CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 749

Brave Old World: On the Historical Roots of Interstate Conflicts

Marcello D'Amato and Francesco Flaviano Russo

April 2025

University of Naples Federico II

University of Salerno

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS AND STATISTICS – UNIVERSITY OF NAPLES FEDERICO II
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 749

Brave Old World: On the Historical Roots of Interstate Conflicts

Marcello D'Amato*and Francesco Flaviano Russo †

Abstract
We show empirically that interstate conflicts are less likely among countries that share more of their oral tradition, as enshrined in the folktales. Popular tales and narratives are related to expectations and beliefs held about the other parties behavior: larger similarity in the systems of beliefs in the populations - cultural relatedness- reduces information frictions in dispute resolutions and negotiation failures between states. To validate this interpretation, we show that countries with more oral tradition in common are more likely to form military alliances, more likely to participate to the same international organizations, more likely to vote similarly in the UN general assembly, more likely to trade with each other and, in case a conflicts breaks out, more likely to terminate it with a negotiation.

JEL Classification: F5, N4, Z1.

Keywords: Ethnic Culture; Narratives; Cultural distances, Interstate conflicts.

Acknowledgments: We would like to thank Alberto Bisin and Massimo Morelli for very helpful conversations. We also thank Antonio Acconcia, Tullio Jappelli, Karen Macours, Enrico Spolaore, and Marco Stimolo for comments and suggestions. Albachiara Abate provided excellent research assistance. Financial Assistance by the Italian Ministry of the University and Research, Project: PRIN – CUP D53D23017970001, Title: Cultural Background, Social and Political Identity of individuals and groups, their Dynamics and the Role of Policies, is gratefully acknowledged. All remaining errors are ours.

* University of Naples Suor Orsola Benincasa and CSEF. Email: marcello.damato@unisob.na.it.

† University of Naples Federico II and CSEF. Email: francescoflaviano.russo@unina.it *(Corresponding Author)*

De te fabula narratur.

About you the tale is told.

Orazio, Satire, I, 1, 69-70

The Leontines, invoking their ancient alliance with Athens and their shared Ionian identity, obtained Athenian support.

Thucydides, Peloponnesian War, Book III, Chapter 86, Hobbes trans.

1 Introduction

The quote in the epigraph reports the brief account by Thucidides of the role of shared cultural ties in the negotiations that led the Athenians to embark in what would turn out to be a self-defeating invasion of Sicily[1]. Since then, the role of cultural traits in international relations has been extensively investigated by both economists and political scientists. The central tenet is that different cultural identities determine negotiation failures leading to conflicts, mainly because of their impact on strategic expectations, information asymmetries, scope for commitments, or as frictions at several stages of the dispute resolution process (Russett et al. 2000; Henderson and Tucker 2001; Gartzke and Gleditsch 2006; Dube et al. 2024; Baliga and Sjöström 2024; Acemoglu and Wolitzky 2024; Genicot and Ray 2024).

The objective of this paper is to offer new empirical evidence on the relationship between culture and interstate conflicts by providing a new index of cultural similarity that measures the extent to which oral traditions overlap in ethnic group and country pairs. Starting from

[1]It is unclear whether Thucidides deems the appeal to the common Ionian root by Leontines a rethoric argument to persuade Athenians and motivate their intervention on emotional grounds, an element to account for the benefits perceived by Athenians for their decision, or as an indication of a more fundamental strategic reason for why close cultural ties could cement the alliance between city states in the Attic League. Elsewhere, Thucidides strongly argues for motives and causes of the break of the peace of Nicias with the Sparthans in line with what is today known as the "realist" approach to armed conflicts. For instance: " [...] Although peace was nominally in effect, each side suspected the other of violations. Mutual mistrust, fear, and strategic calculation undermined the treaty, making renewed conflict inevitable".

the Berezkin's (2015) folklore catalog, we compute the similarity between the folklore motifs, or structural elements of the traditional stories in the oral tradition of two ethnic groups, using text analysis, and then aggregate it at the country level, in order to explain interstate conflicts. We argue that this measure is particularly apt to explore the relationship between culture and conflict because, as shown in Bortolini et al. (2017) and in D'Amato and Russo (2025), oral tradition similarities are a marker of the extent to which ethnic groups horizontally shared information, values and beliefs. Given that folktales are, in traditional societies, one of the important ways through which moral systems are inter-generationally transmitted, folklore similarities should reflect similarities in the moral systems, which are fundamental determinants of behavior in social relationships[2] (Enke 2009).

The main idea is that different oral traditions, reflecting divergent cultural histories, represent a credible measure of differences in expectations and beliefs, with bigger distances heightening mistrust, impeding effective communication, exacerbating uncertainty, and fostering misperceptions and misunderstanding (Acemoglu and Wolitzky, 2024; Baliga and Sjöström, 2024), thereby increasing the probability of a conflict. In other words, our folklore-based measure of cultural similarity provides an explicit empirical proxy for historically shaped similarity in "cultural beliefs" (Greif, 1994), that is shared expectations about other players' behavior in strategic interactions, including off-equilibrium beliefs in the face of unexpected events such as defections or commitment violations. Rather than considering cultural identities as inherently antagonistic, as in the Huntington's clash of civilizations hypothesis, we emphasize historically evolved overlaps in systems of expectations and beliefs about reciprocity, trustworthiness, credible enforcement, and cooperation, as fundamental to peaceful equilibria[3].

An important, methodological, advantage of our folklore based measure of cultural similarity is that it is based on motifs or "replication units" in folkloric traditions, which are unlikely to change over time, therefore making it plausibly exogenous to modern and con-

[2]The empirical relationship between traditional stories and contemporary beliefs provided in Michalopolous and Xue (2021) supports this interpretation.

[3]Examples include the selection of focal points and the emergence of cooperation in repeated games. Note that the definition in Grief (1994) encompasses also expectations about off-equilibrium events such as defections or breaches of commitments. This is particularly important in the context of International Relations where cooperative equilibria can be supported in repeated interactions with no third-party enforcement (the "State of Anarchy" of the realist tradition) only if agents coordinate also on systems of off-equilibrium beliefs. In such a setting, cooperative equilibria can be sustained by reputational mechanisms and by tit for tat (Axelrod 2006).

temporary conflicts. This excludes potential reverse causation or simultaneity concerns that may affect other cultural similarity measures such as the ones based on language[4] or current surveys, used, among others, by Bove and Gokmen (2017). Moreover, the advantage of our measure of cultural similarity with respect to alternatives based on genetic markers (Spolaore and Wacziarg 2016) is that the latter captures the time since two population separated (Cavalli-Sforza 2001), rather than the intensity of their cultural relationships in their history.

We deem narratives essential for the exploration of the culture-conflict nexus also because, in a broader perspective, they are key social constructs that shape the way people perceive the social and natural environment around them, influencing the way other people's actions are interpreted and, therefore, affecting the outcomes of bilateral negotiations. Indeed, recent contributions from narrative economics (Akerlof and Snower, 2016; Shiller, 2017) suggest that culturally transmitted narratives shape cognitive frames, biases, stereotypes and collective expectations, influencing behavioral aspects of strategic interactions. For instance, widely shared narratives can foster empathic understanding, reduce biases in threat perception, and promote trust-based behaviors, reinforcing peaceful bargaining outcomes. Thus, our measure of folklore similarity provides empirical grounding for behavioral mechanisms that complement rational-choice explanations of conflict[5].

Our empirical results provide robust evidence for the view that greater oral tradition similarity decreases the likelihood of interstate conflicts and, so, the number of conflicts observed between two countries between 1816 and 2010, as recorded in the Correlates of War (COW) database (Palmer et al. 2020). In our benchmark regression, moving from the first to the third quartile of folklore similarity between two countries implies a reduction of the probability of a conflict by 34%. The results are robust to the inclusion of various potential confounding factors in the regressions, including: other measures of bilateral cultural similarity computed starting from genetic, linguistic and religious markers; measures of geographic barriers; controls for similar historical exogenous shocks; country fixed effects, which account for specific country characteristics such as the political systems or the presence of natural resources and disputed territories; time fixed effects, which account for sample periods with peaks in the

[4]On the endogeneity of linguistic distance measures due to loan-words, see Blouin and Dyer (2023).

[5]For the reasons why narratives may matter in the context of international relations see Mercer (2005) and Herrmann and Fischerkeller (1995), among others. For the importance of misperceptions in international relations see Jervis (1976).

number of conflicts such as those around the World Wars. We also show that our results are robust across different possible alternative measures of folklore similarity. One empirical challenge we face is the potential endogeneity of country-level measures of folklore similarity, aggregated from the ethnic-group level measures using contemporary population shares by ethnicity, which can be the result of past conflicts. To overcome this challenge, we consider an alternative measure of ancestral folklore similarity between the ethnic groups with centroid, or ancient locations, within the contemporary country borders, finding the same empirical results.

To validate our interpretation of the empirical findings, we provide additional evidence on the effects of cultural similarity between country pairs on other indicators of their ability, or willingness, to cooperate and coordinate internationally. In particular, we find that countries sharing more of their oral traditions are also more likely to: form military alliances, participate in the same international organizations, vote similarly at the United Nations General Assembly, and- importantly- more likely to sustain stronger and more stable bilateral trade relationships. Moreover, we also show that conflicts are more likely to end with a negotiated settlement, rather than imposed, in case of more oral tradition shared. Finally, we find that a concept-specific measure of folklore similarity, computed using only the folklore motifs associated with the concepts related to beliefs and strategies that might sustain cooperation in repeated interactions, i.e. those related to punishment, cooperation, threats, reputation, etc., turns out to be negatively and significantly associated to the probability of a conflict. Taken together, these findings suggest that cultural similarities measured with folklore significantly reduce transaction costs, strategic uncertainty, and misperceptions in international negotiations, facilitating coordination (voting patterns, trade, participation to international organizations), helping bilateral monitoring, fostering reputational equilibria and facilitating commitment (alliances), so to make peaceful solutions more readily available, consistently with established theoretical frameworks in the international relations literature.

There are also several historical examples that illustrate the role of cultural similarities in bilateral negotiations, in line with our interpretation of the empirical results. For instance, in the classic age cultural differences between Athens and Persia contributed to persistent mistrust, causing difficulties in establishing credible peace agreements despite mutual inter-

ests in avoiding costly conflicts (Thucydides, Book 1.102). Conversely, culturally similar Greek city-states benefited significantly from shared religious festivals, myths, and institutional frameworks, allowing credible diplomatic commitments, fostering alliances, and, more generally, reducing uncertainty (Thucydides, Book 1.85). The relationships between European crusaders and culturally different local Middle Eastern populations lacked flexible bargaining frameworks, without a clearer mutual understanding of the symbolic value of specific offers and counter-offers in a bargaining process over indivisible contested religious sites, without the possibility of reaching a compromise with mutual concessions. Similarly, the long-term interactions between medieval Genoese and Maghribi traders, discussed by Greif (1994), where differences in cultural beliefs shaped distinct institutional solutions to disputes in the context of commerce and trade. Analogously, repeated conflicts India and Pakistan after 1947 exemplify how divergent narratives shaped strategic mistrust and how they impede credible commitment, despite shared economic interests. Indo-Greek encounters under Alexander's successors demonstrated that, despite profound differences in cultural traditions, initial interactions could lead either toward mutual trade and understanding or persistent hostility—depending precisely on whether culturally shaped expectations and informal institutions emerged to reduce uncertainty, mistrust, and bargaining failures.

The rest of the paper is organized as follows. Section 2 discusses the related literature; Section 3 briefly describes the measures of cultural similarity based on folklore; Section 4 explains the empirical strategy; Section 5 summarizes the main results and their robustness; Section 6 discusses the mechanisms linking folklore similarity to conflict; Section 7 concludes.

2 Related Literature and Contribution

The main contribution of this work is to show, empirically, that cultural similarities, measured with text analysis on folklore motifs, decrease the probability of interstate conflicts in the modern era. We therefore contribute to the literature on the cultural determinants of conflicts, at the intersection of political relations, political science and economics (Russett et al., 2000; Henderson and Tucker, 2001).

Within the international relations literature, classical realism (Carr, 1939; Morganthau,

1948) considers the existence of human traits such as power seeking or security concerns an important determinant of international conflicts, but these traits are considered universal and inherent to the human nature, not to specific cultures, although some scholars, such as Aron (1966), explicitly acknowledge the existence of variation, both across cultures and across different historical periods, in the way such traits determine the emergence of group identities within political organizations (Connor, 1978; Smith, 1986; Huntington, 1993). Structural realism (Waltz, 1979; Mearsheimer, 2001), on the other hand, focuses the attention on the distribution of power and on its determinants in the absence of a legal framework and of third party enforcement in dispute resolutions. Finally the rationalist approach (Fearon, 1995; Powell, 1999, 2006) nests structural elements into game-theoretic settings based on bargaining frictions, commitment problems, asymmetric information, and indivisibilities. Our interpretation of the relationships between cultural differences and armed conflicts explicitly aligns with rationalist perspectives (Fearon, 1995; Powell, 2006; Gartzke and Gleditsch, 2006; Acemoglu and Wolitzky, 2024), in which interstate violence arises strictly from bargaining failures (Kant 1795, Shelling 1966, Dube et al. 2024) due to information asymmetries, commitment problems, or issue indivisibilities. Cultural similarity, as we measure it, significantly reduces misunderstandings, stereotypes, biases, and inaccurate threat assessments, thus mitigating informational asymmetries.

Our paper is especially motivated by Acemoglu and Wolitzky (2024), who provide a general framework for analyzing the role of information frictions in explaining the onset, persistence, and termination of international conflicts. They explicitly argue that "misperceptions may be more likely between individuals or groups with less shared history or culture". The index we construct, based on the similarity of motifs in the folkloric traditions, is aimed at measuring precisely the shared cultural history of group and country pairs. The evidence we provide in this paper is consistent with this hypothesis.

Although firmly grounded in the rationalist paradigm as a conceptual framework for the interpretation of our results, our approach remains open to complementary behavioral explanations[6]. Insights from behavioral models suggest that culturally shaped cognitive frames and biases could influence interstate interactions independently of strategic rationality. For

[6]See Jervis (1976) and Axelrod (2006), among others.

instance, cultural similarity might enhance empathic understanding, reduce cognitive biases in perception of threat, and foster trust-based behaviors, thereby reinforcing peaceful bargaining outcomes. This complementarity further strengthens the relevance of our cultural-beliefs interpretation, without detracting from its rational-choice foundations. More generally, our work is related to the emerging field of narrative economics (Akerlof and Snower 2016). According to Shiller (2017), narratives are "stories that offer interpretations of events", thereby making them crucial in bilateral negotiations, and folklore similarities measure the extent to which this narratives are similar between individuals[7].

On the empirical side, Henderson (1997) was among the first studies in the post-Cold War period to examine the relationships between cultural factors and the onset of interstate war in state dyads. Although religious distances rather than ethnic differences (based on the shares of ethnic groups in the dyads) appear to the empirically relevant in his regressions, he concludes that cultural variables should be recognized as important correlates of war. Bove and Gokmen (2017) show how cultural similarity, measured starting from language, from Hofstede's national cultural scales (power distance, uncertainty avoidance, masculinity/femininity, and individualism) and from World Values Survey answers, decreases the probability of interstate conflicts. They interpret the evidence in favor of "Clash of Civilizations" hypothesis (Huntington, 1993), that argues for fixed, deeply rooted, or primordial and fundamentally irreconcilable cultural identities as direct determinants of conflict. Spoloare and Wacziarg (2016), in open contrast with the deep traits hypothesis, find that a higher measure of ancestral relatedness between populations, genetic similarity, actually increases the probability of a conflict, because of both a higher probability of having disputes over similar, rival interests, goods and resources, and because of a lower cost of managing an integrated country in case of annexation (the so-called "winning the peace" argument).

Our contribution to this empirical literature is twofold. First, the use of a new measure of cultural similarity that evaluates the extent to which the oral tradition is shared in the history of the groups and countries. This has two main advantages: it is well-suited to capture information frictions in negotiations that might lead to failures and, hence, conflict; being determined in the past, it is exogenous to modern conflicts. Second, we provide several

[7]Following this interpretation, folklore similarities can be thought of as a metric to measure the similarity of what Durkheim called "Collective Representations", that help make sense of the World.

transmission mechanisms linking cultural similarities to conflict, including: a higher probability of forming military alliances, the participation to the same international organizations, similar voting patterns in the UN general assembly, and bigger bilateral trade flows, all of which highlight a higher propensity and/or capability to coordinate internationally in case of oral tradition similarity, thereby strengthening the interpretation of the results in terms of information frictions.

The main measure of oral tradition similarity adopted in our regressions has been developed by D'Amato and Russo (2025), who find that it is associated with income per capita differences as a consequence of shared cultural heritage on the diffusion of innovations. In the present context we argue that the measure of cultural similarity matters because it tracks the emergence of common systems of beliefs that reduce informational frictions and misperceptions in repeated interactions.

Our paper is also related to the abundant literature linking trade and, more generally, economic interdependence, to military conflict (Copeland 2015; Gartzke and Westerwinter 2016). Martin et al. (2008), in an influential paper, show that bilateral trade decreases the probability of conflict because of the expected loss of trade gains, while multilateral trade openness increases the probability of conflict because of the lower dependence on any single trade link. Morelli and Sonno (2017), along similar lines, show that bilateral trade dependence decreases the probability of conflict. All of these contributions candidate international trade as a leading mechanisms linking cultural similarities to international conflict. We contribute to this literature by showing how a larger index of folklore similarities is associated with more bilateral trade.

Since we study interstate conflict our paper is related but different from those works that study the effect of diversity on civil conflict, among populations living in the same countries, such as, among others, Fearon and Laitin (1993) or the recent contribution by Arbatli et al. (2020). Our approach is also different from Bakaki et al. (2015), who study the impact of cultural differences on mediation between parties at war, rather than focusing on the cultural differences between the parties at war themselves.

In general, the findings reported in this paper contribute to the ongoing debate on the cultural determinants of interstate conflicts, providing consistent and robust empirical evidence

for the role that similarity of cultural traits have in bilateral international relations between states. The very nature of our measure, based on cultural identities and distances molded by historical exchanges between populations, questions interpretations of the culture-conflict nexus based on primordial, immutable, traits.

3 Folklore-Based Cultural Similarity

We use one of the measures of oral tradition similarity developed in D'Amato and Russo (2025), which is a marker of the amount and quality of bilateral horizontal exchanges that took place in the history of two ethnic groups, then aggregated at the country level. In the following we report the main elements for the construction of such a measure, referring to D'Amato and Russo (2025) for further details.

Ethnic Group Level Measure. The starting point of the computations of the ethnic group level index of cultural similarity is the Berezkin folklore catalog (Berezkin 2015), which lists a total of about 2500 folklore *motifs* for 958 distinct ethnic groups. These motifs, or "replication units", are brief descriptions of the main structural elements of a traditional story that is part of the oral tradition of a group, and include images, narrative episodes, plots, or group of characters.

The benchmark measure evaluates the extent to which folklore motifs are similar, between two ethnic groups, leveraging on Latent Semantic Analysis[8] (LSA), a lexicostatistics algorithm designed to assess text similarity. In particular, the measure is computed as follows:

$$\chi_{kl} = \frac{1}{|M^k \cup M^l|} \sum_{i=1}^{N_k} \max_{m_j} [C(m_i, m_j)] \qquad |M^k| \leq |M^l| \tag{1}$$

where $M^k = \{m_1, m_2, \ldots, m_{N_k}\}$ is the set of motifs of the ethnic group k (shared with at least another ethnic group in the catalog), $M^l = \{m_1, m_2, \ldots, m_{N_l}\}$ is the set of motifs of the ethnic group l (shared with at least another ethnic group in the catalog), $|M|$ is the cardinality of the set M, that is the total number of motifs that are coded in the catalog, and the function C is the cosine similarity of the motifs computed with LSA, which is akin to a

[8]Deerwester et al. (1990) for details.

correlation coefficient for text strings.

The great advantage of this measure is that it is able to account for imperfect cultural transmission as recorded in the text (mutations), that is for the possibility of transmission of only some of the elementary components of the narratives across groups over the history of their cultural encounters. Therefore small variations of the stories are taken as evidence of a past cultural exchange or doorway, that is as a footprint of the travel of the narratives between ethnic groups and, therefore, of the mutual influences accumulated over time. In other words, the measure accounts for the fact that motifs can be transmitted across groups, and then vertically preserved, possibly in slightly different textual variants. Although we believe that this measure is better suited to evaluate folklore similarities, we will also explore the robustness of the results to potential alternatives (see section 5.4).

Crucially for our empirically strategy, as the folklore catalog does not provide information on the year since a motif appeared in a specific oral tradition, we assume, consistently with the ethnographic literature, that they originated in the distant past, meaning that they are exogenous to modern conflicts. In practice, we are excluding the possibility that the motifs were transmitted across groups as a result of conflicts that took place after 1816, an assumption satisfied by the criteria used by Berezkin to collect the popular oral tales described in his catalog. This is also the great advantage of our measures as compared, for instance, to the survey-based measures of contemporary cultural similarity used, for example, by Bove and Gokmen (2011), since very recent information on survey answers are likely to be endogenous to modern and contemporary conflicts. For instance, ethnic groups with a recent history of violent conflict ending in defeat, possibly resulting in several casualties, might be more prone to report values related to vengeance or punishment, or place more emphasis on religious matters.

Country-Level Aggregation. For the aim of our analysis, we need to compute folklore similarity at the country-pair level, so the ethnic-level index at eq. (1) must be aggregated at the level of country pairs, given their own ethnic composition. Following the economic literature on cultural diversity (Spolaore and Wacziarg 2009, among others) we compute weighted averages using population shares by ethnicity:

$$F_{ab} = \sum_{k \in K} \sum_{l \in L} \omega_k^a \, \omega_l^b \, \chi_{kl} \tag{2}$$

where ω_k^a is the population share of the ethnic group k in country a, ω_l^b is the population share of the ethnic group l in country b, and where K and L are, respectively, the sets of all ethnic groups in country a and b, with $\sum_{k \in K} \omega_k^a = 1 = \sum_{l \in L} \omega_l^b$. We use the population shares data in Michalopolous and Xue (2021), relative to the beginning of 2000's, although we find similar results when using the shares reported in Alesina et al. (2003), that refer to the 1990s.

There are two main issues associated with this aggregation strategy. The first is that, since there are no time-varying population shares available at the relevant time frequency, we cannot correctly measure the folklore similarity at the onset time of the conflict, perhaps with the only exception of a small time window around the population shares observation date. Second, currently observed population shares are potentially endogenous to conflict, for instance as a result of forced migration or simply as a consequence of the casualties, change of state borders after the war.

To overcome these issues, we consider an alternative aggregation procedure using ancestral, rather than actual, population shares. For each country we identify "native" ethnic groups among the ones that are currently living in the country, defined as those whose centroid, or ancestral location, lies within the current political boundaries of the countries. Starting from the population shares by ethnicity in each country, we drop the groups whose centroid is not within the current country borders, and, in the absence of reliable and comparable historical data on the size of these native groups, we assign equal population shares to all of them. The measure of ancestral folklore similarity at the country-pair level is then computed as follows:

$$\bar{F}_{ab} = \frac{1}{|K^a| \, |L^b|} \sum_{k \in K^a} \sum_{l \in K^b} \chi_{kl} \tag{3}$$

where $K^a \subseteq K$ and $L^b \subseteq L$ are the sets of all native ethnic groups to countries a and b. Given that the folklore motifs were determined in the past, this measure is therefore exogenous with respect both to modern conflicts and to their consequences, such as forced migrations and mass killings. To avoid overestimating the role of relatively small ethnic groups, we delete the

ones with actual population shares below 0.1%, ending up with a total of 593 ethnic groups that are part of the current population of some country and 452 native or ancestral groups[9].

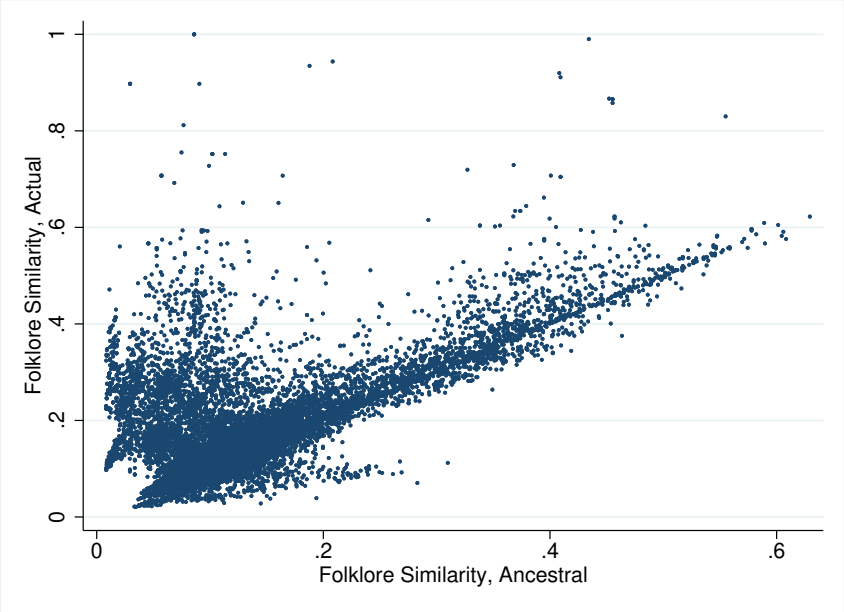
Results. The F measure of country level folklore similarity with actual population shares ranges from 0.020 (Latvia and Central African Republic), meaning almost completely different oral traditions, to 1, which happens in case of countries with the same population shares by ethnicity (Dominica and Grenada). Very high values are found, among others, for Australia and the UK (0.691), Austria and Switzerland (0.857), Bahrain and Qatar (0.704) and Russia and Kazakhstan (0.615). The sample median is 0.164, with mean 0.195 and standard deviation 0.121. This has been already discussed in D'Amato and Russo (2025), and we refer to that paper for further details and for a comment on the magnitudes. As for the $\bar{F}$ measure of ancestral folklore similarity, it ranges from 0.009, meaning almost completely different oral traditions for the native ethnic groups (Cuba and Bulgaria as an example) to 0.629, meaning very similar oral traditions for the native groups (Ukraine and Latvia as an example) with a median of 0.115, a mean of 0.145 and a standard deviation of 0.104. The high standard deviation, in turn, stems from the presence of few country pairs with very similar ancestral oral traditions (1% of the sample of country pairs with similarity above 0.5).

Figure 1 shows graphically the relationship between the two measures of folklore similarity computed between actual and ancestral ethnic groups. The correlation between the two is actually very high, 0.619, with significant differences observed for country pairs where at least one of the two countries experienced mass migrations that determined a significant population re-composition, typically with the former ancestral groups that slowly or abruptly became a minority. Examples include, among others, country pairs involving the US, Argentina, Brazil, Australia, New Zealand, and many more "New World" countries.

Notice that our construction of the ancestral measure of folklore similarity is that of a

[9]There is a total of 894 ethnic groups covered in the Berezkin folklore catalog that are currently part of the population of some country, meaning that there are 301 groups with very small population shares in all countries and 141 groups that are currently living in a territory different from their ancestral location. The number of ethnic groups per country ranges from 1 (Denmark and Burundi, among others) to 90 (Brazil and the United States), with a median of 5, a mean of 10.6 and a quite high standard deviation of 16.5 determined mostly by the presence of immigration countries such as the US and Canada, among others. The number of ancestral groups per country ranges instead from 1 (Netherlands, Portugal, South Korea and Madagascar, among others) to 37 (Mexico), with a median of 2, a mean of 3.78 and a high standard deviation of 4.72 as a result of the presence of few countries with a lot of native groups (beside Mexico, we have Colombia, Peru, US, Tanzania and India above 12 native ethnic groups).

Figure 1: **Folklore Similarity, Actual and Ancestral**



Notes: Folklore similarity in country pairs computed with actual (Y axis) or ancestral (X axis) population shares. 12797 observations.

simple counterfactual for what folklore similarity in country pairs would have looked like in the absence of modern migrations. Therefore, from figure 1 it is also possible to assess whether migrations increased or decreased the similarity between countries. By inspection, many country pairs are clustered around the 45 degrees line, meaning that modern migrations did not affect their oral tradition similarity much. A substantial amount of country pairs are instead above the 45 degrees line, meaning that their post-migrations folklore similarity is actually larger than the ancestral, whereas very few pairs have seen their similarity reduced[10].

4 The Empirical Strategy

In this section we discuss our empirical strategy. Starting from a description of the estimated model (section 4.1), we propose a thorough discussion of the control variables used (section 4.2) in the regression, and we end with a discussion of the identification assumptions (section

[10]This evidence, combined with the results that we are going to discuss later, points towards an effect of modern migrations on conflict, which however we are not going to analyze here and it will be the focus of future work. On the importance of modern migrations for historical comparative development, see Putterman and Weil (2010).

4.3).

4.1 Model

Our baseline empirical specification entails regressing a dummy in case of a conflict between two countries in a given year, on folklore similarity and controls, including two-way fixed country effects and year effects:

$$W_{ijt} = \beta_0 + \beta_1 F_{ij} + \Gamma X_{ijt} + \theta_i + \theta_j + \tau_t + \varepsilon_{ijt} \tag{4}$$

where W_{ijt} is an indicator variable for a conflict between country j and country i in year t, F_{ij} is the folklore-based cultural similarity between country i and country j, X_{ijt} are control variables (discussed in great detail in section 4.2), θ_i and θ_j are country fixed effects for the two sides of the conflict, τ_t are year effects, and ε_{ijt} the error term. The coefficient of interest is β_1, which captures the effect of folklore similarity on conflict. We estimate[11] the model using a probit estimator with clustered standard errors at the level of the country-pair i, j (the standard errors are numerically very similar in case of two-way clustering at the level of the two countries in the pair).

Following the literature, we construct the interstate conflict dummy W starting from the Correlates of War (COW) database, which lists all Militarized Interstate Disputes (MID) between 164 countries from 1816 to 2010, assigning to them a score on the basis of the maximum hostility level reached in a given year. We code a conflict a dummy equal to one in case of hostility of level 3 or higher, corresponding to: display use of force, use of force, and war, thereby excluding the cases of dispute without military actions and with the threat of using force. Since this dataset is widely used in the literature (see Spolaore and Wacziarg 2016, among others), we do not describe it in detail here.

We will report results for the two aggregated measures of folklore similarity between countries described in section 3, that is computed using actual or ancestral population shares. An alternative strategy would be to use the folklore similarity measure computed with ancestral

[11] Major differences with respect to the empirical model by Spolaore and Wacziarg (2016) are: the use of the full panel of MID data, the inclusion of country fixed effects for both sides of the conflict and of time effects, mostly to account for the effects of the World Wars.

shares as an instrument for folklore similarity computed using contemporary shares, within an instrumental variables probit regression framework. When implementing this strategy, we obtained the same empirical results (details available upon request) although, due to the size of the dataset, to the high number of fixed effects, and to the standard error clustering, the computational time to obtain marginal effects on a medium-scale PC is extremely long.

4.2 Control Variables

In all regression specifications, we control for genetic (Pemberton et al. 2013), linguistic and religious distance, that is, for other available measures of relatedness between populations and countries. These additional similarity measures are determined by correlated, yet different, historical cultural evolution mechanisms, so they indeed contain different information (Cavalli Sforza 2001; Greenhill 2021; D'Amato and Russo 2025), which we exploit in order to isolate the effect of cultural transmission captured by folkloric similarity (Bortolini et al. 2017). Moreover, the inclusion of these variables allows us to compare our results to the literature on culture and conflict (Spolaore and Wacziarg 2016, among others), and also to isolate a new mechanism not highlighted so far. We use country-level weighted averages of the ethnic-level measures computed using actual population shares, to better compare our results to the literature, although we obtained similar results when using distances between ancestral groups.

We include geographic distance between the capitals of the countries because conflicts are much more frequent between close countries (Bremer 1992). To give a sense, the median distance between countries in the sample is equal to about 6.98 thousand kilometers, whereas the median distance between the capitals of the country pair that ever experienced a conflict is 3.23 thousand kilometers and, excluding WWI and WWII, the median is actually 1.86. As an additional geographic control we use a dummy equal to one in case of contiguous countries, since we observe many conflicts occurring among countries that share borders, for instance because of competition about natural resources and territorial claims.

We include the (time-varying) Polity2 scores for both countries in the pair, which are aggregate measures of the degree of democracy in a country (on a scale between 0 and 10). The reason is that democracies, are less likely to engage in conflicts - in particular against

other democratic regimes, the so-called democratic peace - as shown in Bremer (1992, 2000), Henderson (1998), Gartzke (2007) and Jackson and Morelli (2007), among others. We also include the historical GDP at PPP from Maddison and the population size for both countries, as a generic control for country power.

We also control for the colonial and political history of a country pair. To this aim we code a dummy equal to one in case the two countries were part, in the past, of the same political or administrative area, and a dummy equal to one in case of a colonial relationship.

All regressions feature country fixed effects, which allows to control for economic, political and geographical characteristics that can be associated with bargaining failures and, therefore, conflicts, such as: the availability of indivisible natural resources (oil, fertile land, etc.), the nature of the political system, which might determine agency problems between the population and the ruling class; ethnic fractionalization, as multilateral bargaining between several ethnic groups on both sides is intrinsically more complex; the size of the country, because countries with growing populations might fight to acquire new resources; individual preferences over policies; national capabilities in general, in the sense of Singer (1988). Moreover, all regression feature time effects to control for the clustering of conflict during the World Wars and during other episodes of general World turmoil (Cold War, Yom Kippur etc.), and for the effect of increasing market globalization that might affect the incentives to fight (Gartzke 2007; Martin et al. 2008).

In addition to controlling for spatial proximity, following Spolaore and Wacziarg (2016), we also consider several additional control variables to account for the effect of similar geo-climatic conditions[12]: the difference of the longitude and the difference of the latitudes of the capital cities; a climatic difference measure computed as the average difference in the percentage of the country surface in the 12 Koppen-Geiger climate zones, to control for climatic barriers; the average difference in elevation and the average difference in terrain ruggedness, to control for topographical barriers; the average difference in the percentage of tropical area, the average difference in the percentage of desert area, and the average difference in the percentage of fertile soil, to control for the presence of indivisible resources to fight over; the difference in the average distance to coast, and the difference in the percentage of ice-free coast, as a further

[12]See Nunn and Puga (2012) for details on the construction of these measures.

control for the access to transportation by water; the human mobility index by Ozak (2013), that, measuring travel time between two points on the globe, is a good summary measure of a geographical barriers. All of these controls we deem important because cultural similarity is, in principle, heavily influenced by geographic proximity.

4.3 Identification

The identification of our empirical model rests on the exogeneity of folklore similarity to interstate conflicts that took place in the last two centuries. This assumption is well grounded in the procedure used by Berezkin to build the folklore catalog that, instead of listing traditional stories, collects motifs, which are the structural elements, or replication units, of the stories, that originated in the distant past and that persist over time (Thomson 1946), so that they are unlikely to be influenced by current events such as modern-day conflicts and/or socioeconomic conditions[13]. It is important to notice that folklore similarities between ethnic groups and countries, are not strictly exogenous to ancient conflicts, being determined by a myriad of ancient shocks such as trading relationships, invasions, demographic admixture etc., which determined cultural exchanges recorded in the oral tradition. But whatever is the nature of those ancient shocks that have composed the oral tradition of a group pair, we assume that their influence on the probability of conflict in modern times, conditional on observables, is only through the oral tradition similarity, and the consequent reduction of information frictions, that they determined.

For instance, folklore motifs could have been exchanged between ethnic groups for reasons related to ancient trade relationships, but the reasons why two populations traded hundreds or thousands of years ago, conditional on geographic distance and other observable measures of relatedness, are not the basis on which trade is sustained today, because, among others, of sharply different preferences, production and communication technologies, societal organizations, political boundaries, relevance of resource location, etc. Therefore, the effect of ancient trade relationships on conflict is only through the cultural similarity, recorded in folklore, that those ancient exchanges determined, a similarity that reduces frictions in bilateral

[13]See also Toelken (1996) on the persistence of the tradition and Michalopolous and Xue (2021) and D'Amato and Russo (2025) for a discussion on the nature of folklore motifs.

negotiations and not because of an effect of their exogenous determinants on modern trade. Thus, even if ancient trade relationship influenced modern trade, the regression coefficient on folklore similarity would not pick up the effect of modern trade on the probability of conflicts today because of the historical nature of the motifs in the catalog.

As a further example, consider populations with a history of domination and exploitation as a result of ancient invasions, colonizations and conflicts. They might have developed a fundamental luck of bilateral trust, reflected in their oral tradition, that impedes credible commitments and the enforcement of agreements, so to increase the probability of conflict today. The plausible assumption we make is that the effect of those ancient conflicts on current conflict operates through the information frictions determined by past interactions, since the motivations at the basis of ancient conflicts are inherently different as a consequence of different political and societal structures (modern states between which we observe conflict today did not exist in the distant past), different ancient technologies that made different natural resources strategic etc. Moreover, as already stated, the ancient nature of the folklore motifs ensures that modern conflicts do not affect folklore similarity today.

Therefore, following the ethnographic literature, we assume that folklore similarities are the result of several ancient shocks in the historical process of cultural diffusion, but our goal is not to estimate the effect of these deeper determinants of folklore similarity on conflict. Rather, our aim is to isolate, empirically, the impact of the popular cultural similarity, resulting from the myriad of such ancient shocks, on current interstate conflicts. In this respect, the logic of our empirical exercise is similar to a number of recent contributions in the economic literature. Examples include D'Amato and Russo (2025), who study the effect of folklore similarity on long-run growth, under the assumption that the shocks that determined folklore similarity in the are exogenous to current income differentials; Spolaore and Wacziarg (2016) who propose an analysis of the effect of genetic distance on modern-day conflicts based on the exogeneity of those ancient separations; Spolaore and Wacziarg (2009), who study the effect of genetic distance on income differences, building again on the idea that genetic separations are exogenous; Ashraf and Galor (2013), who analyze the impact of expected heterozigosity, as determined by out-of-Africa migrations, on comparative development, based on the idea that - whatever the role of the myriad of geo-climatic and demographic shocks that triggered

and sustained the migration path of different groups out of Africa, and the ensuing within-group and between-groups genetic diversity, these are exogenous to current levels of income per capita.

As already noted, the only remaining potential concern of endogeneity is related to the population shares used to compute country-pair weighted average measures, since they might be the result of, among others, selective migrations determined by past and current conflicts. For this reason and to show robustness, we will report our results separately for a measure of folklore similarity computed using ancestral population shares, unaffected by modern and contemporary migrations. To further validate our results, we will also provide empirical evidence for the relationship between folklore similarity and conflict at the ethnic group level.

5 Folklore and Conflict: Results

In this section we summarize the main empirical results of the paper, organized along the following lines: subsection 5.1, summarize the baseline results; subsection 5.2 proposes a discussion of the relationship between genetic and folkloric similarity and about their different impact on the probability of conflict; subsection 5.3 illustrates the results obtained using alternative, disaggregated, measures of folklore similarity; subsection 5.4 extensively discusses the robustness of these results, to the inclusion of additional controls, to the use of further alternative indicators of folklore similarity, and to various sample restrictions; subsection 5.5 discusses the empirical results obtained on a cross-section of ethnic groups; All in all, the evidence suggests that folklore similarities, through their effect on information frictions, significantly decrease the probability of an interstate conflict.

5.1 Main result

The main empirical results are summarized in table 2. We found that country pairs with more oral tradition overlap have a lower probability to engage in a conflict. To get a sense of these magnitudes, moving from the first quartile of folklore cosine similarity measured with actual population shares to the third (or slightly more than a 1 std deviation change), which is equivalent to a movement from the rather small similarity between and Angola and Australia

or between Indonesia and Belize to the rather high one between Mexico and Spain or between Oman and Lebanon, decreases the probability of conflict[14] by about 34%, which is quite sizable. Similarly, moving from the first to the third quartile of ancestral folklore similarity implies a 23.7% reduction of the conflict probability. As shown in table 2, the results are robust when including the additional geographic controls (see section 4.2 for the list).

Consistently with Spolaore and Wacziarg (2016), we find a negative coefficient on genetic distance, that is country pairs composed of genetically more related populations are more likely to fight with each other. This is rather important and we will comment the result in more details in the next section. We also find that conflicts are less likely between more distant countries, although the coefficient on geodesic distance becomes not significant after adding the full set of additional geographic controls. As for linguistic and religious distance, we do not find similarly robust empirical results. Consistently with the political science literature (Gartzke 2007; Jackson and Morelli 2007), we also find negative and significant coefficients on the time-varying polity2 scores (details available upon request), meaning that more democratic countries are, in general, less prone to conflicts.

5.2 Genes, Folklore, and Conflict

As noticed in the previous section, the coefficient on genetic distance turns out to be negative and statistically significant in all our regressions[15], consistently with the results by Spolaore and Wacziarg (2016). This is important because it adds further credibility to our interpretation of the results: genetic distance, which is a marker of the ancestral processes of genetic differentiation, enters the regressions with a sign that is not consistent with any primitivistic interpretation of the culture-war nexus.

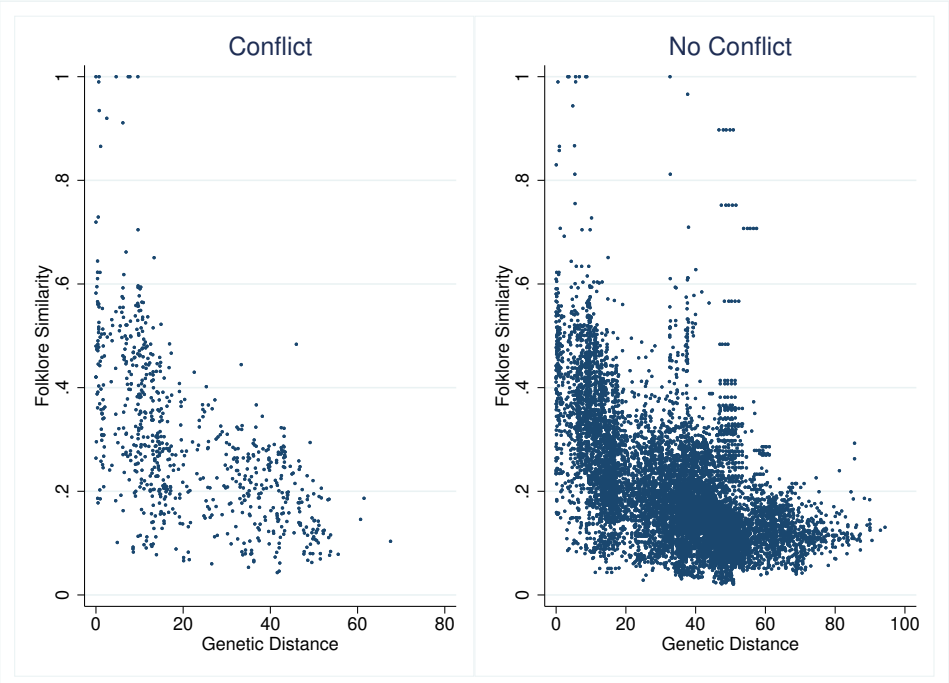
A deeper question is why do we find different effects on conflict of these two different measures of cultural relatedness? The reason is that genetic and folklore based measure of cultural distance, although related, are the results of different evolution processes. Genetic

[14]Note that conflicts are, fortunately, rather rare event: we observe conflicts in only 0.78% of the sample observations.

[15]We find similar results when using the genetic distance computed with dominant ancestral ethnic groups only, that is the variable that Spolaore and Wacziarg (2016) use as exogenous instrument to account for the potential endogeneity of current population shares. This evidence suggests that the results are not driven by the population shares used to compute country-level genetic distances.

distance measures genetic admixture, likely stemming from demic processes (Cavalli-Sforza 2001) whereas folkloric distance is a marker for the occurrence of historical events of cultural exchange across groups (cultural doorways) or, in other words, a measure of cultural transmission and diffusion (Bortolini et al. 2017). This cultural transmission fosters the accumulation of common traits in the cultural heritage of ethnic groups in a pair. Similar systems of beliefs, in turn, favors the emergence of cooperation in repeated interactions between the two groups, with a statistically detectable effect on the probability of conflict. To highlight the fact that genetic and folkloric distances contain different information as a result of different historical evolution processes, D'Amato and Russo (2025) report evidence of remarkable heterogeneity in folklore similarity among ethnic groups with similar genetic distance, especially for those who separated more recently.

Figure 2: **Folklore Similarity and Genetic Distance, with or without Conflicts**



Notes: Observations are for country pairs that experienced a conflict (left panel, 755 pairs) or not (right panel, 11889 observations).

Further information about this point is reported in figure 2, where we do observe many genetically close countries with very different oral traditions[16] (lower left corners of both

[16]This heterogeneity mirrors another heterogeneity that we observe in the relationship between genetic

panels), and some of these country pairs experienced one or several MID over the sample period. Examples include: Russian Federation and Afghanistan between 1982 and 1985; India and Pakistan in various years between 1951 and 1986; China and the Philippines between 1950 and 1953 and between 1967 and 1970; Iran and Russia between 1986 and 1987; Syria and Israel in various years between 1950 and 1981; Ethiopia and Somalia in various years between 1963 and 1985; Iran and Saudi Arabia between 1986 and 1988; Laos and Thailand in various years between 1960 and 1983; Kenya and Uganda between 1973 and 1975; Israel and Jordan between 1960 and 1968; The Republic of Congo and the Democratic Republic of the Congo between 1969 and 1972; Mali and Burkina Faso between 1974 and 1975 and between 1985 and 1986. These are all examples of population pairs that feature a recent genetic separation but who did not exchange much cultural traits in their oral traditions, after separation. Intuitively, results in our regressions emerges from these groups: conditional on recent genetic separation conflict is less or more likely depending on the history of the cultural exchanges they experienced according to our folklore based marker of similarity. Conversely, it is less likely to observe conflicts between genetically close countries with significant oral tradition overlap (upper left corners of both panels).

In conclusion, genetic relatedness, that according to the interpretation in Spolaore and Wacziarg (2016) captures similar preferences for rival goods and resources, and folklore differences, capturing similarity in systems of cultural beliefs, appear to be complementary explanations for the onset of interstate conflicts and, more generally, for the understanding of the culture-conflict nexus.

5.3 Disaggregated Measures of Folklore Similarity: Sustaining Cooperation.

Our interpretation of the empirical relationship between folklore similarity ad conflicts rests on the idea that similar cultural beliefs reflect more frequent information exchanges between populations over their histories, which reduce information frictions, misperceptions and uncertainty about the intentions or resources of the negotiation counterparts today. In addition,

distance and cultural distance measured, as in Desmet et al. (2011), by the difference in the responses to World Values Survey questions. In particular, there are genetically close ethnic groups that provide very different answers, exhibiting different sets of current values and beliefs.

similarity in cultural beliefs favors the emergence of cooperation in repeated interactions sustained by trigger strategies or tit for tat. Berezkin's folklore catalog can be usefully employed to test for this hypothesis: if a larger degree of similarity reduces the emergence of conflict in repeated interactions as a consequence of cooperative behavior emerging from trigger strategies or tit-for-tat, we should observe a relationship between conflicts and the amount of bilateral similarity in folklore motifs related to "punishment" "defectors", "reputation", "patience", and other similar concepts that support mutually efficient outcomes in these settings. This insight will be explored in the following.

To this aim we computed an alternative, concept-specific cosine measure of folklore similarity, implementing the computations summarized in equation 1 but, instead of using the full set of motifs in the Berezkin catalog, we used only those related to the support of cooperative strategies in repeated interactions, according to prevailing theories in this class of games. To isolate such motifs, we followed the procedure by Michalopulous and Xue (2021), that associate each motifs descriptions to several keywords and tags using text analysis techniques, selected only the motifs associated with the following tags: punish, retribution, cooperate, revenge, retaliation, patience, promise, threat, defect, forgive, memory, misunderstand, reputation, commitment, betray. Consistently with the procedure adopted throughout the paper, the measure produced at ethnic-group pair level is aggregated at the country-pair level using either actual or ancestral population shares.

The resulting cooperation-specific cosine measure of folklore similarity turns out to be very highly correlated with the cosine measure computed using the full extent of the folklore catalog. In particular, the correlation is equal to 0.856 in case of aggregation with actual population shares, and 0.943 in case of ancestral shares. The mean cooperation specific similarity in the country-pair cross section is equal to 0.158 with median 0.113 and a quite high standard deviation 0.134. Examples of country pairs with more similar oral traditions relative to this cooperation motifs include: Belarus and Russia; Switzerland and Finland; France and Germany. Conversely, country pairs with very little similarities include, among others: France and Brunei; Afghanistan and the Philippines; Colombia and Sierra Leone. Regressing the conflict dummy on this disaggregated folklore similarity measures and controls, using the same benchmark empirical model described in section 4.1, we found negative and

strongly statistically significant coefficients, as reported in columns 1 and 3 of table 3. This is evidence that similarity in the system of beliefs underlying concepts related to the emergence of cooperation (trigger strategies or tit for tat), are relevant, further validating our interpretation of the empirical results: folklore similarities matter as marker of similarity in motifs related to system of beliefs which theory predicts to be relevant in repeated bilateral interactions.

As a further check, we computed another concept-specific cosine measure of folklore similarity using keywords associated to conflict only. Specifically these keywords are: dispute, expand, hegemony, invade, intrude, attack, war, fight, combat, contend, military, army, troop, soldier, peace, sovereignty. The resulting variable measures the extent to which the oral traditions related to conflict are similar. The idea is to check whether there is something specific to the way conflict are described in the oral tradition that drive countries more often to fight each other. The resulting war specific cosine measure of folklore similarity is also correlated with the overall cosine measure, although less so than the cooperation specific measure described above (0.517 in case of aggregation with actual population shares, and 0.662 in case of aggregation with ancestral shares). Moreover, the correlation between this measure and the cooperation specific is equal to 0.585 in case of actual shares, and 0.676 in case of ancestral, so it seems that they are picking up similar information. However, when using this alternative folklore similarity measure in the benchmark regression, we find non-statistically significant coefficient, both in case of aggregation with current and of ancestral population shares. The results are in columns 2 and 4 of table 3.

Although this empirical exercise is an important placebo test for potentially relevant motifs in pairs of traditions, it cannot be considered as an attempt to precisely identify the specific historical causes of modern interstate conflicts. We take the results as evidence that the shared traits in the oral traditions that matter empirically are those related to themes and concepts that theory suggests to be relevant for the emergence of cooperation in repeated interactions. This evidence lends support to the interpretation of the empirical results described in the previous section.

5.4 Folklore and Conflict: Robustness

We extensively explored the robustness of our empirical results to several possible alternative measures of folklore similarity, to the inclusion of additional controls for potential confounders, and to several sample restrictions. A detailed description of the results follows. The uninterested reader can skip directly to section 6, where we highlight the mechanisms linking folklore based measures of cultural similarity to interstate conflict.

Alternative measures of folklore similarity. Our first robustness check entails the use of an alternative measure of folklore similarity based on content rather than text, also taken from D'Amato and Russo (2025). In particular, this measure evaluates differences in the prominence of several themes/concepts/images/topics in the oral tradition, and it is based on a content extraction. This measure is important because it accounts for the possibility that textually different stories carry similar messages. Building on Michalopolous and Xue (2021), the first step of the computation is the association of each folklore motif to several themes, based on the presence of keywords or tags, to then compute concept intensity as the ratio of the number of motifs associated with the theme over the total number of motifs. Folklore content similarity, for each ethnic group pair, is then measured as 1 minus the average of the absolute value of the differences in concept intensity. In short:

$$\Theta_{kl} = 1 - \frac{1}{S}\sum_{s=1}^{S}\left|\frac{|N_s^k|}{|M^k|} - \frac{|N_s^l|}{|M^l|}\right| \tag{5}$$

where $|N_s^k|$ is the total number of motifs in the folklore catalog of the ethnic group k containing theme s (i.e. that contain a specific keyword or a tag) and S is the total number of selected themes. The index ranges between 0 and 1, with 1 meaning that the two oral traditions feature essentially the same themes. Note that this is not a measure of the absolute importance of a theme in the oral tradition of group: there can be small differences both for very important themes (high absolute importance) and for marginal themes (low absolute importance). As for the case of the text measure of folklore similarity, we aggregated at the country-pair level using both actual and ancestral population shares. When using this alternative measure of folklore similarity in the main regression, we find the same empirical results as in the benchmark. The regression results are summarized in table 4. Note that the

text and content measures of folklore similarity are positively correlated, so having similar motifs also entails sharing a similar content of the oral traditions, although the correlation coefficient is not very high (0.265 for the measure computed with actual population shares and 0.171 for the measures computed with ancestral population shares) stressing that the two measures contain different information[17].

Following the literature (Bortolini et al. 2017), we also tried using an alternative measure of folklore similarity equal to the Jaccard coefficient of the folklore catalog entries, that is the number of common motifs divided by the number of motifs in either one of the traditions, again aggregated at the country-pair level using either actual or ancestral population shares. This measure does not account for the possibility of horizontal transmission of single traits or motifs in the narratives (see D'Amato and Russo 2025 for a more detailed discussion) but it presumes that the entire tale is shared or not in a pair of groups (as per the construction by the catalog by Berezkin). It turns out that this measure is correlated with our benchmark one (at the country level, we get 0.97 correlation if actual population shares are used in the aggregation, and 0.96 in case of ancestral shares). When used in the main regression the same empirical results obtain (details available upon request).

When computing a weighted average of folklore similarities in country pairs, we have- so far- considered either all the ethnic groups currently present in the country, or those with centroids within the current country borders. However, some of those groups are actually small, and one might question whether it is meaningful to consider them in order to explain bilateral conflict, given that they are unlikely to have significant political power, especially in democracy, to affect government decisions. To address this issue, we considered an alternative weighted average, country-level, measure of folklore similarity that excludes minorities. We arbitrarily consider minorities, in a country, the ethnic groups whose population share is below a 10% threshold. However, to avoid excluding small ethnic groups with significant political power, we leverage information from the Ethnic Power Relations (EPR) database (Vogt et al. 2015), which provides a classification and scoring of ethnic groups, by country, according

[17]In particular, we do see counties with a small cosine similarity measure, that is with few similar textual elements, but with narratives that speak about similar themes, that is with a small content differences. Examples are: Australia and Canada (0.092 ancestral cosine similarity and 0.964 ancestral content similarity), Cameroon and Jamaica (0.069 cosine similarity and 0.976 content similarity) and Cuba and Argentina (0.094 cosine similarity and 0.979 content similarity).

to their political relevance, according to the following scale: Monopoly, Dominant, Senior Partner, Junior Partner, Discriminated, Self-Exclusion, Irrelevant, Powerless. We kept small ethnic groups, below the threshold, in case in at least in one year since WWII (the starting point of the the EPR sample), they were either dominant, monopolists, senior partners or junior partners. We then computed a weighted average measure of folklore similarity using only the big, or politically relevant ethnic groups, rescaling the ethnic group shares. When running the baseline regression on this new index, we find very similar empirical results (details available upon request). Note, however, that this empirical exercise involves a restriction over two variables, ethnic group size and political relevance, that may both endogenous to conflict, and the EPR data are available only for the last part of the sample, so the empirical results must be interpreted with caution.

Additional controls. The results are robust when controlling for the common migratory history of the ethnic groups. In particular, we use the area of the migratory triangle defined, for each ethnic group pair, by their centroids or ancestral locations and by the closest out-of-Africa migratory midpoint(see D'Amato and Russo 2025 for details), then we aggregate the information at the country-pair level. The bigger the area, the more ancient the group separation, meaning a shorter migratory history in common[18]. When including this additional variable in the regression, we find robust empirical results. This we interpret as evidence that vertical population separation is unlikely to be behind our empirical results, which are driven by information frictions.

As a further control for modern migrations, that could have influenced bilateral relationship and population compositions (see section 3 for a discussion on the impact of modern migrations on folklore similarity), we used information from the Putterman and Weil (2010) migration matrix. This matrix lists, for any couple of countries, the percentage of the population of one that originates from the other, as of 2000 AD. [19] When including this additional variable, we find robust results, and the coefficient on this control turns out to be not statistically significant.

[18]Notice that genetic distance tends to be higher between ethnic groups with higher migratory distance from East Africa. Those population are also the ones with less genetic variability (serial founder hypothesis, se Ashraf and Galor 2013) and also less cultural variability (see Galor et al. 2023).

[19]The reason why we do not have this additional variable in the baseline regression specification is because it is a "bad" control, as a result of selective migration due to cultural similarity.

We then tried restricting attention to the country pairs where most of the population is indigenous, as a further control for the potential endogeneity of folklore similarity coming from population shares. In particular, starting again from the Putterman-Weil matrix, we selected only the countries with less than 25% of the actual population coming from another country. The empirical results turned out to be robust. This check can also be interpreted as a way to address the concern that the results might be driven western off-shoots.

The results are also robust when controlling for the composite index of national capabilities by Singer (version 6, 1998) for both countries. This index measures, broadly speaking, national power, stemming from six dimensions: population, urban population, iron and steel production, energy consumption, military expenditure and military personnel. Thus, it is a measure of the "balance of power". We also tried including, in the regression, a dummy equal to one in case either one of the countries is an oil producer/exporter, to control for a prominent rival resource that might trigger conflict (See Caselli et al. 2015, among others). We found robust results and a non-significant coefficient on the dummy, evidence for the fact that oil availability is not driving or confounding our empirical results.

Cao et al. (2024) show that traditional herding practices lead to the formation of a "Culture of honor" that emphasizes violence and revenge, increasing the probability and severity of conflict. They also show that ethnic groups that traditionally rely on herding more frequently feature, in their oral traditions, motifs related to revenge and violence. Therefore part of the reason why folklore similarity impact conflict could be a similarity in the subset of motifs related to the honor culture. To control for this source of folklore similarity, and for its direct effect on conflict, we defined a dummy variable equal to one, for an ethnic group pair, in case of a high reliance on herding, and then computing weighted averages.

Following Becker (2025), the reliance on herding is defined as the product between the dependence on animal husbandry from the Ethnographic Atlas and an indicator equal to one in case the main animal in a society is a herding animal (sheep, cattle, horses, reindeer, alpacas, or camels). High reliance, in turn, is defined using various threshold values around the sample median reliance (across all ethnic groups in the sample). We find robust results when including this additional control in the regression, and a control itself that is not statistically significant. This last result, in particular, highlights that similarities in an honor culture, or

the prevalence/acceptance of violence, are not the main drivers of the empirical results in the baseline regression, which depend instead on cultural similarity in general because of its effect on information frictions, and not on specific cultural traits related to violence. This aligns with the previous findings that sharing cultural beliefs related to the emergence of cooperation is associated with a lower probability of conflicts.

Spolaore and Wacziarg (2016) argue that conflict is more frequent between similar population because of the similarity in political preferences, that determines an easier political management of the "peace" after the conflict. To control for this effect, we included an additional variable equal, for a country pair, to the ethnic fractionalization of the virtual country obtained joining the populations of the two countries in each pair. We found robust empirical results, both for folklore similarity and for genetic distance. Moreover, the coefficient on joint folklore fractionalization is not significant, perhaps because genetic distance, that is also a proxy for the same factors, is already included in the regression. These results stress, once again, that our explanation of the relationship between folklore similarity and conflicts is complementary to those provided in Spolaore and Wacziarg (2016) (see also section 5.2).

We also computed, for all country pairs, the share of the joint population made up by ethnic groups that are present in both countries, to control for the possibility of so-called "liberation" wars, whose objective is to re-unite the population, but also to control for the presence of artificial states in the sense of Alesina et al. (2011). When including this additional regressor, we found robust empirical results, and a coefficient on this variable that is positive and statistically significant, as expected. The results turned out to be robust also when including an interaction between this common population share variable and the dummy for contiguous countries, which therefore measures ethnic group commonality only for bordering countries.

As argued in D'Amato and Russo (2025), a shared oral tradition between ethnic groups can also be the results of common, ancient, exogenous shocks, rather than the result of a process of cultural diffusion. For instance, the prominence of Sun and Moon in the traditional story could have been triggered by eclipses observed in the past, or earthquake related stories might be prominent among groups that lived in seismic zones. To check if we are truly estimating the effect of oral traditions similarities on conflict, rather than the effect of common shocks

to the geo-climatic environment of the pair, we included, in the regression, an additional set of controls. In all cases we started from variables defined at the ethnic group pair level, and aggregated at the country pair level using either actual or ancestral population shares. The featured controls are: a dummy equal to one in case of both ethnic groups that experienced at least one solar or lunar eclipse before O CE (ethnic group centroid within a 100 km range from the point of maximum eclipse visibility); a dummy equal to one in case of both ethnic groups with centroids within 100 km of meteorite landing or finding site; a dummy equal to one in case of ethnic groups close to volcanoes (centroids within 100 km); a dummy equal to one in case of ethnic groups close to high-intensity seismic areas (centroids within 100 km); a dummy equal to one in case of ethnic groups with centroid in the same Koppen-Geiger climate zone; the difference in the average temperature in the area around the ethnic group centroid; the difference in the caloric potential from agriculture from Galor and Ozak (2015); a dummy equal to one in case of both groups homeland is close to rivers. In addition, we also included the weighted average geodesic distance between the ancestral ethnic groups centroids, to further account for unobserved geo-climatic shocks. In all cases the empirical results of the baseline regression turned out to be robust.

Additional Robustness Checks. The results are robust also when excluding World War I and II from the sample (both of them or one at a time). When restricting the sample to the post Cold War period, that is after 1990, we do not find any significant effect of folklore similarity on conflict, meaning that there is no evidence in favor of the specific version of the "Clash of Civilizations" hypothesis by Huntington. We also find no empirical results when restricting to the period before WWI, although we have a small number of observations in this case. When focusing on the post-WWII period, we find instead robust result to the benchmark. The results are also robust when focusing on the XIX century. The results are robust when consolidating the COW database into decades, that is coding a dummy variable equal to 1, for a country-pair, in case of conflict in the decade, rather than in a year. We enumerated decades starting from 1 for years between 1816 and 1820 and proceeding onward.

We computed, for all country pairs in the sample, the number of years with conflicts[20], and

[20]It is important to use the number of conflicts as dependent variable because Acemoglu and Wolitzky (2024) show that information frictions in the form of misunderstandings (imperfect monitoring) my determine rich dynamics in the time pattern of conflicts, such as spirals, traps and cycles.

regressed it on folklore similarity, controls and with country fixed effects. Since the dependent variable is a count in this regression, we used a Pseudo-Poisson maximum likelihood estimator, with standard errors clustered at the level of the two countries in the pair. We found negative and significant coefficients on folklore similarity, as shown in table 5, and the results are robust in case we compute the number of years with conflicts rather than the total number of conflicts (details available upon request).

The COW database allows the distinction between territorial and non-territorial wars, by a coding what they call the "Revisionist" status of the countries in conflict, that is whether the war has been associated to the willingness to change the status quo, distinguishing between changes regarding the country borders and changes regarding the political regime. This distinction is important because Spolaore and Wacziarg (2016) argue that cultural similarity should actually increase the probability of conflict over rival goods, and territorial wars fall in this category. We coded a dummy for territorial wars as equal to one in case of a conflict with at least one revisionist state about the borders/territory, and regressed it on folklore similarity and controls with the same benchmark empirical specification. We find the same empirical results as in the baseline model specification, that is folklore similarities decrease the probability of bilateral conflict. So the main effects that we highlighted work also in case of conflicts over rival goods. Conversely, when focusing on non-territorial wars only, we do not find any empirical result.

We also tried focusing on war of the highest intensity with COW hostility level equal to 5, excluding the display and use of force between two countries, which account for 26% of the episodes classified as conflicts in the benchmark empirical model. We found robust empirical results.

We tried running the main regression separately for geographically close and distant countries, using the median distance between capitals among all country pairs as a threshold. We find that folklore similarity only explains conflicts between geographically close countries. However, notice that only about 1/3 of the conflicts are between country pairs with distance above the median: conflicts are, in general, more likely among closer countries, which is especially true if we exclude the two World Wars from the sample. When correctly splitting the sample at the median distance among all country pairs that experienced a conflict, we find a

negative and significant coefficient on both subsamples of high and low distance countries.

We also tried using two alternative outcome variables from the COW database, the raw indicator of the overall hostility level and the number of fatalities caused by the MID (from 0 for no deaths, covering 99% of the sample, to 6 for more than 1000 deaths, covering 0.2% of the sample but 31% of all MID disputes with intensity equal or bigger than 3). When running pseudo-Poisson maximum likelihood regressions, we still find negative and significant coefficient on the folklore similarity measures. Thus cultural similarities in oral traditions turn out to affect also less severe disputes, those with a lower number of fatalities.

5.5 Ethnic Group Level Evidence

Interstate conflicts, to the extent that nation states reflect common ethnic ties of their citizens, can also be thought as conflicts between the ethnic groups that live on the opposite side of borders in the pairs of fighting countries. This logic is indeed behind the computation of weighted average folklore similarity measures at the country level starting from ethnic group level measures to use as explanatory variable for interstate conflicts. Building on this idea, we explored an alternative empirical approach by constructing a panel of ethnic group pairs. We coded a dummy equal to one in case of a conflict is recorded, in a given year, in any of the country pairs in which the groups are present, zero otherwise. We regress this variable on folklore similarity and other controls measured for ethnic-group pairs. Notice that this strategy, by avoiding the use of population shares by ethnicity, can be considered as an important robustness check to tackle the endogeneity problem of the population shares in our baseline regression.

Moreover, this approach allows the inclusion of ethnic group fixed effects in the regression, which control for all ethnic characteristics that can lead to negotiation failures and conflicts such as lack of trust, greed, etc. To avoid overfitting the model, consistently with the procedure outlined in section 5.4, we kept only bigger or politically relevant ethnic groups in the countries at war in which they are present, using a 10% population share threshold and using the EPR data on political relevance. As regression controls, we included geodesic, linguistic, genetic, and religious distances, together with the area of the migratory triangle (see section 5.4 for details on the computation). Genetic distance data are available only for a subset of the

groups, so the regressions could only be run on a much smaller sample.

The results are summarized in columns 1 and 2 of table 6. We find that folklore similarities are negatively and statistically significantly associated with conflict. One potential concern with the result is that the very presence of an ethnic group in a country could be the consequence of modern conflicts, for instance, because of territorial invasions. To check our results, we also performed an alternative empirical test focusing on a sample of native ethnic groups only, that is taking away, from the previous dataset, the groups whose ancestral homeland centroid is not within the borders of the countries at war. The results are summarized in columns 3 and 4 of table 6, and are in line with those already discussed.

6 Folklore and Conflict: Mechanisms and Further Evidence

In interpreting our results, we argued that countries with more similar oral traditions are less prone to conflict because of lower information frictions in bilateral bargaining and more similar cultural beliefs, which make it easier to sustain peaceful equilibria and to enforce agreements. The main goal of this section is twofold. On the one hand, we provide further empirical evidence in favor of this interpretation, that is showing evidence of more coordination, and less bargaining failures, for countries with higher record of oral tradition in common. On the other hand, this evidence is meant to highlight few potential transmission channels through which folklore similarities indirectly affects the probability of conflict.

In particular, we will investigate whether countries with higher record of common oral tradition are also more likely to form military alliances (subsection 6.1), more likely to vote similarly in the UN general assembly (subsection 6.2), more likely to participate in the same intergovernmental organizations (subsection 6.3) and more dependent on their bilateral trade relationship (subsection 6.4), in the sense that they tend to trade more with each other and less with other countries. Finally, we also investigate whether conflicts are more likely to end with a negotiation, instead of an imposition, among countries with more folklore similarity (subsection 6.5).

6.1 Folklore and Military Alliances

We study here the relationship between oral tradition similarity and military alliances, which we take as an indicator of the ability, for two countries, to coordinate to sustain agreements, that are particularly important for the possibility that a dispute may escalate into a war. We again use data from the Correlates of War database, those who register alliances based on their type (Gibler and Sarkees 2004). There, in particular, we find an indicator variable equal to 1 in case of defense pact, 2 in case of neutrality, 3 in case of entente, or 4 in case of no pact. We defined a dummy equal to one in case of any form of agreement in a given year, i.e. of score less than 4. Since our aim is to study the determinants of the probability to sustain agreements, we computed, for all country pairs, the number of years with any form of alliance in the sample, and then regressed this total number of years with an alliance on folklore similarity and controls, including country fixed effects. Because of the preesnce of many zeros in the count data, we used a Pseudo-Poisson Maximum Likelihood estimator. The results are summarized in table 7. We report a positive and significant coefficient on folklore similarity, both in case of aggregation with actual and ancestral population shares. We conclude that higher sharing of historical oral traditions among the people who compose he country has made it easier for their governments to sustain alliances. As for the other measures of relatedness, we find negative and statistically significant coefficients on genetic and religious distances, while we do not find any robust result for linguistic distance.

Summarizing, the evidence suggests that similar oral traditions make it easier to sustain military alliances, thereby reducing the probability of a conflict onset. This strengthens the interpretation of the effect of folklore similarity on conflict in terms of information frictions. In particular, if folklore similarity increases the probability of a military alliance, the repeated interactions that this alliance imply should decrease mistrust and should also decrease the probability of mis-interpreting the actions of the allies.

6.2 Folklore and United Nations Voting Patterns

Gartzke (2007) argues that UN votes correlations signal similar international interests. To the extent that states with similar policy interests are less prone to conflict (Bueno de Mesquita

1985), this is indeed an additional potential transmission channel linking folklore similarities to conflict. Building on Gartzke (2006), we use the database of UN general assembly votes from 1946 to 2002, considering three indicators: the yearly votes correlation; the affinity index of the general assembly votes computed including abstentions, and the same affinity index without abstentions. The affinity index is computed, for each country pair and each year, as one minus twice the ratio of the sum of the metric distances between all UN general assembly votes in a given year over the maximum metric distance of those votes. The index ranges from -1 to 1, with a value of 1 (-1) indicating "Most (least) similar interests". The vote themselves are coded as follows: 1 for approval, 2 for abstention and 3 for disapproval. In case abstentions are excluded, there is only 1 for approval and 2 for disapproval. We compute the median value of the correlation over the sample (1816-2010) to smooth out the time variability, although we obtain the same results when running the regression on the full panel with time fixed effects.

We run OLS regressions with the same set of basic controls that we used in all others empirical models plus an additional control, the composite index of national capabilities proposed by Singer (1988), to account for the (time-varying) status of the countries. The results are summarized in table 8 and show that folklore similarity is associated with more correlated/affine UN votes. Thus coordination is easier in case of folklore similarities. We also find, interestingly, that genetically, linguistically and religiously more distant countries also feature lower affinity of votes at the UN Assembly. Summarizing, country pairs that share more traits in their oral traditions tend to vote similarly in the UN general assembly, thereby showing similar international interests.

6.3 Folklore and Inter-Governmental Organization

We argued that folklore similarities ease information frictions, stem misperceptions and support punishment in bilateral negotiations, in such a way to foster cooperation. Therefore, we expect to see that oral tradition similarities facilitate the participation to institutions whose existence is based on willingness to cooperation. Our next analysis features the relationship between oral oral tradition similarity and the participation to Inter-Governmental Organizations (IGO henceforth).

We use four separate indicators of IGO participation for a country pair: the participation to all types of IGO, the participation to minimal organizations only (no bureaucracy and no enforcement), to structured organizations only (formal procedures and rules, bureaucracy, and policy implementation), interventionist organizations only (mechanisms for mediation, arbitration and adjudication, coercive power over states, enforcement of decisions), and to security organizations only. Note that interventionist organizations include financial institutions (IMF, WB, etc.) with no explicit security mandate. For each country pair and year, the indicators are equal to the number of the corresponding organizations to which both country participate in a given year. We employ a pseudo maximum likelihood possion estimator with both country and year fixed effects, and we cluster the standard errors at the level of the country pair.

The results are reported in table 9, and show that folklore similarity predict the participation to all types of IGO, including security organizations and structured organizations, which have the capability to reduce conflict according to the evidence in Boehmer et al. (2004). We also find that genetic, linguistic, and religious distance are all negatively associated with the joint participation to international organizations.

According Boehmer et al. (2004), there are three main channels that link the participation to IGO to conflict, identified within a bargaining approach to model inter state conflict. First IGO foster peace by revealing private information, for instance about military capabilities, thereby reducing the incentives to bluff. Second, IGO can alter the payoffs from a conflict with their power to assist and punish. Third, IGO ease costly signaling between states through the imposition of (non-military) sanctions, and this is more likely for IGO with security mandates.

6.4 Folklore and Trade

Following the literature on economic interdependence and conflict, in this section we explore the role of international trade as a mechanism linking folklore similarity to MID. Our first empirical test builds on Martin et al. (2008) and Morelli and Sonno (2017), who show that countries that trade more with each other, but less with third countries, have lower incentives to engage in international conflict due to the loss of the gains from trade.

The measure of bilateral trade dependence we use is that introduced by Morelli and Sonno

(2017), computed as follows:

$$T^{dep} = \frac{T_{ij} + T_{ji}}{\sum_k T_{ik} + \sum_k T_{jk}} \tag{6}$$

where T_{ij} are imports in country i from country j. The great advantage of this measure is that it increases with bilateral trade flows (bigger numerator), and decreases with multilateral trade flows (bigger denominator). So it evaluates how much the bilateral trade matters for the two countries in relative terms. We compute a time-varying index using historical trade data from the CEPII database (TRADHIST-Fouquin and Hugot 2016) from 1900 to 2010. One attractive feature of this dataset is that the authors developed a procedure to reconstruct the zero trade flows, distinguishing them from the missing observations.

By analogy with this trade dependence measure, we computed a measure of folklore dependency equal, for a country-pair, to twice their folklore similarity divided by the average folklore similarity with all other countries in the World:

$$D^{dep} = \frac{2F_{ij}}{\sum_k F_{ik} + \sum_k F_{jk}} \tag{7}$$

The idea is that, the more similar are two countries, and the more different from others, the more important is their bilateral relationship because of, among others, trade, and so the lower should be the expected cost of a war. When regressing the conflict dummy on folklore dependence and controls, we find indeed a negative and significant coefficient: wars are less likely to occur among country pairs whose folklore dependence is larger. The results are summarized in columns 1 and 2 of table 11, both in case of folklore similarity aggregated with actual or ancestral population shares.

To test if trade is one of the transmission channels behind this effect, we also regressed trade dependence on folklore dependence and controls, including country and year fixed effects, clustering standard errors at the country pair level. We find a positive and significant coefficient: countries that depend more on the particular trade relationship are also those with more folklore in common and less folklore in common with third countries. The results are summarized in columns 3 and 4 of table 11. Given that Morelli and Sonno (2017) also show that trade dependence decreases the probability of conflict, this evidence suggests that

the effect of folklore similarity on conflict goes also through international trade dependence[21].

Our second empirical test entails the use of the trade agreements data. Trade agreements are part of the international organization that structures international relations among countries, hence they are built on the same principles of self enforcing cooperative agreements in the realm of power relations. Moreover, given the gains from bilateral trade, and given the effects of trade agreements on trade (Baier and Bergstrand 2007; Bergstrand et al. 2015) we expect to see less conflict between countries with trade agreements. This is also a particular case of the results presented in section 6.3, particularly important because of the literature on the link between international trade and conflict. We use the data from Bergstrand et al. (2015), which assigns a (time-varying) score for the strength of the economic relationship between two countries. In particular, the score is equal to 1 in case of non-reciprocal preferential trade agreement (PTA), 2 in case of PTA, 3 in case of a free trade agreement, 4 in case of a customs union, 5 in case of a common market, and 6 in case of an economic union. We regress the scores on folklore similarity and controls, country effects and time effects, clustering the standard errors at the level of the country pair. The results are summarized in table 12, where we report that country pairs with more similar oral traditions feature stronger involvement in trade agreements.

All in all, the results in this section highlight trade as a mechanism through which folklore similarities decrease inter-state conflict. Moreover, they also provide evidence that international trade negotiations are easier to be established, and agreements easier to enforce, among countries with similar oral traditions, further validating our main interpretation of common narratives as a marker of common systems of beliefs that facilitate negotiations.

6.5 Folklore and MID Settlements

By using the data recorded in the COW database, we also have the opportunity to explore the way MID ended. The possible settlement types for MID are categorized as either "negotiated" or "imposed", although, in some cases, the classification is difficult, resulting into an "unclear". For each country pair in the sample that engaged in at least one MID, we computed the share

[21]Note then that bilateral trade dependence would be a "bad control" in the main regression of conflict on folklore similarity.

of negotiated settlement, that is the total number of settlements in the period of observations (1816-2016) divided by the total number of MID in the same period. We then regressed these variables on folklore similarity and controls, as in the main empirical specification. The results are summarized in table 10: countries with more similar oral traditions reach a negotiated settlement more often. Instead, the share of imposed settlements turns out to not to be related to folklore similarity, conditional on controls. These results suggest that information frictions are less relevant, and negotiations easier, among countries with similar oral traditions even conditionally on having an ongoing dispute.

7 Conclusion

We explored empirically the culture-conflict nexus proposing a new measure of cultural similarity that evaluates the extent to which oral traditions and folklore overlap in ethnic group pairs and countries. Compared to those measures based on philogenetic trees, like genetic, religious or linguistic distances, it captures a notion of cultural similarity which is not based on the branching process of ancestral separation, but rather on the idea that cultural traits can migrate across groups with mutation. As such it measures the slow and steady accumulation of common traits as a result of the episodes of cultural exchanges experienced over time.

We documented that country pairs with more similar oral traditions have been less likely to fight with each other in the modern era. The relationship is robust and it holds even after accounting for several other markers of cultural similarity previously used in the literature, and for many potential confounders.

Our interpretation of the above results is not in support of a simple primitivist approach, where interstate conflicts are merely the final destiny of ancestral cleavages. Rather, we interpret them as evidence of the role of narratives as "cultural beliefs", a system of interpretations of the counterpart actions in bilateral negotiations. The fact that the Spolaore and Wacziarg (2016) result of a negative association between conflicts and genetic distance holds in our analysis, with genetic distance being the closest proxy for primordial separations in group identity, lends further support to this interpretation. We also offered additional empirical results to further support this interpretation, linking folklore similarity to trust, cooperation

and coordination in repeated interactions. In particular, we showed that country pairs who share more of their oral traditions are also more likely to form military alliances, to participate more frequently in the same international organizations, to vote more similarly at the United Nations General Assembly, and sustain stronger bilateral trade relationships. Moreover, we showed that a concept-specific measure of folklore similarity, based on a set folklore motifs associated with system of beliefs that might sustain better outcomes in repeated interactions (i.e. those related to punishment, cooperation, threats, reputation, etc.) is negatively and significantly associated to the probability of a conflict.

One of the challenges of our empirical design was to tackle the complex relationship between ethnic groups and the political organization (modern states) that take decisions on disputes and conflicts. In particular, we did not address an important dimension for which ethnic-cultural traits may be important, the culture-war nexus for civil conflicts and its relation with interstate conflicts. This is left for future work. Furthermore, the evidence we reported, i.e. that oral tradition similarity computed with actual population shares is actually bigger, for many countries, than the one computed with ancestral population shares, suggests a potential important role of modern migrations for pacifying tensions and reducing the number and intensity of interstate wars. This definitely calls for more specific analysis, which we also leave to future research.

References

Acemoglu, Daron and Alexander Wolitzky. 2024. "Mistrust,misperception,and misunderstanding: Imperfect information and conflict dynamics" in Dube et al. (eds.), Volume 1, pp. 59-104.

Akerlof George A. and Dennis J. Snower. 2016. "Bread and bullets". *Journal of Economic Behavior & Organization* 126: 58-71.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg. 2003. "Fractionalization". *Journal of Economic Growth* 8: 155-194.

Alesina, Alberto, William Easterly and Janina Matuszeski. 2011. "Artficial States", *Journal of the European Economic Association* 9(2): 246-277.

Arbatli, Cemal Eren, Quamrul H. Ashraf, Oded Galor, and Marc Klemp. 2020. "Diversity and Conflict". *Econometrica* 88(2): 727-797.

Aron, Raymond. 1966. *Peace and War: A Theory of International Relations.* Translated by Richard Howard and Annette Baker Fox. Garden City, NY: Doubleday Company.

Ashraf, Quamrul and Oded Galor. 2013. " The 'Out of Africa' Hypothesis, Human Genetic Diversity and Comparative Economic Development". *American Economic Review* 103(1): 1-46.

Axelrod, Robert. 2006. *The Evolution of Cooperation (Revised edition)*, London, Perseus Books Group.

Baier, Scott L., and Jeffrey H. Bergstrand 2007. "Do Free Trade Agreements Actually Increase Members' International Trade?", *Journal of International Economics*, 71(1): 72-95.

Bakaki, Zorzeta, Tobias Böhmelt and Vincenzo Bove. 2015. "Barriers to Coordination? Examining the Impact of Culture on International Mediation Occurrence and Effectiveness". *Political Studies* 64(3): 492-512.

Baliga Sandeep and Tomas Sjöström. 2024. "Causes of war" in in Dube et al. (eds.), Volume 1: 1-58.

Becker, Anke, Benjamin Enke and Armin Falk. 2020. "Ancient Origin of the Global Variation in Economic Preferences". *American Economic Review Papers and Proceedings* 110: 319-323.

Becker, Anke. 2025. "On the Origins of Restricting Women's Promiscuity". *Review of Economic Studies*, forthcoming.

Bergstrand, Jeffrey H., Mario Larch and Yoto V. Yotov. 2015. "Economic Integation Agreements, Border Effects, and Distance Elasticities in the Gravity Equation", *European Economic Review* 78: 307-327.

Berezkin, Yuri. 2015. "Folklore and Mythology Catalogue: Its Lay-out and Potential for Research". *The Retrospective Methods Network Newsletter* 10: 58–70.

Blouin, Arthur and Julian Dyer. 2023. "How Cultures Converge: An Empirical Investigation of Linguistic Exchange". University of Toronto Working Paper.

Boehmer, Charles, Erik Gartzke, and Timothy W. Nordstrom. 2004 "Do Intergovernmental Organizations Promote Peace?" *World Politics* 57(1): 1-38.

Bortolini, Eugenio, Luca Pagani, Enrico R. Crema, Stefania Sarno, Chiara Barbieri, Alessio Boattini, Marco Sazzini, Sara Graca da Silva, Gessica Martini, Mait Metspalu, Davide Pettener, Donata Luiselli, and Jamshid J. Tehrani. 2017. "Inferring Patterns of Folktale Diffusion Using Genomic Data". *Proceedings of the National Academy of Sciencies* 114(34): 9140-9145.

Bove, Vincenzo and Gunes Gokmen. 2017. "Cultural Distance and Interstate Conflict", *British Journal of Political Science* 47(4): 939-949.

Bueno de Mesquita, Bruce. 1985. "The War Trap Revisited." *American Political Science Review* 79(1):156–77.

Bremer, Stuart A.. 1992. "Dangerous Dyads: Conditions Affecting the Likelihood of Interstate War: 1816-1965", *Journal of Conflict Resolution* 36(2): 309-341.

Bremer, Stuart A.. 2000. "Who Fights Whom, When, Where, and Why?" in J. A. Vasquez, ed., What Do We Know about War?, Rowman and Littlefield, Lanham, (MD)

Cao, Yiming, Benjamin Enke, Armin Falk, Paola Giuliano, and Nathan Nunn. 2024. "Herding, Warfare, and a Culture of Honor: Global Evidence". Working Paper.

Carr, E.H.. 1939. The Twenty Years' Crisis: 1919–1939. London: Macmillan.

Caselli, Francesco, Massimo Morelli and Dominic Rohner. 2015. "The Geography of Interstate Wars". *Quarterly Journal of Economics* 130(1): 267-315.

Cavalli-Sforza, Luigi L.. 2001. "Genes, Languages, People". Penguin Books Press, London, UK.

D'Amato, Marcello and Francesco Flaviano Russo. 2025. "Cultural Doorways in the Barriers to Development", *Journal of Economic Growth*, forthcoming.

Desmet, Klaus, Michel Le Breton, Ignacio Ortuno-Ortin and Sholmo Weber. 2011. "The Stability and Breakup of Nations: a Quantitative Analysis". *Journal of Economic Growth* 16(2): 183-213.

Di Lorenzo, Matthew and Bryan Rooney. 2025. "Leader Similarity and International Conflict", *Journal of Peace Research* 62(1): 102-118.

Dube, Oeindrila, Massimo Morelli and Debraj Ray (Eds.). 2024. "Handbook of the Economics of Conflict", Vol. 1, Elsevier B.V., Amsterdam.

Enke, Benjamin. 2019. "Kinship, Cooperation, and the Evolution of Moral Systems", *The Quarterly Journal of Economics*, 134(2): 953–1019.

Fearon, James D. 1995. "Rationalist Explanations for War", *International Organization* 49(3): 379-414.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency and Civil War", *American Political Science Review* 97(1): 75-90.

Fouquin, Michel and Jules Hugot. 2016. "Two Centuries of Bilateral Trade and Gravity Data: 1827-2014". CEPII Working Paper, N°2016-14.

Galor, Oded, Marc Klemp and Daniel C. Wainstock.. 2023. "The Impact of the Prehistoric Out of Africa Migration on Cultural Diversity", NBER working paper no. 31274, NBER, Cambridge, MA.

Galor, Oded, and Omer Ozak. 2015. "Land Productivity and Economic Development: Caloric Suitability vs. Agricultural Suitability". Mimeo, Brown University.

Gartzke, Erik. 2006. "The Affinity of Nations Index, 1946-2002", Version 4.0, University of California San Diego.

Gartzke, Erik. 2007. "The Capitalist Peace". *American Journal of Political Science* 51(1): 166-191.

Gartzke, Erik, and Kristian S. Gleditsch. 2006. "Identity and conflict: Ties that bind and differences that divide". *European Journal of International Relations*, 12(1): 53–87.

Gartzke, Erik and Oliver Westerwinter. 2016. "The complex structure of commercial peace contrasting trade interdependence, asymmetry, and multipolarity". *Journal of Peace Research* 53(3): 325-343.

Genicot Garance and Debraj Ray. 2024. "Identities in conflict". In Dube et al. (eds.), *Handbook of the Economics of Conflict*, Vol. 1 Ch. 5, pp. 201-247.

Gibler, Douglas M, and Meredith R. Sarkees. 2004. "Measuring Alliances: The Correlates of War Formal Interstate Alliance Dataset, 1816–2000". *Journal of Peace Research* 41(2): 211–222.

Giuliano, Paola, Antonio Spilimbergo and Giovanni Tonon. 2014. "Genetic Distance, Transportation Costs, and Trade". *Journal of Economic Geography* 14: 179-198.

Gokmen, Gunes. 2017. "Clash of Civilization and the Impact of Cultural Differences on Trade." *Journal of Development Economics* 127: 449-458.

Greenhill, Simon J.. 2021. "Do languages and genes share the same cultural evolutionary history?", *Science Advances* 7, eabm2472.

Greif, Avner. 1994. "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies", *Journal of Political Economy*, 102:5, 912-950.

Henderson, Errol A.. 1997. "Culture or Contiguity: Ethnic Conflict, the Similarity of States, and the Onset of War, 1820-1989". *Journal of Conflict Resolution* 41(5): 649-668.

Henderson, Errol A.. 1998. "The Democratic Peace Through the Lens of Culture, 1820-1989". *International Studies Quarterly* 42: 461-484.

Henderson, Errol A. and Richard Tucker. 2001. "Clear and Present Strangers: The Clash of Civilizations and International Conflict". *International Studies Quarterly* 45, 2: 317–338.

Herrmann, R. K. and Fischerkeller, M. P. 1995. "Beyond the Enemy Image and Spiral Model: Cognitive–Strategic Research After the Cold War.", *International Organization* 49(3), 415–450.

Huntington, Samuel P. 1993. "The Clash of Civilizations?", *Foreign Affairs* 72(3): 22–49.

Jackson, Matthew O. and Massimo Morelli. 2007. "Political Bias and War", *American Economic Review* 97(4): 1353-1373.

Jervis, Robert 1976. "Perception and Misperception in International Politics.", Princeton, NJ: Princeton University Press.

Martin, Philippe, Thierry Mayer, and Mathias Thoenig. 2008. "Make Trade Not War?" *Review of Economic Studies* 75: 865–900.

Mearsheimer, John J. 2001. "The Tragedy of Great Power Politics", New York: W.W. Norton Company (NY).

Mercer, J.. 2005. "Rationality and Psychology in International Politics", *International Organization*, 59(1), 77–106.

Michalopoulos, Stelios. 2025. "Ethnographic Records, Folklore, and AI" *National Bureau of Economic Research* Working Paper Series 33700, April.

Michalopoulos, Stelios and Melanie Meng Xue. 2021. "Folklore" *Quarterly Journal of Economics* 136(4): 1993-2046.

Morelli, Massimo and Tommaso Sonno. 2017. "On Economic Interdependence and War". *Journal of Economic Literature* 55(3): 1084-1097.

Morgenthau, Hans J. 1948. "Politics Among Nations: The Struggle for Power and Peace", Alfred A. Knopf, New York (NY).

Nunn, Nathan and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa". *Review of Economics and Statistics* 94: 20-36.

Palmer, Glenn, Roseanne W. McManus, Vito D'Orazio, Michael R. Kenwick, Mikaela Karstens, Chase Bloch, Nick Dietrich, Kayla Kahn, Kellan Ritter, Michael J. Soules. 2020. "The MID5 Dataset, 2011-2014: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science*, 39(4): 470-482.

Pemberton, Trevor J., Michael De Giorgio, and Noah A. Rosenberg. 2013. "Population Structure in a Comprehensive Genomic Data Set on Human Microsatellite Variation". *G3-Genes/Genomes/Genetics*, 3: 903-919.

Putterman, Louis and David N. Weil. 2010. "Post-1500 Population Flows and the Long Run Determinants of Economic Growth and Inequality", *Quarterly Journal of Economics* 125 (4): 1627-1682.

Shiller, Robert. 2017. "Narrative Economics". *American Economic Review* 107(4): 967-1004.

Singer, David J. 1988. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985" *International Interactions* 14: 115-32.

Spolaore, Enrico, and Romain Wacziarg. 2009. "The Diffusion of Development." *Quarterly Journal of Economics* 124 (2): 469-529.

Spolaore, Enrico and Romain Wacziarg. 2016. "War and Relatedness". *Review of Economics and Statistics* 98(5): 925-939.

Thucydides. The History of the Peloponnesian War. Translated by Richard Crawley. London: J. M. Dent Sons, 1910. Project Gutenberg, https://www.gutenberg.org/ebooks/7142.

Toelken, Barre. 1996. "Dynamics of Folklore", University Press of Colorado, Boulder, CO.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker, and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family", *Journal of Conflict Resolution* 59(7): 1327–42.

Waltz, Kenneth. 1979. Theory of International Politics. McGraw-Hill, Boston, Mass.

Wichmann, Søren, Eric W. Holman, and Cecil H. Brown (eds.). 2022. "The ASJP Database version 20" (https://asjp.clld.org/).

Table 1: **Folktales Similarity and Geodeisc Distances, Summary**

	Mean	Median	STD	1^{st} Qrt	3^{3d} Qrt	Min	Max
Folksim	0.1145	0.1115	0.0635	0.0663	0.1547	0.0005	0.6841
Folkcntdif	0.0581	0.0401	0.0533	0.0253	0.0671	0.0004	0.3168
Geo	8.9391	9.0407	4.4961	5.4240	12.2263	0.0141	19.9928

Notes: Observations are for ethnic group pairs. Folk Cosine is folklore similarity based on the cosine similarity of the motifs. Folk Cont is the folklore similarity measure based on difference in content. Geo is the geodesic distance (in thousand km) between the ancestral locations (centroids) of the ethnic groups.

Table 2: **Folklore Text Similarity and Conflict**

	Militarized Interstate Dispute (MID)			
	Actual Pop Shares		Ancestral Pop Shares	
	(1)	(2)	(3)	(4)
Folksim	-0.8595***	-1.0959***	-0.4785***	-0.7083***
	(0.2369)	(0.2891)	(0.1637)	(0.1962)
Geo	-1.2905***	-0.5221	-1.5013***	-0.5227
	(0.2237)	(0.3740)	(0.2489)	(0.3998)
Gen	-1.5043***	-1.4789***	-1.1225**	-1.1390***
	(0.3064)	(0.3295)	(0.2854)	(0.2942)
Lang	-1.5074*	-1.9165***	0.4743	0.2336
	(0.8259)	(0.9203)	(0.7113)	(0.8392)
Rel	0.2215	0.7917	0.0564	0.6720
	(0.4863)	(0.5096)	(0.4691)	(0.5062)
P. R^2	0.376	0.370	0.366	0.359
Obs	286828	218893	281983	216729
Geo Contr	no	yes	no	yes

Notes: Table entries are elasticities from a Probit estimation. Observations are for country pairs in a given year (1816-2010). Dependent variable is a dummy in case equal to one in case of a Militarized Interstate Dispute in the year (conflict intensity - COW - score equal to 3 or more. Folksim is the weighted average text measure of folklore similarity. In columns 1 and 2, weighted averages are computed with actual population share by ethnicity; in columns 3 and 4, weighted averages are computed using ancestral population shares. Geo is geodesic distance between the capitals of the countries. Gen is weighted average genetic distance from Pemberton et al. (2013). Lang is weighted average linguistic distance from Wichmann et al. (2022), ASJP. Rel is weighted average religious distance from the World Chrisitan database. All regressions include: a dummy for contiguous countries, a dummy for a common sea or ocean, a dummy for a past colonial link, a dummy in case the countries were ever part of the same country, the log of the GDP of both countries at PPP, the log of the population of both countries, the polity2 scores for both countries, country fixed effects for both sides, and time effects. Geo Contr is yes in case of additional control variables in the regression (see section 4 for the list). Clustered standard errors at the level of the country pair in brackets. P R^2 is the pseudo R^2 *** significant at 1% level. ** significant at 1% level. * significant at 10% level.

Table 3: **Disaggregated Folklore Text Similarity and Conflict**

	Militarized Interstate Dispute (MID)			
	Actual Pop Shares		Ancestral Pop Shares	
	(1)	(2)	(3)	(4)
Folksim Coop	-0.8731***		-0.4107***	
	(0.2082)		(0.1093)	
Folksim War		-0.0388		-0.0294
		(0.0634)		(0.0494)
Geo	-0.5831	-0.4939	-0.4820	-0.4907
	(0.3815)	(0.3776)	(0.4010)	(0.3784)
Gen	-1.7147***	-1.0246***	-1.2166***	-0.9788***
	(0.3414)	(0.3040)	(0.2992)	(0.2939)
Lang	0.5343	-0.0011	0.2065	0.4129
	(0.8210)	(0.7897)	(0.8325)	(0.8348)
Rel	0.3856	0.7442	0.6085	0.6504
	(0.5059)	(0.5040)	(0.5042)	(0.4994)
P. R^2	0.371	0.369	0.360	0.358
Obs	218893	218893	216729	216729

Notes: Table entries are elasticities from a Probit estimation. Observations are for country pairs in a given year (1816-2010). Dependent variable is a dummy in case equal to one in case of a Militarized Interstate Dispute in the year (conflict intensity - COW - score equal to 3 or more. Folksim coop is the weighted average text measure of folklore similarity computed using only the motifs related to concepts that might sustain cooperation in repeated interactions (tags: punish, retribution, cooperate, revenge, retaliation, patience, promise, threat, defect, forgive, memory, misunderstand, reputation, commitment, betray). Folksim war is the weighted average text measure of folklore similarity computed using only the motifs related to conflicts (tags: dispute, expand, hegemony, invade, intrude, attack, war, fight, combat, contend, military, army, troop, soldier, peace, sovereignty). In columns 1 and 2, weighted averages are computed with actual population share by ethnicity; in columns 3 and 4, weighted averages are computed using ancestral population shares. Geo is geodesic distance between the capitals of the countries. Gen is weighted average genetic distance from Pemberton et al. (2013). Lang is weighted average linguistic distance from the ASJP. Rel is weighted average religious distance from the World Christian database. All regressions include: a dummy for contiguous countries, a dummy for a common sea or ocean, a dummy for a past colonial link, a dummy in case the countries were ever part of the same country, the log of the GDP of both countries at PPP, the log of the population of both countries, the polity2 scores for both countries, the geographic controls listed in section 4, country fixed effects for both sides, and time effects. Clustered standard errors at the level of the country pair in brackets. P R^2 is the pseudo R^2 *** significant at 1% level. ** significant at 1% level. * significant at 10% level.

Table 4: **Folklore Content Similarity and Conflict**

	Militarized Interstate Dispute (MID)			
	Actual Pop Shares		Ancestral Pop Shares	
	(1)	(2)	(3)	(4)
Folkcnt	-4.9914***	-5.3016***	-3.0948**	-3.8782**
	(1.7076)	(1.9077)	(1.4693)	(1.6713)
Geo	-1.2613***	-0.4849	-1.4040***	-0.5149
	(0.2253)	(0.3774)	(0.2426)	(0.3899)
Gen	-1.3474***	-1.2712***	-1.0443***	-1.0039***
	(0.2971)	(0.3117)	(0.2791)	(0.2878)
Lang	-0.2441	-0.4412	0.5059	0.3017
	(0.7024)	(0.7883)	(0.7072)	(0.8362)
Rel	0.2508	0.8632*	0.1201	0.7390
	(0.4774)	(0.5038)	(0.4717)	(0.5078)
P. R^2	0.375	0.369	0.366	0.358
Obs	286828	218893	281983	216729
Geo Contr	no	yes	no	yes

Notes: Table entries are elasticities from a Probit estimation. Observations are for country pairs in a given year (1816-2010). Dependent variable is a dummy in case equal to one in case of a Militarized Interstate Dispute in the year (conflict intensity - COW - score equal to 3 or more. Flkcnt is the weighted average content measure of folklore similarity. In columns 1 and 2, weighted averages are computed with actual population share by ethnicity; in columns 3 and 4, weighted averages are computed using ancestral population shares. Geo is geodesic distance between the capitals of the countries. Gen is weighted average genetic distance from Pemberton et al. (2013). Lang is weighted average linguistic distance from the ASJP. Rel is weighted average religious distance from the World Chrisitan database. All regressions include: a dummy for contiguous countries, a dummy for a common sea or ocean, a dummy for a past colonial link, a dummy in case the countries were ever part of the same country, the log of the GDP of both countries at PPP, the log of the population of both countries, the polity2 scores for both countries, country fixed effects for both sides, and time effects. Geo Contr is yes in case of additional control variables in the regression (see section 4 for the list). Clustered standard errors at the level of the country pair in brackets. P R^2 is the pseudo R^2 *** significant at 1% level. ** significant at 1% level. * significant at 10% level.

Table 5: **Folklore Text Similarity and Years with Conflict**

	Number of Militarized Interstate Dispute (MID) in the sample			
	Actual Pop Shares		Ancestral Pop Shares	
	(1)	(2)	(3)	(4)
Folksim	-0.4291***	-0.5831***	-0.3155**	-0.5117**
	(0.1639)	(0.1902)	(0.1625)	(0.2188)
Geo	-0.7319***	-0.4653**	-0.8721***	-0.4748**
	(0.1925)	(0.2139)	(0.1847)	(0.2419)
Gen	-1.0869***	-0.9407***	-0.9060**	-0.7752***
	(0.2634)	(0.2383)	(0.2672)	(0.2366)
Lang	-1.5264***	-1.7564***	0.4743	-0.9685**
	(0.5168)	(0.5424)	(0.7113)	(0.5047)
Rel	0.2685	0.7704*	0.2790	0.7827*
	(0.5127)	(0.4410)	(0.5204)	(0.4237)
P. R^2	0.679	0.686	0.673	0.681
Obs	8338	5946	8027	5945
Geo Contr	no	yes	no	yes

Notes: Table entries are elasticities from a Pseudo-Poisson Maximum likelihood estimation. Observations are for country pairs. Dependent variable is total number of Militarized Interstate Disputes between 1816 and 2010. Folksim is the weighted average text measure of folklore similarity. In columns 1 and 2, weighted averages are computed with actual population share by ethnicity; in columns 3 and 4, weighted averages are computed using ancestral population shares. Geo is geodesic distance between the capitals of the countries. Gen is weighted average genetic distance from Pemberton et al. (2013). Lang is weighted average linguistic distance from the ASJP. Rel is weighted average religious distance from the World Chrisitan database. All regressions include: a dummy for contiguous countries, a dummy for a common sea or ocean, a dummy for a past colonial link, a dummy in case the countries were ever part of the same country, a dummy in case both countries are democracies, country fixed effects for both sides, and time effects. Geo Contr is yes in case of additional control variables in the regression (see section 4 for the list). Clustered standard errors at the level of the two countries in brackets. P R^2 is the pseudo R^2 *** significant at 1% level. ** significant at 1% level. * significant at 10% level.

Table 6: **Folklore Text Similarity and Conflict: Ethnic Groups**

	Militarized Interstate Dispute (MID)			
	Actual Groups		Ancestral Groups	
	(1)	(2)	(3)	(4)
Folksim	-0.6413***	-0.4172**	-0.5789***	-1.3689***
	(0.1044)	(0.1963)	(0.1998)	(0.5257)
Geo	-3.8150***	-4.6990***	-5.1344***	-6.7410***
	(0.2063)	(0.3649)	(0.3512)	(0.7896)
Lang	-1.8370***	-3.4842***	-1.7008*	-4.6773***
	(0.6707)	(0.8739)	(0.9059)	(1.2661)
Rel	-0.4460***	0.0005	0.1302	0.2140
	(0.1461)	(0.2561)	(0.2410)	(0.4797)
Gen		-0.7166**		-0.9442
		(0.3756)		(0.9308)
P. R^2	0.332	0.353	0.365	0.425
Obs	641981	128057	215422	26631

Notes: Table entries are elasticities from a Probit estimation. Observations are for ethnic group pairs in a given year (1816-2010). Dependent variable is a dummy equal to one in case of a Militarized Interstate Dispute in the year between the countries where the groups are present. Folksim is the text measure of folklore similarity. Geo is log geodesic distance between the ethnic groups centroids. Lang is linguistic distance from the ASJP. Rel is religious distance from the World Chrisitan database. Gen is genetic distance from Pemberton et al. (2013). All regressions include: the area of the migratory triangle defined by the centroids and the closest out-of-Africa migratory midpoint, ethnic group fixed effects for both sides, and time effects. Clustered standard errors at the level of the ethnic group pair in brackets. P R^2 is the pseudo R^2 *** significant at 1% level. ** significant at 1% level. * significant at 10% level.

Table 7: **Folklore and Military Alliances**

	Years with and Alliance			
	Actual Pop. Shares		Ancestral Pop. Shares	
	(1)	(2)	(3)	(4)
Folksim	0.8433***	0.4496**	0.6359***	0.3713**
	(0.2160)	(0.2284)	(0.1615)	(0.1805)
Geo	-0.21173**	0.1195	-0.2083***	0.1068
	(0.0867)	(0.0905)	(0.0821)	(0.0916)
Gen	-1.1940***	-1.4669***	-1.3145***	-1.5415***
	(0.2921)	(0.2853)	(0.2832)	(0.2707)
Lang	0.0881	-0.5624***	-0.0556	-0.6762***
	(0.1158)	(0.2112)	(0.1245)	(0.2284)
Rel	-1.6907***	-1.4882***	-1.8399***	-1.5552***
	(0.3305)	(0.2953)	(0.3543)	(0.3002)
P. R^2	0.748	0.761	0.748	0.514
Obs	7800	6652	7800	6695
Geo Contr	no	yes	no	yes

Notes: Table entries are coefficients from a Poisson Pseudo Maximum Likelihood estimation. Observations are for country pairs. Dependent variable is the number of years with an alliance in the period 1816-2010. Folksim is the log of the weighted average text measure of folklore similarity. In columns 1 to 4, weighted averages are computed with actual population share by ethnicity; in columns 5 to 8, weighted averages are computed using ancestral population shares. Geo is the log of geodesic distance between the capitals of the countries. Gen is the log of the weighted average genetic distance from Pemberton et al. (2013). Lang is the log of the weighted average linguistic distance from the ASJP. Rel is the log of the weighted average religious distance from the World Chrisitan database. All regressions include a dummy for contiguous countries, a dummy for a common sea or ocean, a dummy for a past colonial link, a dummy for both countries democratic (polity2 scores above 5), a dummy equal to one in case the two countries were ever part of the same political or administrative entity, same continent dummies, and country fixed effects for both countries in the pair. Geo Contr is yes in case of additional control variables in the regression (see section 4 for the list). Two-way clustered standard errors at the level of the two countries in brackets. P R^2 is the pseudo R^2 *** significant at 1% level. ** significant at 5% level. * significant at 10% level.

Table 8: **Folklore and UN Votes**

	UN Vote Correlation		Affinity w/o Abst		Affinity with Abst	
	(1)	(2)	(3)	(4)	(5)	(6)
Folksim Act	0.2747***		0.1827***		0.2264***	
	(0.0533)		(0.0399)		(0.0461)	
Folksim Anc		0.2308***		0.1942***		0.2468***
		(0.0487)		(0.0430)		(0.0459)
Geo	-0.0422**	0.0267	-0.0601***	-0.0029	-0.0583***	0.0163
	(0.0209)	(0.0274)	(0.0205)	(0.0262)	(0.0195)	(0.0231)
Gen	0.0368*	-0.0268	0.0123	0.0026	0.0091	-0.0017
	(0.0223)	(0.0214)	(0.0189)	(0.0179)	(0.0209)	(0.0194)
Lang	-0.0763**	-0.1377***	-0.0277	-0.1083***	-0.0238	-0.1187***
	(0.0276)	(0.0269)	(0.0278)	(0.0270)	(0.0297)	(0.0247)
Rel	-0.0290	-0.0580	-0.0478*	-0.0651**	-0.0684**	-0.0924***
	(0.0373)	(0.0391)	(0.0268)	(0.0291)	(0.0305)	(0.0337)
R^2	0.613	0.633	0.699	0.701	0.691	0.692
Obs	8917	8389	12072	10919	12072	10919

Notes: Table entries are standardized coefficients from an OLS regressions. Dependent variable is in column: median correlation of UN general assembly votes, median affinity index of UN general assembly votes in a year excluding abstentions and mean affinity index including abstentions. Folksim Act is weighted average folklore similarity computed with actual population shares. Folksim Anc is weighted average folklore similarity computed with ancestral population shares. Geo is geodesic distance between the capitals of the countries. Gen the weighted average genetic distance from Pemberton et al. (2013). Lang is weighted average linguistic distance from ASJP; Rel is weighted average religious distance from the World Christian Database. All regressions include: a dummy for contiguous countries, a dummy for a common sea or ocean, a dummy for a past colonial link, a dummy for both countries democratic based on the polity2 scores, a dummy equal to one in case the two countries were ever part of the same political or administrative entity. Two-way clustered standard errors at the level of the two countries in a conflict in brackets. *** significant at 1% level. * significant at 10% level.

Table 9: **Folklore and Inter-Governmental Organizations (IGO)**

	IGO all		IGO min		IGO str		IGO int		IGO sec	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Folksim Act	0.1063***		0.2378***		0.2074***		0.0468***		0.1654***	
	(0.0079)		(0.0238)		(0.0165)		(0.0034)		(0.0189)	
Folksim Anc		0.0694***		0.1843***		0.1862***		0.0179***		0.1831***
		(0.0056)		(0.0171)		(0.0109)		(0.0023)		(0.0135)
Geo	-0.1513***	-0.1266***	-0.4510***	-0.3911***	-0.2202***	-0.1507***	-0.0281***	-0.0219***	-0.3198***	-0.2711***
	(0.0056)	(0.0059)	(0.0155)	(0.0175)	(0.0099)	(0.0104)	(0.0022)	(0.0023)	(0.0126)	(0.0136)
Gen	-0.0920***	-0.1579***	-0.2191***	-0.3677***	-0.2398***	-0.3333***	-0.0227***	-0.0538***	-0.0410**	-0.1107***
	(0.0095)	(0.0075)	(0.0283)	(0.0212)	(0.0181)	(0.0136)	(0.0042)	(0.0033)	(0.0207)	(0.0161)
Lang	-0.0323	-0.0806**	-0.0331	-0.1523**	-0.0067	-0.1183***	-0.0198*	-0.0367***	-0.0936**	-0.1558**
	(0.0282)	(0.0350)	(0.0524)	(0.0703)	(0.0354)	(0.0459)	(0.0113)	(0.0132)	(0.0475)	(0.0501)
Rel	-0.1037***	-0.1122***	-0.4825***	-0.4930***	-0.2317***	-0.2477***	-0.0022	-0.0068*	-0.5819***	-0.5894**
	(0.0096)	(0.0102)	(0.0286)	(0.0298)	(0.0189)	(0.0198)	(0.0037)	(0.0039)	(0.0213)	(0.0214)
P. R^2	0.428	0.437	0.436	0.448	0.207	0.213	0.306	0.311	0.264	0.270
Obs	264245	251407	264051	251213	261868	249030	264245	251407	261868	249030

Notes: Table entries are coefficients from a Poisson Pseudo Maximum Likelihood estimation. Observations are for country pairs in a year. Dependent variable is in column. IGO all is the total number of IGO to which both country participate in a year; IGO min is the number of minimal IGO to which both country participate in a year; IGO str is the number of structured IGO to which both country participate in a year; IGO int is the number of interventioninsts IGO to which both country participate in a year; IGO sec is the number of security IGO to which both country participate in a year. Folksim Act is the log of the weighted average folklore-based cultural similarity computed with actual population shares. Folksim Anc is the log of the weighted average folklore-based cultural similarity computed with ancestral population shares. Geo is log geodesic distance between the capitals. Gen is the log of the weighted average genetic distance from Pemberton et al. (2013). Lang is the log of the weighted average linguistic distance from the ASJP. Rel is the log of the weighted average religious distance from the WCD. All regressions include: a dummy for contiguous countries, a dummy for a common sea or ocean, a dummy equal to one in case the two countries were ever part of the same political or administrative entity, a dummy for a colonial link, the polity2 scores of both countries, the GDP of both countries, the population of both countries, country fixed effects and year effects. Clustered standard errors at the level of country pair in brackets. P R^2 is the pseudo R^2 *** significant at 1% level. * significant at 10% level.

Table 10: **Folklore and MID Settlements**

	MID Settlements			
	Share Negotiated		Share Imposed	
	(1)	(2)	(3)	(4)
Folksim Act	0.0588***		-0.0055	
	(0.0222)		(0.0152)	
Folksim Anc		0.0621**		-0.0039
		(0.0298)		(0.0121)
Geo	-0.0835***	-0.0909***	0.0131	0.0089
	(0.0328)	(0.0363)	(0.0141)	(0.0151)
Gen	0.0233	0.0099	-0.0123	-0.0096
	(0.0339)	(0.0353)	(0.0221)	(0.0202)
Lang	0.0007	-0.0118	-0.0108	-0.0054
	(0.0109)	(0.0098)	(0.0096)	(0.0069)
Rel	0.0255	0.0219	-0.0737***	-0.0776***
	(0.0180)	(0.0206)	(0.0213)	(0.0222)
P. R^2	0.343	0.336	0.308	0.308
Obs	973	932	973	932

Notes: Table entries are coefficients from a OLS regression. Observations are for country pairs. Dependent variable is the share of negotiated MID settlements in the period 1816-2010 (columns 1 and 2) or the share of imposed settlements (columns 3 and 4). Folksim Act is the log of the weighted average folklore-based cultural similarity computed with actual population shares. Folksim Anc is the log of the weighted average folklore-based cultural similarity computed with ancestral population shares. Geo is the log of geodesic distance between the capitals of the countries. Gen is the log of the weighted average genetic distance from Pemberton et al. (2013). Lang is the log of the weighted average linguistic distance from the ASJP. Rel is the log of the weighted average religious distance from the World Chrisitan database. All regressions include a dummy for contiguous countries, a dummy for a common sea or ocean, a dummy for a past colonial link, a dummy for both countries democratic (polity2 scores above 5), a dummy equal to one in case the two countries were ever part of the same political or administrative entity, same continent dummies, and country fixed effects for both countries in the pair. Two-way clustered standard errors at the level of the two countries in brackets. P R^2 is the pseudo R^2 *** significant at 1% level. ** significant at 5% level. * significant at 10% level.

Table 11: **Folklore Dependence, Conflict and Trade Dependence**

	Conflict		Trade Dependence	
	(1)	(2)	(3)	(4)
Folkdep Act	-0.4638**		0.0391***	
	(0.2377)		(0.0096)	
Folkdep Anc		-0.3361**		0.0389**
		(0.1621)		(0.0170)
Geo	-1.2693***	-1.5939***	0.0558***	0.0551***
	(0.2253)	(0.2712)	(0.0203)	(0.0193)
Gen	-1.3118***	-1.1149***	0.0059	-0.0058
	(0.3038)	(0.2935)	(0.0128)	(0.0158)
Lang	-0.6854	0.5568	-0.0554***	-0.0721***
	(0.7947)	(0.7107)	(0.0152)	(0.0158)
Rel	0.1988	-0.0290	0.0360***	0.0408***
	(0.4807)	(0.4701)	(0.0107)	(0.0114)
R^2	0.375	0.366	0.306	0.471
Obs	286828	281983	571193	338002

Notes: Observations are for country pairs in a given year. Dependent variable in columns 1 and 2 is a dummy in case of a conflict; in columns 3 and 4 is the index of bilateral trade dependence (see section 6.4). Folkdep Act is folklore dependence computed with actual population shares. Folkdep Anc is folklore dependence computed with ancestral population shares. Geo is geodesic distance between the capitals. Gen is weighted average genetic distance. Lang is weighted average linguistic distance from the ASJP. Rel is weighted average religious distance. All regressions include: a dummy for contiguous countries, a dummy for a common sea or ocean, a dummy for a past colonial link, the number of common ancestral ethnic groups in the pair, the log of the gdp of the two countries, country fixed effects and time fixed effects. Regressions in columns 1 and 2 also include the polity2 scores of the two countries and the log of the population of the two countries. Regressions in columns 3 and 4 also include the transportation cost from Giuliano, Spilimbergo and Tonon (2014). Columns 1 and 2: Probit regression, table entries are elasticities. Columns 3 and 4: Ols regression, table entries are standardized coefficients. Clustered standard errors are the country-pair in brackets. *** significant at 1% level. * significant at 10% level.

Table 12: **Folklore and Trade Agreements**

	Trade Agreements	
	(1)	(2)
Folksim Act	0.3523***	
	(0.0548)	
Folksim Anc		0.1007***
		(0.0367)
Geo	0.0538	0.0608
	(0.0509)	(0.0529)
Gen	-0.5423***	-0.6750***
	(0.0556)	(0.0477)
Lang	-0.0334	-0.2385**
	(0.0751)	(0.1033)
Rel	-0.4654***	-0.4674***
	(0.0672)	(0.0681)
P. R^2	0.471	0.471
Obs	338062	338002

Notes: Observations are for country pairs in a given year. Dependent variable is the strength of the trade agreement (see section 6.4 for the rankings) Folksim Act is weighted average folklore similarity computed with actual population shares. Folksim Anc is weighted average folklore similarity computed with ancestral population shares. Geo is the log of the geodesic distance between the capitals of the Country. Gen is log weighted average genetic distance. Lang is log weighted average linguistic distance from the ASJP. Rel is log weighted average religious distance. All regressions include: a dummy for contiguous countries, a dummy for a common sea or ocean, a dummy for a past colonial link, the number of common ancestral ethnic groups in the pair, the log of the gdp of the importer, the log of the gdp of the exporter, the transportation cost from Giuliano et al. (2014), the polity2 scores, the geographic controls described in section 6.4, importer dummies, exporter dummies and time dummies. Pseudo-maximum likelihood poisson estimation. Clustered standard errors are the country-pair in brackets. *** significant at 1% level. * significant at 10% level.