Appendices

for Wimmer, Lee, and LaViolette, "Diffusion through multiple domains. The spread of romantic nationalism across Europe, 1770-1930", in *American Journal of Sociology*, forthcoming

Table of Content

Appendix A. Data sources, descriptive statistics, and illustrative figures

Appendix A Table A1. Data sources and coverage

Variable	Type of variable	Sources	Temporal coverage
Number of nationalist works produced in a city	Dependent	ERNiE	1770-1920
Number of nationalist works produced by an artist/writer	Dependent	ERNiE	1770-1920
Number of first nationalist works by artists in a city	Depend. for robustness	ERNiE	1770-1920
Protestant majority dummy	Independent	Times (1900); Andrees (1887)	Ca. 1900
Majority language	Independent	Rand McNally Atlas (1897); Troinitskii (1905)	Ca. 1900
Majority religion	Independent	Times (1900); Andrees (1887)	Ca. 1900
Foreign ruled dummy	Independent	Wimmer (2023) and new coding	1770-1920
Ruled by Napoleon	Independent	Acemoglu et al. (2011)	1770-1920
Number of letters received in previous decade	Independent	ERNiE	1770-1920
Number of nationalist writings near letter sender	Independent	ERNiE	1770-1920
Distance to universities with prev. nat. prod.	Independent	Rüegg and Briggs (1996; 2004); ERNiE	1770-1920
Distance to universities	Independent	Rüegg and Briggs (1996; 2004); ERNiE	1770-1920
Distance to newspapers with previous nationalist production	Independent	Zeitschriftendatenbank of the German National Library, ERNiE	1770-1920
Distance to newspapers	Independent	Zeitschriftendatenbank of the German National Library	1770-1920
Distance to nationalist production on Roman roads	Independent	McCormick et al (2013), ERNiE	1770-1920
Number of nationalist works in Roman road cluster	Independent	McCormick et al (2013), ERNiE	1770-1920

|			
Distance to nationalist production on stagecoach roads	Independent	Franz Güssefeld (1793); Auguste-Henri Dufour (1848), ERNiE	1793, 1848
Distance to nationalist production on railway roads	Independent	Martí-Henneberg (2013); Cima (1998-2008); ERNiE	1770-1920
Political instability dummy	Independent for alt. expl.	Wimmer (2023); PolityIV (Marshall et al. 2017)	1770-1920
Lost independent statehood dummy	Independent for alt. expl.	Wimmer (2023)	1770-1920
Distance to battle fields	Independent for alt. expl.	Wimmer (2023)	1770-1920
Distance to coal mining	Independent for alt. expl.	Fernihough and O'Rourke (2021)	Time invariant
Distance to textile production	Independent for alt. expl.	International Committee for the Conservation of the Industrial Heritage (2013)	Time invariant
Industrializing region dummy	Independent for alt. expl.	Pollard (1981)	1815; 1875
Population size	Control	Bosker et al. (2013)	1770-1920 (interpolated)
Capital city dummy	Control	Bosker et al. (2013)	1770-1920 (interpolated)
Bishop seat dummy	Control	Bosker et al. (2013)	1770-1920 (interpolated)
Catholic majority city dummy	Control	Times (1900); Andrees (1887)	Ca. 1900
Distance to the nearest river	Control	Global Runoff Data Centre	Time invariant
Distance to the nearest sea	Control	Patterson and Kelso (2012)	Time invariant
Distance to the nearest renowned artist/writer	Control	Wikipedia	1770-1920
Number of renowned artists/writers in a city	Control for robustness	Wikipedia	1770-1920
Dummies for five most frequent letter writers	Control for robustness	ERNiE	1770-1920
Number of letters received prior to previous decade	Control for robustness	ERNiE	1770-1920
Number of letters sent from a city	Control for robustness	ERNiE	1770-1920
Number of letters sent from and received in the city	Control for robustness	ERNiE	1770-1920

Note: For references see the bibliography in the main text

Descriptive statistics are presented in Appendix A Table 2 and Appendix A Figure 1. Appendix Table 2 shows the distributions of categorical variables and Appendix A Figure 1 shows the histograms of continuous variables. The spatial distributions of some of the key variables are shown in Appendix A Figures 2 to 5.

Appendix A Table A2. Descriptive statistics of categorical variables

Variable	No		Yes		Total
Capital city	35,776	(98.50%)	544	(1.50%)	36,320
Bishop city	31,648	(87.14%)	4,672	(12.86%)	36,320
Catholic city	15,392	(42.38%)	20,928	(57.62%)	36,320
Protestant city	27,984	(77.05%)	8,336	(22.95%)	36,320
Received a letter from a nationalist	35,292	(97.17%)	1,028	(2.83%)	36,320
Roman empire city	10,992	(30.26%)	25,328	(69.74%)	36,320
Napoleon					
- Country not occupied by Napoleon	25,311	(69.69%)	11,009	(30.31%)	36,320
- Country occupied; city not occupied	24,875	(68.49%)	11,445	(31.51%)	36,320
- Both city and country occupied	27,398	(75.44%)	8,922	(24.56%)	36,320
- French city	31,376	(86.39%)	4,944	(13.61%)	36,320
Foreign ruled	23,367	(64.34%)	12,953	(35.66%)	36,320

Appendix A Figure A1. Distributions of continuous variables
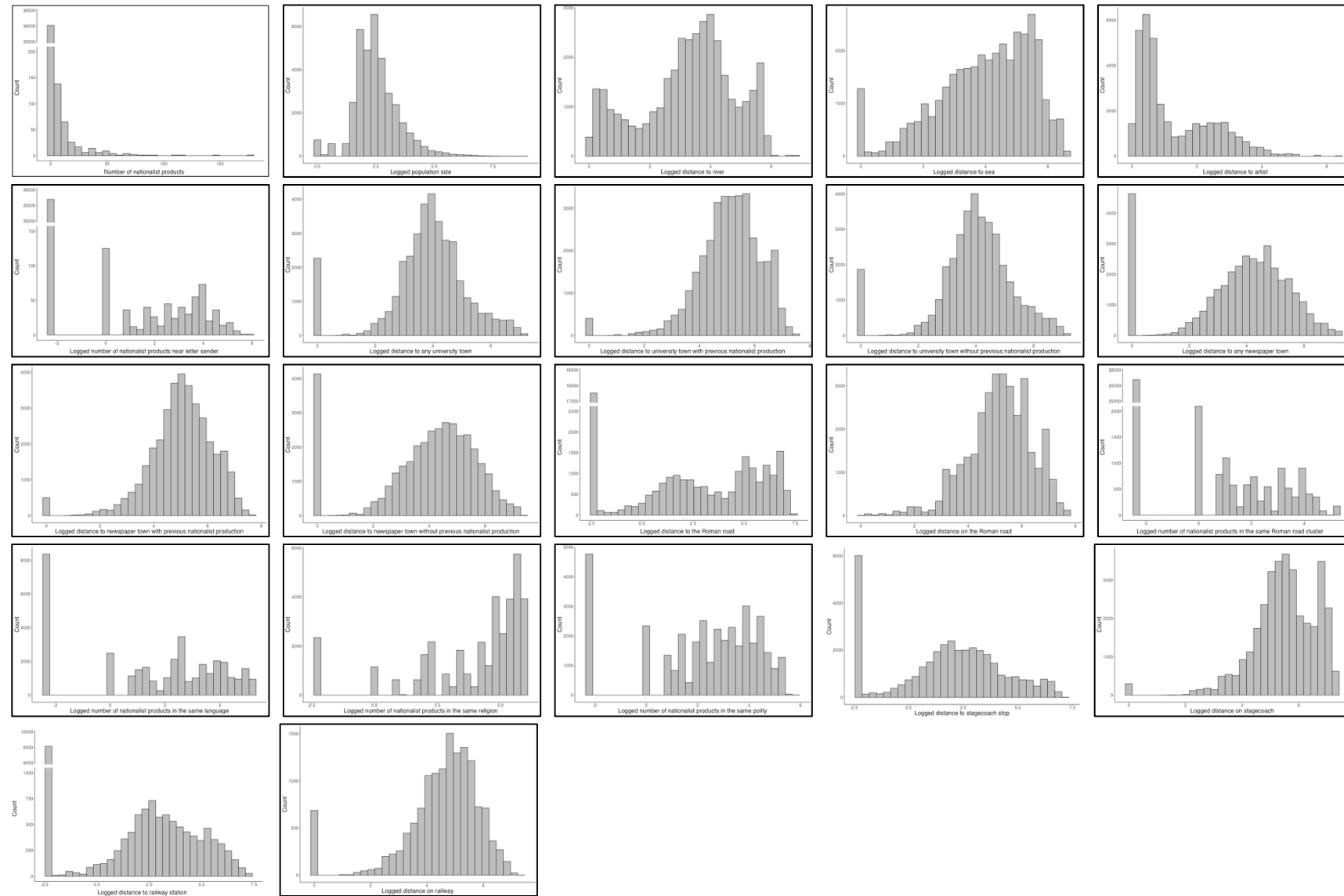
Appendix A Figure A2. Spatial distribution of the Romantic nationalist art works
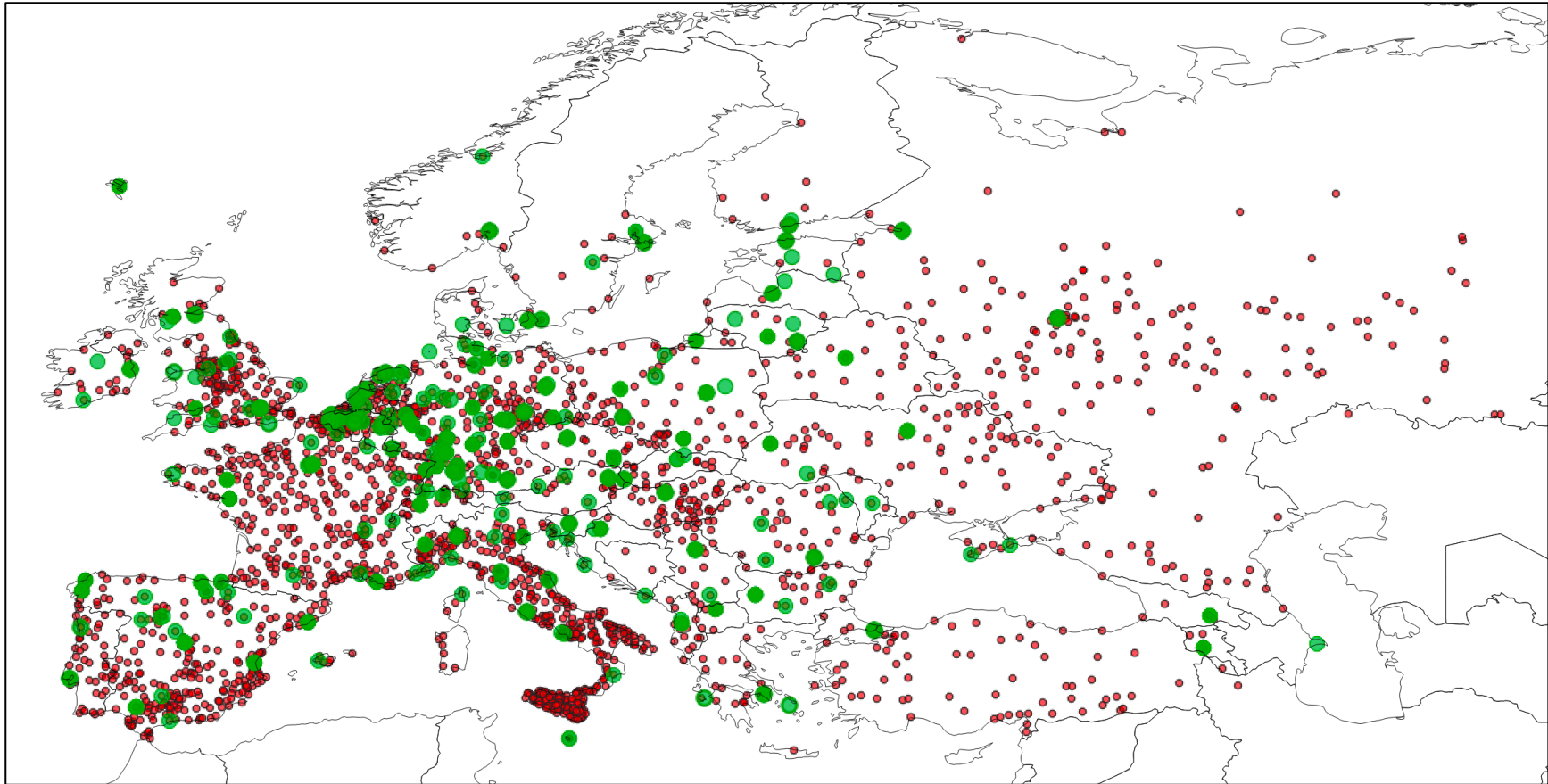
Note: Red dots represent cities and green circles represent nationalist works.

Appendix A Figure A3. Spatial distribution of Roman roads
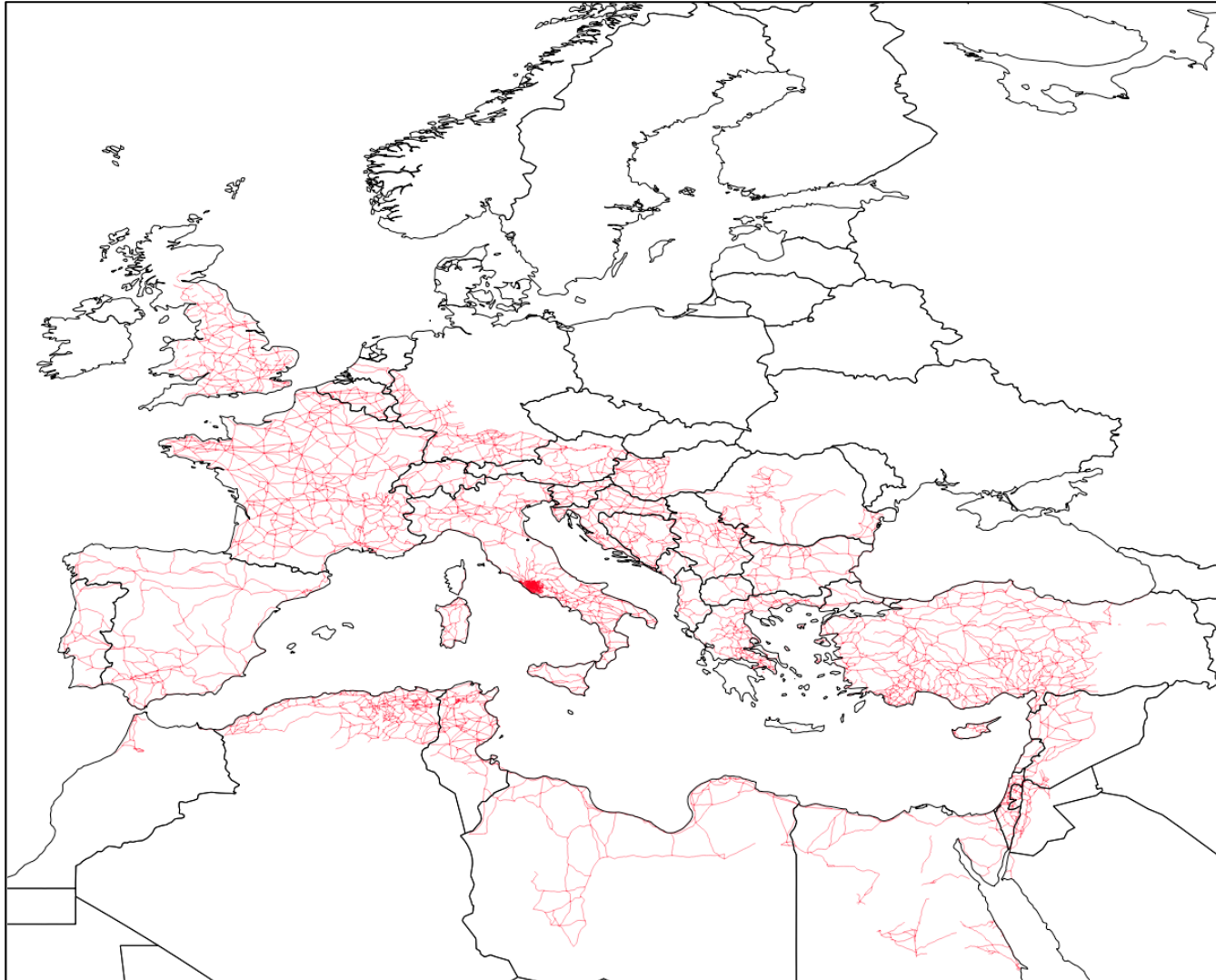
Appendix A Figure A4. Spatial distribution of stagecoach routes

Stagecoach routes in 1793
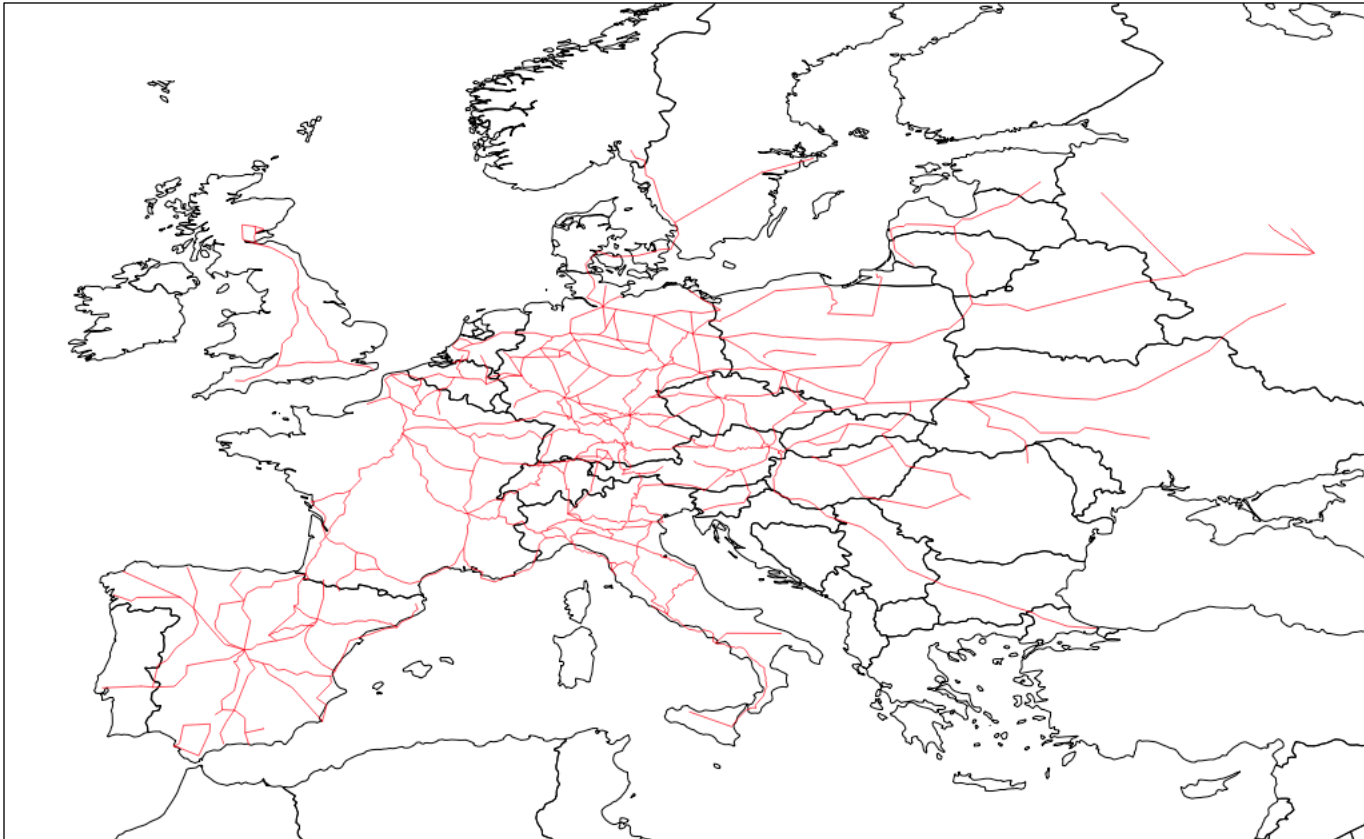
Stagecoach routes in 1848

Appendix A Figures A5. Spatial distribution of railways in different periods

Railways in the 1860s

Railways in the 1870s

Railways in the 1880s

Railways in the 1890s

Railways in the 1900s
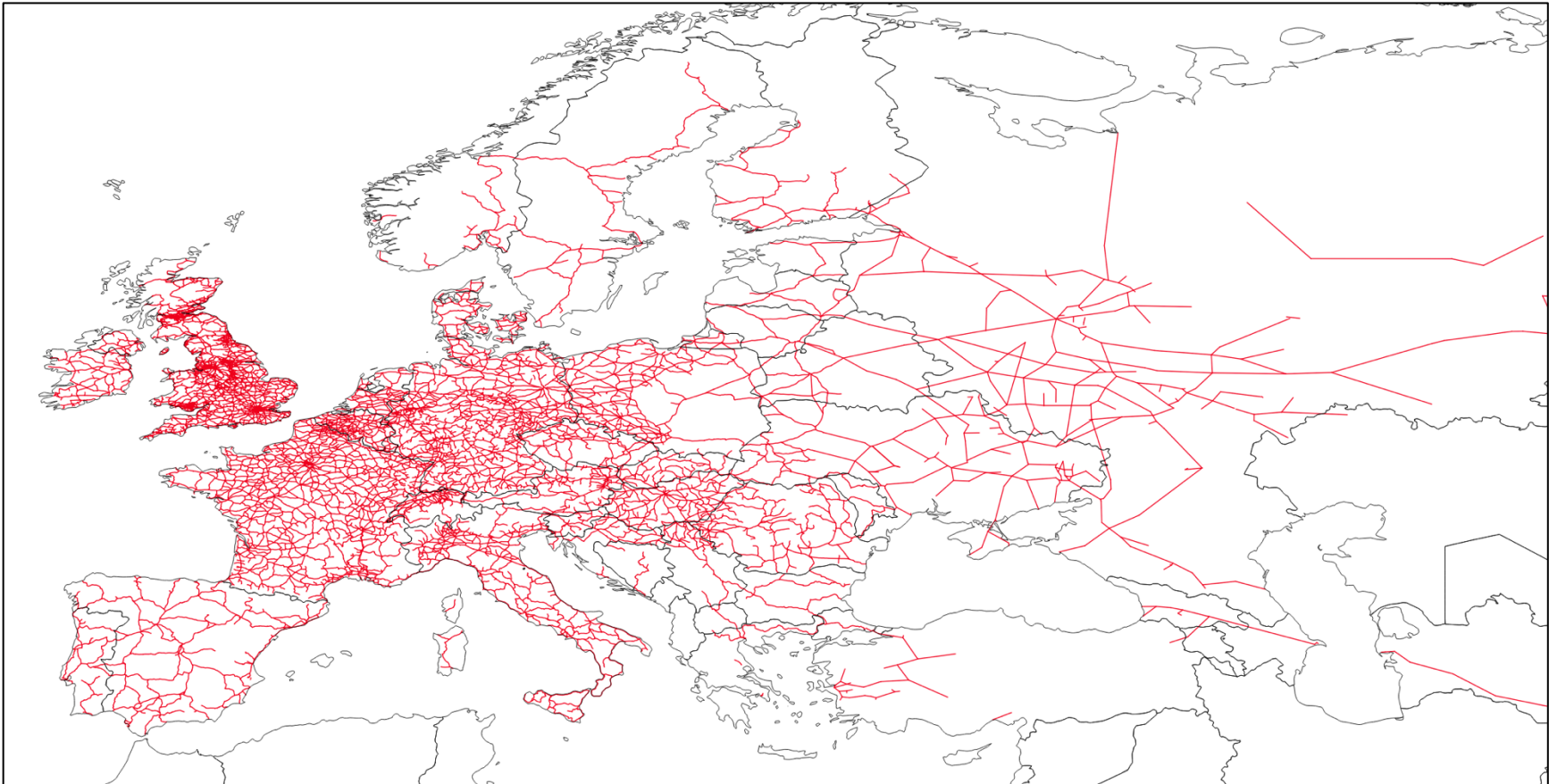
Railways in the 1910s
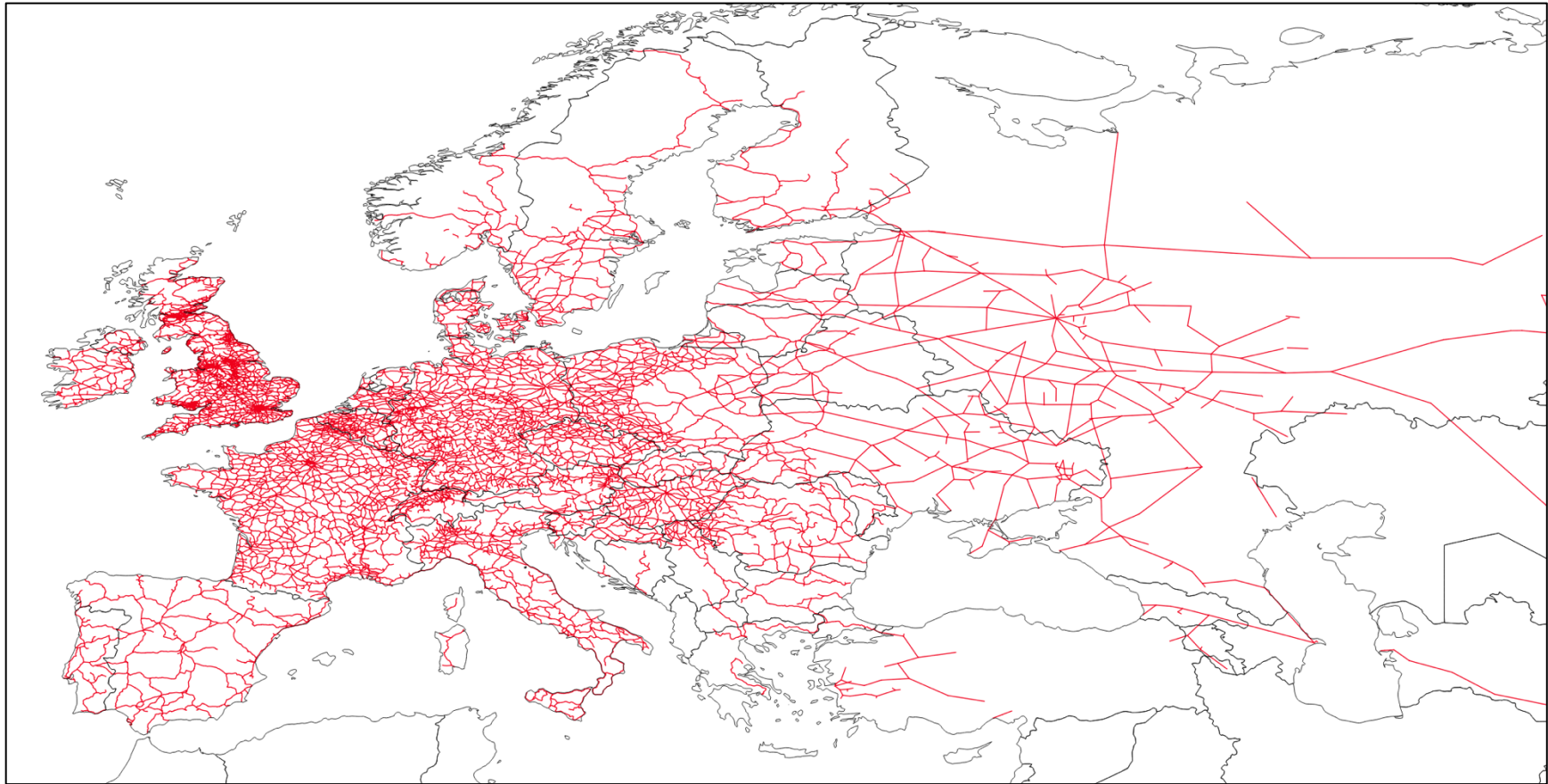
Railways in the 1920s
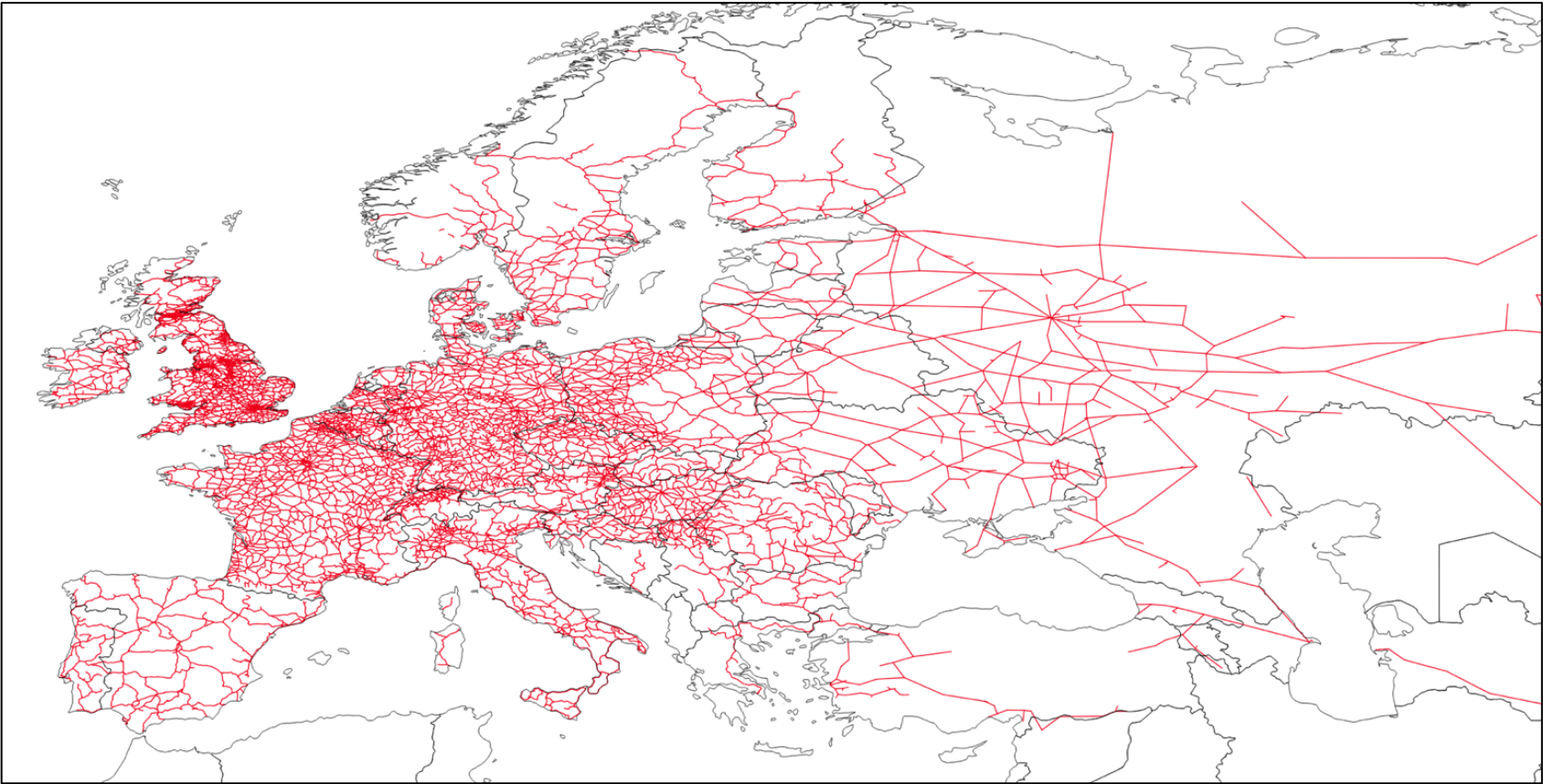
Appendix B. Robustness tests for shared language, religion, and polity

In total, 57 languages were identified on the basis of the two maps. They can be subsumed into 16 language families. Languages and families are listed below.
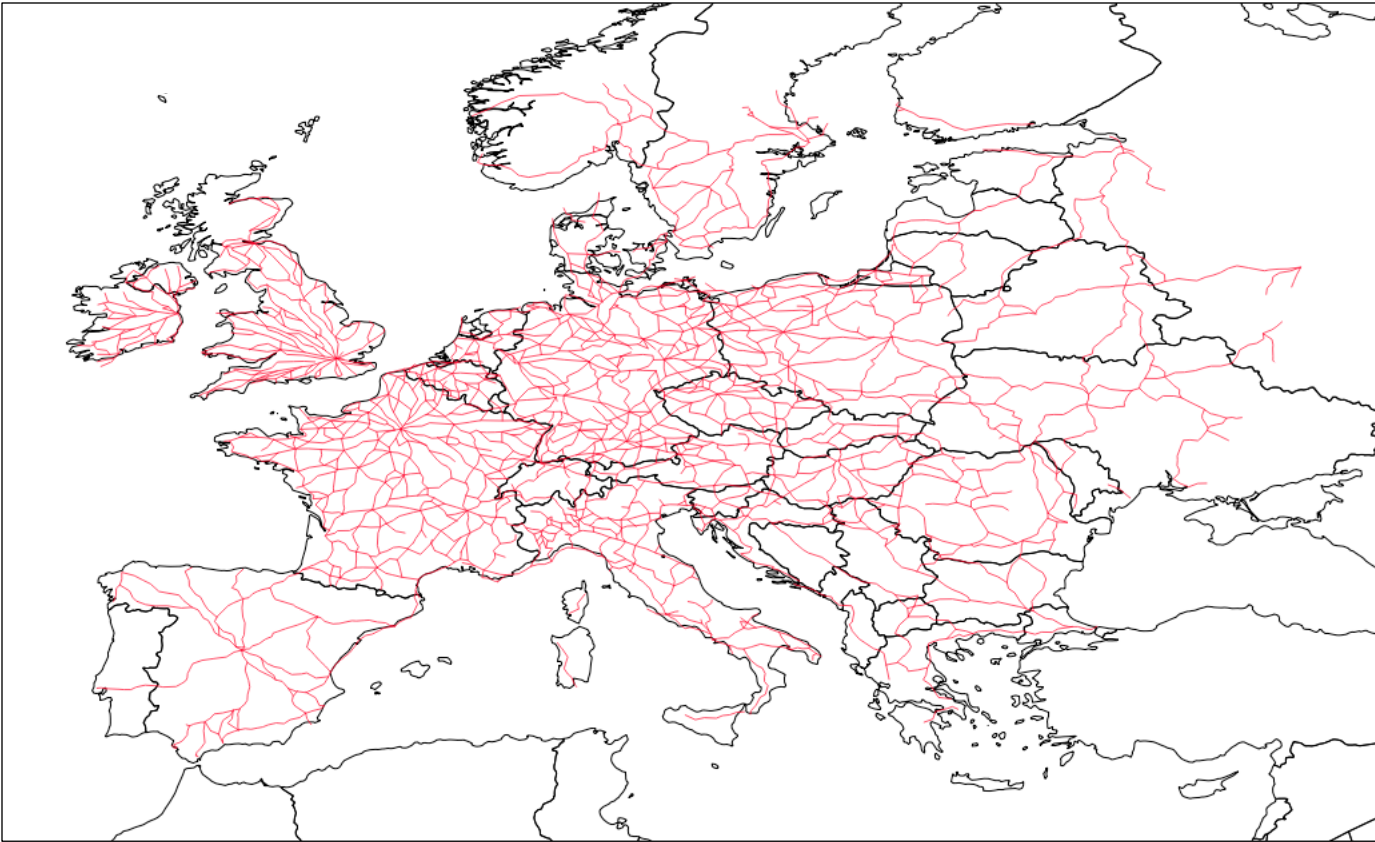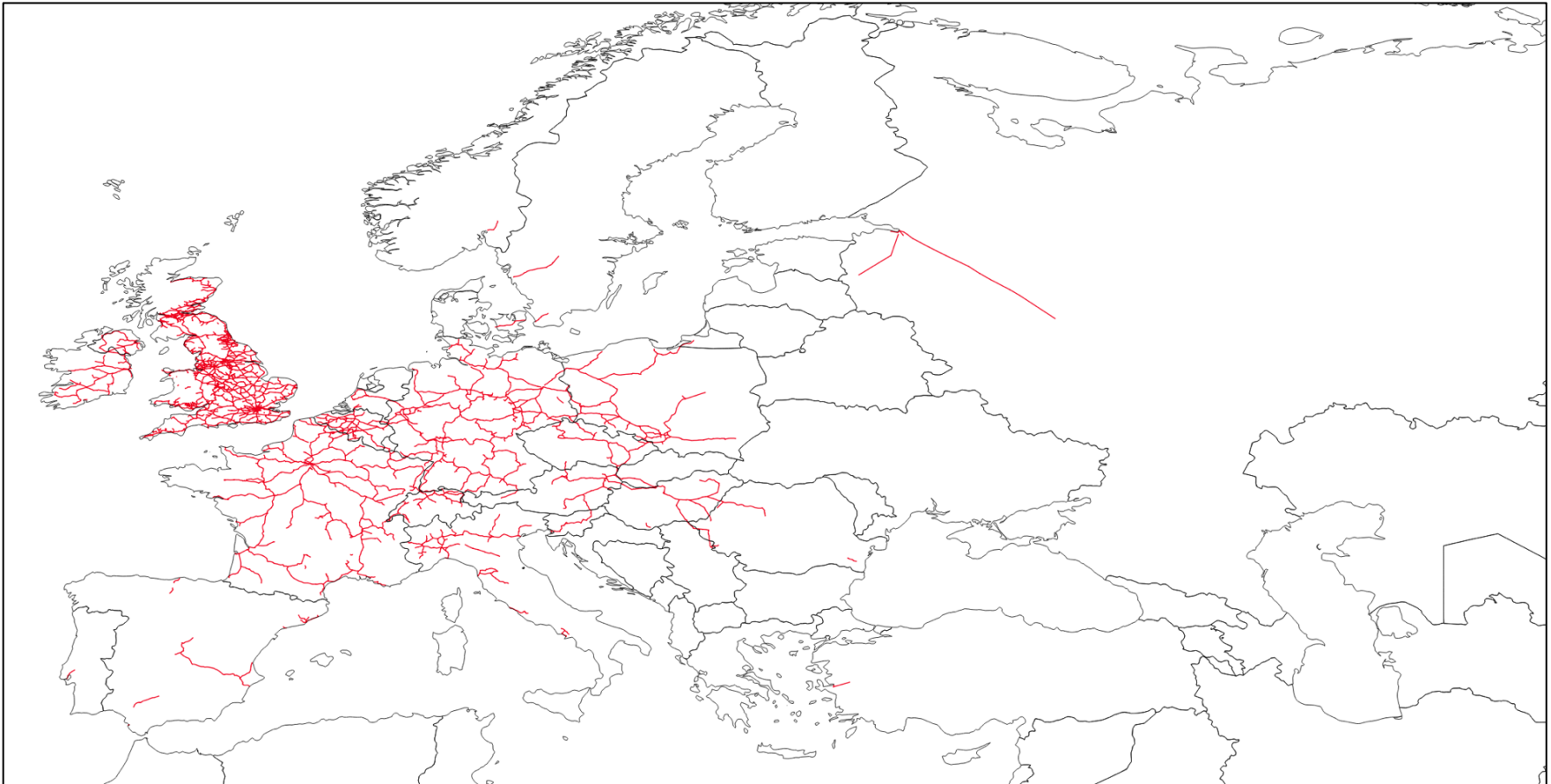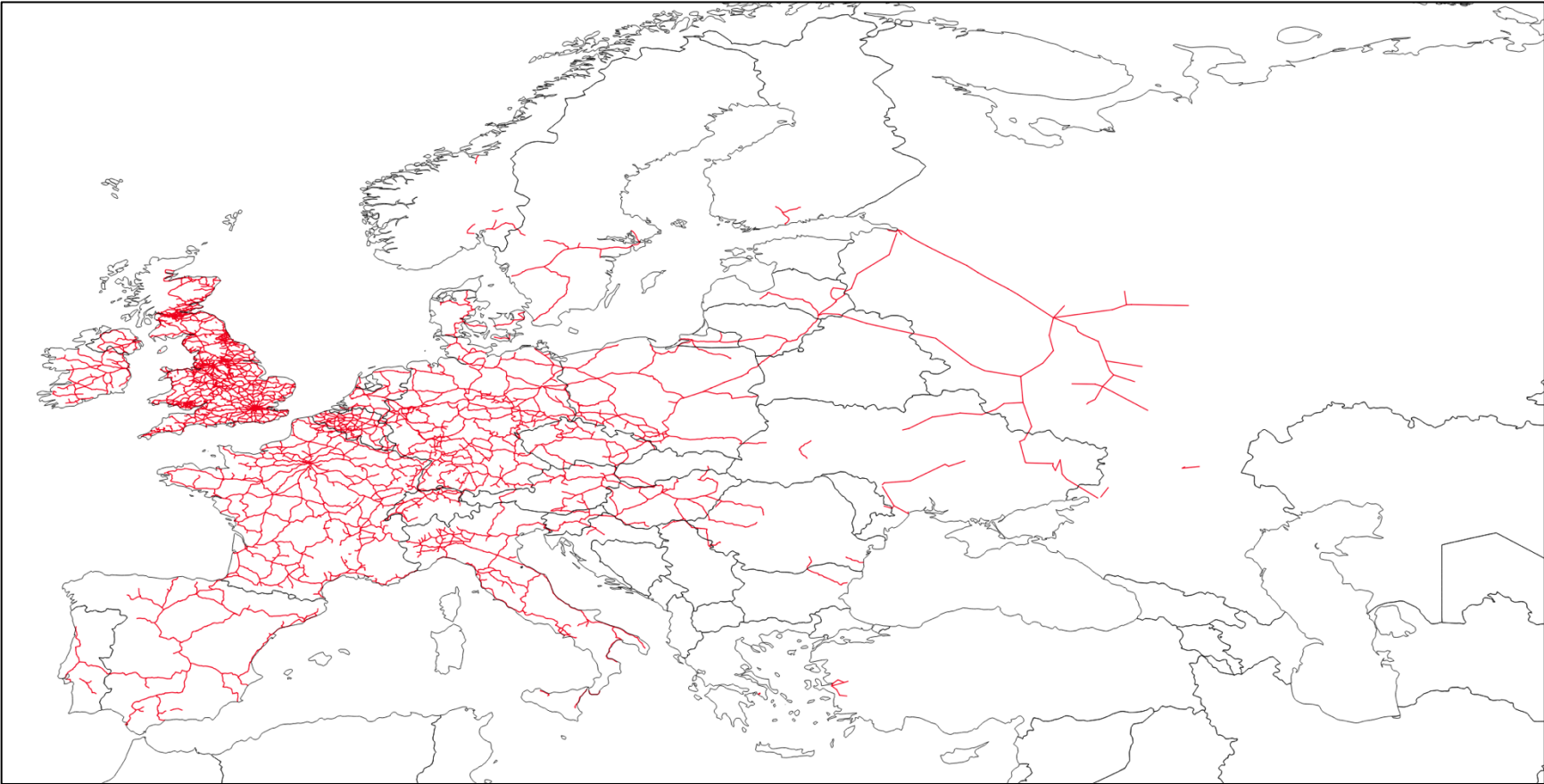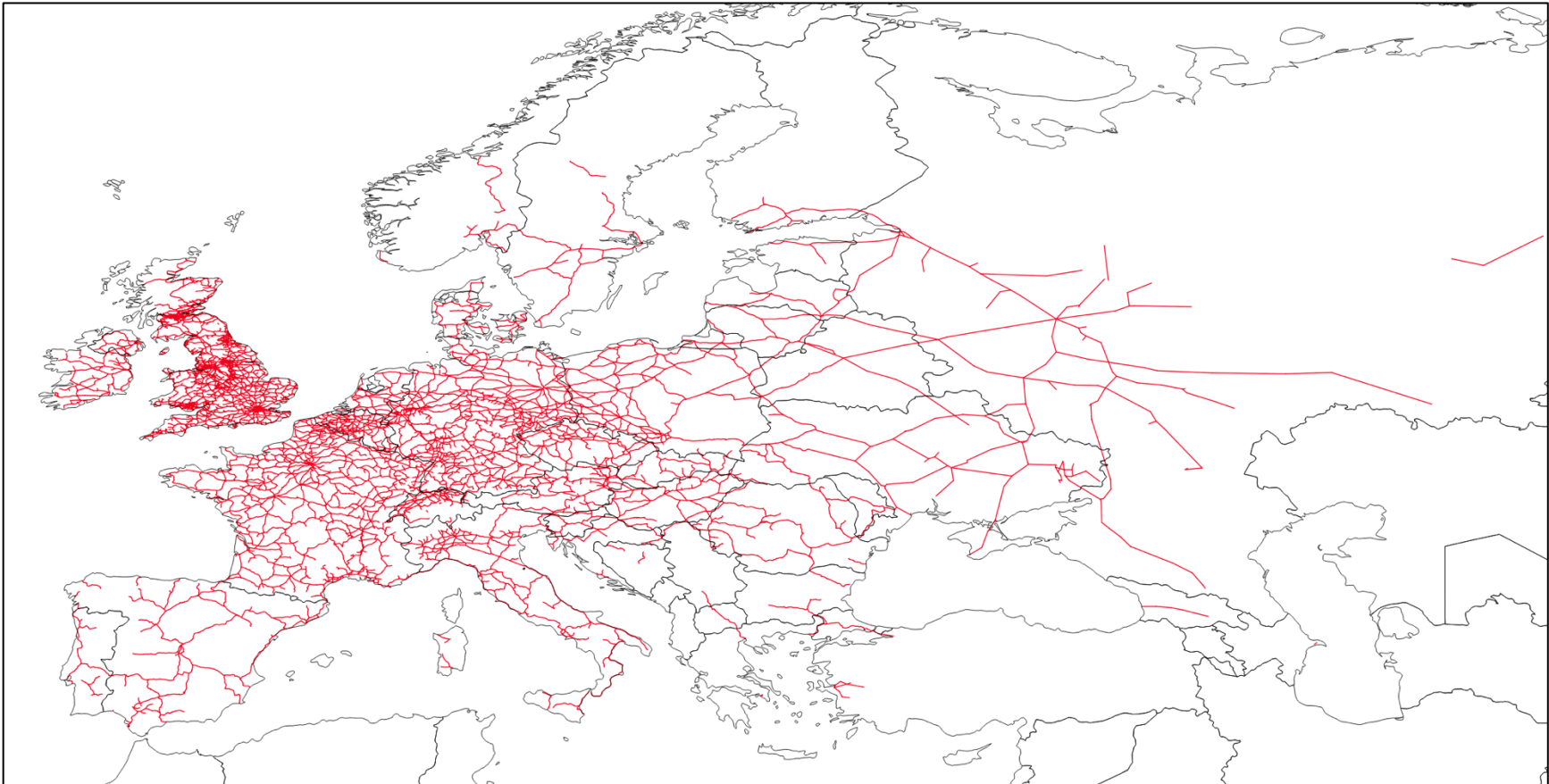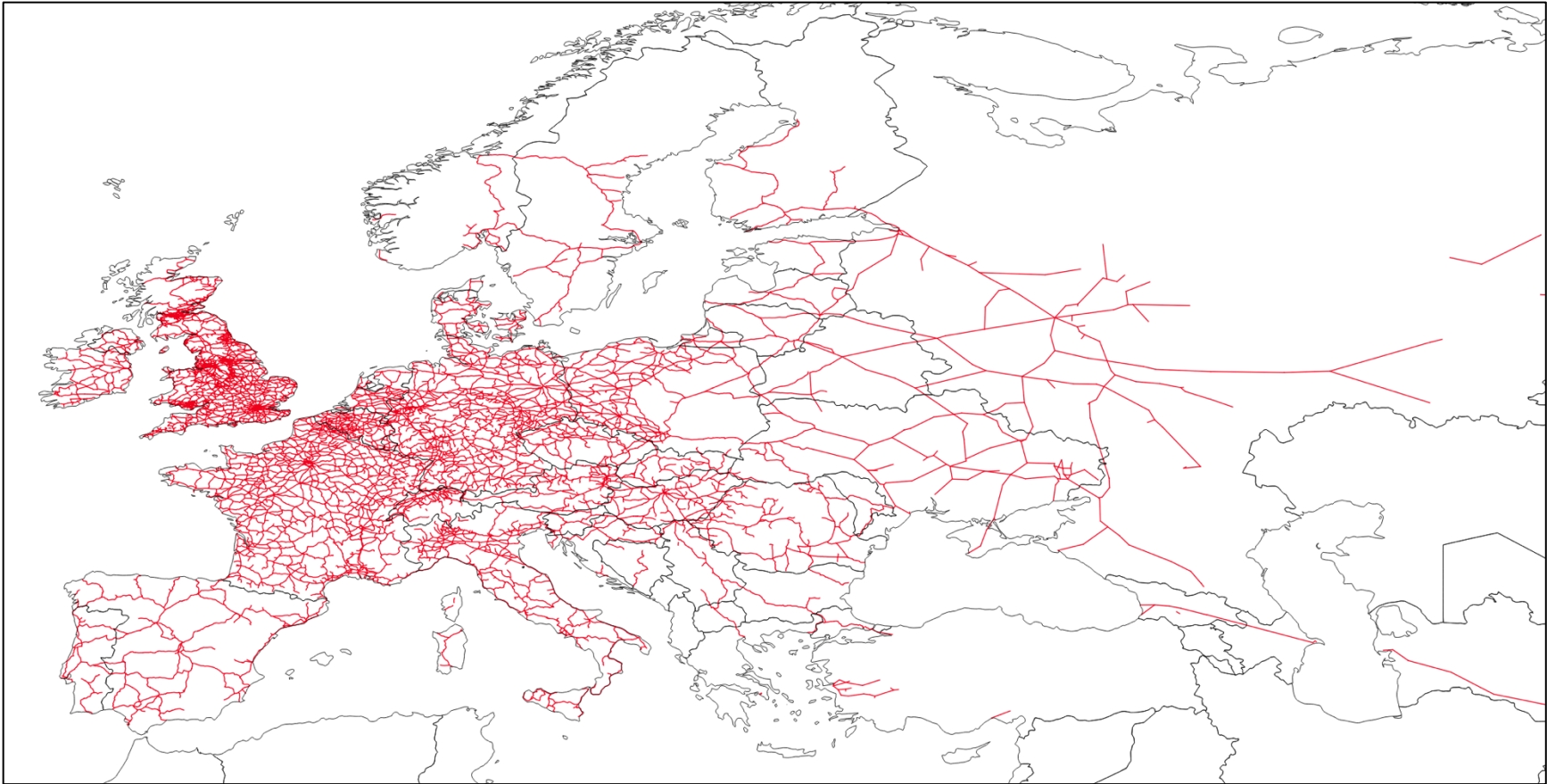
Appendix B Table B1. List of languages and language families

Language family	Language	Language family	Language
Armenian	Armenian	Romance	Catalan
Baltic	Lettish	Romance	French
Baltic	Lithuanian	Romance	Friulian
Basque	Basque	Romance	Gallegan
Caucasian	Caucasian	Romance	Italian
Celtic	Breton	Romance	Portuguese
Celtic	Irish	Romance	Romanian
Celtic	Welsh	Romance	Spanish
Finno-Ugrian	Cheremissian	Semitic	Arabic
Finno-Ugrian	Finnish	Semitic	Maltese
Finno-Ugrian	Karelian	Slovanic	Bulgarian
Finno-Ugrian	Livonian	Slovanic	Czech
Finno-Ugrian	Magyar	Slovanic	Macedonian Slavs
Finno-Ugrian	Ostyak	Slovanic	Polish
Finno-Ugrian	Samoyedic	Slovanic	Russian
Germanic	Danish	Slovanic	Serbo-Croatian
Germanic	Dutch	Slovanic	Slovakian
Germanic	English	Slovanic	Slovenian
Germanic	Flemish	Slovanic	Ukrainian
Germanic	Frisian	Slovanic	White Russian
Germanic	German	Thraco-Illyrian	Albanian
Germanic	Norwegian	Turkish-Tataric	Bashkirian
Germanic	Swedish	Turkish-Tataric	Chuvashian
Hellenic	Greek	Turkish-Tataric	Karachaic
Iranic	Ossetic	Turkish-Tataric	Kirghizic
Kamchadal	Kamchadal	Turkish-Tataric	Kumykian
Mongolian	Buryat	Turkish-Tataric	Tataric
Mongolian	Kalmuckian	Turkish-Tataric	Turkish
		Turkish-Tataric	Yakut

Seven religions were identified in the maps: Catholic, Protestant, Greek Orthodox, Armenian Orthodox, Muslim, Buddhism, and "Heathen". Some regions have overlapping religions (e.g. Muslim and Armenian Orthodox in mid-to-eastern Turkey or Muslim and Greek Orthodox in the Balkan region). In these cases, nationalist production were counted for both religions.

Appendix B Table B2. List of religions and religious groups

Religious group	Religion
Western Christian	Catholic
Western Christian	Protestant
Orthodox Christian	Greek Orthodox
Orthodox Christian	Armenian Orthodox
Muslim	Muslim
Buddhist	Buddhist
Heathen	Heathen

Appendix B Table B3. Robustness test results for different levels of aggregation of language and religion

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Negative binomial model Controls for decades, population size, capital city, bishop seat, catholic city, distance to river, and distance to sea	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Logged number of nationalist productions in the same religion	0.0953 (0.134)							
Logged number of nationalist productions in the same religious group		0.182 (0.122)						
Logged number of nationalist productions in the same religious group (with a broader time window)			0.122 (0.146)					
Logged number of nationalist productions in the same religious group (with a broader time window)				0.163 (0.115)				
Logged number of nationalist productions in the same language					0.0212 (0.0609)			
Logged number of nationalist productions in the same language family						-0.0466 (0.0821)		
Logged number of nationalist productions in the same language (with a broader time window)							0.0103 (0.0575)	
Logged number of nationalist productions in the same language family (with a broader time window)								-0.0350 (0.0844)
Zero-Inflation model Distance to a renowned artist (logged miles)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	36320	36320	36320	36320	36320	36320	36320	36320

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Appendix B Table B4. Robustness test results for polity variables

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Logged number of nationalist productions in the same polity (numbers added when multiple polities)	0.0336 (0.0842)							
Logged number of nationalist productions in the same polity (numbers added when multiple polities & broader time window)		-0.0311 (0.0771)						
Logged number of nationalist productions in the same polity (numbers averaged when multiple polities)			0.0162 (0.0862)					
Logged number of nationalist productions in the same polity (numbers averaged when multiple polities & broader time window)				-0.0497 (0.0793)				
Number of nationalist productions in the same polity (numbers added when multiple polities)					0.00149 (0.00246)			
Number of nationalist productions in the same polity (numbers added when multiple polities & broader time window)						-0.000536 (0.000475)		
Number of nationalist productions in the same polity (numbers averaged when multiple polities)							0.00101 (0.00270)	
Number of nationalist productions in the same polity (numbers averaged when multiple polities & broader time window)								-0.000749 (0.000542)
Observations	36320	36320	36320	36320	36320	36320	36320	36320

Standard errors in parentheses

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Controls for decades, population size, capital city, bishop seat, catholic city, distance to river, and distance to sea included in the negative binomial part, distance to a renowned artists included in the zero-inflated part

Appendix B Table B5: Models disaggregated by genre

Each cell or group of cells in a column (separated by lines) shows the results of a different model.

Variables	Column 1 DV: Writings	Column 2 DV: Paintings	Possible explanations for divergence from aggregate results (in italics)
Frame resonance			
Protestant city	0.642 (0.465)	0.339 (0.563)	
Foreign ruled	0.857** (0.272)	1.030** (0.355)	
Napoleon			
- Country not occupied by Napoleon	Ref.	Ref.	
- Country not occupied by Napoleon	1.021** (0.391)	1.519** (0.482)	
- Country occupied; city not occupied	0.396 (0.343)	1.225 (0.685)	
- French cities	-1.330* (0.628)	-0.176 (0.743)	
Logged number of previous nationalist writings (C1) or paintings (C2) in the same language group	*0.209** (0.0696)*	-0.00400 (0.0726)	Aligns with the argument about domain specificity; Writing is more language bound than painting, thus linguistic domains should be more relevant for writings.
Logged number of previous nationalist writings (C1) or paintings (C2) in the same religion group	0.181 (0.171)	0.0274 (0.138)	
Personal ties			
Received at least one letter	1.738*** (0.269)	0.762* (0.365)	
Logged number of nationalist writings (C1) or paintings (C2) near letter sender	0.387*** (0.0557)	*0.0682 (0.0534)*	The dummy above is significant, which tests the same mechanism.
Cultural domains			
Distance to university town with previous nationalist writings (C1) or paintings (C2)	-0.513*** (0.0707)	-0.358*** (0.0838)	

Distance to university town without previous nationalist writings (C1) or paintings (C2)	*-0.217** *(0.0857)*	0.0392 (0.0807)	The size of the coefficient is about a half compared to distance to a university town with previous nationalist writings, in line with the overall argument.
Distance to newspaper town with previous nationalist writings (C1) or paintings (C2)	-0.372*** (0.0532)	-0.534*** (0.0778)	
Distance to newspaper town without previous nationalist writings (C1) or paintings (C2)	*-0.125*** *(0.0482)*	0.0967 (0.0676)	The size of the coefficient is about a third compared to distance to a university town with previous nationalist writings, in line with the overall argument.
Distance to the next Roman road (logged miles)	0.0788 (0.0429)	0.0272 (0.0601)	
Distance to the nearest previous nationalist writing (C1) or painting (C2) on the Roman road network (logged miles)	-0.442*** (0.113)	*0.134* *(0.129)*	The variable below is significant, which tests the same mechanism.
Logged number of previous nationalist writings (C1) or paintings (C2) in the same Roman road cluster	0.349*** (0.0955)	0.266*** (0.0798)	
Political domains			
Logged number of previous nationalist writings (C1) or paintings (C2) in the same polity	*0.230**** *(0.0505)*	-0.00792 (0.137)	Aligns with the argument about domain specificity; Writing is more closely reflecting political issues (e.g. in political essays), thus political domains become relevant.
Economic / infrastructural domains			
Distance to the nearest stagecoach station (logged miles)	-0.055 (0.052)	-0.214** (0.0685)	
Distance to the nearest nationalist writing (C1) or painting (C2) on the stagecoach network (logged miles)	-0.111 (0.073)	0.198 (0.192)	
Distance to the railway station (logged miles)	Did not converge	Did not converge	
Distance to the nearest nationalist writing (C1) or painting (C2) on the railway network (logged miles)	Did not converge	Did not converge	

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001
Controls for decades, population size, capital city, bishop seat, catholic city, distance to river, and distance to sea included in the negative binomial part, distance to a renowned artist included in the zero-inflated part

Appendix C. Robustness tests for main models

We conduct two types of robustness test, reported in the two tables below. The first table (Appendix C Table 1) shows the coefficients from the main analyses reported in the article (Column 1) and the coefficients from five different robustness tests. Each row represents a different model, of which we only present the coefficients for the main variable of interest. Columns 2 represents results with bootstrapped standard errors, which are substantially identical to our main results. Column 3 summarizes a series of models with two additional control variables: the geodesic distance to the nearest nationalist production and a lagged outcome variable. The third robustness test (see Column 4) uses a logistic regression model instead of a zero-inflated negative binomial and includes only cities that had at least one artist according to the Wikipedia data described in the main text (generating an N of 24,928 artist-decades). The results are substantially identical. Column 5 shows the results from yet another modeling strategy: an event history set-up where the dependent variable is the first nationalist work produced, after which the city is dropped from the dataset. Results are substantially similar. Exceptions are protestant city, foreign-ruled, and institutionalized cultural channels. Regarding the latter, distance to university or newspaper towns are no longer affecting nationalist production due to the fact that such towns are dropped from the analysis after the first nationalist production appears, thus obscuring the influence they might have had on other cities. Column 6 shows the results of a difference-in-difference specification, which we model as a two-way-fixed effects OLS regression. All results hold up except the variables related to Napoleon's occupation and the Roman road variables, for which we have to drop control variables that don't vary over time. Column 7 shows the results from models where capital cities were dropped. These analyses are important because ERNiE assigns capital cities as location when the exact locations are unknown. The results show the robustness of our findings even when the cases were dropped where the ERNiE could possibly have assigned the locations incorrectly.

The second table (Appendix C Table 2) reports results from additional robustness checks that address concerns about causal inference and influential observations (for details see section 5.8 of the main text). Column 1 again reports the main results as rendered in the main tables. Column 2 adds a control for the number of artists who lived in a city and who had already produced a nationalist work. Column 3 addresses concerns about the city-level analysis that may be confounded by nationalist artists moving

from city to city. In this model, only the first nationalist work of artists is taken into account when constructing the dependent variable. In Column four, the inflated part of the ZINB regression includes the number of writer or artist who lived in the city, whether or not they had produced nationalist work. The final column omits the influential decade of 1810 to see if these observations dominate the results.

Appendix C Table C1. Five robustness tests with different model specifications, controls, and samples

Each cell or group of cells in a column (separated by lines) shows the results of a different model.

Variables	Column 1 Main results	Column 2 Bootstrapped SE	Column 3 Additional covariates	Column 4 Logistic regression	Column 5 Event history	Column 6 Two-way-fixed effects	Column 7 Without capital cities
Frame resonance							
Protestant city	0.447	0.447	0.551	0.216	0.693**	NA	0.565
	(0.404)	(0.503)	(0.388)	(0.418)	0.266		(0.420)
Foreign ruled	0.964***	0.964**	0.814**	0.844***	0.0774	0.0839*	1.097***
	(0.230)	(0.366)	(0.286)	(0.179)	(0.158)	(0.0351)	(0.267)
Napoleon							
- Country not occupied by Napoleon's empire	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
- Country occupied; city not occupied	1.253***	1.253*	1.103*	0.590*	0.553**	-0.0306	1.311***
	(0.357)	(0.574)	(0.444)	(0.246)	(0.202)	(0.0849)	(0.386)
- Both country and city occupied	0.634	0.634	0.496	0.267	0.552**	-0.0496	0.722
	(0.393)	(0.495)	(0.380)	(0.321)	(0.208)	(0.0857)	(0.442)
- French city	-0.772	-0.772	-0.766	-0.705	-0.470	NA	-0.795
	(0.480)	(0.613)	(0.550)	(0.485)	(0.291)		(0.504)
Logged number of previous nationalist productions in same language group	0.0212	0.0212	-0.0459	-0.0345	0.0230	0.0366*	0.0113
	(0.0609)	(0.0728)	(0.0553)	(0.0529)	(0.0399)	(0.0149)	(0.0691)
Logged number of previous nationalist productions in same religious group	0.0953	0.0953	0.0362	0.0896	0.258	0.0159	0.102
	(0.134)	(0.152)	(0.148)	(0.185)	(0.147)	(0.0326)	(0.136)
Personal ties							
Received a letter from a nationalist	1.738***	1.738**	1.495***	1.381***	0.782**	0.316**	1.990***
	(0.269)	(0.492)	(0.250)	(0.237)	(0.246)	(0.104)	(0.256)
Logged number of nationalist writings near letter sender	0.387***	0.387**	0.306***	0.320***	0.153*	0.0913*	0.436***
	(0.056)	(0.105)	(0.0621)	(0.0475)	(0.0644)	(0.0423)	(0.060)

Cultural domains

Distance to university town with previous nationalist production (logged miles)	-0.412*** (0.0580)	-0.412*** (0.104)	-0.323*** (0.0893)	-0.481*** (0.0757)	-0.182 (0.106)	-0.228*** (0.069)	-0.428*** (0.0630)
Distance to university town without previous nationalist production (logged miles)	-0.0872 (0.0798)	-0.0872 (0.0779)	-0.161* (0.0802)	0.0674 (0.0677)	-0.150* (0.0691)	0.00201 (0.034)	-0.138 (0.034)
Distance to newspaper town with previous nationalist production (logged miles)	-0.470*** (0.0535)	-0.470*** (0.141)	-0.432*** (0.122)	-0.503*** (0.0704)	-0.0433 (0.117)	-0.240*** (0.041)	-0.459*** (0.0667)
Distance to newspaper town without previous nationalist production (logged miles)	0.0160 (0.0480)	0.0160 (0.0545)	-0.0675 (0.0548)	0.0369 (0.0430)	-0.181*** (0.0347)	0.0653* (0.028)	-0.00109 (0.0501)
Distance to the next Roman road (logged miles)	0.0666 (0.0455)	0.0666 (0.0580)	0.0580 (0.0393)	0.0745 (0.0387)	0.0534 (0.0292)	NA	0.0646 (0.0505)
Distance to the nearest previous nationalist production on the Roman road network (logged miles)	-0.232* (0.0973)	-0.232 (0.1222)	-0.179 (0.118)	-0.194** (0.0744)	-0.247*** (0.0519)	-0.0105 (0.012)	-0.308* (0.127)
Logged number of previous nationalist productions in the same Roman road cluster	0.277*** (0.0720)	0.277** (0.0888)	0.244** (0.066)	0.222** (0.0837)	0.102 (0.0627)	0.027 (0.038)	0.294*** (0.0712)

Political domains

Logged number of previous nationalist productions in the same polity	0.0336 (0.0842)	0.0336 (0.0970)	0.0111 (0.0833)	0.0998 (0.0656)	0.0526 (0.0515)	0.0149 (0.014)	0.00735 (0.0907)

Economic / infrastructural domains

Distance to the nearest stagecoach station (logged miles)	-0.0723 (0.0451)	-0.0723 (0.0572)	-0.0624 (0.0407)	-0.0555 (0.0331)	-0.0882** (0.0310)	0.009 (0.014)	-0.0538 (0.0538)
Distance to the nearest nationalist production on the stagecoach network (logged miles)	0.0886 (0.122)	0.0886 (0.125)	0.209 (0.124)	0.0628 (0.0745)	-0.0390 (0.0541)	0.0028 (0.015)	0.0698 (0.015)

Distance to the nearest railway station (logged miles)	-0.0892 (0.104)	-0.0892 (0.0920)	-0.0243 (0.0684)	0.0298 (0.0694)	0.0578 (0.0794)	-0.025** (0.009)	-0.159 (0.134)
Distance to the nearest nationalist production on the railway network (logged miles)	0.146 (0.138)	0.146 (0.147)	0.265 (0.194)	0.0278 (0.0926)	0.212 (0.223)	0.0284 (0.029)	0.112 (0.165)

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Controls for decades, population size, capital city, bishop seat, catholic city, distance to river, and distance to sea included in the negative binomial part, distance to a renowned artist included in the zero-inflated part

Appendix C Table C2. Four robustness tests with additional control variables, different outcomes and samples

Each cell or group of cells in a column (separated by lines) shows the results of a different model.

Variables	Column 1 Main results	Column 2 With control for No of nationalist artists in the city	Column 3 Outcome is first nationalist works of artists only	Column 4 With control for number of artists in the city	Column 5 Without the 1810 decade
Frame resonance					
Protestant city	0.447	0.673*	-0.216	-0.367	0.286
	(0.404)	(0.331)	(0.358)	(0.488)	(0.402)
Country is foreign ruled	0.964***	0.712***	1.016***	1.063***	1.144***
	(0.230)	(0.204)	(0.204)	(0.254)	(0.231)
Country not occupied by Napoleon	Ref.	Ref.	Ref.	Ref.	Ref.
	.	.	.	.	.
Country occupied; city not occupied	1.253***	1.103***	0.619	1.054**	1.204***
	(0.357)	(0.287)	(0.335)	(0.335)	(0.360)
Both country and city occupied	0.634	0.591	0.335	0.286	0.672
	(0.393)	(0.316)	(0.340)	(0.360)	(0.402)
French city	-0.772	-0.675	-1.004	-0.877	-0.635
	(0.480)	(0.415)	(0.521)	(0.476)	(0.485)
Logged number of previous nationalist productions in same language group	0.0212 (0.0609)	0.0173 (0.0524)	-0.104* (0.0451)	0.0142 (0.0606)	-0.0183 (0.0617)
Logged number of previous nationalist productions in same religious group	0.0953 (0.134)	0.0864 (0.124)	-0.0588 (0.102)	-0.0336 (0.106)	0.0347 (0.126)
Personal ties					
Received at least one letter	1.738***	1.634***	0.922***	0.910***	1.760***
	(0.269)	(0.229)	(0.210)	(0.202)	(0.282)

Logged number of nationalist writings near letter sender	0.387*** (0.0557)	0.351*** (0.0542)	0.232*** (0.0694)	0.189*** (0.0544)	0.402*** (0.0567)
Cultural domains					
Distance to university town with previous nationalist production	-0.412*** (0.0580)	-0.277*** (0.0582)	-0.276*** (0.0531)	-0.272*** (0.0390)	-0.383*** (0.0620)
Distance to university town without previous nationalist production	-0.0872 (0.0798)	-0.157* (0.0781)	0.0451 (0.0676)	0.0811 (0.0698)	-0.0249 (0.0849)
Distance to the nearest newspaper town with previous nationalist production	-0.470*** (0.0535)	-0.370*** (0.0812)	-0.358*** (0.0545)	-0.337*** (0.0453)	-0.458*** (0.0519)
Distance to the nearest newspaper town without previous nationalist production	0.0160 (0.0480)	-0.0338 (0.0497)	0.0621 (0.0499)	0.131* (0.0518)	0.0953* (0.0465)
Distance to the next Roman road	0.0666 (0.0455)	0.0679 (0.0378)	0.0338 (0.0382)	0.108** (0.0408)	0.0686 (0.0452)
Distance to the nearest previous nationalist production on the Roman road network	-0.232* (0.0973)	-0.258** (0.0806)	-0.0444 (0.0790)	0.0184 (0.0892)	-0.156 (0.100)
Distance to the next Roman road (Roman cities only)	0.0809 (0.0738)	0.0933 (0.0683)	0.0861 (0.0724)	0.127 (0.0734)	0.0804 (0.0761)
Distance to the nearest previous nationalist production on the Roman road network (Roman cities only)	-0.353*** (0.105)	-0.348*** (0.0945)	-0.0642 (0.105)	-0.104 (0.148)	-0.260* (0.103)
Logged number of previous nationalist productions in the same Roman road cluster	0.277*** (0.0720)	0.239*** (0.0610)	0.262** (0.0806)	0.246*** (0.0730)	0.306*** (0.0764)
Political domains					
Logged number of previous nationalist productions in the same polity	0.0336 (0.0842)	0.0344 (0.0741)	0.0484 (0.0773)	0.0680 (0.0726)	0.0736 (0.0884)
Economic / infrastructural domains					
Distance to the nearest stagecoach station	-0.0718 (0.0569)	-0.0827* (0.0414)	-0.0670 (0.0345)	-0.0165 (0.0415)	-0.0621 (0.0447)

Distance to the nearest nationalist production on the stagecoach network	0.0929 (0.130)	0.0556 (0.112)	0.180* (0.0740)	0.220* (0.109)	0.0845 (0.128)
Distance to the nearest stagecoach station (Before 1870)	-0.0871 (0.0547)	-0.0908 (0.0482)	-0.0685 (0.0438)	-0.0348 (0.0506)	-0.0845 (0.0523)
Distance to the nearest nationalist production on the stagecoach network (Before 1870)	0.0247 (0.0755)	-0.0046 (0.0690)	0.0889 (0.0744)	0.143 (0.0868)	0.00925 (0.0822)
Distance to the nearest railway station	-0.0892 (0.104)	-0.0572 (0.0824)	Did not converge	-0.0919 (0.0844)	-0.0892 (0.104)
Distance to the nearest nationalist production on the railway network	0.146 (0.138)	0.142 (0.134)	Did not converge	0.367*** (0.110)	0.146 (0.138)

Standard errors in parentheses

Controls for decades, population size, capital city, bishop seat, catholic city, distance to river, and distance to sea included in the negative binomial part, distance to a renowned artist include in the zero-inflated part

* p<0.05, ** p<0.01, *** p<0.001

Appendix D. Robustness tests for the transportation networks

Appendix D1: Using different distance thresholds and clustering algorithms

In the main analyses, we use 5 miles as a threshold to define which cities were within reach of the road, stagecoach, or railway networks. In other words, if a nationalist work of art or writing was produced farther than 5 miles away from these networks, they were not considered as part of an influence network. Given that it takes around 1 hour and 40 minutes for an adult to walk 5 miles, we believe it is a reasonable threshold. However, we also use 10 miles and 50 miles as a threshold to test the robustness of our findings. The results are not too different from the main findings.

This appendix also shows the results of using different clustering algorithms to identify regions of dense connectivity through the Roman road networks.

Appendix D Table D1. Roman road networks using different thresholds

Variables	Model 1	Model 2	Model 3	Model 4
Negative binomial model Controls for decades, population size, capital city, bishop seat, catholic city, distance to river, and distance to sea	Yes	Yes	Yes	Yes
Distance to the next Roman road (logged miles)	0.0692 (0.0456)	0.0996 (0.0749)	0.0686 (0.0456)	0.0999 (0.0751)
Distance to the nearest production on the Roman road network (logged miles) (10 miles threshold)	-0.220* (0.0933)	-0.308* (0.128)		
Distance to the nearest production on the Roman road network (logged miles) (50 miles threshold)			-0.161 (0.0893)	-0.255 (0.139)
Zero-Inflation model Distance to a renowned artist (logged miles)	Yes	Yes	Yes	Yes
Observations	33284	22756	33286	22756

Standard errors in parentheses; * p<0.05, ** p<0.01, *** p<0.001

Appendix D Table D2. Roman road clusters using different algorithms

Variables	Model 1	Model 2	Model 3
Negative binomial model			
Controls for decades, population size, capital city, bishop seat, catholic city, distance to river, and distance to sea	Yes	Yes	Yes
Logged number of previous nationalist productions in the same Roman road cluster (*Louvain*)	0.277*** (0.0720)		
Logged number of previous nationalist productions in the same Roman road cluster (*Girvan-Newman*)		0.251*** (0.0699)	
Logged number of previous nationalist productions in the same Roman road cluster (*Clauset-Newman-Moore*)			0.182* (0.0766)
Zero-Inflation model			
Distance to a renowned artist (logged miles)	Yes	Yes	Yes
Observations	18240	18240	18240

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Note: The Louvain algorithm automatically chooses the number of clusters based on the modularity of the resulting clusters (i.e. how strong the within-cluster connections are compared to between-cluster connections). In contrast, the Girvan-Newman algorithm is hierarchical in structure, so researchers need to set a number of clusters.

Since we don't have any theory as to how many clusters we expect, we ran the Louvain algorithm and then used the number of clusters we got for the Girvan-Newman algorithm as well.

Appendix D Table D3. Transportation variables using different distance threshold

Variables	Model 1	Model 2	Model 3 Years before 1870	Model 4 Years before 1870	Model 5 Years after 1870	Model 6 Years after 1870
Negative binomial model Controls for decades, population size, capital city, bishop seat, catholic city, distance to river, and distance to sea	Yes	Yes	Yes	Yes	Yes	Yes
Distance to the Stagecoach station (logged miles)	-0.0520 (0.0436)	-0.0251 (0.0435)	-0.0763 (0.0578)	-0.0575 (0.0562)		
Distance to the nearest production on the Stagecoach (logged miles) (10 miles threshold)	-0.0534 (0.0786)		-0.1228 (0.0788)			
Distance to the nearest production on the Stagecoach (logged miles) (50 miles threshold)		-0.0481 (0.0352)		-0.0884* (0.0396)		
Distance to the nearest railway station (logged miles)					-0.0916 (0.104)	-0.166 (0.132)
Distance to the nearest production on the railway (logged miles) (10 miles threshold)					0.120 (0.146)	
Distance to the nearest production on the railway (logged miles) (10 miles threshold)						0.0902 (0.0741)
Zero-Inflation model Distance to an artist (logged miles)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	34600	35023	21987	22164	11593	12296

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Appendix D2: Top-coding transportation variables with missing values

Appendix D Table D4. Top-coded transportation variables

Variables	Model 1	Model 2 Only cities with a Roman past	Model 3	Model 4 Years before 1870	Model 5
Control variables					
Decade dummies, controls for population size, capital city, bishop seat, catholic city, distance to river, and distance to sea	Yes	Yes	Yes	Yes	Yes
Roman roads					
Distance to the next Roman road	0.0792	0.0710			
	(0.0433)	(0.0731)			
Distance to the nearest previous nationalist production on the Roman road network (top-coded)	-0.234**	-0.383***			
	(0.0881)	(0.0938)			
Stagecoach routes					
Distance to the nearest stagecoach station			-0.0635	-0.0921	
			(0.0440)	(0.0537)	
Distance to the nearest nationalist production on the stagecoach network (top-coded)			0.139	0.0371	
			(0.119)	(0.0784)	
Railways					
Distance to the nearest railway station					-0.179
					(0.108)
Distance to the nearest nationalist production on the railway network (top-coded)					0.290*
					(0.117)
Observations	36320	25328	36320	22700	13620

Standard errors in parentheses; a control for distance to a renowned artist is included in the zero-inflated part; * p<0.05, ** p<0.01, *** p<0.001

Appendix E. Robustness tests for letters in artist-level analyses

The main results of the writer/artist-level analyses presented in Table 2 are based on negative binomial regression models. To test the robustness of our findings, we report the results from survival models as well as an OLS specification. They are substantially identical.

Appendix E Table E1. Writer/artist-level analyses using event history and OLS regression models

Variables	Model 1 Event history model	Model 2 Event history model	Model 3 OLS	Model 4 OLS
Controls				
Decade dummies	Yes	Yes	Yes	Yes
Writer/artist-level IVs				
Number of letters received during the last decade	0.0147*** (0.002)		0.0085** (0.0026)	
Artist received at least one letter in the last decade		2.215*** (0.255)		0.331** (0.113)
Observations	26,326	26,326	20,509	20,509

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

Appendix E Table 2 present results disaggregated by genre. The results show that receiving letters only affected the number of writings (these results are identical to those in Table 2 in the main text), but not musical works or paintings. We arrive at similar results if we disaggregate by type of writer/artist, distinguishing between the nationalist works produced by writers, musicians, or painters (results not shown). This is not surprising as written words can more easily transmit ideas that manifest themselves in other written words than in music or paintings. These results thus further support our argument about the domain specificity of diffusion processes.

Appendix E Table E2. Writer/artist-level analyses disaggregated by genre

Variables	Model 1 DV: Writings	Model 2 DV: Writings	Model 3 DV: Music	Model 4 DV: Music	Model 5 DV: Paintings	Model 6 DV: Paintings
Controls						
Decade dummies	Yes	Yes	Yes	Yes	Yes	Yes
Writer/artist-level IVs						
Number of letters received during the last decade	0.0720* (0.0366)		-0.339 (0.735)		-0.127 (0.190)	
Writer/artist received at least one letter in the last decade		2.125*** (0.562)		-0.255 (1.137)		-0.0707 (0.598)
Observations	20509	20509	20509	20509	20509	20509

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

Appendix E Table 3 addresses concerns about the validity of the writer/artist level analysis with additional controls: the number of letters received in the time span before the previous decade (measuring the overall productivity / activity of an artist/writer); and fixed effects for the most prominent and prolific letter writers to account for unobserved heterogeneity across correspondences.

Appendix E Table E3. Writer/artist-level analyses with additional controls

Variables	Model 1	Model 2	Model 3	Model 4
Controls				
Decade dummies	Yes	Yes	Yes	Yes
Number of letters received prior to the preceding decade	Yes	Yes	No	No
Fixed effects for the 5 most frequent letter writers	No	No	Yes	Yes
Writer/artist -level IVs				
Number of letters received during the last decade	0.0737*		0.0383	
	(0.0372)		(0.0372)	
Received at least one letter during the last decade		2.214***		2.495**
		(0.594)		(0.891)
Observations	20509	20509	20509	20509

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

Appendix F. Robustness tests for letters in city-level analyses

The following table reports the results from models with additional covariates to alleviate concerns about the inference gained from city-level analysis of the influence of letters written by prominent romantic nationalists. They include, in various combinations, controls for the number of letters that originated in a city, where addressee and writer lived in the same city, as well as the total number of writers / artists who had already produced nationalist work in the city. They all account for the possibility that cities with more active or more nationalist writers /artists will generate more letters written by nationalists, thus raising doubt of whether receiving such letters stimulates nationalist production.

Appendix F Table F1. Robustness tests for letter analyses at the city-level

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Basic control variables								
Decade dummies, controls for population size, capital city, bishop seat, catholic city, distance to river, and distance to sea	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional control variables								
Number of letters sent from the city	Yes	Yes	No	No	Yes	Yes	No	No
Number of letters circulated within the city	No	No	Yes	Yes	No	No	Yes	Yes
Number of writers / artists living in a city who had already produced a nationalist work	No	No	No	No	Yes	Yes	Yes	Yes
Independent variables of interest								
City received at least one letter	1.714*** (0.264)		1.750*** (0.268)		1.612*** (0.227)		1.651*** (0.228)	
Logged number of nationalist writings near letter sender		0.370*** (0.0545)		0.397*** (0.0557)		0.334*** (0.0600)		0.363*** (0.0563)
Observations	36320	36320	36320	36320	36320	36320	36320	36320

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001
A control for distance to a renowned artist is included in the zero-inflated part

Appendix G. Temporal heterogeneity

Given the long duration of our observation period, it is worth asking to what extent the findings vary over time. Appendix G Table 1 shows the results of linear interactions with time (in the first column) and non-linear interactions with each decade (the remaining columns), allowing us to test whether our conclusions are driven by particular sub-periods or display other non-linear patterns.

The results suggest that diffusion through specific channels is the main driver of romantic nationalism throughout our observation period. It further highlights the unique nature of the 1810s (highlighted in gray), either because otherwise insignificant variables are relevant during this period, or because the coefficient reverses its sign compared to other decades. We attribute this to the effect of the Napoleonic occupations and wars, which represent the equivalent of a "super-spreader" event, in the language of the epidemiology of contagious diseases, that brought romantic nationalism to many corners of Europe.[1] The main results of our analysis hold up, however, when dropping the 1810 decade from the sample (see Appendix C Table 2, Column 5). Finally, there is no consistent pattern for generic domains, especially if we take coefficients on the borderline of significance into account as well. The only exception is nationalist production within language communities, for which there is a more consistent trend already discussed in the preceding analysis.

---

[1] In a subsample analysis for 1810 that includes all variables that are significant in individual models, as reported in Table 7, we find that the standardized coefficient for the Napoleon variable is 6 to 8 times larger than those of the other variables.

Appendix G Table G1. Temporal heterogeneity

		Sign. linear interactions with time, flipping signs	Interactions with individual decades															
			1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920
Diffusion																		
Personal	Letters	No		−	+	+	+	+***	+**	+**	+**	+**	+					
Cultural	Newspaper (nat.)	No	+	+	−	+	+***	+	+	+	+	+	+*	+	+	+	+	+
Cultural	University (nat.)	No	−	+	−	+***	+***	+	+*	+*	+	+	+	+	.	−	+	−
Cultural	Cultural region	Yes	+	+	.	+	+***	+	+*	+*	+	+	+***	+	−	+*	−	−
Political	Polity	Yes	+	+**	.	+*	.	.	.	−*	−*	.	.	.	.	.	+*	−*
Infrastructural	Stagecoach	Yes	+	+	−*	+	+***	.	+*	+	−	−	−	−	.	−	−	−*
Infrastructural	Railway	No									+	−	+	+	−	−	+	−
Frame resonance																		
Contrasting	Foreign ruled	Yes	−	−	+	−	−***	+	+*	+	+***	+***	+**	+**	+***	+***	+**	+
Contrasting	Napoleon	No			+	+**	+***	+	+**	+**	+	+	+*	+	+***	+	+**	+
Compatibility	Protestant	No	+	+	+	+	+	+	-	+	+	−	+	-	+	+	+*	-
Credibility	Language	Yes	+	+	+	+*	+***	.	.	.	.	.	.	.	.	.	.	−***
Credibility	Religion	Yes	+	+	.	.	+**	.	.	.	.	.	.	.	.	.	.	.

Note: Each cell of the right-hand side panel reports the results from a ZINB model with interactions between the main independent variable and decades; + indicates that the coefficient points in the right direction in that decade; – indicates that the coefficient points in the wrong direction; a dot (.) indicates that the coefficient is minuscule
* p<0.05, ** p<0.01, *** p<0.001