Geography and the Formation of Exclusive Communities in Europe: A Long Run View

Sascha O. Becker (Monash & Warwick)
Kalle Kappner (LMU Munich)
Marvin Suesse (Trinity College Dublin)
Nikolaus Wolf (HU Berlin)

IZ D2MCM, HU Berlin, Economics

February 2024

Motivation

- How does Physical Geography ('topology') shape Social Space ('openness towards outsiders')?

Motivation

- How does Physical Geography (‘topology’) shape Social Space (‘openness towards outsiders’)?
- Is people’s ‘social horizon’ shaped by their physical environment, beyond travel costs?

Motivation

- How does Physical Geography (‘topology’) shape Social Space (‘openness towards outsiders’)?
- Is people’s ‘social horizon’ shaped by their physical environment, beyond travel costs?
- Economics focus on
  - endowments
  - disease environments
  - cost of mobility

Motivation

- How does Physical Geography (‘topology’) shape Social Space (‘openness towards outsiders’)?
- Is people’s ‘social horizon’ shaped by their physical environment, beyond travel costs?
- Economics focus on
  - endowments
  - disease environments
  - cost of mobility
- Less focus on the effect of ‘invisible’ topological features
  - watersheds

This Project

Figure 1: Stylized example: Two watersheds, four cities

This Project

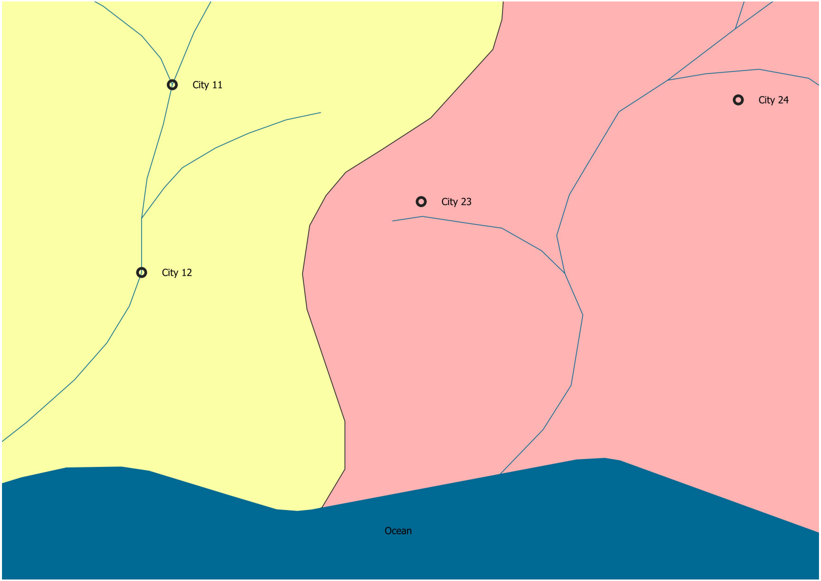

Figure 2: Stylized example: Two watersheds, four cities

This Project

- Can watersheds affect the intensity of social interaction across space and shape openess attitudes?

This Project

- Can watersheds affect the intensity of social interaction across space and shape openess attitudes?
- Are people in the same watershed more alike than those ‘on the other side of the divide’, even conditional on the same travel time?

This Project

- Can watersheds affect the intensity of social interaction across space and shape openess attitudes?
- Are people in the same watershed more alike than those ‘on the other side of the divide’, even conditional on the same travel time?
- How large and how persistent are these effects?

This Project

- Can watersheds affect the intensity of social interaction across space and shape openess attitudes?
- Are people in the same watershed more alike than those ‘on the other side of the divide’, even conditional on the same travel time?
- How large and how persistent are these effects?
- Gravity models for cultural outcomes: Grosjean (2011, AER P&P)

This Project

- Can watersheds affect the intensity of social interaction across space and shape openess attitudes?
- Are people in the same watershed more alike than those 'on the other side of the divide', even conditional on the same travel time?
- How large and how persistent are these effects?
- Gravity models for cultural outcomes: Grosjean (2011, AER P&P)
- Watersheds affect dialects ('linguistic watersheds'): Coblin (2002), Davison (2006), Chamberlain (2015); and dialects affect trade: Lameli et al., 2015

Data: European Value Study (EVS)

- 57k individuals nested within 935 European NUTS3 regions.

Data: European Value Study (EVS)

- 57k individuals nested within 935 European NUTS3 regions.
- Wave 5 (2017): ca. 200 questions on attitudes on religion, family, work, life etc.

Data: European Value Study (EVS)

- 57k individuals nested within 935 European NUTS3 regions.
- Wave 5 (2017): ca. 200 questions on attitudes on religion, family, work, life etc.
- Focus on questions 22–30 on openness towards 'outsiders'.
  - Q: 'On this list are various groups of people. Could you identify any that you would not like to have as neighbours?'
  - People of a different race, heavy drinkers, immigrants/foreign workers, drug addicts, homosexuals, Christians, Muslims, Jews, Gypsies
  - binary answers: 0 = mentioned ; 1 = not-mentioned

Data: European Value Study (EVS)

- 57k individuals nested within 935 European NUTS3 regions.
- Wave 5 (2017): ca. 200 questions on attitudes on religion, family, work, life etc.
- Focus on questions 22–30 on openness towards ‘outsiders’.
  - Q: ‘On this list are various groups of people. Could you identify any that you would not like to have as neighbours?’
  - People of a different race, heavy drinkers, immigrants/foreign workers, drug addicts, homosexuals, Christians, Muslims, Jews, Gypsies
  - binary answers: 0 = mentioned ; 1 = not-mentioned
- Calculate average dyadic similarity between regions i, j:
- $OpenSim_{ij} = 1 - \frac{\sum_{q=22}^{30} |q_i - q_j|}{9} * 100 \in [0, 100]$

Data: Watersheds

- European Commission's CCM2 (Catchment Characterisation Model): Consistently defined watersheds

Data: Watersheds

- European Commission's CCM2 (Catchment Characterisation Model): Consistently defined watersheds
- Exploit Pfafstetter system to define 4 nested hierarchy levels.
  - level 0: Top-level watersheds, e.g. Danube vs. Rhine
  - levels 1–3: Increasingly local watersheds, e.g. tributaries to Danube and Rhine

Data: Watersheds

- European Commission's CCM2 (Catchment Characterisation Model): Consistently defined watersheds
- Exploit Pfafstetter system to define 4 nested hierarchy levels.
  - level 0: Top-level watersheds, e.g. Danube vs. Rhine
  - levels 1–3: Increasingly local watersheds, e.g. tributaries to Danube and Rhine
- Calculate dyadic watershed similarity for NUTS-3 regions:
  - (1) For regions i, j, compute the area shares of each encompassed watershed at hierarchy levels $h \in \{0, 1, 2, 3\}$.
  - (2) i and j are similar at h if at least 50% of their respective areas are covered by shared watershed(s).
  - (3) Add up the 4 h-specific similarity dummies.

Data: Watersheds

- European Commission's CCM2 (Catchment Characterisation Model): Consistently defined watersheds
- Exploit Pfafstetter system to define 4 nested hierarchy levels.
  - level 0: Top-level watersheds, e.g. Danube vs. Rhine
  - levels 1–3: Increasingly local watersheds, e.g. tributaries to Danube and Rhine
- Calculate dyadic watershed similarity for NUTS-3 regions:
  - (1) For regions i, j, compute the area shares of each encompassed watershed at hierarchy levels $h \in \{0, 1, 2, 3\}$.
  - (2) i and j are similar at h if at least 50% of their respective areas are covered by shared watershed(s).
  - (3) Add up the 4 h-specific similarity dummies.
- $WatershedSim_{ij} \in \{0, 1, 2, 3, 4\}$

Data: Travel time

- Control for the geodetic distance between NUTS-3 regional centroids.

Data: Travel time

- Control for the geodetic distance between NUTS-3 regional centroids.
- Additionally, take historical and modern infrastructure into account.

Data: Travel time

- Control for the geodetic distance between NUTS-3 regional centroids.
- Additionally, take historical and modern infrastructure into account.
- Dyadic travel time estimates:
  - Pre-roman: Human Mobility Index with Seafaring from Özak (2010)
  - Roman roads and shipable rivers: Flückiger et. al. (2022)
  - Railroads in 1900: Marti-Henneberg (2023)
  - Individual car travel on modern roads: ESPON EGTC via OpenStreetMap (2016)

Data: Travel time

- Control for the geodetic distance between NUTS-3 regional centroids.
- Additionally, take historical and modern infrastructure into account.
- Dyadic travel time estimates:
  - Pre-roman: Human Mobility Index with Seafaring from Özak (2010)
  - Roman roads and shipable rivers: Flückiger et. al. (2022)
  - Railroads in 1900: Marti-Henneberg (2023)
  - Individual car travel on modern roads: ESPON EGTC via OpenStreetMap (2016)
- Compute travel time saved relative to more primitive transport means, deflated by overall time savings (‘connectivity’).

Data: NUTS-3 sample

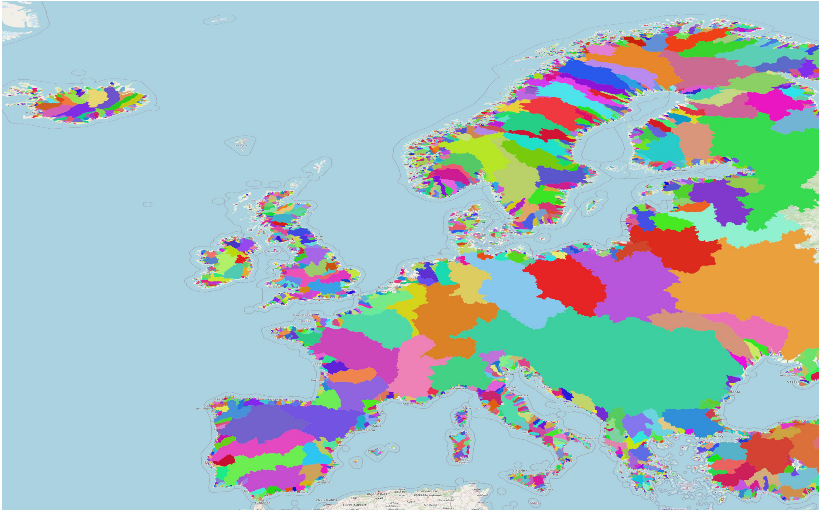

Figure 3: European top-level watersheds

Data: NUTS-3 sample

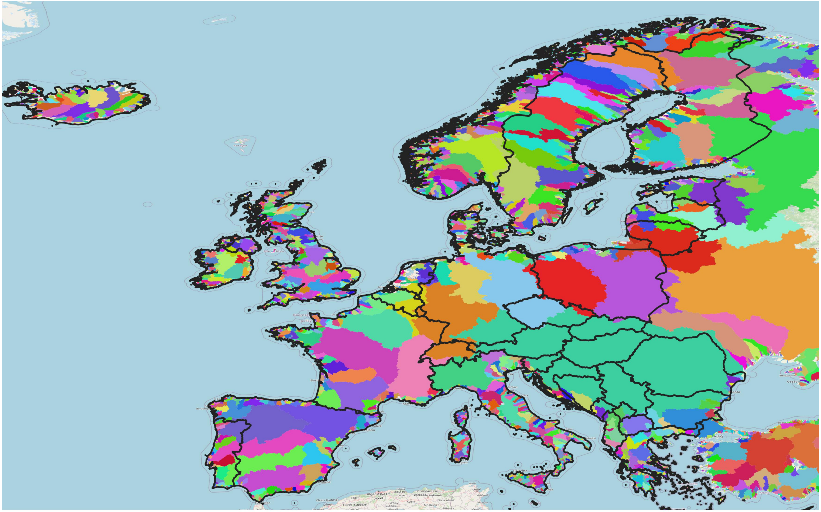

Figure 4: European top-level watersheds & Country borders

Data: NUTS-3 sample

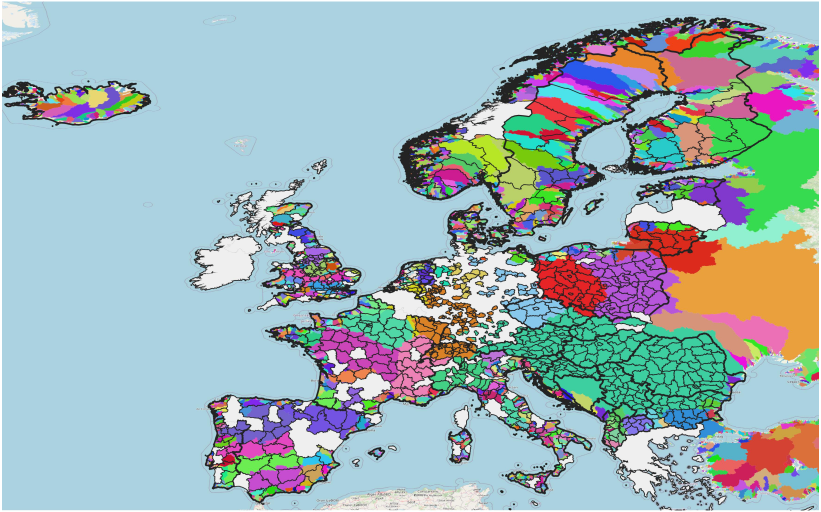

Figure 5: European top-level watersheds & NUTS-3 sample

Data: Dyade-level summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
OpenSim	437580	8.13	1.07	0	100
ExpSim	437580	-2	17.6	-97.7	100
WatershedSim	437580	.129	.563	0	4
Distance	437580	1123.2	618.5	3.4	4163.2
Int. border	437580	.925	.264	0	1
Coast	437580	.091	.288	0	1
Inland	437580	.326	.469	0	1
Urban	437580	.411	.345	0	1
time saved Roman	435365	.001	.001	0	.003
time saved rail	432471	.002	.581	-9.56	230.6
time saved car	435506	.001	.008	0	2.07

Rivers and openess attitudes

Basic monadic correlations at NUTS-3

	(1) Benchmark	(2) with controls	(3) no islands	(4) no landlocked
	Dep. var.: Pro-openness attitudes [0-100]			
river connectivity	0.0203**	0.0228**	0.0248**	0.0517***
	(0.0098)	(0.0092)	(0.0096)	(0.0185)
Roman road connectivity	0.0847	0.0792	0.0735	0.0182
	(0.0544)	(0.0530)	(0.0530)	(0.0519)
1900 rail connectivity	0.0016	0.0035	0.0037	-0.0099***
	(0.0032)	(0.0033)	(0.0033)	(0.0023)
modern motorway connectivity	-0.0055	-0.0027	-0.0012	0.0282***
	(0.0144)	(0.0144)	(0.0146)	(0.0108)
ruggedness		-0.0010	-0.0020	0.0015
		(0.0069)	(0.0070)	(0.0080)
elevation		0.0002	0.0003	0.0004*
		(0.0002)	(0.0002)	(0.0002)
rural		-0.4427**	-0.5801***	0.6977***
		(0.1884)	(0.2187)	(0.2344)
inland		-0.3088	-0.0779	
		(0.3404)	(0.3030)	
Country FE	Y	Y	Y	Y
N	935	935	806	400
r2	0.008	0.013	0.014	0.061

Standard errors in parentheses, spatial clustering with Bartlett kernel

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Watersheds and openess attitudes

Pairwise similarity regressions

$$OpenSim_{ij} = \beta_1 WatershedSim_{ij} + \beta_2 Dist_{ij} + \beta_3 Border_{ij} + X_{ij} + \mu_i + \eta_j + \epsilon_{i,j}$$

$OpenSim$:	Similarity of openness attitudes $\in [0, 100]$
$WatershedSim$:	Similarity of associatied watersheds $\in \{0, 1, 2, 3, 4\}$
$Dist$:	Geodetic distance
$Border$:	Indicator for international border
X :	Vector of dyadic controls: both coast, inland, urban
μ :	Region i fixed effects
η :	Region j fixed effects
ϵ :	errors clustered at i-j level

Note: Frequency weights using no. of respondents $min(N_i, N_j)$

Watersheds and openess attitudes

Pairwise similarity regressions

$$OpenSim_{ij} = \beta_1 WatershedSim_{ij} + \beta_2 Dist_{ij} + \beta_3 Border_{ij} + X_{ij} + \mu_i + \eta_j + \epsilon_{i,j}$$

	(1) unconditional	(2) fixed effects	(3) controls	(4) no weighting
	Dependent variable: Similarity in openness attitudes (0-100)			
river similarity (1-4)	1.037***	1.922***	0.508***	0.326***
	(0.186)	(0.178)	(0.140)	(0.091)
geodetic distance			-0.004***	-0.004***
			(0.000)	(0.000)
international border			-2.778***	-1.530***
			(0.351)	(0.236)
County i FE	-	Y	Y	Y
County j FE	-	Y	Y	Y
N	8870768	8870768	8870768	437578
r2	0.004	0.646	0.684	0.723
F	31	117	76	82

Additional dyadic controls in (3) and (4): coasts, inland, rural; s.e. clustered at i and j (in parentheses)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Watersheds and general value attitudes

Pairwise similarity regressions

Define outcome over (almost) all attitudes measured in the EVS.

	(1) unconditional	(2) fixed effects	(3) controls	(4) no weighting
	Dependent variable: Similarity in openness attitudes (0-100)			
river similarity (1-4)	0.995***	1.131***	0.192***	0.133***
	(0.086)	(0.074)	(0.051)	(0.029)
geodetic distance			-0.003***	-0.002***
			(0.000)	(0.000)
international border			-2.324***	-1.973***
			(0.125)	(0.099)
County i FE	-	Y	Y	Y
County j FE	-	Y	Y	Y
N	8870768	8870768	8870768	437578
r2	0.024	0.619	0.737	0.783
F	133	233	248	265

Additional dyadic controls in (3) and (4): coasts, inland, rural; s.e. clustered at i and j (in parentheses)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Watershed hierarchy and openess attitudes

Pairwise similarity regressions

Use dummies for association with the same watersheds at level h, conditional on being in the same watershed at level $h-1$.

	(1) 1st level river	(2) 2nd level river	(3) 3rd level river	(4) 4th level river
	Dependent variable: Similarity in openness attitudes (0-100)			
same 1st level river	1.696***			
	(0.433)			
same 2nd level river		-0.226		
		(0.357)		
same 3rd level river			1.670***	
			(0.541)	
same 4th level river				1.556**
				(0.732)
geodetic distance	-0.004***	-0.005***	-0.003***	-0.002***
	(0.000)	(0.001)	(0.001)	(0.001)
international border	-2.728***	-3.187***	-3.466***	-5.773***
	(0.351)	(0.601)	(0.964)	(0.991)
Sample	all	same 1st level	same 2nd level	same 3rd level
County i FE	Y	Y	Y	Y
County j FE	Y	Y	Y	Y
N	8870768	784952	307149	196667
r2	0.685	0.742	0.807	0.830
F	77	12	13	19

Additional dyadic controls: coasts, inland, rural; s.e. clustered at i and j (in parentheses)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Watersheds, openness and transport infrastructure

Pairwise similarity regressions

	(1) infrastructure controls	(2) interaction	(3) no islands
	Dependent variable: Similarity in openness attitudes (0-100)		
river similarity (1-4)	0.454***	0.723***	0.707***
	(0.140)	(0.165)	(0.201)
geodetic distance	-0.004***	-0.005***	-0.004***
	(0.000)	(0.000)	(0.000)
international border	-2.761***	-2.858***	-3.279***
	(0.350)	(0.352)	(0.394)
rel. time saved Roman roads	-484.831	-2537.572***	-2383.526***
	(344.855)	(491.322)	(514.297)
rel. time saved 1900 railways	-5.118	-4.055	-3.754
	(3.317)	(2.946)	(2.789)
rel. time saved modern motorways	-4.180**	3319.691***	3216.404***
	(2.129)	(701.303)	(776.648)
river similarity × time saved Roman roads		310.621	-171.753
		(202.821)	(431.728)
river similarity × time saved railways		-3.680***	-3.669***
		(0.855)	(0.831)
river similarity × time saved motorways		-829.918***	-311.446
		(175.527)	(560.991)
County i FE	Y	Y	Y
County j FE	Y	Y	Y
N	8745333	8745333	7375308
r2	0.686	0.687	0.669

Additional dyadic controls: coasts, inland, rural; s.e. clustered at i and j (in parentheses)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Are upstream or downstream regions more open?

- Trade on rivers is easier in *one* direction: downstream.
- For each pair, there is one geographically induced upstream 'exporter' and one downstream 'importer'.
- For each pair i, j, denote i as the upstream region if elevation$_i$ $>$ elevation$_j$.
- Calculate average degree to which upstream openness exceeds downstream openess:
- $OpenExp_{ij} = \frac{\sum_{q=22}^{30}(q_i - q_j)}{9} * 100$ where elevation $i > j$

Watersheds and upstream openness

Pairwise directionality regressions

Redefined outcome variable measures the difference between upstream and downstream regions.

	(1) basic	(2) controls	(3) 1st level river	(4) elevation
	Dependent variable: Upstream openness - downstream openness (0-100)			
river similarity (1-4)	1.043***	1.045***		
	(0.274)	(0.275)		
same 1st level river basin			2.519***	0.901
			(0.734)	(0.817)
same 1st level river basin × elevation difference				3.393***
				(0.997)
elevation difference				-1.328*
				(0.722)
geodetic distance	-0.004***	-0.004***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
international border	1.814***	1.754***	1.633***	1.457 **
	(0.617)	(0.610)	(0.609)	(0.604)
ruggedness (pairwise product)		-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
County i FE	Y	Y	Y	Y
County j FE	Y	Y	Y	Y
N	8870768	8870768	8870768	8870768
r2	0.296	0.296	0.296	0.297
F	12	11	10	9

Additional dyadic controls: coasts, inland, rural; s.e. clustered at i and j (in parentheses)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Conclusion

- Physical topology shapes the transmission of openness attitudes across space, beyond simple distance effects.
- Watersheds matter in a complex way.
  - More homogeneous attitudes within watersheds
  - Within watersheds, upstream places have more open attitudes than downstream places.
  - Possible non-linearities across the hierarchy of watersheds
- Need to understand correlated fault lines, e.g. historical political borders and mountain chains