced regional traditions through their own parents, they have lower costs of transmitting those traditions to her children. For instance, if someone used to sing a traditional song or perform a dance with her parents, or cook a regional dish with them, they can transmit that knowledge and skills to their own children more easily. For simplicity, we assume that they can transmit it for free instead of facing the fixed costs, but it is only necessary that they face lower costs. Moreover, the model requires that at some point policies converged again with regard to the teaching of regional and national traditions after the treatment period is over. In fact, public schooling policy was slowly adapted after WW2 and permitted the teaching of regional culture and dialect again. The Bordeaux Trial in 1953, with the convictions of soldiers from Alsace-Lorraine who fought for the German side, can be thought of as a last event potentially reactivating memories of suppressive policies. Today, the treated area uses the same curricula and is fully accepted as a part of France.<sup>6</sup>

### 3 Data, measures, and identification strategy

#### 3.1 Data

France is divided into 22 regions, which consist of 96 départements. The départements are further divided into 323 arrondissements and 1995 cantons, but those two sub-units are of lesser importance and do not possess the status of a legal entity. We use survey results at the département level from a large scale survey, the *Observatoire Interrégional du Politique* carried out in 1999, 2001 and 2003. In contrast to all other French surveys, it offers a sufficient number of observations at the département level. For our regression discontinuity estimations, we focus on the smallest unit, which is the municipality level. Shapefiles for the 3320 municipalities in Alsace and Lorraine are from

---

<sup>6</sup> Note that the equilibrium level of national and regional identity in both areas depends on the objective functions of the parents, as well as the cost of transmitting traditions. There can be functional forms and costs, for which it is optimal to give up regional culture altogether. Also note there is one remaining difference with regard to schooling. Students in the treated area still receive a few hours of religious classes in school today. We will demonstrate that this seems orthogonal to our outcome variables.

[www.data.gouv.fr](http://www.data.gouv.fr). The *National Institute of Statistics and Economic Studies* (INSEE) provides data on municipality characteristics like age composition, commercial activity and education. Electoral data, such as voter turnout, and election and referenda results, are obtained from the *Center for Socio-Political Data* (CDSP).

We present both results on differences in stated versus revealed preferences. Stated preferences have the obvious advantage that we can use direct questions asking people about the strength of their respective identities. However, those are "costless" answers, and might thus exaggerate existing differences or yield biased estimates. A measure of revealed preferences is ideally a costly decision, where a representative sample of the population in the treated and untreated area face a decision that signals the strength of regional identity. In addition to survey evidence, we benefit from the fact that in France repeated nationwide referenda were held that directly touched upon questions relating to the political influence of regions, the recognition of regional culture, and regional decision-making.

Our main measures of regional identity at the municipal level is the agreement in three referenda in 1969, 1992 and 2005. The referenda are a good measure as they were important decisions with a political cost to them, but there was no binding monetary constraint preventing certain groups or parts of the population from voting. We use data on voter turnout to verify that the results are representative of the underlying population. With regards to national identity, we can use a clear measure of stated national identity in surveys, but for revealed preferences the best available measure is the vote share for the candidate of the right-wing National Front, Jean-Marie Le Pen. Regarding mechanisms, we present results about regionalist parties and regional newspaper subscriptions. The following paragraphs describe the indicators capturing revealed preferences.

### **Referendum on Regionalization (De Gaulle), 1969**

First, we use a referendum that President Charles De Gaulle called in 1969, which was explicitly about decentralization and the establishment of regions as an important political unit in the French constitution (Bon, 1970). Regions were supposed to take control of public utilities, housing and urbanization and be enabled to borrow money. Furthermore, they would be independent contractual parties, be able to set up public organizations, and be part of an adapted second chamber representing the territorial collectivities. De Gaulle justified the referendum by saying that wherever possible decision-making should happen closer to the citizens, and that the regions cultural importance should be reflected politically. In the end, 52.4 percent of French voters rejected the proposal and De Gaulle resigned immediately afterwards. Thanks to the help of the director of the Lorrainian departmental archive, Jean-Eric Jung, we got access to digitized newspapers from April 1969, which we transcribe and match to the current municipalities for the three Lorrainian départements. For Alsace, we rely on département level results, which are available for free nation-wide.

### **Treaty of Maastricht, 1992**

The Maastricht Treaty included several reform proposals about the institutional and political structure of the European Union (EU). The crucial aspect is that it changed the role of regions in the

EU by fostering regional decision-making and the expression of regional identity. The treaty was a huge step forward for regions in the institutional landscape in Europe. It formally introduced the principle of subsidiarity, which codified the aim of decision-making at the lowest feasible level of authority in the EU (Treaty on the European Union, 1992), often the regional instead of the national level. In addition, it established a *Committee of the Regions* as part of the European institutional structure. This “created a political space for regions” (Fitjar, 2010, p.528). Research in political science describes how it allows regions to “seek a greater voice in EU affairs” and “reignite calls for decentralization and regional autonomy” (Chacha, 2013, p.209; Scott et al., 1994).

### **Constitution for Europe, 2005**

The referendum about the so-called *Constitution for Europe* was again perceived as helping regions in their scope of decision-making and possibility to express regional identity. Official assessments of the regional and local authorities associations, which were publicly available and communicated to voters (CEMR, 2004, source of all following citations), regard it as a “an achievement for regional and local authorities”, which would “strengthen the role of local and regional governments”. An important point was the reinforcement of the subsidiarity principle and “greater recognition to the role of regional authorities” as well as “respect for regional and local self-government as part of national identities”. Cross-border regions were introduced as a new way of representing common regional interests formerly divided by nation states.

The widespread opinion in 1992 and 2005 was that the EU is “moving towards a Europe of the regions” (Chacha, 2013 p.208). EU integration was seen as reducing the costs of regional autonomy and allowing regions to bypass national governments and deal with Brussels directly. For that reason, regionalist parties “favor European integration because it creates a more favorable political opportunity structure for their subnational autonomy movements.” and “perceive the EU ‘as an ally against the central state’” (Jolly, 2007, p.110 & 124). The moderate regionalist Alsatian party *Le parti Alsacien*, for instance, campaigns on its website for an “independent Alsace in a federal European Union”.

### **Regional newspaper subscriptions and regionalist parties**

In addition to those measures, we use data on regional newspaper subscriptions and regionalist parties to examine potential mechanisms. Although we use them to measure mechanisms, the share of households that subscribe to regional newspapers and the vote share of regionalist parties is itself of course also a proxy for regional identity. We managed to get access to municipal level data for subscriptions to the Lorrainian newspaper “Le Republicain Lorraine”, but only for the year 2014, ten years after the third referendum. Regionalist party results are from the 2015 regional elections, where all moderate regionalist parties in Alsace and Lorraine ran on a joint list.

## National identity/nationalism

In addition to département level survey results, we use the strength of the extreme right-wing populist *Front National* candidate Jean-Marie Le Pen at an election close to our two main outcome variables as a proxy variable for differences in national identity. Clearly, a stronger national identity does not necessarily lead to higher support for a right-wing party. Nevertheless, for it to be an informative proxy requires only that voters with a stronger national identity are, all else equal, more likely to vote for the nationalistic *Front National*. We use 2007 because for the first-round data are available for all municipalities; for previous elections the data cover only municipalities larger than 3500 inhabitants. In our robustness section, we also use data from tweets supporting the French national team during the World Cup 2014 as an alternative.

Out of all measures, the first referendum in 1969 has the advantage that it clearly focuses on fostering regions as an important political unit in France, thus clearly relating and measuring differences in regional identity. As explained above, the two referenda in 1992 and 2005 were both also clearly related to the political recognition of regions and more regional decision-making, and perceived as such by the population. It is reassuring that there the strong overlap between regional and European identity is not only documented by other studies (Chacha, 2013) but also visible in our data (see Table A9). Moreover, département level data show that already in 1972 a referendum about EU expansion yielded a comparable vote pattern in the region than the 1969 referendum on the establishment of regions (Figure A9d). Nonetheless, as those two referenda also relate to broader questions about the European Union, we show that treated and control municipalities at the border do not differ in other factors that might make European integration more or less beneficial.

We also compute geographical characteristics to evaluate the exogeneity of the border. The data on terrain ruggedness is from Nunn and Puga (2012), but we use it at a more disaggregated level. Raw elevation data comes from the NASA Shuttle Radar Topography Mission (SRTM) data set. Data on potato and soil suitability, which we choose as the two crops which are likely to be the most important ones, comes from the FAO’s Global Agro-Ecological Zones database (GAEZ). To best approximate pre-“Green Revolution” growing conditions in 19th and early 20th-century Europe we choose a medium input intensity and irrigation.<sup>7</sup>

Most of the outcome measures are available for both Alsace and Lorraine, except the referendum in 1969 and the regional newspapers. Table 2 provides the reader with an overview of all data, and its availability and usage. Table A10 shows summary statistics for our variables of interest in the full sample of municipalities in Alsace and Lorraine. Tables A2 and A11 show definitions and sources, as well as descriptive statistics for the variables. The next section begins by focusing on Lorraine, which provides a better counter-factual and the exact location of the border is more clearly exogenous. Later, we show that the results hold for Alsace as well. This is reassuring with regard to the external validity of the results, and allows us to compare the two regions with regard to the mechanisms.

---

<sup>7</sup> Ruggedness: <http://diegopuga.org/data/rugged/>. Elevation: accessed through the web page of ESRI. Soil Suitability: <http://www.fao.org/nr/gaez/en/>.

Table 2: Variables, level of aggregation and region in order of appearance.

#	Content	Preferences	Year(s)	Level	Paper	Appendix <sup>a</sup>
1	Geographic variables	-	-	Mun <sup>b</sup>	L <sup>c</sup>	A+L <sup>c</sup>
2	Cahiers de doléances	Stated	1789	Dep <sup>b</sup>	L	A+L
3	Survey I	Stated	1999, '01, '03	Dep	L	A+L
4	Referendum	Revealed	1969	Mun	L	-
5	Referenda, Le Pen	Revealed	1992,'05,'07	Mun	L, A+L	-
6	Regional Newspaper	Revealed	2014	Mun	L	-
7	Regionalist parties	Revealed	2015	Mun	A+L	-
8	Survey II	Stated	1999, '01, '03	Dep	L, A+L	L, A+L

*Notes:* This table provides an overview about the main variables in order of appearance throughout the paper. Preferences are distinguished between revealed and stated preferences, the level of aggregation is either département or municipality. The last two columns show where we use variables for either only Lorraine or for both Alsace and Lorraine. The table refers to the main regressions, robustness tests for different variables at various levels are shown in the Appendix in addition.

<sup>a</sup> Appendix = Online Appendix. Figures and Tables with "A"+Number.

<sup>b</sup> Levels are either Mun = municipality or Dep = département

<sup>c</sup> Regions refer to L = Lorraine or A+L = Alsace and Lorraine, respective to which data is presented in the main paper or the online appendix

### 3.2 Identification strategy

Our treatment variable in the municipal level regression is a deterministic function of the geographical location of a municipality, testing for a discontinuity in treatment at the threshold defined by the former border dividing Alsace and Lorraine. The causal interpretation draws on studying municipalities close to the former border using a RDD. Formally, the following regression model:

$$y_c = \alpha + \beta Treatment_c + p(\text{distance to border}_c) + \mathbf{z}'_c \gamma + \epsilon_c, \quad (1)$$

where  $y_i$  is the outcome variable of interest for municipality  $c$ ,  $Treatment_c$  is a dummy taking the value 1 for municipalities in the formerly occupied region. The linear term measures the direct distance from the municipality centroid to the former national border.  $\mathbf{z}_c$  comprises the distances to the city of Metz, city of Strasbourg, city of Nancy, and to the current French-German border. As suggested by Gelman and Imbens (2017), we include a linear term for the distance, allowing its coefficient to vary on either side of the border. This means that we estimate a local linear regression model according to (1) close to the former border, using a uniform kernel density function, for different bandwidths. Figures A17 through A26 present estimates from plausible alternative specifications.<sup>8</sup> All results are in line with those presented here.

<sup>8</sup> Dell (2010) discusses why a semi-parametric approach could be superior when the geospatial data is not precise in terms of geographical location. In our case, we do not have data on individuals and, for instance, their addresses. Instead, our outcome variables measure the municipality level aggregate of individual actions, and we approximate their location in relation to the former border by the distance from the municipality centroid. We find no discontinuity for any of these measures, suggesting that they are orthogonal to our treatment variable. Figure A13 to A23 show that the results are not affected when (i) omitting controls, controlling for (ii.) border segments and (iii.) distance to the language border, controlling for (iv.) longitude, latitude, as well as (v.) both and their interaction to compare only actual neighboring municipalities.

The treatment effect in (1),  $\beta$  is given by

$$\beta = \lim_{x_c \rightarrow 0^+} \mathbf{E}[y_c | x_c] - \lim_{x_c \rightarrow 0^-} \mathbf{E}[y_c | x_c], \quad (2)$$

where  $x_c$  is the distance to the border normalized at 0, meaning that the distance for municipalities in the treated region is equal to the actual distance, while it is equal to the actual distance multiplied by minus one for municipalities in the untreated region. Under the assumption of the conditional expectation function,  $\mathbf{E}[y_c | x_c]$ , being continuous, the treatment effect is equal to the difference in the outcomes at the border between municipalities in the treated and untreated area. This assumes that all other factors relevant in explaining the outcome are continuous at the border and that the treatment is orthogonal to potential outcomes. We address this by formally testing for discontinuities in covariates and looking for discontinuities in geographic factors, which are plausibly not affected by the treatment and thus capture potential pre-treatment imbalances. Specifically, we show that there is no discontinuity in terrain ruggedness, elevation, and soil suitability for production of potatoes, wheat and barley (Table A12).

To get a sense of identity before 1871, we make use of the fact that Louis XVI, shortly before the French revolution, felt the need to assess the loyalty of his citizens. These data, known as the “Cahiers de doléances”, specifically ask about the relative strength of regional compared to national identity. They were aggregated and transformed to a scale between 1 and 3 originally by Hyslop (1968) and have recently been used to assess the effect of state capacity on identity formation (Johnson, 2015). Following Johnson, we exclude the clergy, which was more driven by religious policy, and include the second (nobility) and third (other citizens) estate as well as the category “unified orders”. The average response for all four départements within Lorraine is exactly or very close to 2, as Table 3 shows, and there is no statistically significant difference between Moselle and the rest of Lorraine.

Table 3: National vs. Regional identity in Lorraine in 1789 (Cahiers de doléances)

	Mean	Std. dev.	Obs.
Lorraine (average)	2.021	0.541	24
Moselle (treated)	2.000	0.816	7
Meurthe-et-Moselle	2.000	0.598	8
Meuse	2.000	0.000	4
Vosges	2.100	0.224	5
	Difference	Std. dev. <sup>a</sup>	Obs.
Moselle vs. untreated	-0.029	0.349	24

*Notes:* National identity in 1789 based on Cahiers de doléances for each département in Lorraine (and Vosges). The measures are based on an index created by Hyslop (1934), where the value 3 corresponds to "National patriotism strongest", 2 corresponds to "Mixed loyalties: national patriotism combined with regionalism or class spirit, or both.", and 1 corresponds to "Other loyalties, regional, or class, or both, outweigh national patriotism". Hyslop (1934) computed these values at the level of selected important city areas based on more disaggregate reports in verbal form. We assign the city areas to current départements.

<sup>a</sup> Heteroscedasticity-consistent standard errors.

## 4 Main results

### 4.1 Survey evidence

We begin by considering survey evidence on stated differences in identity. The OIP surveys have the big advantage that they include direct questions proxying for the perceived common identity of the average individual. The parameter of interest  $\Delta$  comes from the equation:

$$y_i = \pi + \Delta Treatment_i + \Gamma'_i \lambda + \eta_i, \quad (3)$$

where  $Treatment_i = \mathbf{1}$ [individual in treated area] and  $\Gamma$  contains controls for (reported) age, education, employment status and sex. As the geographic precision of the survey is the département, our estimation compares the conditional means of regional and national identity in the treated and the untreated area.

According to Table 4, people in the treated area today clearly express a significantly stronger common regional identity. In contrast, there is no difference in common French identity. We also compute the ratio of regional relative to national identity, and standardize this variable to ease interpretation. People in the treated areas of Alsace and Lorraine exhibit a ratio that is 14 percent of a standard deviation higher. It is interesting that there is no difference with regard to national identity, emphasizing the importance of our model and definition where identities need not to be substitutes. This is the comparison within Lorraine, Table A7 shows that the differences are similar but larger when including Alsace as well.

Table 4: Survey results, Lorraine

Survey question	Mean, control	$\Delta$	P-value	No. obs.
Feel close to region (Regional identity)	3.362	0.154	<0.001	1314
Feel close to nation (National identity)	3.635	0.028	0.409	1313
Regional identity/National identity (standardized)	-0.138	0.138	0.011	1311

*Notes:* Sources are the Observatoire Interrégional du Politique (OIP) 1999, 2001, and 2003, using respondents in Lorraine, at the département level. Identity is measured on a 4-point Likert-scale. Table A7 shows similar results for all of Alsace and Lorraine. A positive  $\Delta$  indicates that people in the treated region agree more with the statement.

Another interesting result supports our interpretation of the 1992 and 2005 referenda, which touch upon issues relating to regional-decision making as well as assigning more competences to the European Union. There is a highly significant correlation between stating a stronger regional and stronger European identity in the sample regions and in France overall (Table A9), and 85% of respondents stating a stronger European identity also express a stronger regional identity.

Note that the survey results measure differences in stated instead of revealed preferences, and could be affected by omitted variables. If, for instance, the proximity to neighboring countries correlates with regional identity, this could affect the results as the treated area is closer to neighboring countries. The next section proceeds with municipal level data on three outcomes revealing regional identity, and resolves concerns about omitted variables through a geographic RD design.

## 4.2 Referenda and nationalism

Figure 4 provides maps of the election and referenda results from 1969, 1992, 2005 and 2007. Figures 4b (available within Lorraine only), 4c and 4d show higher agreement to the referenda that would strengthen regional decision-making powers in the treated area to the right of the former border.<sup>9</sup> Figure 4a shows no clear pattern regarding nationalism. It is important to note that there are no obvious visible differences in turnout for all four votes (see Figure A10). Moreover, Figure A9c shows no comparable pattern of support for De Gaulle in the 1968 presidential election, suggesting that preferences about him as a person do not explain the differences in 1969.

Table 5 shows ordinary least squares estimates of  $\beta$  from equation (1) including controls to get a first feel for the data, as well as enabling us to assess the external validity of the RD estimates. Although the RDD has advantages in most dimensions, potential sorting is more likely to be an issue at the border because the costs of moving to neighboring municipality are lower. The OLS specifications avoid this problem. If both approaches yield similar results, we can not only disregard worries about sorting, but also have more faith in the causal interpretation of the département level survey results.

For *Share Le Pen 2007*, *Share Yes 1969*, *Share Yes 1992* and *Share Yes 2005*, the coefficients indicate lower levels of national identity and a stronger regional identity in the treated area. The interpretation of the regression coefficient for the treatment variable is the average difference in percentage points between treated and untreated municipalities. It is illustrative to relate them to the average vote share of the whole region. For instance, the coefficient of *Share Yes 1969* is 14.1 percentage points, which equates to almost 24 percent of the average yes-vote share of 59.2 in all of Lorraine. The coefficient of *Share Yes 1992* is 4.2 points, almost 10 percent of the average yes-vote share, and the 6.2 in *Share Yes 2005* correspond to about 15 percent. *Share Le Pen 2007*, our proxy for nationalism is about 6 percent lower in the treated region than the average vote share.

Table 5: OLS estimates using municipalities in Lorraine

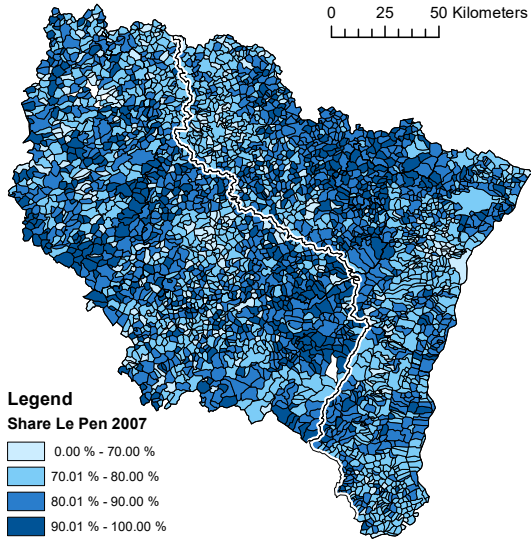
Dep. Variable:	Nationalism		Regional Identity	
	Share Le Pen 2007 (1)	Share Yes 1969 (2)	Share Yes 1992 (3)	Share Yes 2005 (4)
Treatment	-1.088 (0.507) [0.032]	14.100 (2.058) [<0.001]	4.222 (0.870) [<0.001]	6.247 (0.913) [<0.001]
Adj. R-squared	0.046	0.045	0.143	0.057
Obs.	2240	1622	2237	2240

*Notes:* OLS estimates using whole sample of municipalities in all départements in Lorraine. The outcomes are the share of Le Pen votes in the 2007 presidential election (first round), the share of Yes votes in the 1969 referendum, in the 1992 referendum, and in the 2005 referendum. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, distance to Nancy. Conley standard errors are displayed in parentheses and  $p$ -values in brackets.

<sup>9</sup> Figure A9a shows at the département level that the comparison for 1969 using Alsace suggests a similar, if not larger, difference. The yes-vote share out of all eligible voters was above 50 percent in Alsace and between 40 and 45 percent in neighboring Vosges.

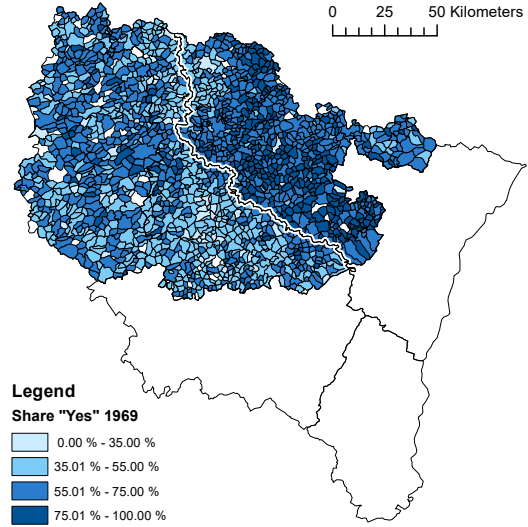


Figure 4: Maps of municipal level outcomes of referenda in 1969, 1992 and 2005; and the presidential election of 2007



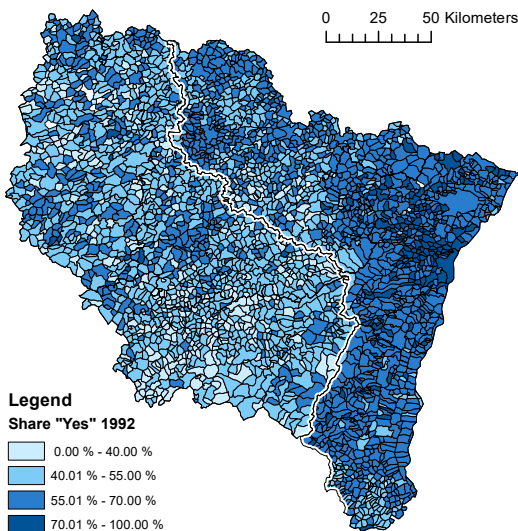
*Notes (a):* Share of votes for the nationalist leader Le Pen in the presidential election 2007. The former border dividing the area is highlighted in white. Darker shades reflect higher values.

(a) Share Le Pen 2007 (Nationalism)



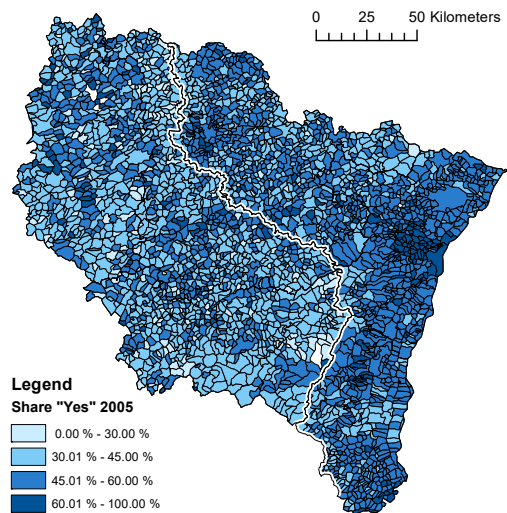
*Notes (b):* Share of yes votes in the constitutional referendum in 1969. Areas where data is not available are shown in white. Data is available for the départements of Meuse, Meurthe-et-Moselle and Moselle. The former border dividing the area is highlighted in white. Darker shades reflect higher values.

(b) Share "Yes" 1969 (Regional identity)



*Notes (c):* Share of yes votes in the referendum in 1992. The former border dividing the area is highlighted in white. Darker shades reflect higher values.

(c) Share "Yes" 1992 (Regional identity)



*Notes (d):* Share of yes votes in the referendum in 2005. The former border dividing the area is highlighted in white. Darker shades reflect higher values.

(d) Share "Yes" 2005 (Regional identity)

Accordingly, the OLS estimates are in line with our predictions and the survey results concerning to regional identity, and slightly deviate with regard to nationalism. Now we turn to the RDD results. Our baseline RD estimation shows estimated treatment effects on all four outcome variables from Figure 4 for bandwidths of 10, 15 and 20 kilometer. In addition, we include one specification using the optimal IK bandwidth (Imbens and Kalyanaraman, 2012). For all outcomes, this is larger than 20 kilometers, suggesting that smaller bandwidths are rather conservative. The closest choice of 10 kilometers basically compares only municipalities directly at the border with their direct neighbors. This should eliminate all concerns regarding comparability, as distances to neighboring countries or cities are virtually identical.

Table 6 shows that the estimated treatment on regional identity effect remains positive and statistically significant for *Share Yes 1969*, *Share Yes 1992* and *Share Yes 2005* across all bandwidths. The coefficient for 1969, which provides the clearest measure of regional identity, is about 13 percentage points at the smallest bandwidth of 10km. This reflects the medium term reaction, in a population still composed of people who themselves consciously experienced repression, and of those who grew up later and were affected through their parents or other groups in society. The significant and strong effects in the two referenda later documents persistence and indicates that the stronger regional identity is indeed transmitted across generations. It ranges from 4.4 percentage points to 5.4 percentage points in 1992, and 3 to 3.9 percentage points in 2005. Thus, the simple OLS estimation seems to have overestimated the actual effect, but not by much. This supports the causal interpretation of the survey results that relied on a comparison of group means at the département level. The occupation and the associated attempts to suppress regional culture lead to a persistent increase in the suppressed group’s regional identity.

With regards to nationalism, the RDD results are largely in line with the survey results on national identity. There is no significant difference for the three of the four bandwidth, and the point estimate is much smaller than with OLS. Figure 5 shows these discontinuities graphically in RD plots using a first-order polynomial that varies across cutoff. In Panel A, the clear jump at the border is visible in all three referenda, while Panel B indicates no jump in nationalism. In the following regressions, we thus concentrate on the three referenda results about regional identity. Figure A11 shows the plots with a 50km bandwidth, Figure A12 with a second order polynomial; the jump at the border is always clearly visible.

Historical evidence suggests that Alsace and Lorraine were as well integrated into France prior to the Franco-Prussian War as other regions. We implement a *placebo* test using the 1992 and 2005 referenda results in the regions geographically west of the control area, and check for a discontinuity at the border between this western part of Lorraine and the rest of France to further validate this. If the complete region was already exhibiting a stronger regional identity previously, we would expect a discontinuity here. Figure 5, Panel B, however, clearly illustrates that in contrast to Panel A there is no discontinuity at this placebo border. Table A24 shows the absence of discontinuities in the corresponding regression table. Table A25 shows another placebo test using the pre-1870 département border within Moselle, and also finds no significant discontinuities.

Table 6: Discontinuities in referenda results and share of Le Pen votes, using municipalities in Lorraine

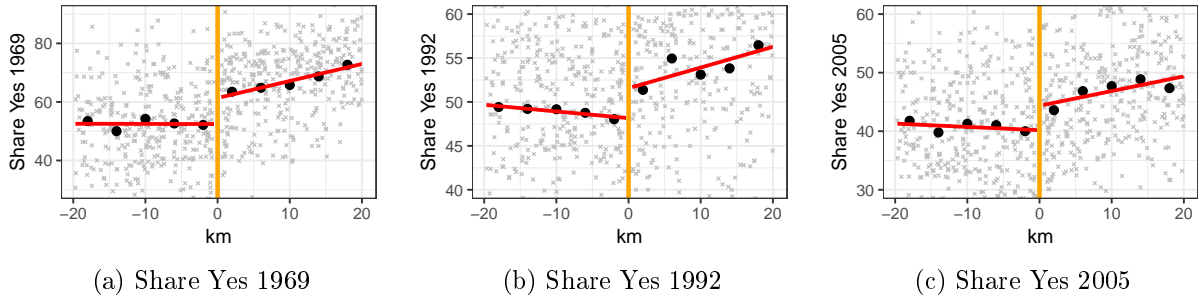
Dep. Variable:	Le Pen 2007 (nationalism)				Share Yes 1969			
	(1)	(2)	(3)	(4) <sup>a</sup>	(5)	(6)	(7)	(8) <sup>a</sup>
Treatment	-0.486	-0.385	-0.482	-1.590	13.195	13.844	11.781	15.711
	(0.961)	(0.808)	(0.774)	(0.711)	(2.836)	(3.438)	(3.156)	(2.787)
	[0.613]	[0.634]	[0.534]	[0.026]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Obs.	394	583	744	1016	375	550	693	1362
Dist	10 km	15 km	20 km	29.12 km	10 km	15 km	20 km	50.29 km
Dep. Variable:	Share Yes 1992				Share Yes 2005			
	(1)	(2)	(3)	(4) <sup>a</sup>	(5)	(6)	(7)	(8) <sup>a</sup>
Treatment	3.752	5.026	4.346	5.751	3.810	3.757	4.892	7.448
	(1.841)	(1.611)	(1.440)	(1.133)	(2.092)	(1.775)	(1.646)	(1.392)
	[0.042]	[0.002]	[0.003]	[<0.001]	[0.069]	[0.035]	[0.003]	[<0.001]
Obs.	394	583	744	1517	394	583	744	1109
Dist	10 km	15 km	20 km	53.22 km	10 km	15 km	20 km	32.86 km

*Notes:* Discontinuity at the treatment border using Municipalities in Lorraine. The outcomes are the share of Le Pen votes in the 2007 presidential election (first round), the share of Yes votes in the 1969 referendum, in the 1992 referendum, and in the 2005 referendum. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, distance to Nancy. Conley standard errors are displayed in parentheses and *p*-values in brackets.

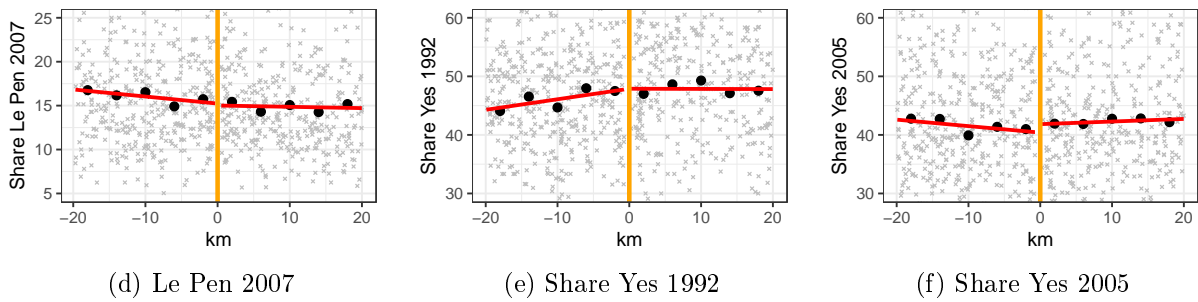
<sup>a</sup> Estimates from using the optimal IK bandwidth.

Figure 5: RD plots for voting outcomes 1969, 1992, 2005 and 2007

Panel A: Regional identity at the treatment border

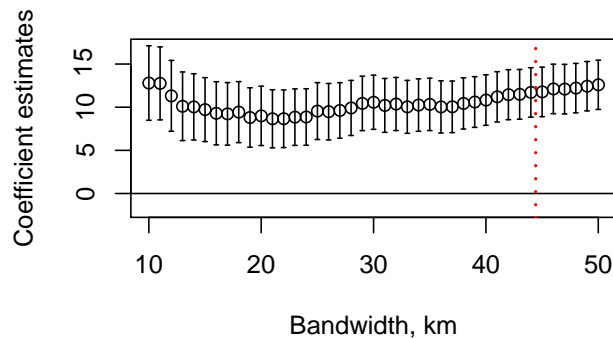


Panel B: National identity at the treatment border, and regional identity at a placebo border



Notes: (a), (b), (c), (d): RD plot using municipalities in Lorraine, 20 kilometer distance to the treatment border, with first degree polynomial fit varying on each side. Dots represent binned means using 4 kilometer bins. (e), (f): RD plots using municipalities within 50 kilometers of the border separating non-annexed Lorraine from rest of France.

Figure 6: Estimation plots for 1969 referendum, within Lorraine



Notes: Estimates of treatment effect, bandwidths ranging between 10 and 50 kilometers, within Lorraine, fitted using first degree polynomial. Dashed vertical line at optimal IK bandwidth. Solid vertical lines represent 90 percent confidence intervals (based on Conley standard errors).

Although Table 6 indicates already that the choice of the bandwidth does not affect our results, Figure 6 clarifies this further. It depicts the individual coefficients and confidence intervals across all plausible bandwidths ranging from 10 to 50 kilometers for the 1969 referendum. The effect size varies little and is always positive. As we would expect, the estimation becomes more precise as we increase the bandwidth, and the coefficient remains remarkably stable. Figure A13 to A23 show the same plot for the 1992 and 2005 referenda. Moreover, they show that the results are not affected when (i) omitting controls, controlling for (ii.) border segments, (iii.) distance to the language border, (iv.) longitude, latitude, as well as (v.) longitude, latitude and their interaction.

The causal interpretation of the coefficients rests on the assumption that the untreated municipalities can be viewed as counter-factuals for the treated communes. We want to remedy one potential concern by considering potential post-treatment discontinuities in socio-demographic characteristics. Note that those factors might be affected by the treatment and act as channels through which the treatment affects the outcome. Based on the literature on the determinants of voter preferences and turnout (e.g., Franklin, 2004), we examine potential discontinuities in yearly median income, mean age, education and occupation. Table A17 shows that there are no discontinuities within Lorraine.

In a next step, we want to examine whether we can find the same results when including Alsace in the analysis. Remember that the border dividing Alsace is partly following linguistic differences, so that the counterfactual municipalities in Vosges might differ with regard to the traditional dialect and potentially culture. Table A17 indicates that at the border the treated municipalities in Alsace are richer and slightly younger on average. However, when looking at the extended RDD results in Table 7, we find that the treatment effects when including Alsace are of a similar magnitude and remain statistically highly significant.

Table 7: Discontinuities in referenda results and share of Le Pen votes, Alsace and Lorraine

Dep. Variable:	Le Pen 2007		Share Yes 1992		Share Yes 2005	
	(1)	(2) <sup>a</sup>	(3)	(4) <sup>a</sup>	(5)	(6) <sup>a</sup>
Treatment	-0.236 (0.852) [0.782]	-1.099 (0.548) [0.045]	4.353 (1.748) [0.013]	3.787 (0.859) [<0.001]	2.957 (1.742) [0.090]	6.443 (1.090) [<0.001]
Obs.	603	1707	604	2781	603	1849
Dist	10 km	30.37 km	10 km	60.88 km	10 km	33.37 km

*Notes:* Discontinuity at the treatment border in Lorraine. The outcomes are the share of Le Pen votes in the 2007 presidential election (first round), the share of Yes votes in the 1969 referendum, in the 1992 referendum, and in the 2005 referendum. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, distance to Nancy. Conley standard errors are displayed in parentheses and *p*-values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

## 5 Mechanisms, persistence and policy preferences

### 5.1 Mechanisms

After documenting that there is a causal effect of occupation, repression and the associated homogenization policies on regional identity, we are interested in potential mechanisms. It is plausible that the exposure during the treatment period unconsciously changed the attitudes of treated individuals, which affects potentially unobservable aspects of raising their children (Dohmen et al., 2012). Moreover, people in the treated area could also consciously have decided to invest in instruments that help to express, strengthen and transmit regional identity. Conceptually, although the theoretical model in the appendix focuses on educational investments by parents in the tradition of Bisin and Verdier (2000), an extended model could incorporate specific instruments as well as investments by groups of individuals. In contrast to shifts in unobservable aspects of raising children, differences in some of those instruments can be observed.

Potential instruments, based on existing literature and our own considerations, could be local festivities, clubs and organizations, political parties and media outlets.<sup>10</sup> We cannot reliably compare clubs as the legal rules for establishing and registering a club differ between the treated and control area, and the available information about festivities does not reliably identify those related to regional culture. However, we can exploit information about regionalist parties and regional media usage. Note that logically all discontinuities reflecting potential mechanisms can also be considered as outcomes (and are thus bad controls in a regression using the other outcomes), which is why we are not able to estimate how much any particular mechanism has contributed to the differences.

Regionalist parties are also interesting with regard to the chronological order of cause and effects. Fouka (2017) finds that in her sample of German immigrants exposed to repressive policies in the US, the observed increase in common group identity occurs only in the medium run (Fouka, 2016). Accordingly, we are also interested in whether the effects that we document begin to materialize during or after the treatment. Historical evidence indicates that in contrast to the reaction of a minority immigrant group, the repressive policies already triggered an increase in regional identity during the treatment period (Goodfellow, 1993; Harvey, 1999). This could be observed through periods of public protest, the establishment of regional organizations and newspapers (Callender, 1927), but also politically through the emergence and success of regionalist parties.

#### Regionalist parties

Regionalist parties emerged and enjoyed great electoral success during German occupation, with a vote share of between 30.2 percent and 56.6 percent (Hiery, 1870). The success of regionalist parties continued during the interwar period under French rule. The Independent Regional Party for Alsace-Lorraine, for instance, received 11.5 percent of the votes in Bas-Rhin in 1928, and Zanoun (2009)

---

<sup>10</sup> Ochsner and Roesel (2017) suggests that war memorials and statues also function as a technology to transmit a common history. There are some well-known statues in Lorraine that might reactivate the memory of repressive policies, but they are mostly related to WWI or WWII, which makes a distinction difficult.

suggests that “autonomists were also present in the Moselle and like their Alsatian counterparts demanded autonomy for Alsace-Lorraine”. Accordingly, the historical evidence indicates that both German and French policies triggered investments in the creation of regionalist parties that then enjoyed electoral success. There were no comparable successful parties during the treatment period in the untreated area. It is also important to note that historians classify the vast majority of these parties after WWI as aiming for more regional autonomy, rather than for a return to Germany (Rothenberger, 1975). After the end of WWI, the regional parliament even proclaimed a sovereign region of Alsace-Lorraine on November 11, 1918. However, this was not accepted by France.

Support for regionalist parties collapsed in the built-up of WWII, as the parties were perceived as being associated with Nazi-Germany. These accusations under French rule were apparently more widespread in Lorraine and less in Alsace, where a larger share suffered under the intrusive French language policies and saw regionalist parties as fighting to reestablish bilingualism. Up until today, political regionalism is much stronger in Alsace than in Lorraine, where support for regionalist parties never recovered to pre-war levels. Alsace features two regionalist parties, the right-wing “Alsace d’abord” and the moderate “Le Parti Alsacien/Unser Land”. Both are rather successful, the former winning about 9 and the latter around 15 percent of the votes in the 2010 regional elections, while the party “Vosges d’abord” in the neighboring untreated département has enjoyed little electoral success. In upper Lorraine, the “Parti des Mosellans” and the more established “Parti Lorrain” are the remaining regionalist parties, campaigning for a strong Lorraine region in a “Europe of the Regions”.

The 2015 regional elections allow us to evaluate remaining differences between the treated and untreated area systematically, because all moderate regionalist parties in the region formed a joint list. Within Lorraine, the combined average vote share is 2.1 percent, nearly twice as much compared to the 1.1 percent in the untreated neighboring Meurthe-et-Moselle. The difference in the averages is also visible in using the RDD. With a bandwidth of 15 or 20 kms and with the efficient bandwidth, the causal effect is about 0.4-0.5 percentage points. It remains positive but becomes insignificant with the 10km bandwidth, however. When taking account of Alsace as well, the differences are more pronounced. This is in line with the historical evidence cited above about the difficulties initially strong regional parties faced in Lorraine. The share is between 1.2 and 2.5 percentage points higher in the treated area, with p-values smaller than 0.05 for all bandwidths. Hence, regionalist parties are one plausible mechanism through which the stronger regional identity in the treated area has been maintained.

### **Regional newspaper subscriptions**

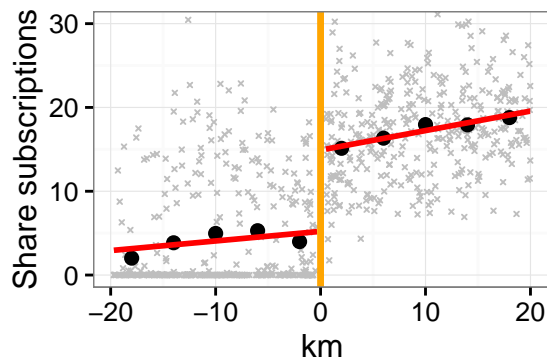
In addition to regionalist parties, we analyze the share of households subscribing to the regional newspaper “Le Republicain Lorraine”. We were able to gather data for one (Lorrainian) regional newspaper for the year 2014, and thus focus on the treated and the untreated parts of Lorraine. There is no Alsatian regional newspaper that is also widely enough read in the département of Vosges to allow a meaningful comparison. Newspapers are particularly interesting as a transmission

mechanism. They do not only provide information to the parents within a household and work as a signal of regional attachment to other households, but can also be used as a useful instrument to transmit regional culture to children.

Table 8 shows a clear discontinuity in subscription rates at the treatment border. At the 10 kilometer bandwidth, the share of subscribers out of all households is around 10 percentage points higher on the treated side. The result is highly significant with p-values below 0.01 in all specifications, and the clear discontinuity is also graphically visible in Figure 7 and in the map in Figure A28). Table A26 shows that the effect size is not driven by differences in the spoken dialect. Figures A24 and following document the robustness to bandwidth choices and alternative specifications. We can also try to disentangle supply and demand side explanations by controlling for the number of points of sale in 2014 (there are more regional offices in the treated area). Conditioning on supply side differences barely affects the point estimates, suggesting that demand side differences dominate (Table A27).

To sum up, there are identifiable differences in two plausible and relevant mechanisms. Regionalist parties are an important instrument to express regional identity and maintain and popularize the importance of regional culture. This channel seems relevant for both regions, but stronger for Alsace. Within Lorraine, we find strong and sizeable differences in subscription rates to regional newspapers. We cannot estimate which share of the differences in the survey questions and in the referenda can be explained by those mechanisms in a precise econometric way, but the size of the effects is sufficiently high to be relevant transmission channels.

Figure 7: RD plot, share households with subscription of “Le Republicain Lorraine”



Notes: RD plots using only municipalities within Lorraine, with first degree polynomial fit varying on each side. Black dots represent means using 4 km bins, our most conservative strategy is also relevant with regard.



Table 8: RD results: Regional newspaper subscription shares, and regionalist parties

Panel A: Share households with subscription of “Le Republicain Lorraine”, within Lorraine				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	10.155	10.132	9.872	11.005
	(1.417)	(1.234)	(1.129)	(0.964)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Obs.	394	583	744	1392
Dist	10 km	15 km	20 km	46.23 km
Panel B: Regionalist parties, within Lorraine				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	0.082	0.429	0.421	0.553
	(0.262)	(0.230)	(0.214)	(0.178)
	[0.755]	[0.062]	[0.050]	[0.002]
Obs.	394	583	744	1666
Dist	10 km	15 km	20 km	65.42 km
Panel C: Regionalist parties, All of Alsace and Lorraine				
Variable	(1)	(2)	(3)	(4) <sup>a</sup>
Treatment	1.153	2.340	2.232	2.497
	(0.583)	(0.535)	(0.496)	(0.411)
	[0.049]	[<0.001]	[<0.001]	[<0.001]
Obs.	604	887	1150	1885
Dist	10 km	15 km	20 km	34.01 km

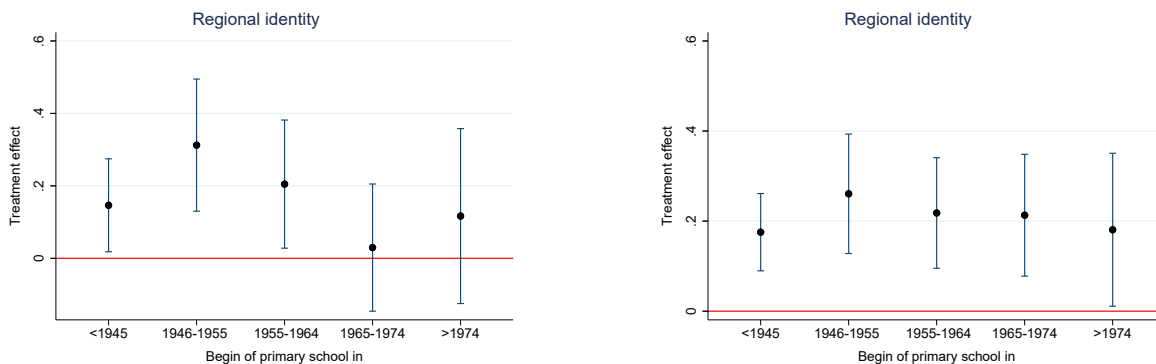
*Notes:* Discontinuity at the treatment border. The outcome in Panel **A** is the share of households subscribing to the regional newspaper “Le Republicain Lorraine”, within Lorraine for 2014. The vote share for regionalist parties is the outcome in both Panel **B** and **C** for the regional elections 2015. The former uses municipalities only in Lorraine, while the latter uses all municipalities in Alsace and Lorraine. Included controls: distance to Germany (border), distance to Metz, distance to Strasbourg, and distance to Nancy. Conley standard errors in parentheses and  $p$ -values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

## 5.2 Regional identity over time

In our model, treated citizens were more likely to privately build up the skills to teach their own children regional traditions during the treatment period; after public schooling returns back to similar levels, this leads to a difference in regional identity. To understand this mechanism and persistence over time better, and as the RDD results provide no reason to expect a systematic bias, we return to the survey results from section 4.1. We re-estimate regression models on regional identity, but now interact the treatment effect with dummy variables for different age cohorts, with the untreated subjects as the left-out reference category. The age cohorts are selected so that the second group started primary schooling after WWII. The model we use to explain persistence makes no clear predictions regarding the net difference for those experiencing the treatment period themselves, but predicts differences for later age cohorts if public schooling returns to comparable levels in the control and treated area.

Figure 8: Identity differences by age cohort



(a) Treatment effect in Lorraine

(b) Treatment effect in Alsace and Lorraine

*Notes:* The treatment effects refer to the parameter  $\Delta$  in the equation:  $y_{ig} = \pi + \sum_g \Delta_g \times Age_g \times Treatment_{ig} + \Gamma'_i \lambda + \eta_{ig}$ , where  $Treatment_{ig} = \mathbf{1}[\text{individual in treated region}]$  and  $\Gamma$  comprises controls for (reported) age, employment status and sex.  $g$  indicates to which age cohort an individual belongs, the group of untreated participants act as the baseline category. Age cohorts are selected such that the second group started schooling after the end of treatment and the end of WWII. A positive  $\Delta$  indicates that people in the treated region exhibit a higher value compared to the control area. Sources are the Observatoire Interrégional du Politique (OIP) 1999 and 2001.

The left-hand side of Figure 8 shows the results within Lorraine, and the right-hand side graph within Alsace and Lorraine combined. The results show several interesting patterns. First, the treatment effect is positive for all age cohorts. Second, it is already positive for the age cohort who began primary schooling prior to 1945, and thus clearly experienced repression themselves. Third, it is strongest for the age cohort who began attending primary school between 1946 and 1964 and declines for later cohorts. Fourth, it remains stable and statistically significant when considering Alsace and Lorraine, but is much smaller for the last two age cohorts within Lorraine. Given that the local dialect is also barely used anymore among younger age cohorts today, our prior results suggest that the stronger presence of regionalist parties in Alsace might explain the stronger persistence when including Alsace.<sup>11</sup>

### 5.3 Effects on policy preferences

In contrast to studies assessing the effect of, for instance, exposure to the rule of law (Lowes et al., 2017), differences in regional identity should not generally result in strong discrepancies in factors like rule-following behavior or risk aversion. We would however, expect differences with regard to regional decision-making and preferences about the allocation about political competences that relate to regional culture. Models on the size of nations like Alesina and Spolaore (1997) suggest that besides economic concerns (Boix et al., 2011; Gehring and Schneider, 2016), the (perceived)

<sup>11</sup> Figure A34 shows similar results when measuring regional identity relative to national identity. Note that a potential dynamic extension of the model, where parents also face a variable cost of teaching with a time-varying  $\alpha$  parameter for the relative return to identity, could explain a decline over time. If parents reduce the value they assign to regional culture over time, it can become no longer optimal to teach it at home even without the fixed costs component: the differences between treated and control area would disappear over time. Reasons could, for instance, be a larger share of children moving out of the region to study or work, increasing the economic returns to national identity.

preference heterogeneity is the major factor influencing preferences about secession or autonomy. We also use the OIP surveys to measure the consequences of a stronger regional identity. Table 9 provides clear evidence that the identity differences in Alsace and Lorraine also affect policy preferences in line with size-of-nation models. People in the treated area feel better informed about regional policies and have a more positive perception of the functioning of regional democratic processes. Asked whether they would be concerned that more regional autonomy would increase inequality between regions, a significantly lower share of people is concerned.

We also create three comprehensive proxy variables regarding the transfer of policy competences to the regional level, more regional autonomy and the allocation of responsibility for education policy. Each proxy is the average of several survey items in the OIP survey, to make sure differences are not caused by different understandings of one particular question. Figures A3 through A6 list the individual questions in each sub-category. The average individual in the treated area favors transferring policy competences from the national to the regional level as well as more regional autonomy significantly more often. Education policy is particularly interesting, as common state education is a major mechanism to impose an identity and influence which and how traditions and culture are taught. Again, treated subject express clearly more favorable views towards setting educational policy and standards at the regional level. Table A30 shows very similar results focusing only on Lorraine.

Table 9: Survey results: policy preferences

Survey question	Mean, control	$\Delta$	P-value	No. obs.
Democracy works well in France	2.536	-0.035	0.324	2606
Democracy works well within region	2.630	0.188	<0.001	2575
Well informed about regional policies	2.704	0.172	<0.001	2604
In favor: transfer policy competence to region (avg. 10)	3.031	0.078	0.002	1218
In favor: allow more autonomy at reg. level (avg. 5)	2.134	0.132	<0.001	2619
Educ. policy should be set at reg. level (avg. 5)	2.855	0.124	0.002	1204
Concerned reg. admin. would increase interreg. inequality	3.208	-0.314	<0.001	1204

*Notes:* Sources are the Observatoire Interrégional du Politique (OIP) 1999, 2001, and 2003, using respondents in all of Alsace and Lorraine, on département level. The Online Appendix shows similar results for within-Lorraine only. The parameter  $\Delta$  comes from the equation:  $y_i = \pi + \Delta Treatment_i + \Gamma_i' \lambda + \eta_i$ , where  $Treatment_i = \mathbf{1}$ [individual in treated region] and  $\Gamma$  comprises of controls for (reported) age, employment status and sex. A positive  $\Delta$  indicates that people in the treated region agree more with the statement. Avg. "x" indicates that the factor is composed of "x" underlying survey items. The underlying survey questions are shown in Table A3.

## 6 Robustness: Alternative explanations

This section discusses alternative explanations to this interpretation, including threats to identification and the interpretation of what constitutes the treatment.

### 6.1 Results are due to linguistic differences

One concern regarding the interpretation of our results is that the border (mostly in Alsace, though) coincides with differences between German (mostly Alsatian and Moselle Franconian) and French dialect speakers. German dialect speakers might develop a stronger regional identity due to the linguistic divide between them and the rest of France, could be exposed to a larger extent to German media, or exhibit different trading patterns (Egger and Lassmann, 2015). Although linguists describe the use of the German Alemannic dialect as steadily declining and now being mostly used by older generations (Vajta, 2013), we trace back the historical language border to separate the treatment effect from linguistic differences. We rely on Harp (1998) and overlay his map with the municipality boundaries to georeference the border along the French municipality boundaries. Figure 2c shows the resulting language border.<sup>12</sup>

To address a potential correlation between spoken (or formerly spoken) dialect and agreement as our proxy for regional identity, we then exclude all German-dialect speaking municipalities and re-estimate the treatment effect within Lorraine. The estimates in Table 10 remain comparable in size and highly significant and reinforce our hypothesis of a persistently stronger regional identity. Accordingly, the results hold even when comparing only directly neighboring municipalities in the same historical region speaking the same dialect.

Table 10: Discontinuities in referenda results, within Lorraine, excluding German-speaking municipalities

Dep. Variable:	Share Yes 1969		Share Yes 1992		Share Yes 2005	
Variable	(1)	(2) <sup>a</sup>	(3)	(4) <sup>a</sup>	(5)	(6) <sup>a</sup>
Treatment	13.019 (2.645) [<0.001]	9.407 (1.960) [<0.001]	4.126 (1.850) [0.026]	4.089 (1.234) [0.001]	3.830 (2.117) [0.071]	5.015 (1.592) [0.002]
Obs.	366	923	385	1269	385	720
Dist	10 km	39.41 km	10 km	61.96 km	10 km	21.48 km

*Notes:* Discontinuity at the treatment border using municipalities in Lorraine, excluding German-speaking municipalities. The outcomes are the share of Yes votes in the 1969 referendum, in the 1992 referendum, and in the 2005 referendum. Conley standard errors in parentheses and *p*-values in brackets.

<sup>a</sup> Estimates from using the optimal IK bandwidth.

<sup>12</sup> See also a similar maps in Callender, 1927. The border was formed in the 8th century and barely moved until the 19th century. Callender (1927, p.430) cites the Count Jean de Pange who traces the border back to barbaric invasions and stated that “in Lorraine the limits of the languages bear no relation to the topography of the country. They form an irregular fringe, [...] these limits, arbitrarily traced by historical accident, have not appreciably altered in fifteen centuries.” We provide the best approximation of the border with the municipality polygons and choose the shortest path around the municipality.

## 6.2 The role of World War II

It is not absolutely clear how to interpret the role of WWII. During most of the war, the treated and untreated area were occupied by Germany. German policies were surely repressive, but the suppression of regional identity and traditions was not the main objective and a potential suppression of French identity took place in all occupied parts of France. Neither the treated nor control area belonged to the self-governed Vichy part of France. We are thus reluctant to emphasize the role of WWII, even though it was clearly a drastic shock influencing the lives of many people.

Nonetheless, one concern is that the shock was stronger in the treated area, as a sizable number of young men were drafted into the German military and exposed to different and potentially more intense war experiences. This difference in exposure probably led to a final phase of perceived alienation and repression, because the French central government sentenced some of these so-called *malgré-nous* who were in the *Waffen-SS* to death in the Bordeaux Trial in 1953 for their involvement in war crimes. This punishment was perceived as unfair and caused massive public outrage and protest, because it did not take the historical circumstances into account.<sup>13</sup> It was probably the last major part of a set of policies which was imposed by the national majority in disregard of the local preferences and opinions. By 1964, all French citizens who had collaborated with the Nazis including the convicts from the Bordeaux trials had benefited from a general amnesty.

Based on the results in Vlachos (2017), using variation within Alsace, the only outcome correlated significantly with a higher share of war veterans is higher support for candidates of the right-wing National Front. As there is no difference in support for nationalist leader Jean-Marie Le Pen, there does not seem to be a problematic discontinuity with regard to WWII exposure at the border we exploit. Finally, the composition of the population might have been affected differently, but Table A17 and A18 indicate no problematic differences.

## 6.3 Migration into and out of the treated area

Another concern is the role of migration to the treated area, and emigration to other parts of France or destinations like the US. Migration mostly happened at two distinct points in time; when Germany annexed the area and when France took it back. First, after 1870, the Germans imposed a requirement that everyone who wanted to remain in the area had to give up her French nationality and opt for German citizenship. Earlier expectations of a large exodus of more than 130,000 people (Vajta, 2013) declined to less than 50,000 when it became clear that this would mean having to leave the region. In addition, Germans migrated or were sent to work in the area between 1870 and WWI. However, as mentioned above, a large share of those immigrants were forced to leave again after the French re-annexation (Harvey, 1999). Nevertheless, a certain share of those Germans or their offspring remain in the area. Conceptually, this should bias against our results as German immigrants are less likely to exhibit a strong Alsatian or Lorrainian identity.

---

<sup>13</sup>Nearly all mayors of towns in Alsace attended a public protest walk in Strasbourg. For alternative versions and views about the actions and historical circumstances see <http://www.scrapbookpages.com/Oradour-sur-Glane/Story/index.html>.

Second, there was a (smaller in magnitude) inflow of French people from other regions after WWI and the re-annexation, to some degree with the purpose to take up posts in the local administration and schools to replace regional traditions and culture with a strong national identity. Again, as these were French citizens from other regions, they should exhibit a weaker regional identity and bias against our results. In terms of migration affecting the composition of the treated and control group, it is reassuring to remember that there are no discontinuities in the socio-economic structure of the population today. Nevertheless, we use a digitized version of census data for the years 1916 to 1946 to compute changes in population at the municipal level. The results in Table A22 show no significant discontinuities for any of those measures at the border. We also employ the population changes as additional control variables in our main specification. Table A23 shows that this does not affect our results.

#### 6.4 Local laws and their effects

The treated areas in Alsace and Lorraine since 1924 enjoy, to a slight and diminishing degree, the freedom to deviate from certain rules imposed by the central state, known as the *local laws*. Glenn (1974, p.772) stated already in the 1970s that “local doctrine is generally of declining importance. There are few, if any, local jurists remaining (...) and the local law is taught only in two or three optional courses (...)”. Moreover, French courts refused to make any reference to German jurisprudence and interpret local laws according to French standards and principles. Accordingly, the visibility of the laws and their potential influence on the salience of regional “uniqueness” was most likely much higher for the first generations after WWII than for more recent generations. Some differences still exist with regard to a small number of welfare policies (including payments to sick employees), which remain more generous in Alsace-Lorraine and include two additional days of vacation. Other differences exist with regard to personal bankruptcy law and voluntary associations.

The sheer existence of this set of local rules can work as a mechanism to maintain regional identity. In terms of our model, they could increase the salience of items that all people in the treated area have in common. A potential concern would be if the local laws decisively influence a third factor that drives the measured differences in regional identity and is unrelated to occupation and the suppression of group identity. To test the extent to which the remaining exceptions led to potentially problematic differences in the socio-economic environment, we run RD regressions on all variables for which we possess measures at the local level and that could plausibly be influenced by the local laws. This includes items in the categories work occupations, economic activity, public goods and population density. In a second step, we assess how these are correlated with our main outcome in the RDD. Table A18 shows that for the about 25 tests of covariates, only one turns out to be significant when using the 10 km bandwidth: There seems to be a somewhat smaller number of industrial companies in the treated area. The last two columns show that industrial companies are positively correlated with agreement in the referenda. Accordingly, while the one significant difference is probably coincidence, it would bias against our main results.<sup>14</sup>

---

<sup>14</sup> Another potentially biasing factor in the referenda could be differences in European Union fund receipts if the

## 6.5 Other

We discuss four alternative explanations in more detail in the Online Appendix C to save space. Outliers within Lorraine, more specifically, the large urban agglomeration of Metz could be an issue as historically, cities enjoyed greater autonomy and might have developed a stronger identity. Moreover, people residing in cities are often diverse and likely to support more European integration for reasons unrelated to regional identity. Even though we already control for distance to major cities, we also show that excluding municipalities belonging to the metropolitan area of Metz does not affect our most conservative estimates within Lorraine.

Another concern and one distinct feature in which the local laws differ from the rest of France is religion. Historically, the church played a larger role in the average citizens life in the treated area until after WWI and still does to some degree until today. In contrast to the rest of France, pupils in the area are still subjected to compulsory religious classes at school (usually two hours per week). We show that in France (for both referenda) there is no relationship between religiosity as well as religious denomination and regional identity or support for the European Union. We also explain why differences in the benefits from trade are not a plausible explanation. Finally, discuss the relative importance of homogenization policies in strengthening regional identity, and the unsuccessful attempts to "Germanize" of the treated individuals

## 7 Concluding remarks

Our paper uses a unique natural experiment that offers variation in the exposure to occupation and the suppression of regional identity within historically homogeneous regions. The setting allows us to observe both treated and control area today in the same institutional environment. To the best of our knowledge, this is the first causal evidence of occupation and the often associated repression and forced assimilation on the identity of the suppressed group in their home region. Studying minority groups within their home region in larger nation states is relevant for a large range of regions like the Kurdish part of Turkey, the Xinjiang Uyghur region in China, Chechnya and Crimea in Russia, but also minority regions in established democracies like the Basque country or Catalonia.

Our results clearly show that regional identity, both using stated and revealed preferences, is stronger today in the treated part of the region. This is in line with and complements evidence by Fouka (2016) on the negative effect of intrusive homogenization policies on German immigrants in the United States. In contrast to her study, we can observe not only immigrants as a selected share but the average of the population in their home region. We show historical evidence that conscious

---

treated area would receive significantly more money which could directly affect the likelihood to vote yes or indirectly through potential growth effects (Becker et al., 2010). However, the funds are allocated to regions, not départements (the respective categories in the 2014-2020 period are "Lorraine et Vosges - ERDF/ESF" and "Lorraine - Rural Development"). The whole region is responsible for the within-region allocation and there is no reason to assume that municipalities just right of the former border in the treated area would be awarded more funds. In the 2007-2013 period, neither Lorraine nor Alsace were eligible under the convergence, competitiveness or employment objective. For the 2000-2006 period receipts per capita in the treated part of Alsace Lorraine were 100€ compared to 180€ in the untreated area.

investments in regional identity in the form of newspapers and parties began already during the repressive period. Our data then allows us to trace the medium term effect in 1969, about 15 years after the end of the treatment, as well as about half a century later. Our survey data also suggest that a positive treatment effect is already visible for age cohorts who were themselves exposed, as well as for later generations.

Our evidence on potential mechanisms suggest that regionalist parties played an important role, somehow more so in Alsace than in Lorraine. Within Lorraine, we show that in treated municipalities households more often subscribe to a regional newspaper, which signals regional attachment and provides information about regional traditions and culture to parents and children. Moreover, we show that a stronger regional identity has important policy implications in line with size-of-nations models. Treated individuals express more satisfaction with regional democracy, and would prefer more regional-decision making and a shift of policy competences about policies like education to the regional level.

What can we learn from these results for policies and future research? It is important to take into account to what degree identities constitute substitutes and are perceived as (mis-)aligned. Our study demonstrates that people with a stronger regional identity do not necessarily have a weaker national identity. We show how this can be integrated into theoretical models using our adapted conceptualization of common identity, which relies on the salience or weights put on attributes that an individual has in common with the rest of the group. This definition can also explain why there are strong existing group identities even though actual within group heterogeneity is larger than between group differences (Desmet et al., 2017). When people hold multiple identities, whether the state can impose a new identity depends on the degree to which it is perceived as oppositional to the existing identity (relating to, e.g., Benjamin et al., 2010; Carvalho and Koyama, 2016).

The results are also important for analyzing separatism and the number and size of nations (Alesina and Spolaore, 1997), where separatist tendencies are explained by economic (e.g. regional resources Gehring and Schneider, 2016) and cultural reasons relating to preference heterogeneity. We argue that a common group identity is best modeled as *perceived* preference homogeneity. Cases like Catalonia, where central government policies are perceived as discriminatory or repressive towards a particular region and fuel existing separatist tendencies suggest a similar mechanism.

Finally, it is important to stress that the strengthening of group identity is not necessarily the deterministic outcome or natural reaction to suppressive policies. Our model provides some guidance in that respect. Whether parents or generally members of the suppressed groups are willing to invest in the skills to maintain their traditions depends on the relative utility they derive from their group identity and from an overarching common national identity. Policies can be so intrusive or the disadvantages of not teaching children the national identity instead can be so high that existing group identities disappear. To end on a more positive note, our results suggest that a joint identity embracing existing groups can be built up without necessarily replacing existing identities. This, however, requires the central authority to accept regional identities and an institutional setup allowing sufficient regional autonomy.



## References

- AKERLOF, G. A. AND R. E. KRANTON (2000): “Economics and Identity,” *The Quarterly Journal of Economics*, 115, 715–753.
- ALESINA, A. AND B. REICH (2018): “Nation Building and Education,” *mimeo*.
- ALESINA, A. AND E. SPOLAORE (1997): “On the Number and Size of Nations,” *The Quarterly Journal of Economics*, 112, 1027–1056.
- ANDERSON, M. (1972): “Regional Identity and Political Change: The Case of Alsace from the Third to the Fifth Republic,” *Political Studies*, 20, 17–30.
- ASPACHS-BRACONS, O., I. CLOTS-FIGUERAS, J. COSTA-FONT, AND P. MASELLA (2008): “Compulsory Language Educational Policies and Identity Formation,” *Journal of the European Economic Association*, 6, 434–444.
- BANDIERA, O., M. MOHNEN, I. RASUL, AND M. VIARENGO (2017): “Nation-Building Through Compulsory Schooling During the Age of Mass Migration,” *mimeo*.
- BECKER, S. O., K. BOECKH, C. HAINZ, AND L. WOESSMANN (2015): “The Empire is Dead, Long Live the Empire! Long-run Persistence of Trust and Corruption in the Bureaucracy,” *The Economic Journal*.
- BECKER, S. O., P. H. EGGER, AND M. VON EHRLICH (2010): “Going NUTS: The Effect of EU Structural Funds on Regional Performance,” *Journal of Public Economics*, 94, 578–590.
- BÉNABOU, R. AND J. TIROLE (2011): “Identity, Morals, and Taboos: Beliefs as Assets,” *The Quarterly Journal of Economics*, 126, 805–855.
- BENJAMIN, D. J., J. J. CHOI, AND A. J. STRICKLAND (2010): “Social Identity and Preferences,” *American Economic Review*, 100, 1913–1928.
- BESLEY, T. AND M. REYNAL-QUEROL (2014): “The Legacy of Historical Conflict: Evidence from Africa,” *American Political Science Review*, 108, 319–336.
- BISIN, A., E. PATAACCHINI, T. VERDIER, AND Y. ZENOU (2011): “Formation and Persistence of Oppositional Identities,” *European Economic Review*, 55, 1046–1071.
- BISIN, A. AND T. VERDIER (2000): “A Model of Cultural Transmission, Voting and Political Ideology,” *European Journal of Political Economy*, 16, 5–29.
- (2010): “The Economics of Cultural Transmission and Socialization,” *Handbook of Social Economics*, 1, 339–416.
- BOIX, C., B. CODENOTTI, AND G. RESTA (2011): “War, Wealth and the Formation of States,” in *Political Economy of Institutions, Democracy and Voting*, Springer, 45–74.
- BON, F. (1970): “Le Référendum du 27 avril 1969: Suicide Politique ou Necessité Stratégique?” *Revue Française de Science Politique*, 20.
- BORDALO, P., K. COFFMAN, N. GENNAIOLI, AND A. SHLEIFER (2016): “Stereotypes,” *The Quarterly Journal of Economics*, 131, 1753–1794.
- BRANSCOMBE, N. R., M. T. SCHMITT, AND R. D. HARVEY (1999): “Perceiving Pervasive Discrimination among African Americans: Implications for Group Identification and Well-being,” *Journal of Personality and Social Psychology*, 77, 135.
- CALLENDER, H. (1927): “Alsace-Lorraine Since the War,” *Foreign Affairs*, 5, 427–437.
- CARROL, A. (2010): “Socialism and National Identity in Alsace from Reichsland to République, 1890-1921,” *European History Quarterly*, 40, 57–78.

- CARROL, A. AND L. ZANOUN (2011): "The View from the Border: A Comparative Study of Autonomism in Alsace and the Moselle, 1918–29," *European Review of History*, 18, 465–486.
- CARVALHO, J.-P. (2013): "Veiling," *The Quarterly Journal of Economics*, 128, 337–370.
- CARVALHO, J. P. AND M. KOYAMA (2016): "Resisting Education," *GMU Working Paper in Economics No. 13-18*.
- CEMR (2004): "Hearing on the Draft Treaty Establishing a Constitution for Europe in the European Parliament's Committee on Constitutional Affairs," Tech. rep., European Associations representing regional and local authorities (AEBR, AER, CEMR, CPMR, EUROCITIES, CALRE and REGLEG).
- CHACHA, M. (2013): "Regional Attachment and Support for European Integration," *European Union Politics*, 14, 206–227.
- CLOTS-FIGUERAS, I. AND P. MASELLA (2013): "Education, Language and Identity," *The Economic Journal*, 123, F332–F357.
- CONVERSI, D. (2004): *Ethnonationalism in the Contemporary World: Walker Connor and the Study of Nationalism*, vol. 21, Psychology Press.
- (2008): "'We Are All Equals!' Militarism, Homogenization and 'Egalitarianism' in Nationalist State-building (1789-1945)," *Ethnic and Racial Studies*, 31, 1286–1314.
- DELL, M. (2010): "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 78, 1863–1903.
- DELL, M. AND P. QUERUBIN (2017): "Nation Building Through Foreign Intervention: Evidence from Discontinuities in Military Strategies," *Quarterly Journal of Economics*, 1, 64.
- DEPETRIS-CHAUVIN, E., R. DURANTE, AND F. R. CAMPANTE (2018): "Building Nations Through Shared Experiences: Evidence from African Football," .
- DESMET, K., I. ORTUÑO-ORTÍN, AND R. WACZIARG (2017): "Culture, Ethnicity and Diversity," *American Economic Review*, 107(9), 2479–2513.
- DOEPKE, M. AND F. ZILIBOTTI (2017): "Parenting With Style: Altruism and Paternalism in Intergenerational Preference Transmission," *Econometrica*, 85, 1331–1371.
- DOHMEN, T., A. FALK, D. HUFFMAN, AND U. SUNDE (2012): "The Intergenerational Transmission of Risk and Trust Attitudes," *The Review of Economic Studies*, 79, 645–677.
- EGGER, P. H. AND A. LASSMANN (2015): "The Causal Impact of Common Native Language on International Trade: Evidence from a Spatial Regression Discontinuity Design," *The Economic Journal*, 125, 699–745.
- FITJAR, R. D. (2010): "Explaining Variation in Sub-state Regional Identities in Western Europe," *European Journal of Political Research*, 49, 522–544.
- FÖRSTER, S. (1990): "General-Feldmarschall Helmuth von Moltke und die Kriegsgeschichtlichen Folgen des Deutsch-französischen Krieges," *Pariser Historische Studien*, 29, 86–91.
- FOUKA, V. (2016): "Backlash: The Unintended Effects of Language Prohibition in US Schools After World War I," *Stanford Center for International Development Working Paper No. 591*.
- (2017): "How do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I," Tech. rep., Mimeo.
- FRANKLIN, M. N. (2004): *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies Since 1945*, Cambridge University Press.

- FRYER JR., R. G. AND P. TORELLI (2010): “An Empirical Analysis of Acting White,” *Journal of Public Economics*, 94, 380 – 396.
- GEHRING, K. AND S. SCHNEIDER (2016): “Regional Resources and Democratic Secessionism,” *ETH and University of Zurich CIS Working Paper Series No. 90*.
- GELLNER, E. AND J. BREUILLY (2008): *Nations and Nationalism*, Cornell University Press.
- GELMAN, A. AND G. IMBENS (2017): “Why High-order Polynomials Should not be Used in Regression Discontinuity Designs,” *Journal of Business & Economic Statistics*, (forthcoming).
- GENNAIOLI, N. AND I. RAINER (2007): “The Modern Impact of Precolonial Centralization in Africa,” *Journal of Economic Growth*, 12, 185–234.
- GIULIANO, P. AND N. NUNN (2016): “Understanding Cultural Persistence and Change,” *NBER Working Paper No. 23617*.
- GLENN, H. P. (1974): “Local Law of Alsace-Lorraine, A Half Century of Survival,” *Int'l & Comp. LQ*, 23, 769–781.
- GOODFELLOW, S. (1993): “From Germany to France? Interwar Alsatian National Identity,” *French History*, 7, 450–471.
- GRASSER, J.-P. (1998): *Une Histoire de l'Alsace*, vol. 1, Éditions Jean-Paul Gisserot.
- GROSJEAN, P. (2014): “A History Of Violence: The Culture Of Honor And Homicide In The US South,” *Journal of the European Economic Association*, 12, 1285–1316.
- GUISSO, L., P. SAPIENZA, AND L. ZINGALES (2016): “Long-term Persistence,” *Journal of the European Economic Association*, 14, 1401–1436.
- HARP, S. L. (1998): *Learning to be Loyal: Primary Schooling as Nation Building in Alsace and Lorraine, 1850-1940.*, De Kalb: Northern Illinois University Press.
- HARVEY, D. A. (1999): “Lost Children or Enemy Aliens? Classifying the Population of Alsace After the First World War,” *Journal of Contemporary History*, 34, 537–554.
- HIERY, H. (1870): “Wahlen und Wahlverhalten im Reichsland Elsaß-Lothringen,” *Angelo Ara/Eberhard Kolb (Hg.): Grenzregionen im Zeitalter der Nationalismen: Elsaß-Lothringen/Trient-Triest*, 1914, 71–89.
- HOBBSBAWM, E. (1994): “The Nation as an Invented Tradition,” in *Nationalism*, ed. by J. H. . A. D. S. (Eds.), Oxford: Oxford University Press, 76–82.
- HÖPEL, T. (2012): “The French-German Borderlands: Borderlands and Nation-building in the 19th and 20th Centuries,” *European History Online (EGO)*.
- HYSLOP, B. F. (1968): *French Nationalism in 1789, According to the General Cahiers*, New York: Octagon Books, 1968 [c1934].
- IMBENS, G. AND K. KALYANARAMAN (2012): “Optimal Bandwidth Choice for the Regression Discontinuity Estimator.” *The Review of Economic Studies*, 79, 933 – 959.
- JOHNSON, N. D. (2015): “Taxes, National Identity, and Nation Building: Evidence from France, mimeo,” *GMU Working Paper in Economics No. 15-33*.
- JOLLY, S. K. (2007): “The Europhile Fringe? Regionalist Party Support for European Integration,” *European Union Politics*, 8, 109–130.
- KRANTON, R. E. (2016): “Identity Economics 2016: Where Do Social Distinctions and Norms Come From?” *American Economic Review: Papers and Proceedings*, 106, 405–09.

- LEED, E. J. (1981): *No Man's Land: Combat and Identity in World War 1*, Press Syndicate of the University of Cambridge.
- LIPGENS, W. (1964): "Bismarck, die Öffentliche Meinung und die Annexion von Elsaß und Lothringen 1870," *Historische Zeitschrift*, 199, 31–144.
- LOTT, JR, J. R. (1999): "Public Schooling, Indoctrination, and Totalitarianism," *Journal of Political Economy*, 107, 127–157.
- LOWES, S., N. NUNN, J. A. ROBINSON, AND J. L. WEIGEL (2017): "The Evolution of Culture and Institutions: Evidence From the Kuba Kingdom," *Econometrica*, 85, 1065–1091.
- MESSERSCHMIDT, M. (1975): *Militär und Politik in der Bismarckzeit und im Wilhelminischen Deutschland*, vol. 43, Wiss. Buchges.
- MICHALOPOULOS, S. AND E. PAPAIOANNOU (2014): "National Institutions and Subnational Development in Africa," *The Quarterly Journal of Economics*, 129, 151–213.
- (2016): "The Long-run Effects of the Scramble for Africa," *American Economic Review*, 106, 1802–1848.
- NUNN, N. AND D. PUGA (2012): "Ruggedness: The Blessing of Bad Geography in Africa," *Review of Economics and Statistics*, 94, 20–36.
- NUNN, N. AND L. WANTCHEKON (2011): "The Slave Trade and the Origins of Mistrust in Africa," *The American Economic Review*, 101, 3221–3252.
- OCHSNER, C. AND F. ROESEL (2017): "Activating History – The Turkish Sieges of Vienna, Anti-Turkish Campaigns, and the Rise of Right-wing Populism," *mimeo*.
- ORTEGA, J. AND T. P. TANGERÅS (2008): "Unilingual versus Bilingual Education: a Political Economy Analysis," *Journal of the European Economic Association*, 6, 1078–1108.
- ROHNER, D., M. THOENIG, AND F. ZILIBOTTI (2013): "Seeds of Distrust: Conflict in Uganda," *Journal of Economic Growth*, 18, 217–252.
- ROTHENBERGER, K.-H. (1975): *Die Elsass-lothringische Heimat- und Autonomiebewegung Zwischen den Beiden Weltkriegen*, vol. 42, Herbert Lang.
- SCOTT, A., J. PETERSON, AND D. MILLAR (1994): "Subsidiarity: A Europe of the Regions v. the British Constitution?" *Journal of Common Market Studies*, 32, 47–67.
- SHAYO, M. (2009): "A Model of Social Identity With an Application to Political Economy: Nation, Class, and Redistribution," *American Political Science Review*, 103, 147–174.
- SILVERMAN, D. P. (1966): "Political Catholicism and Social Democracy in Alsace-Lorraine, 1871-1914," *The Catholic Historical Review*, 52, 39–65.
- TABELLINI, G. ET AL. (2008): "The Scope of Cooperation: Values and Incentives," *The Quarterly Journal of Economics*, 123, 905–950.
- TAJFEL, H., M. G. BILLIG, R. P. BUNDY, AND C. FLAMENT (1971): "Social Categorization and Intergroup Behaviour," *European Journal of Social Psychology*, 1, 149–178.
- TILLY, C. (1975): "Reflections on the History of European State-making," *The Formation of National States in Western Europe*, 38.
- TURNER, J. C. (1982): "Towards a Cognitive Redefinition of the Social Group," *Social Identity and Intergroup Relations*, 15–40.
- VAJTA, K. (2013): "Linguistic, Religious and National Loyalties in Alsace," *International Journal of the Sociology of Language*, 2013, 109–125.

VLACHOS, S. (2017): “Inglorious Wars and Political Radicalisation,” *mimeo*.

VOIGTLÄNDER, N. AND H.-J. VOTH (2012): “Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany,” *The Quarterly Journal of Economics*, 127, 1339–1392.

WAWRO, G. (2005): *The Franco-Prussian War: The German Conquest of France in 1870-1871*, Cambridge University Press.

ZANOUN, L. (2009): “Interwar Politics in a French Border Region: the Moselle in the Period of the Popular Front, 1934-1938,” Ph.D. thesis, The London School of Economics and Political Science (LSE).