

Supplementary Material:
Shifting Routes, Shifting Attitudes

Anonymized

April 2025

Contents

1	Creating a Democratic Roots Index for Pre-industrial Societies and Tolerance Index	3
2	Outcome Measures in the Pre-Industrial Analysis	5
2.1	GPT-based Tolerance Index	6
2.2	GPT-based Democracy Roots Index	6
2.3	Default Democracy Index	7
2.4	Default Tolerance Roots Index	7
3	Control Variables: Pre-Industrial Analysis	7
4	Support for Democracy: Integrated Values Survey (IVS) Measure	8
5	Geolocations of Choke Points and Natural Harbours	8
5.1	Choke Points	8
5.2	Natural Ports	8
6	Locations of Integrated Values Survey Respondents	8
6.1	Ethnolinguistic Geolocations of Integrated Values Survey Respondents	9
6.2	Geocoding Integrated Values Survey Respondents	9
7	Distances from Integrated Values Survey Locations	10
7.1	Ethnolinguistic distances	10
7.2	Contemporary distances	10
8	Conflict Occurrences	10

9	Geographic Controls in the Contemporary Attitudes Analysis: Human Mobility and Caloric Suitability	10
10	Data	11
11	Regression Robustness Checks - Pre-industrial Values	11
12	Regression Robustness Checks - Contemporary Support for Democracy	15
13	Conflicts and Distance to Choke Points	24

1 Creating a Democratic Roots Index for Pre-industrial Societies and Tolerance Index

Building on Inglehart’s argument that cultures conducive to democracy are characterized by postmaterialist and self-expression values—such as tolerance, trust, universal values, emancipation from authority, and civic engagement—I construct an index that captures these underlying traits without explicitly referencing democracy [5]. Societies with stronger democratic foundations tend to exhibit these values more prominently, shaping their political and social norms. Folklore motifs reflect societal values [2], offering a historical and cultural lens to assess the extent to which these norms align with democratic principles without the limitations associated with direct survey questions.

To operationalize these cultural dimensions, I construct two indices: the *Tolerance Index* and the *Democratic Roots Index*. Each index is derived using two distinct methods: keyword-based classification approach [6], to identify relevant concepts within folklore motifs. The second method leverages ChatGPT-4o in an API environment to classify motifs based on thematic relevance. By employing both approaches, I ensure robustness in capturing the cultural foundations of democracy and tolerance across societies.

To assess the extent to which a society’s folklore aligns with democratic norms, I calculate the fraction of motifs reflecting pro-democratic values, based on the framework of postmaterialist and self-expression values [5]. These values include universal values (such as tolerance, justice, equality, and fairness), trust, and emancipation from authority. Instead of relying on civic participation as a measure, I use prosocial behavior as a broader category that captures any actions benefiting others [4].

To classify motifs according to these values, I employ a context-sensitive approach using the ChatGPT-4o API, which analyzes each motif description and determines whether it relates to prosocial behavior, trust, tolerance, emancipation from authority, or universal values. If a motif is flagged as corresponding to any of these categories, it is counted as reflecting a pro-democratic value.

Conversely, I also compute the fraction of motifs representing “anti-democratic” values, which, according to Inglehart’s framework, include intolerance, communal dependence, authority, and selfishness (as opposed to prosocial behavior) and interpersonal distrust. The denominator for both measures is always the total number of motifs for each society, ensuring comparability across groups.

After calculating the fractions of motifs reflecting pro-democratic and anti-democratic values, I standardize both measures by computing their z-scores to ensure comparability across societies. I then subtract the anti-democratic z-score from the pro-democratic z-score and standardize the resulting difference to construct the Democratic Roots Index.

The Tolerance Index is constructed in a similar manner. I first identify concepts related to tolerance, such as tolerance and tolerant as well as those associated with intolerance using ChatGPT-4o. Next, I calculate the total number of motifs reflecting tolerance and

intolerance. To standardize the measure, I compute the z-score of tolerance-related motifs and subtract the z-score of intolerance-related motifs. Finally, I apply standardization to the final value to ensure comparability across societies.

To construct an alternative measure of democratic cultural foundations, I rely on keyword-based classification, following the method employed by [6], rather than motif classification using ChatGPT. This approach systematically identifies specific concepts—predefined themes that appear in motif descriptions—allowing for a structured analysis of cultural traits. These concepts were later utilized by [2] to measure key dimensions such as universal values, trust, prosociality, and tolerance. I adopt Enke’s framework for these democratic values while also incorporating distrust (including anti-trust and corruption-related concepts) and emancipation from authority, as they are integral to democratic culture.

I classify motifs into pro-democratic and anti-democratic categories based on their prevalence in folklore. Following [5], pro-democratic values are those associated with self-expression and postmaterialism, including trust, tolerance, prosocial behavior, and emancipation from authority. Anti-democratic values include communal dependence, distrust, selfishness (as opposed to prosociality), and submission to authority. This classification provides a structured way to compare societies’ cultural orientations toward democracy.

- **Trust:** trust
- **Universal values (justice, fairness, and equality):** tolerance, tolerant, reciprocal, equality, equitable, fairness, justice
- **Emancipation from authority:** independence, autonomy, freedom, resistance, liberty
- **Prosocial behavior** (acts benefiting others, as categorized in [4]):
 - share, sharing, give, donation, charitable
 - help, helping, generosity, generous, generously, kindness
 - compassion, empathy, compassionate
 - cooperate, collaborate, caring, shared
 - given, donate, gift, goodwill, helper
 - helpful, assist, assistance, aid, kindly
 - pity, sympathy, sympathetic, cooperative, collaborative
 - collaborator, contribute, contributing, contribution, contributor
 - caregiver, support, supporting, supportive
- **Communal dependence:**
 - group, together, collectively, collective, cohesion, solidarity

- unite, unity, home, homeland, hometown, clan
- family, cousin, kin, kinship
- father, mother, brother, brother-in-law, sister
- kid, parent, wife, husband, son, daughter
- mom, relative, dad, grandmother, grandfather
- uncle, aunt, ancestor, sibling, grandchild
- grandson, granddaughter, lineage, ancestral, familial
- mother-in-law, tribe, stepfather, motherhood, infant
- child, baby, parental, parenting, spouse
- mommy, daddy, niece, nephew, ancestry
- descendant, offspring, loyal, loyalty, faithful
- friend, buddy, friendship, communal, community-based, community

- **Submission to authority:**

- order, orderly, rule, norm, tradition
- traditional, traditionally, convention, conventional
- heritage, custom, follow, obey, obedience
- authority, comply, compliance

- **Selfishness (opposite of prosociality):** selfish, self-interest

- **Intolerance (opposite of tolerance):** intolerance

- **Distrust and corruption (opposite of trust):** corruption, corrupt, fraud, betray, betrayal, cheat

2 Outcome Measures in the Pre-Industrial Analysis

Folklore motifs are classified using a constrained large-language-model (LLM) labeling protocol. For each motif description, a fixed prompt is submitted consisting of (i) a system message that restricts the model’s role to a binary/ternary classifier and (ii) a user message that asks whether the text is related to a target concept. To minimize stochastic variation and improve replicability, the model’s sampling temperature is set to 0, which enforces deterministic decoding given the same model, prompt, and input text. The prompt also constrains the output space to three discrete labels—**Yes**, **No**, and **Not sure**—thereby reducing free-form responses and limiting ambiguity in downstream coding. In the empirical analysis, the classification is operationalized as a binary indicator by treating **Yes** as concept-relevant

and No as not relevant; **Not sure** is used only as a diagnostic category for low-confidence cases and is not interpreted as evidence of concept presence. This design—fixed prompt + deterministic decoding + restricted label set—aims to maximize reproducibility of the automated coding pipeline and to reduce idiosyncratic variability in motif-level annotations.

Below is the prompt used to classify motifs with respect to the concept of “communal”.

```
{'role': 'system',
  'content': 'You are a helpful assistant that answers questions with "Yes" or "No". '
            'Return "Not sure" if you are very unsure.'},
{'role': 'user',
  'content': f'Please analyze the following text and determine whether it relates to '
            f'the concept of "communal". Respond with "Yes" if the text is related, '
            f'"No" if it is not. Here is the text: "{motif_desc}".'
```

2.1 GPT-based Tolerance Index

$$Tolerance_Motifs_i = tolerance_i + tolerant_i \quad (1)$$

$$Intolerance_Motifs_i = intolerance_i \quad (2)$$

$$Tolerance_Share_i = \frac{Tolerance_Motifs_i}{Total_Number_Motifs_i} \quad (3)$$

$$Intolerance_Share_i = \frac{Intolerance_Motifs_i}{Total_Number_Motifs_i} \quad (4)$$

$$GPT_Tolerance_Index_i = z(z(Tolerance_Share_i) - z(Intolerance_Share_i)) \quad (5)$$

2.2 GPT-based Democracy Roots Index

$$Pro_Democracy_Motifs_i = emancipation_i + prosocial_i + universal_i + interpersonal_trust_i + tolerance_i \quad (1)$$

$$Anti_Democracy_Motifs_i = communal_dependence_i + selfish_i + authority_i \\ + intolerance_i + interpersonal_distrust_i \quad (2)$$

$$Pro_Democracy_Values_i = \frac{Pro_Democracy_Motifs_i}{Total_Number_Motifs_i} \quad (3)$$

$$Anti_Democracy_Values_i = \frac{Anti_Democracy_Motifs_i}{Total_Number_Motifs_i} \quad (4)$$

$$\text{Democracy Roots Index}_i = z(z(\text{Pro_Democracy_Values}_i) - z(\text{Anti_Democracy_Values}_i)) \quad (5)$$

2.3 Default Democracy Index

$$\text{Pro_Democracy_Motifs}_i = \sum_{k \in \mathcal{P}} m_{k,i} \quad (1)$$

$$\text{Anti_Democracy_Motifs}_i = \sum_{k \in \mathcal{A}} m_{k,i} \quad (2)$$

$$\text{Frac_Pro_Democracy}_i = \frac{\text{Pro_Democracy_Motifs}_i}{\text{Total_Number_Motifs}_i} \quad (3)$$

$$\text{Frac_Anti_Democracy}_i = \frac{\text{Anti_Democracy_Motifs}_i}{\text{Total_Number_Motifs}_i} \quad (4)$$

$$\text{Default_Democracy_Roots_Index}_i = z(z(\text{Frac_Pro_Democracy}_i) - z(\text{Frac_Anti_Democracy}_i)) \quad (5)$$

2.4 Default Tolerance Roots Index

$$\text{Tolerance_Motifs}_i = \text{tolerance motifs}_i + \text{tolerant motifs}_i \quad (1)$$

$$\text{Tolerance_Share}_i = \frac{\text{Tolerance_Motifs}_i}{\text{Total_Number_Motifs}_i} \quad (2)$$

$$\text{Intolerance_Motifs}_i = \text{intolerance motifs}_i \quad (3)$$

$$\text{Intolerance_Share}_i = \frac{\text{Intolerance_Motifs}_i}{\text{Total_Number_Motifs}_i} \quad (4)$$

$$\text{Tolerance_Index}_i = z(z(\text{Tolerance_Share}_i) - z(\text{Intolerance_Share}_i)) \quad (5)$$

3 Control Variables: Pre-Industrial Analysis

Two control variables are obtained from [6]; detailed definitions and discussion are provided in [6].

Year of first publication (logged): the earliest year in which any folklore title is recorded for ethnolinguistic group i , capturing differences in when groups entered the archival

record and potential shifts in collection practices over time [6].

Number of titles (logged): the total number of distinct folklore titles available for ethnolinguistic group i , capturing cross-group differences in documentation intensity and sampling effort [6].

4 Support for Democracy: Integrated Values Survey (IVS) Measure

The Support for Democracy index is constructed from the following questions in the IVS[1]: ‘Having a strong leader who does not have to bother with parliament and elections,’ ‘Having experts, not government, make decisions according to what they think is best for the country,’ ‘Having the army rule,’ and ‘Having a democratic political system.’ Each question is rated on a scale from 1 to 4, with higher values indicating greater support for democracy. The index is calculated as the average of these four variables.

5 Geolocations of Choke Points and Natural Harbours

5.1 Choke Points

To compute distances from ethnolinguistic locations to the nearest choke point for (i) respondents in the Integrated Values Survey (IVS) and (ii) ethnolinguistic groups in the Folklore Catalogue, I first construct a geolocation dataset for the six baseline choke points—Bosphorus, Suez Canal, Strait of Gibraltar, Bab-el-Mandeb, Strait of Hormuz, and Strait of Malacca—and an extended seven-choke point variant that additionally includes the English Channel. I obtain the coordinates using Python’s ‘geopy’ (Nominatim) geocoder and save the resulting point layers as GeoJSON files.

5.2 Natural Ports

Natural harbour locations are obtained from the World Port Index (WPI).¹ I retain only ports classified as CN (Coastal Natural) and keep all shelter categories except N (None).

6 Locations of Integrated Values Survey Respondents

Two location measures are used in the analyses: (i) respondents’ ethnolinguistic locations and (ii) their contemporary locations.

¹https://msi.nga.mil/api/publications/download?key=16694622/SFH00000/WPI_Shapefile.zip&type=view

6.1 Ethnolinguistic Geolocations of Integrated Values Survey Respondents

The Integrated Values Survey does not report geocoded information for respondents' ethnolinguistic locations. Ethnolinguistic locations are therefore proxied using the language spoken at home and merging the Integrated Values Survey with Glottolog, which provides geolocations for language communities.²

To implement this merge, a dictionary is constructed that matches language labels in the IVS *language spoken at home* variable to the corresponding language names in Glottolog.

Moreover, a rule-based Old World macro-region indicator is constructed for robustness checks. Each ethnolinguistic homeland coordinate is assigned to a coarse *origin macro-region* using a transparent coordinate-based partition that approximates broad Afro–Eurasian geographic separations. When an ethnolinguistic label is observed in multiple macro-regions (i.e., multiple homeland coordinates map to different regions), the macro-region with the highest frequency is assigned to that label to preserve observations and ensure a unique origin classification. The Old World domain is defined as locations with longitude in $[-30^\circ, 180^\circ]$ and latitude in $[-35^\circ, 90^\circ]$, which captures Afro–Eurasia. Within this domain, two fixed cut-offs are applied to obtain three macro-regions. First, a longitude threshold at 50°E separates *East_OldWorld* ($\text{lon} \geq 50^\circ\text{E}$) from *West_OldWorld* ($\text{lon} < 50^\circ\text{E}$), serving as a parsimonious proxy for the conventional Europe–Asia separation running toward the Caspian/Ural corridor. Second, within western longitudes ($\text{lon} < 50^\circ\text{E}$), a latitude threshold at 36°N separates *South_WestOldWorld* ($\text{lat} < 36^\circ\text{N}$) from *North_WestOldWorld* ($\text{lat} \geq 36^\circ\text{N}$), providing an approximation of the Mediterranean divide. These assignments are used as macro-region fixed effects and as a robustness design that restricts comparisons to languages within the same coarse macro-region.

6.2 Geocoding Integrated Values Survey Respondents

To locate respondents' contemporary place of residence, the ISO 3166-2 subnational region identifier reported in the survey (IVS) is used. Each unique country–region pair is geocoded in Python, and the resulting coordinates are assigned to all respondents residing in that region.

Distances (in kilometers) from each contemporary location to the nearest choke point and to the nearest natural harbour are then computed in QGIS. These two distance measures are merged into a single contemporary-distance, and log-transformed versions are created.

Finally, an Old World indicator is constructed for these locations: observations are coded as Old World (= 1) if the country is not in Oceania or the Americas, and as New World (= 0) otherwise.

²Glottolog can be accessed at <https://glottolog.org/glottolog/language>.

7 Distances from Integrated Values Survey Locations

I compute distance measures using two location concepts: ethnolinguistic locations and contemporary locations.

7.1 Ethnolinguistic distances

Ethnolinguistic locations are defined at the level of the respondent’s reported *language spoken at home*. Because a given language can correspond to multiple ethnolinguistic groups (and, conversely, a survey language label can map to more than one Glottolog language when IVS categories are less granular), the resulting distance measures may vary across matched Glottolog entries. I therefore compute distances at the observation level and then average them within each `language_spkn_home` category. All ethnolinguistic-distance calculations are restricted to languages whose matched geolocations are in the Old World.

7.2 Contemporary distances

Contemporary locations reflect respondents’ current place of residence (geocoded from ISO 3166-2 regions). To align the sample with the Old World focus of the paper, I restrict contemporary-distance calculations to respondents residing in Old World countries; in practice, this excludes observations in the Americas as well as Australia and New Zealand.

8 Conflict Occurrences

Conflict exposure is measured as the *number of conflict events within a 100 km radius*. This measure is constructed for (i) IVS respondents’ ethnolinguistic homelands, (ii) IVS respondents’ contemporary locations, and (iii) the ethnolinguistic group locations in the Folklore Catalogue.

For the IVS sample, the conflict-exposure control is computed for the historical periods 1000–1500, 1501–1800, and 1801–1989 based on respondents’ ethnolinguistic locations, and a contemporary measure is computed as the number of conflict events within 100 km of the respondent’s contemporary location in the survey year.

For the Folklore Catalogue ethnolinguistic groups, conflict exposure is computed for the periods 1000–1500 and 1501–1800.

9 Geographic Controls in the Contemporary Attitudes Analysis: Human Mobility and Caloric Suitability

Pre-1500 caloric suitability. I control for pre-1500 caloric suitability, which captures the local agro-ecological potential for caloric production under pre-industrial conditions [3]. For

each ethnolinguistic homeland location, I compute the average value within a 100 km radius and use its logarithm in the regressions.

Human Mobility Index (HMI). I also control for the Human Mobility Index (HMI), a cost-surface measure of terrain-based travel frictions that proxies how easily people could move across space [7, 8]. I compute the mean HMI within a 100 km radius around each homeland location and include its logarithm as a control.

10 Data

Main datasets are as follows:

- **Folklore data:** Berezkin Folklore Catalogue, obtained from the replication files of Michalopoulos and Xue [6].
- **Contemporary attitudes:** Integrated Values Surveys (IVS).
- Human Mobility Index (Sea) [7, 8] ³
- Pre-1500 Caloric Suitability Index [3] ⁴
- **Conflict data:** UCDP/PRIO Armed Conflict Dataset (Uppsala Conflict Data Program, UCDP), supplemented with data from the Harvard Dataverse replication package. ⁵
- **Natural harbour locations:** World Port Index (WPI), published by the U.S. National Geospatial-Intelligence Agency (NGA). ⁶
- **Language locations:** Glottolog data. ⁷

11 Regression Robustness Checks - Pre-industrial Values

Table 1 reports the full set of regression results analyzing the relationship between proximity to the closest historical choke points and two cultural indices: the *Tolerance Index* and the *Democracy Roots Index*. The main independent variable is the logarithm of distance to the nearest strategic trade choke point. ⁸

³Available at: <https://human-mobility-index.github.io/>.

⁴Available at: <https://ozak.github.io/Caloric-Suitability-Index/>.

⁵Available at: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/6ZFCOV>

⁶Available at: <https://msi.nga.mil/Publications/WPI>

⁷Available at: <https://glottolog.org/glottolog/language>

⁸The specification with country fixed effects includes 429 observations (compared to 477 in the specification with continent fixed effects). The smaller sample reflects the exclusion of singleton observations (i.e., countries represented by only one observation).

Columns (1)–(4) present estimates for the Tolerance Index using two different motif classification methods (Default and GPT) and different fixed effects. Columns (5)–(8) report analogous specifications for the Democracy Roots Index. In each case, Columns (1), (3), (5), and (7) include country fixed effects, while Columns (2), (4), (6), and (8) include continent fixed effects.

Across all specifications, the coefficient on $\text{Ln}(\text{Distance to Choke Point})$ is negative and statistically significant, indicating that ethnolinguistic groups located closer to historical trade choke points tend to display more democratic or prosocial cultural motifs. The magnitude and significance of the coefficients are robust to the use of alternative motif classification methods and different sets of fixed effects.

All models control for the logarithm of the year of first publication and the logarithm of the number of titles available per group. Standard errors are clustered at the language level. This table supplements the main text by providing a complete view of the empirical robustness of the findings beyond the baseline models shown in the paper.

Table 1: Regression Results: Pre-industrial Values (Berezkin) — Distance to Choke Points

	Tolerance Index				Democracy Roots Index			
	Default		GPT		Default		GPT	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Dist. to Closest Choke Point)	-0.267** (0.112) [0.026]	-0.400*** (0.119) [0.003]	-0.349*** (0.092) [0.001]	-0.182** (0.067) [0.012]	-0.226** (0.091) [0.021]	-0.108* (0.056) [0.067]	-0.208** (0.085) [0.023]	-0.157** (0.072) [0.039]
Ln(Year of First Publication)	1.511 (5.210) [0.774]	2.585 (4.560) [0.576]	5.045 (3.332) [0.144]	3.065 (2.939) [0.308]	-3.912 (4.415) [0.385]	-5.404 (3.607) [0.148]	1.905 (3.744) [0.616]	0.375 (3.799) [0.922]
Ln(Number of Titles)	-0.191** (0.091) [0.047]	-0.233*** (0.076) [0.006]	-0.047 (0.056) [0.411]	-0.009 (0.039) [0.811]	-0.038 (0.069) [0.582]	-0.018 (0.071) [0.802]	-0.020 (0.103) [0.848]	0.004 (0.095) [0.966]
Country FE	✓		✓		✓		✓	
Continent FE		✓		✓		✓		✓
Adjusted R^2	0.150	0.115	0.101	0.073	0.053	0.026	0.178	0.131
Observations	429	477	429	477	429	477	429	477

Notes: Coefficients are reported in the first row. Robust standard errors clustered at the language level are in parentheses. Exact p-values are reported in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The distance measure is the log of distance to the closest choke point computed from Berezkin ethnolinguistic locations.

Next, I account for the possibility that proximity to historical choke points may also be associated with a higher incidence of conflict. To address this, I include a control for the logarithm of the number of conflicts that occurred between 1000 and 1500 CE within a 100 km radius of each ethnolinguistic group’s geographic center. As shown in Table 2, the inclusion of this conflict exposure variable does not meaningfully alter the magnitude or significance of the estimated relationship between distance to choke points and either the Tolerance Index or the Democracy Roots Index. The coefficients on $\text{Ln}(\text{Distance to the Closest Choke Point})$ remain consistently negative and statistically significant across all specifications, suggesting that the results are not driven by historical conflict exposure.

Table 2: Regression Results: Controlling for Historical Conflict Exposure (1000–1500, 100km)

	Tolerance Index				Democracy Roots Index			
	Default		GPT		Default		GPT	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Dist. Choke Points, 6)	-0.255** (0.117) [0.040]	-0.352*** (0.110) [0.004]	-0.341*** (0.103) [0.003]	-0.165** (0.074) [0.035]	-0.222** (0.094) [0.027]	-0.107* (0.059) [0.084]	-0.210** (0.095) [0.037]	-0.139* (0.081) [0.097]
Ln(Conflicts 1000–1500, 100km)	0.028 (0.050) [0.585]	0.078** (0.036) [0.038]	0.019 (0.042) [0.659]	0.028 (0.027) [0.318]	0.009 (0.032) [0.794]	0.001 (0.020) [0.948]	-0.005 (0.040) [0.907]	0.029 (0.028) [0.306]
Ln(Year of First Publication)	1.407 (5.221) [0.790]	2.259 (4.508) [0.621]	4.975 (3.413) [0.158]	2.950 (2.999) [0.336]	-3.944 (4.395) [0.379]	-5.410 (3.587) [0.145]	1.922 (3.810) [0.619]	0.254 (3.862) [0.948]
Ln(Number of Titles)	-0.189** (0.091) [0.048]	-0.236*** (0.075) [0.005]	-0.046 (0.052) [0.390]	-0.011 (0.038) [0.783]	-0.038 (0.071) [0.598]	-0.018 (0.071) [0.800]	-0.020 (0.102) [0.844]	0.003 (0.096) [0.977]
Country FE	✓		✓		✓		✓	
Continent FE		✓		✓		✓		✓
Adjusted R^2	0.148	0.126	0.099	0.073	0.051	0.024	0.176	0.131
Observations	429	477	429	477	429	477	429	477

Notes: Standard errors are reported in parentheses and p-values are reported in brackets. Standard errors are clustered at the language level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Finally, as previously discussed, natural harbors are indicative of post-industrial trade dynamics, whereas the folklore catalogue reflects pre-industrial cultural values. Therefore, proximity to the closest natural harbor is not expected to be systematically associated with the tolerance- or democracy-related motifs captured in the dataset. Table 3 supports this expectation. While a few coefficients on $Ln(\text{Distance to the Closest Natural Harbour})$ are weakly significant in the Tolerance Index models (e.g., Columns 1 and 2), the results are inconsistent across specifications and generally small in magnitude. In the Democracy Roots Index models (Columns 5 to 8), the coefficients are either statistically insignificant or even slightly positive in sign, further indicating the absence of a meaningful relationship. Overall, these results confirm that exposure to post-industrial trade geography, as captured by distance to the closest natural harbor, does not explain variation in pre-industrial democratic values.

Table 3: Regression Results: Pre-industrial Values — Distance to Natural Harbour and Pre-industrial Indices

	Tolerance Index				Democracy Roots Index			
	Default		GPT		Default		GPT	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Dist. Natural Harbour)	-0.187*** (0.058) [0.004]	-0.148** (0.056) [0.014]	-0.068 (0.068) [0.330]	-0.063 (0.042) [0.145]	0.091 (0.061) [0.150]	0.078* (0.041) [0.068]	-0.074 (0.061) [0.241]	-0.027 (0.041) [0.509]
Ln(Year of First Publication)	-0.739 (4.132) [0.860]	0.568 (3.954) [0.887]	3.331 (3.123) [0.297]	2.155 (2.844) [0.456]	-4.087 (4.632) [0.387]	-5.771 (3.712) [0.134]	0.651 (3.581) [0.857]	-0.372 (3.588) [0.918]
Ln(Number of Titles)	-0.185* (0.090) [0.051]	-0.241*** (0.079) [0.006]	-0.049 (0.056) [0.391]	-0.014 (0.035) [0.702]	-0.047 (0.075) [0.540]	-0.037 (0.078) [0.639]	-0.019 (0.097) [0.845]	-0.004 (0.088) [0.968]
Country FE	✓		✓		✓		✓	
Continent FE		✓		✓		✓		✓
Adjusted R^2	0.163	0.088	0.083	0.064	0.051	0.031	0.175	0.122
Observations	429	477	429	477	429	477	429	477

Notes: Standard errors are reported in parentheses and p-values are reported in brackets. Standard errors are clustered at the language level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The inclusion of the English Channel does not materially affect the main findings (Table 4). Across specifications, the estimated association between pre-industrial trade exposure—proxied by log distance to the nearest choke point—and the pre-industrial indices remains negative and of comparable magnitude to the baseline six-choke point measure (Table 1). The only notable change is in column (6), where the coefficient is attenuated and becomes statistically marginal when continent fixed effects (rather than country fixed effects) are used for the default democracy-roots outcome (coefficient = -0.091 , $p = 0.107$). Overall, however, the pattern of results is stable, indicating that the core relationship is not driven by the inclusion of the English Channel.

Table 4: Regression Results: Pre-industrial Values — Distance to All Choke Points and Pre-industrial Indices

	Tolerance Index				Democracy Roots Index			
	Default		GPT		Default		GPT	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Dist. All Choke Points)	-0.285** (0.108) [0.014]	-0.435*** (0.113) [0.001]	-0.339*** (0.090) [0.001]	-0.190** (0.071) [0.014]	-0.223** (0.091) [0.022]	-0.091 (0.054) [0.107]	-0.197** (0.084) [0.029]	-0.167** (0.075) [0.036]
Ln(Year of First Publication)	1.628 (5.143) [0.754]	3.038 (4.499) [0.506]	5.068 (3.316) [0.140]	3.232 (2.933) [0.282]	-3.884 (4.413) [0.388]	-5.416 (3.641) [0.150]	1.900 (3.743) [0.617]	0.530 (3.794) [0.890]
Ln(Number of Titles)	-0.190** (0.091) [0.048]	-0.226*** (0.076) [0.007]	-0.046 (0.057) [0.420]	-0.007 (0.039) [0.853]	-0.038 (0.068) [0.586]	-0.018 (0.071) [0.797]	-0.020 (0.103) [0.851]	0.006 (0.095) [0.949]
Country FE	✓		✓		✓		✓	
Continent FE		✓		✓		✓		✓
Adjusted R^2	0.151	0.125	0.100	0.074	0.053	0.025	0.178	0.132
Observations	429	477	429	477	429	477	429	477

Notes: Standard errors are reported in parentheses and p-values are reported in brackets. Standard errors are clustered at the language level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

12 Regression Robustness Checks - Contemporary Support for Democracy

Table 5: OLS Regression Results: Support for Democracy (Ethnolinguistic Distances)

	Distance to Closest Choke Points					Distance to Closest Natural Harbors				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ln(Dist. Closest Choke Points)	0.174*	0.175*	0.197**	0.185*	0.195**					
	(0.098)	(0.098)	(0.093)	(0.101)	(0.096)					
	[0.081]	[0.077]	[0.038]	[0.071]	[0.046]					
Ln(Dist. Closest Natural Harbors)						-0.102**	-0.101**	-0.105**	-0.113**	-0.108*
						(0.048)	(0.048)	(0.052)	(0.055)	(0.057)
						[0.037]	[0.037]	[0.047]	[0.045]	[0.061]
Town size				0.005					0.005*	
				(0.003)					(0.003)	
				[0.122]					[0.093]	
Male		0.006	0.006	0.008	0.005		0.006	0.006	0.008	0.005
		(0.005)	(0.005)	(0.005)	(0.005)		(0.005)	(0.005)	(0.005)	(0.005)
		[0.237]	[0.188]	[0.106]	[0.288]		[0.231]	[0.221]	[0.109]	[0.330]
Age		0.001***	0.001***	0.001***	0.001***		0.001***	0.001***	0.001***	0.001***
		(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)
		[0.001]	[0.001]	[0.001]	[0.000]		[0.001]	[0.001]	[0.001]	[0.001]
Employed		0.022***	0.019***	0.021***			0.020***	0.018***	0.020***	
		(0.006)	(0.006)	(0.006)			(0.006)	(0.006)	(0.006)	
		[0.001]	[0.004]	[0.001]			[0.001]	[0.004]	[0.001]	
Married		-0.001	0.000	-0.000	0.001		0.001	0.001	0.001	0.002
		(0.006)	(0.006)	(0.006)	(0.006)		(0.005)	(0.005)	(0.006)	(0.006)
		[0.895]	[0.960]	[0.979]	[0.817]		[0.887]	[0.861]	[0.890]	[0.751]
Atheist					-0.017					-0.016
					(0.025)					(0.024)
					[0.498]					[0.497]
Muslim					-0.051					-0.029
					(0.040)					(0.033)
					[0.213]					[0.371]
Christian					-0.028					-0.028
					(0.023)					(0.022)
					[0.227]					[0.223]
Hindu					-0.020					-0.033
					(0.034)					(0.030)
					[0.562]					[0.272]
Buddhist					-0.031					-0.022
					(0.028)					(0.025)
					[0.272]					[0.387]
Town size (included)				✓					✓	
Individual controls		✓	✓	✓	✓		✓	✓	✓	✓
Religion dummies					✓					✓
Origin macro-region FE			✓	✓	✓			✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
R ²	0.188	0.189	0.190	0.195	0.191	0.190	0.191	0.192	0.198	0.194
Observations	101,345	100,894	100,894	97,295	97,899	101,345	100,894	100,894	97,295	97,899

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses; p-values in square brackets. Standard errors are clustered at the country level. Columns (1)–(5) use distance to choke points; columns (6)–(10) use distance to natural harbors. Individual controls include male, age, employment status, and marital status. Religion coefficients are reported only in specifications that include religion dummies.

Furthermore, these results highlight the enduring historical influence of trade on cultural and political attitudes. The significance of geographic proximity to choke points or natural harbors is evident when measured based on the respondent’s ethnic group’s historical location. However, this significance disappears when using the respondent’s contemporary location. While the positive impact of pre-1500 trade on pre-industrial values was disrupted by the 16th-century geographical discoveries and further reshaped by advancements in sea-faring during the 19th century—leading to the emergence of natural harbors as new trade hubs—the historical location of an ethnic group relative to these key trade routes remains a strong predictor of support for democracy. In contrast, contemporary proximity to these geographical features does not exhibit a statistically significant effect, as demonstrated in the analysis below in the table 6

Table 6: OLS Regression Results: Support for Democracy and Contemporary Distances (Old World, Excluding Immigrants)

	Contemporary Distance to Choke Points				Contemporary Distance to Natural Harbors			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Contemp. Dist. to Choke Points)	0.011 (0.022) [0.610]	0.011 (0.022) [0.610]	0.010 (0.029) [0.726]	0.002 (0.023) [0.932]				
Ln(Contemp. Dist. to Natural Harbors)					-0.016 (0.011) [0.148]	-0.016 (0.011) [0.147]	-0.015 (0.011) [0.189]	-0.016 (0.011) [0.158]
Town size (coefficient)			0.008** (0.004) [0.035]				0.007* (0.004) [0.064]	
Male		0.009 (0.006) [0.159]	0.012* (0.006) [0.067]	0.009 (0.006) [0.162]		0.009 (0.006) [0.155]	0.011* (0.006) [0.069]	0.009 (0.006) [0.153]
Age		0.001*** (0.000) [0.000]	0.001*** (0.000) [0.000]	0.001*** (0.000) [0.000]		0.001*** (0.000) [0.000]	0.001*** (0.000) [0.000]	0.001*** (0.000) [0.000]
Employed		0.025*** (0.008) [0.002]	0.022*** (0.008) [0.008]	0.025*** (0.007) [0.001]		0.024*** (0.007) [0.002]	0.021*** (0.008) [0.008]	0.024*** (0.007) [0.001]
Married		0.002 (0.007) [0.824]	0.001 (0.007) [0.893]	0.004 (0.007) [0.618]		0.002 (0.007) [0.729]	0.002 (0.007) [0.831]	0.004 (0.007) [0.539]
Atheist				-0.082* (0.043) [0.065]				-0.083* (0.043) [0.058]
Muslim				-0.131** (0.052) [0.015]				-0.130** (0.051) [0.013]
Christian				-0.071 (0.046) [0.127]				-0.072 (0.045) [0.114]
Hindu				-0.105* (0.061) [0.089]				-0.105* (0.060) [0.085]
Buddhist				-0.096** (0.041) [0.021]				-0.097** (0.039) [0.017]
Town size (included)			✓				✓	
Individual controls		✓	✓	✓		✓	✓	✓
Religion dummies				✓				✓
Country FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
R ²	0.190	0.192	0.204	0.193	0.191	0.193	0.205	0.194
Observations	93,023	92,682	86,225	90,362	93,023	92,682	86,225	90,362

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses; p-values in square brackets. Standard errors are clustered at the country level. The sample is restricted to respondents located in Old World countries and excludes immigrants. Individual controls include male, age, employment status, and marital status. Religion coefficients are reported only in specifications that include religion dummies.

Finally, I control for historical and contemporary exposure to conflict to further assess

robustness (Table 7). Specifically, in models (1)–(3), I account for historical conflict exposure measured at respondents’ ethnolinguistic homeland locations, separately for the periods 1000–1500, 1501–1800, and 1801–1989. This addresses the concern that past conflict could shape contemporary support for democracy through culturally persistent legacies. In model (4), I additionally control for contemporary conflict exposure measured at respondents’ current locations. Across specifications, the results remain stable: historical proximity to major choke points continues to positively predict contemporary democratic attitudes, independent of both historical and contemporary conflict exposure.⁹ In both cases, exposure is operationalized as the (log) number of conflict events within a 100 km radius of the relevant location, with conflicts assigned to the corresponding year(s) of activity.

⁹Historical conflict exposure is based on the Historical Conflict Event Dataset (HCED), using the HCED Data v2 shapefile available via Harvard Dataverse (DOI: 10.7910/DVN/6ZFC0V). Contemporary conflict exposure is constructed from the Uppsala Conflict Data Program (UCDP) georeferenced event data, available from the UCDP website <https://ucdp.uu.se/>

Table 7: Effect of Choke Point Distance and Conflict Exposure on Support for Democracy

	Support for Democracy			
	(1)	(2)	(3)	(4)
Ln(Distance to Closest Choke Point)	0.212** (0.097) [0.032]	0.195** (0.093) [0.038]	0.197* (0.101) [0.055]	0.214** (0.098) [0.032]
Ln(Conflicts 1000–1500, 100km)	0.035** (0.016) [0.035]			
Ln(Conflicts 1501–1800, 100km)		0.003 (0.011) [0.793]		
Ln(Conflicts 1801–1989, 100km)			-0.001 (0.013) [0.960]	
Ln(Contemporary Conflicts, 100km)				0.003 (0.003) [0.278]
Male	0.006 (0.005) [0.180]	0.006 (0.005) [0.189]	0.006 (0.005) [0.185]	0.005 (0.005) [0.263]
Age	0.001*** (0.000) [0.001]	0.001*** (0.000) [0.001]	0.001*** (0.000) [0.001]	0.001*** (0.000) [0.002]
Employed	0.022*** (0.006) [0.001]	0.021*** (0.006) [0.001]	0.021*** (0.006) [0.001]	0.021*** (0.007) [0.002]
Married	0.000 (0.006) [0.939]	0.000 (0.006) [0.956]	0.000 (0.006) [0.965]	0.001 (0.006) [0.869]
Year FE	✓	✓	✓	✓
Country FE	✓	✓	✓	✓
Origin macro-region FE	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓
R^2	0.191	0.190	0.190	0.193
Observations	100,894	100,894	100,894	94,082

Notes: Standard errors are clustered at the country level and reported in parentheses. P-values are reported in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

As an additional robustness check, I control for pre-1500 caloric suitability [3] and the Human Mobility Index [7, 8], which proxy for agricultural potential and mobility frictions, respectively. Since higher HMI indicates higher movement costs and stronger geographic isolation, it also captures persistent barriers to interaction at ethnolinguistic homelands. Including these controls leaves the main coefficients on distance to choke points and natural harbors essentially unchanged in sign and magnitude, and the interpretation remains the same (Table 8).

Table 8: OLS Regression Results: Support for Democracy (Geographic Controls Added)

	Distance to Closest Choke Points					Distance to Closest Natural Harbors				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ln(Dist. Closest Choke Points)	0.174*	0.175*	0.226**	0.213*	0.224**					
	(0.101)	(0.100)	(0.099)	(0.118)	(0.102)					
	[0.090]	[0.083]	[0.025]	[0.075]	[0.032]					
Ln(Dist. Closest Natural Harbors)						-0.103**	-0.102**	-0.104*	-0.114**	-0.107*
						(0.048)	(0.047)	(0.053)	(0.056)	(0.058)
						[0.035]	[0.036]	[0.052]	[0.044]	[0.068]
Ln(Caloric Suitability Index, pre-1500)	0.007	0.007	-0.002	-0.002	-0.002	0.013	0.014	0.007	0.007	0.007
	(0.013)	(0.013)	(0.011)	(0.014)	(0.011)	(0.014)	(0.014)	(0.010)	(0.012)	(0.010)
	[0.609]	[0.604]	[0.824]	[0.902]	[0.819]	[0.349]	[0.343]	[0.511]	[0.586]	[0.509]
Ln(Human Mobility Index, Sea10)	0.253	0.260*	0.288**	0.278	0.304**	-0.003	0.007	0.021	-0.038	0.034
	(0.152)	(0.150)	(0.144)	(0.174)	(0.152)	(0.110)	(0.109)	(0.112)	(0.107)	(0.119)
	[0.101]	[0.089]	[0.048]	[0.114]	[0.049]	[0.978]	[0.947]	[0.853]	[0.726]	[0.776]
Town size (coefficient)				0.005					0.005	
				(0.003)					(0.003)	
				[0.114]					[0.101]	
Male		0.006	0.006	0.008	0.005		0.006	0.006	0.008	0.005
		(0.005)	(0.005)	(0.005)	(0.005)		(0.005)	(0.005)	(0.005)	(0.005)
		[0.232]	[0.197]	[0.112]	[0.306]		[0.201]	[0.211]	[0.104]	[0.316]
Age		0.001***	0.001***	0.001***	0.001***		0.001***	0.001***	0.001***	0.001***
		(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)
		[0.001]	[0.001]	[0.001]	[0.000]		[0.001]	[0.001]	[0.001]	[0.001]
Employed		0.022***	0.022***	0.019***	0.022***		0.020***	0.020***	0.018***	0.020***
		(0.006)	(0.006)	(0.006)	(0.006)		(0.006)	(0.006)	(0.006)	(0.006)
		[0.001]	[0.001]	[0.003]	[0.001]		[0.001]	[0.001]	[0.004]	[0.001]
Married		-0.001	0.000	-0.000	0.001		0.001	0.001	0.001	0.002
		(0.006)	(0.006)	(0.006)	(0.006)		(0.005)	(0.005)	(0.006)	(0.006)
		[0.915]	[0.957]	[0.991]	[0.810]		[0.839]	[0.847]	[0.884]	[0.735]
Atheist					-0.017					-0.016
					(0.025)					(0.024)
					[0.492]					[0.509]
Muslim					-0.051					-0.030
					(0.039)					(0.033)
					[0.199]					[0.360]
Christian					-0.030					-0.027
					(0.023)					(0.022)
					[0.198]					[0.227]
Hindu					-0.047					-0.036
					(0.037)					(0.031)
					[0.212]					[0.247]
Buddhist					-0.024					-0.021
					(0.027)					(0.026)
					[0.367]					[0.411]
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Individual controls		✓	✓	✓	✓		✓	✓	✓	✓
Origin macro-region FE			✓	✓	✓			✓	✓	✓
Town size				✓					✓	
Religion dummies					✓					✓
R ²	0.188	0.189	0.190	0.195	0.192	0.191	0.192	0.192	0.198	0.194
Observations	101,345	100,894	100,894	97,295	97,899	101,345	100,894	100,894	97,295	97,899

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are reported in parentheses and p-values are reported in brackets (both clustered at the country level where applicable). All specifications include year and country fixed effects. Columns (1)–(5) use ln distance to the closest choke point; Columns (6)–(10) use ln distance to the closest natural harbor. Individual controls include male, age, employment status, and marital status. “Ln(Caloric Suitability Index, pre-1500)” is the log of the 100km-buffer mean of pre-1500 caloric suitability around ethnolinguistic homelands; “Ln(Human Mobility Index, Sea10)” is the log of the 100km-buffer mean of the Human Mobility Index with seafaring around ethnolinguistic homelands.

The results presented in Table 9 indicate that while the overall pattern of cultural revision persists, the statistical significance of the findings weakens when the English Channel is included as a choke point.

In this study, distance to the closest choke point serves as a proxy for exposure to pre-1500 trade. In the European context, however, the English Channel is less representative of pre-1500 intercontinental exchange than the Bosphorus or Gibraltar, which historically functioned as major gateways connecting the Mediterranean to wider Eurasian and African trade networks. The English Channel, by contrast, is more closely associated with the post-1500 shift toward Atlantic commerce following the discovery of the Americas and the opening of the Cape of Good Hope route. Consistent with this historical interpretation, when the English Channel is added to the choke-point set (7 choke points), the estimated coefficient on log distance is attenuated in the specifications where the core identification is implemented: the coefficient ranges from 0.166 to 0.180 in models (3)–(5). By contrast, using the baseline set of six pre-1500–relevant choke points yields somewhat larger estimates over the same specifications, ranging from 0.185 to 0.197. This pattern suggests that including a choke point whose economic salience is primarily post-1500 introduces measurement noise into the pre-1500 trade-exposure proxy, weakening both the magnitude and precision of the estimated association with contemporary democratic attitudes.

Table 9: OLS Regression Results: Distance to Choke Points (7) and Support for Democracy

	Support for Democracy				
	(1)	(2)	(3)	(4)	(5)
Ln(Distance to 7 Choke Points)	0.145 (0.097) [0.140]	0.148 (0.097) [0.131]	0.180* (0.092) [0.054]	0.166* (0.099) [0.097]	0.178* (0.095) [0.064]
Male		0.006 (0.005) [0.242]	0.006 (0.005) [0.190]	0.008 (0.005) [0.107]	0.005 (0.005) [0.291]
Age		0.001*** (0.000) [0.001]	0.001*** (0.000) [0.001]	0.001*** (0.000) [0.001]	0.001*** (0.000) [0.000]
Employed		0.022*** (0.006) [0.001]	0.021*** (0.006) [0.001]	0.019*** (0.006) [0.004]	0.021*** (0.006) [0.001]
Married (binary)		-0.001 (0.006) [0.895]	0.000 (0.006) [0.953]	-0.000 (0.006) [0.987]	0.001 (0.006) [0.808]
Town size				0.005 (0.003) [0.125]	
Atheist					-0.016 (0.025) [0.507]
Muslim					-0.051 (0.041) [0.212]
Christian					-0.028 (0.023) [0.231]
Hindu					-0.021 (0.034) [0.528]
Buddhist					-0.031 (0.028) [0.278]
Year FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Origin Macro-Region FE	No	No	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	Yes	Yes
Town size	No	No	No	Yes	No
Religion dummies	No	No	No	No	Yes
R^2	0.187	0.188	0.190	0.195	0.191
Observations	101,345	100,894	100,894	97,295	97,899

Notes: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Standard errors are clustered at the country level

13 Conflicts and Distance to Choke Points

Historically important choke points have frequently been contested due to their strategic value in global trade. Using conflict data from the Uppsala Conflict Data Program (1989–2023) and the Historical Conflict Event Dataset (covering periods prior to 1989), Tables 10 and 11 show that regions closer to major choke points—such as the Bosphorus, Suez, and the Strait of Malacca—experienced significantly higher frequencies of conflicts. This pattern highlights how the geopolitical and economic importance of these locations historically attracted conflict.

The country-level analysis controls for important geographic factors—including % of fertile soil, terrain ruggedness, and distance to coast—and includes continent-fixed effects to account for unobserved regional heterogeneity. The ethnolinguistic group-level analysis employs either country or continent fixed effects (examined separately to test robustness). Both analyses consistently underline the pivotal role of choke points as centers of historical contestation.

Table 10: Effect of Distance to Choke Points on Conflict (Country-Level)

	1000–1500		1501–1800		1801–1989		1990–2023	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Distance to Choke Points)	-1.044*** (0.331) [0.002]		-0.731** (0.327) [0.027]		-0.952*** (0.321) [0.004]		-1.007*** (0.308) [0.001]	
Ln(Distance to Choke Points: English Channel)		-1.206*** (0.340) [0.001]		-0.915*** (0.326) [0.006]		-0.925*** (0.322) [0.005]		-1.201*** (0.310) [0.000]
Absolute Latitude	0.079*** (0.022) [0.000]	0.075*** (0.021) [0.001]	0.069** (0.027) [0.011]	0.067** (0.026) [0.012]	0.035 (0.026) [0.176]	0.030 (0.025) [0.241]	-0.022 (0.024) [0.349]	-0.025 (0.023) [0.279]
Ln(Soil Productivity)	0.388*** (0.078) [0.000]	0.403*** (0.076) [0.000]	0.360*** (0.092) [0.000]	0.376*** (0.090) [0.000]	0.473*** (0.093) [0.000]	0.471*** (0.093) [0.000]	0.215*** (0.077) [0.006]	0.233*** (0.076) [0.003]
Ln(Ruggedness)	-0.079 (0.165) [0.632]	-0.048 (0.158) [0.761]	-0.267 (0.214) [0.215]	-0.248 (0.209) [0.238]	-0.316 (0.211) [0.137]	-0.282 (0.210) [0.181]	-0.120 (0.219) [0.584]	-0.091 (0.215) [0.673]
Ln(Distance to Coast)	0.510*** (0.163) [0.002]	0.556*** (0.162) [0.001]	0.366* (0.195) [0.062]	0.402** (0.196) [0.042]	0.553** (0.214) [0.011]	0.586*** (0.215) [0.007]	0.824*** (0.189) [0.000]	0.871*** (0.185) [0.000]
Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Adjusted R^2	0.485	0.500	0.375	0.388	0.262	0.258	0.332	0.352
Observations	144	144	144	144	144	144	144	144

Notes: Robust standard errors are reported in parentheses and p-values are reported in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Effect of Distance to Choke Points on Conflict (Berezkin Ethnolinguistic Groups)

	1000–1500		1501–1800	
	(1)	(2)	(3)	(4)
Ln(Distance to 6 Choke Points)	−0.575*** (0.153) [0.000]	−0.389** (0.164) [0.018]	−0.271* (0.163) [0.098]	−0.336 (0.228) [0.142]
Continent FE	Yes	No	Yes	No
Country FE	No	Yes	No	Yes
Observations	487	438	487	438
Adjusted R^2	0.366	0.470	0.276	0.313

Notes: Robust standard errors are reported in parentheses and p-values are reported in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

References

- [1] Daron Acemoglu, Nicolás Ajzenman, Cevat Giray Aksoy, Martin Fiszbein, and Carlos Molina. (successful) democracies breed their own support. *Review of Economic Studies*, 92(2):621–655, 2025.
- [2] Benjamin Enke. Market exposure and human morality. *Nature Human Behaviour*, 7(1): 134–141, 2023.
- [3] Oded Galor and Ömer Özak. Land productivity and economic development: Caloric suitability vs. agricultural suitability. *Agricultural Suitability (July 12, 2015)*, 2015.
- [4] Natalie Henrich and Joseph Patrick Henrich. *Why humans cooperate: A cultural and evolutionary explanation*. Oxford University Press, 2007.
- [5] Ronald Inglehart. *Christian Welzel Modernization, Cultural Change, and Democracy The Human Development Sequence*. Cambridge: Cambridge university press, 2005.
- [6] Stelios Michalopoulos and Melanie Meng Xue. Folklore. *The Quarterly Journal of Economics*, 136(4):1993–2046, 2021.
- [7] Ömer Özak. The voyage of homo-economicus: Some economic measures of distance, 2010.
- [8] Ömer Özak. Distance to the pre-industrial technological frontier and economic development. *Journal of Economic Growth*, 23(2):175–221, 2018.