Culture, Ethnicity and Diversity*

Klaus Desmet
SMU and CEPR

Ignacio Ortuño-Ortín
Universidad Carlos III

Romain Wacziarg
UCLA, NBER
and CEPR

February 2015

Abstract

We investigate the empirical relationship between ethnicity and culture, defined as a vector of traits reflecting norms, attitudes and preferences. Using surveys of individual values in 76 countries, we find that ethnic identity is a significant predictor of cultural values, yet that within-group variation in culture trumps between-group variation. Thus, in contrast to a commonly held view, ethnic and cultural diversity are unrelated. We explore the correlates of cultural diversity and of the overlap between culture and ethnicity, finding that the level of economic development is positively associated with cultural diversity and negatively associated with the overlap between culture and ethnicity. Finally, although only a small portion of a country's overall cultural heterogeneity occurs between groups, this does not imply that cultural differences between groups are irrelevant. Indeed, we find that civil conflict becomes more likely when there is greater overlap between ethnicity and culture.

*Desmet: Department of Economics, Southern Methodist University, 3300 Dyer, Dallas, TX 75205, kdesmet@smu.edu; Ortuño Ortín: Department of Economics, Universidad Carlos III de Madrid, 28903 Getafe, Madrid, Spain, iortuno@eco.uc3m.es; Wacziarg: UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles CA 90095, wacziarg@ucla.edu. We thank Georgy Egorov, James Fearon, Paola Giuliano, Wolfgang Keller, Keith Krehbiel, Giacomo Ponzetto and seminar participants at Stanford, Northwestern, Alicante, Pompeu Fabra, Bocconi, the London School of Economics, the Université Catholique de Louvain, the Paris School of Economics and the NBER Summer Institute for useful comments. Any remaining errors are ours.

1 Introduction

Are ethnic cleavages associated with deep differences in culture between groups? Many people think so. In poor countries, often characterized by a high level of ethnic diversity, concerns arise that groups with heterogeneous values, norms and attitudes - the broad set of traits that we will refer to as "culture" - may be unable to agree on policies, the provision of public goods and the broader goals of society. In rich countries, debates rage over multiculturalism and whether population movements brought about by globalization and modernity will result in cultural divisions and the breakdown of social consensus. Underlying these debates is an assumption that people agree within groups and disagree across groups, so that cultural heterogeneity and ethnic heterogeneity are two sides of the same coin. Yet, there is little quantitative research on the links between ethnicity and culture.

In this paper we conduct a systematic investigation of the links between culture and ethnicity. In doing so, we aim to answer the following questions: Is an individual's ethnolinguistic identity a predictor of his norms, values and preferences? Are ethnolinguistic heterogeneity and cultural heterogeneity highly correlated? What is the degree of overlap between both measures of diversity? Finally, is the relationship between ethnicity and culture important to understand salient political economy outcomes, such as civil conflict?

We start by exploring the relationship between ethnolinguistic identity and culture, using individual-level data from various surveys such as the World Values Survey. We seek to explain answers on norms, values and preferences using a respondent's economic and demographic characteristics, among which are ethnic and linguistic indicators, and to evaluate the joint statistical significance of the latter. We find that ethnicity dummy variables are jointly significant predictors of responses for about half of the questions, although this average masks significant heterogeneity across countries. Thus, ethnic identity appears to be an important determinant of cultural norms, values and preferences.

Although this suggests a strong link between ethnicity and culture, a very different picture emerges when we analyze the relation between cultural fractionalization and ethnic fractionalization. To get a measure of cultural fractionalization, we compute the probability that two randomly drawn individuals answer a randomly drawn question from the World Values Survey differently. In contrast to many observers' priors, we find that heterogeneity in norms, values and preferences is uncorrelated with ethnolinguistic fractionalization across countries. Taken together, these results show that even though culture does differ across ethnolinguistic groups, cultural fractionalization and ethnolinguistic fractionalization are not related. Ethnic fractionalization can therefore not readily be taken as a proxy for overall cultural and preference heterogeneity.

How can these seemingly contradictory results be reconciled? Within-group heterogeneity in culture may account for the low correlation between cultural heterogeneity and ethnolinguistic diversity, without precluding the possibility that ethnic identity has predictive power for cultural attitudes: the degree

of between-ethnic group cultural heterogeneity could be small relative to total heterogeneity, yet have significant predictive power for various political economy outcomes. To explore this possibility, we propose new indices of the degree of overlap between ethnicity and culture, derived from a simple model of social antagonism. The first is a χ^2 index that captures the average distance between the answers of each ethnic group and the answers in the overall population. A low value of the index indicates that groups reflect the countrywide distribution of answers, while a high value indicates a lot of group-specificity. The second index, developed in the context of population genetics, is known as a *fixation index*, or F_{ST}. It captures the between-group variance in answers to survey questions as a share of the overall variance. A value of zero indicates that there is no informational content to knowing an individual's ethnic identity, while a value of one indicates that answers can be perfectly predicted based on knowing an individual's ethnic identity.

Using χ^2 and F_{ST}, we find that the degree to which cultural and ethnic cleavages overlap is very small. In particular, we find that only on the order of $1-2\%$ of the variance in cultural norms is between groups. That is, the vast share of the variance is within groups. This explains the close-to-zero correlation between cultural heterogeneity and ethnic heterogeneity. The low share of between-group variance is not a simple consequence of the type of questions asked in the World Values Survey: when taking countries, rather than ethnicities, as the relevant groups, we find that the between-country share of the variance in cultural values is about five times larger. Furthermore, in spite of the small degree of overlap between culture and ethnicity, there is substantial variation across countries in the F_{ST} and χ^2 measures, and this variation is related in meaningful ways to some salient cross-sectional characteristics of countries.

Does cultural diversity between ethnic groups, though of a small magnitude, matter for our understanding of political economy outcomes? To analyze whether the overlap between culture and ethnicity is relevant, we explore the effect of ethnic heterogeneity, cultural heterogeneity and the degree of overlap between the two on the onset and incidence of civil conflict. In principle, civil wars could arise when there is a high degree of cultural heterogeneity, when there is a high degree of ethnic diversity, or when culture and ethnicity reinforce each other. Empirically, we find that both cultural and ethnic diversity have weak effects on civil conflict. If anything, cultural diversity has a pacifying effect. However, the degree of overlap between cultural attitudes and ethnic identity has a strong and robust effect on civil wars: when culture and ethnicity reinforce each other (i.e. for high values of F_{ST} or χ^2) violent conflict becomes more likely.

This paper is related to various strands of the literature on ethnolinguistic diversity. The first strand studies the relationship between ethnolinguistic diversity and socioeconomic outcomes, using conventional measures of diversity such as fractionalization (for instance, Easterly and Levine, 1997, Alesina, Baqir and Easterly, 1999, Alesina et al., 2003, Alesina and La Ferrara, 2005, among many others). Our paper is related to this literature as we examine the effect of ethnic and cultural fractionalization on a particular outcome, civil conflict. By explicitly considering cultural diversity and its relation with ethnic

heterogeneity, we cast light on the mechanisms that led to the empirical regularities uncovered in the earlier literature.

The second strand seeks to advance the measurement of diversity by considering alternative measures that improve on simple fractionalization. These measures take different forms, accounting for distance between groups (Esteban and Ray, 1994, 2004, Bossert, d'Ambrosio and La Ferrara, 2011), looking at income inequality between ethnic groups (Alesina, Michalopoulos and Papaioannou, 2012) or the historical depth of ethnic cleavages (Desmet, Ortuño-Ortín and Wacziarg, 2012). Our paper is related to this measurement literature because we propose a new measure of heterogeneity in cultural attitudes and new measures of the degree of overlap between culture and ethnicity. These measures shed new light on the complex empirical relationship between culture and ethnicity.

A third strand of the literature examines the relationship between culture and economic outcomes. This literature usually examines the effect of a particular historically-determined trait on current outcomes, rather than the effect of cultural *diversity* as we do. This is, again, a vast literature, but salient examples include Alesina, Giuliano and Nunn (2013) on the historical legacy of the heavy plough on values affecting fertility and female labor force participation; Giuliano (2007) on the effect of culture on living arrangements; Fernandez and Fogli (2009) on culture, fertility and female labor force participation; Luttmer and Singhal (2011) on culture and the taste for redistribution; Tabellini (2010) on cultural traits and economic performance across the regions of Europe; and Guiso, Sapienza and Zingales (2009) on trust and bilateral trade. In contrast to this literature, we study the effect of cultural heterogeneity rather than the effect of a specific cultural trait.

Finally, a recent literature seeks to relate genetic differences - a measure associated with cultural differences - with political and economic outcomes, including conflict. For instance, Spolaore and Wacziarg (2009) look at the effect of genetic distance between countries on the diffusion of the Industrial Revolution and Spolaore and Wacziarg (2013) study the effect of genetic distance between countries on interstate conflict and war. While these two studies also use F_{ST} as a measure of distance between groups, this F_{ST} is based on genetic rather than cultural data, and it is used to study interactions between pairs of countries rather than between groups within countries. Ashraf and Galor (2013) investigate the effect of genetic diversity, used as a broader measure of diversity in both cultural and biological traits within countries, on historical and contemporary economic performance. In Arbatli, Ashraf and Galor (2013), the same measure of genetic diversity is found to have a positive effect on the probability of civil conflict. In contrast to these papers, we measure cultural diversity directly using responses to surveys on norms, attitudes and preferences, rather than using genetic data.

The rest of the paper is organized as follows. In Section 2, we use individual level data from surveys of cultural attitudes to explore the relationship between ethnic identity and cultural attitudes. In Section 3, we introduce a simple model of social antagonism leading to three classes of measures of heterogeneity hypothesized to affect socioeconomic outcomes. We show how to operationalize these theoretically derived

measures using data on ethnicity and cultural traits. In Section 4, we introduce our new measures of heterogeneity, compute them using the World Values Survey, and describe their interrelationships and determinants. In Section 5, to illustrate the uses of these new measures, we explore empirically the effect of cultural and ethnic heterogeneity on civil conflict. Section 6 concludes.

2 Identity and Culture

2.1 Methodology

In this section we use the World Values Survey to examine the relationship between ethnic identity and cultural attitudes. The exercise requires individual level data on answers to questions on norms, values and preferences, and corresponding data on the respondent's ethnic or linguistic identity. We examine the joint statistical significance of indicators of ethnolinguistic identity as determinants of survey responses, proceeding question by question and country by country and controlling for observable individual characteristics. In principle, 5% of the questions should feature a significant joint effect of ethnic identity if the statistical criterion is 95% confidence and there was in fact no association between cultural attitudes and ethnicity. We ask whether the share of questions for which there is a significant effect of ethnicity is actually higher, and find that it is in fact much higher. We also examine whether the importance of identity for culture varies in systematic ways across question types, countries, continents, etc.

For each question and each country, we estimate the following specification:

$$Q_m = \alpha + \sum_{s=1}^{S} \beta_s D_m^s + \boldsymbol{\gamma}' \mathbf{X}_m + \varepsilon_m \tag{1}$$

where m denotes a respondent, $s = 1, ..., S$ indexes ethnolinguistic groups, Q_m is individual m's answer to the question under consideration, D_m^s is equal to one if respondent m is part of group s, zero otherwise, and $\mathbf{X}_m$ is a vector of controls. Estimation is by least squares.

We test for the joint significance of the β_s parameters using conventional F-tests. We do so for all questions and countries, and then examine the share of questions for which ethnolinguistic identity is a significant predictor of cultural attitudes at the 5% level. We compute these shares over different categories of questions, for each country separately, and for different regions. We also examine how much additional explanatory power ethnicity dummies bring to the regression, by comparing the simple R^2 statistic from running the specification in (1) to the one obtained when running the same regression without ethnicity dummies. This is meant to capture the magnitude of the joint effect of ethnicity on answers to cultural questions.

2.2 Data

Our main source is the Integrated World Values Survey-European Values Survey (WVS-EVS) dataset covering 1981 to 2008 and five survey waves. In order to examine the relationship between ethnicity and culture in a systematic way, we choose to focus on the broadest set of available questions without casting judgment on which ones are more representative of attitudes and preferences: we let the dataset largely guide our choice of questions, as opposed to making *ad hoc* choices ourselves. In the WVS-EVS integrated dataset, there is a total of $1,031$ fields, or questions. Some of these fields are not survey questions but instead refer to socio-demographic characteristics of the respondent or the interviewer, and some have zero observations. We confine attention to survey questions identified by the survey itself as pertaining to norms, values and attitudes (these come grouped by the survey organization into question categories labelled from A to G), and with a nonzero number of respondents. Among those, in very rare cases some questions were asked in a slightly different manner in some countries (Colombia, Hong Kong, Mexico, Iraq), and those were dropped (19 questions). We also dropped questions that asked about circumstances specific to a given country, i.e. questions that could not conceivably be asked in more than one country (74 questions). In the end we were left with 808 questions.

Among these remaining questions, there were three types: those with a binary response (yes/no, agree/disagree: 252 questions), those with an ordered response (where answers are on a scale of, say, 1 to 10: 496 questions), and those with strictly more than two possible responses that are not naturally ordered (60 questions). The first two categories can be used readily as dependent variables. For the third category, we cannot directly estimate the joint effect of ethnicity on unordered responses, so we transformed each possible response into a series of binary response questions.[1] Thus, the 60 questions with unordered responses resulted in 193 new binary questions, leading to a total of 941 questions. Of course, not every one of these questions was asked in every country, or in every wave. We keep all questions irrespective of where or when they were asked. In the end, out of 941 questions, on average 294 were asked in each country (the number of questions per country varied between 81 and 447 - Appendix Table A1 provides the exact count, country by country). When combined across all waves, the average number of respondents across the countries in the sample, and across all questions, was $1,497$. There is some heterogeneity around this number as some questions were asked in more waves than others, and the number of surveyed individuals varies across countries and waves.

An important aspect of our exercise is to correctly code ethnolinguistic identity in order to estimate the joint effect of ethnicity dummies on responses. To do so, we have to define ethnicity. The WVS/EVS asks

[1] For instance, question C009 asks "Regardless of whether you're actually looking for a job, which one would you, personally, place first if you were looking for a job?" and offers the following choices: "a good income", "a safe job with no risk", "working with people you like", "doing an important job", "do something for community". We define 5 binary response questions, where, for instance, for "a good income", the response value is 1 if the respondent answered "a good income" to question C009, and zero otherwise, and so on for the other answer categories.

respondents to report both their ethnicity and language. In some cases, the reported ethnic categories do not appropriately capture ethnic identity. For many African countries the WVS/EVS integrated survey reports ethnicities as White / Black. For instance in Zambia, 99.47% of respondents are Black, while there are 0.27% Asians and 0.27% Whites. Most ethnographers agree that for Africa, language is a better measure of ethnic identity than race. For Zambia, WVS/EVS respondents speak 18 separate languages, the largest of which (Bemba) represents 36.6% of the respondents. The opposite problem exists in Latin America, where language is not usually used as a measure of ethnic affiliation, and race defines ethnic identity instead. For instance, in Venezuela 100% of respondents report speaking Castilian. However the largest racial group is coded as "Colored (light)", representing 42.7% of respondents.

To correctly characterize ethnic identity in a systematic way, we again rely on existing classifications rather than on our own judgement. We examine the ethnic and linguistic classifications in the integrated WVS/EVS file and see which one is closest to either the Alesina et al. (2003) or the Fearon (2003) classifications, which are widely used in the literature.[2] We choose either ethnic identity or linguistic identity depending on which one gives us a classification and a distribution of individuals across groups that most resembles the Alesina et al. and Fearon classifications. In the above example, ethnic identity in Zambia is coded using the language spoken at home variable, while ethnic identity in Venezuela is coded as the ethnic group to which a respondent belongs. The idea is that a measure of ethnolinguistic fractionalization computed from the resulting group shares in the WVS/EVS dataset should be highly correlated with fractionalization measures derived in Alesina et al. (2003), and Fearon (2003). Indeed our ethnic classification results in fractionalization measures that are 74% correlated with fractionalization in Alesina et al., and 73% correlated with fractionalization according to Fearon - this despite the data coming from very different sources (a survey for WVS/EVS, mostly census for the other two sources). Finally, control variables in the WVS/EVS dataset consist of the respondent's age (question x003), sex (x001), highest educational level attained (x025) and household income (x047).

2.3 Results

The results for the WVS/EVS dataset are presented in Tables 1 and 2. Table 1 presents the overall share of regressions where ethnicity dummies are jointly significant at the 5% level and the R^2 with and without ethnicity dummies, breaking down these results by region. Table 2 presents a breakdown by question category (using the classification of questions provided by the WVS/EVS) and by question type (binary, scale, and binary constructed from multiple response questions). Additionally, Appendix Table A1 presents the results country by country.

Interesting findings emerge. First, the average number of questions for which ethnicity dummies are

[2]The WVS/EVS question on ethnic group is question x051 while the language spoken at home question is g016. These are the two questions we use to code a respondent's ethnic identity.

jointly significant, across all countries, is 43%. Thus, ethnic identity is an important determinant of responses to many questions.

Second, this average masks interesting variation across regions. In South Asia, East Asia and Sub-Saharan Africa, the shares are much higher, respectively 67%, 63% and 62%. In Latin America and Western Europe, the shares are much lower, at 17% and 31% respectively. The small share in Latin America could be due to the fact that, despite racial heterogeneity, linguistic and religious identity in Latin America is much more homogeneous than in places where ethnic identity is a stronger predictor of culture, for instance Africa. The Latin American exception does not extend to the New World as a whole, as North America (defined here as Canada and the US) displays a relatively high share (51%). The results for Latin America and Sub-Saharan Africa are confirmed when analyzing alternative datasets for these regions - Latinobarómetro and Afrobarometer, respectively (these results appear in Appendix 1).

Third, the breakdown by question category shows little variation. We find that ethnic identity matters a bit more for questions pertaining to religion and morals, as well as (predictably) for those pertaining to national identity, and a bit less for questions related to work. Otherwise, there is substantial homogeneity across categories. We conducted the same breakdown by question category continent by continent, finding again little variation in the share of regressions with significant ethnic dummies. These findings suggest that the choice of questions is not very material to the issue of whether ethnic identity affects norms, values and preferences, as regional patterns are stable across question categories.[3]

Fourth, the explanatory power of the regressions is usually quite low. The average R^2 when excluding the ethnicity dummies is only 2.7%, and when including the ethnicity dummies it rises to 4.1%. Thus, it is usually difficult to predict a person's response to WVS/EVS questions using the most obvious observables, yet the addition of ethnic dummies does increase the explanatory power of the regression by about 50%. These averages again mask interesting heterogeneity across regions, which largely mirrors heterogeneity in the share of significant joint F-tests across countries. Moreover, these results suggest that the extent to which ethnic identity can explain cultural attitudes is a small share of the overall cultural variance, a theme to which we will return at length below.

3 Measuring Heterogeneity

This section is about measurement. We present a simple model of social antagonism to guide the choice of functional forms for the heterogeneity measures used in the empirical investigation that follows. Starting from various assumptions about the source of heterogeneity giving rise to antagonism, the model yields

[3] Similarly, we find little variation across types of questions - binary, scale or binary constructed from unordered response questions. Ethnicity predicts answers to scale questions slightly more frequently than for binary questions, but the difference is not large. This again suggests that the specific choice of questions is not very material to our results.

indices of ethnic diversity, cultural diversity and their overlap: χ^2. We also propose a closely related index, F_{ST}. We then show how to operationalize these theoretical indices using data.

In a nutshell, we assume that individuals feel antagonism towards people who are different from them. *Social antagonism* is the sum of all the individual levels of antagonism in society, as in the alienation framework of Esteban and Ray (1994). We adopt a broad interpretation of what antagonism captures. It could represent feelings of alienation felt toward groups with different cultures or different ethnicities. Antagonism could also stem from barriers that prevent fruitful interactions between groups, for instance due to an inability to communicate or trust each other. We consider three distinct types of societies depending on how various dimensions of heterogeneity give rise to antagonism. For each type of society we derive an index measuring the level of social antagonism. Later, we will calculate these indices and relate them to the probability of civil conflict.

It is useful to start with some notation. A country is composed of n individuals characterized by the ethnic group to which they belong and by their cultural values or preferences. There are S ethnic groups, indexed by $s = 1, ..., S$. The share of each ethnic group in the population is w^s. Cultural values (or traits) are the answers to the q questions in the WVS (or any other survey of cultural attitudes), each indexed by $i = 1, ..., q$. Each question i has $r(i)$ different possible answers, indexed by j. Focusing on a given country, w_{ij} is the share of the population that gives answer j to question i. Finally, w_{ij}^s is the share of individuals from ethnic group s that gives answer j to question i.

The type of an individual, k, is given by his ethnic group s and his answers to the q questions. We define a vector ω_k of dimension $1+q$ where the first component is a number from $\{1, 2, ..., S\}$ and denotes his ethnic group, and the remaining q components represent the answers to each of the q questions. For example, if there are two ethnic groups, $S = 2$, and three questions, $q = 3$, and each question has two answers, $r(i) = 2$, the vector $\omega_1 = \{1, 2, 1, 1\}$ characterizes the type of an individual (i.e. type 1) who belongs to the first ethnic group and who gives answers $2, 1, 1$ to the first, second and third question, respectively. Since we have a finite number of individuals, n, as well as a finite number of questions and answers, the total number of different types of individuals is finite. We denote by K the number of different types and by n_k the number of individuals of type k, so $\sum_{k=1}^{K} n_k = n$. The population share of individuals of type k is denoted by $w_k = n_k/n$, where of course $\sum_{k=1}^{K} w_k = 1$. We denote as $\xi(k, i)$ the answer given by an agent of type k to question i, and as $s(k)$ the ethnic group of a type k agent: $\omega_k = (s(k), \xi(k, 1), \xi(k, 2), ..., \xi(k, q))$.

3.1 The Cultural Heterogeneity Channel

We first assume that only cultural values matter for antagonism. Belonging to a different ethnic group s does not generate any antagonism *per se*. An individual's antagonism is given by the share of individuals in society with preferences different from his. More formally, the antagonism of an agent of type k depends,

for each of the q questions, on how many people respond in the same way as him. The population share of individuals that give the same answer to question i as agent of type k is $w_{i,\xi(k,i)}$.

We give the same weight to all the q questions. Thus, for an agent of type k the average share, over the q traits, of individuals giving the same answer as him is $\frac{1}{q}\sum_{i=1}^{q} w_{i,\xi(k,i)}$. Hence, his level of antagonism, v_k, is given by:

$$v_k = 1 - \frac{1}{q}\sum_{i=1}^{q} w_{i,\xi(k,i)} \tag{2}$$

Here, individuals feel antagonism if they live in the same society as other individuals who have different cultural characteristics. Ethnicity does not matter. In this case v_k measures the average probability, over all questions, that a randomly chosen citizen disagrees with the answers given by an agent of type k. Social antagonism v is the summation of all the individual levels of antagonism v_k, normalized by the population size n. Appendix 2A shows that v can be rewritten as the following easy-to-calculate index of cultural fractionalization (CF):

$$CF = \frac{1}{q}\sum_{i=1}^{q}\left(1 - \sum_{j=1}^{r(i)} w_{ij}^2\right). \tag{3}$$

The cultural fractionalization (CF) index measures the average probability that two randomly drawn individuals from a population give different answers to a randomly drawn question from the WVS/EVS. Thus, if we believe that antagonism is driven exclusively by differences in preferences and cultural values, the index of cultural fractionalization, CF, should matter for political economy outcomes.

3.2 The Ethnic Heterogeneity Channel

Alternatively, we assume that antagonism stems only from ethnic differences, not from cultural differences. This antagonism could come from animosity *vis-à-vis* other ethnic groups (racial hatred and prejudice) or from barriers that impede interactions between groups because of lack of communication or trust. The probability that a randomly chosen individual belongs to the ethnic group $s(k)$ is $w^{s(k)}$. We postulate that in this society the level of antagonism of an individual of type k is:

$$v_k = 1 - w^{s(k)}. \tag{4}$$

Thus, under this assumption individual antagonism is just the probability that a person meets or is matched with another person from a different ethnic group. Social antagonism, v, is the average of this probability over all individuals. Appendix 2B shows that v in this case is just the common ELF index of ethnic fractionalization:

$$ELF = 1 - \sum_{s=1}^{S}(w^s)^2. \tag{5}$$

Thus, if we believe that antagonism is driven purely by ethnic animosity or barriers between ethnic groups, without any role for cultural differences, the conventional index of ethnolinguistic fractionalization, ELF,

should matter for socioeconomic outcomes. Such is the case in research where antagonism can stem from ethnic differences *per se* rather than any underlying cultural differences.

3.3 The Overlap Channel

3.3.1 Deriving a Measure of Overlap Between Ethnicity and Culture

As a third alternative, we assume that an individual's antagonism depends on how culturally different her group is from other ethnic groups. An individual does not experience any antagonism if people from other ethnic groups answer the questions in the WVS in the same way as people in her own ethnic group. Ethnicity only matters if ethnic groups differ in their cultural values. In this society cultural differences between the members of the same ethnic group do not increase the level of antagonism.

Take agent of type k and question i. Suppose first that type k only interacts with agents of her own ethnic group $s(k)$. Remember that we denote by $\xi(k,i)$ the answer that agent of type k gives to question i. In this case, by definition the share of people within group $s(k)$ with the same answer to question i as agent of type k is $w^{s(k)}_{i,\xi(k,i)}$. In other words, this is the probability that a randomly chosen agent from the ethnic group $s(k)$ agrees with an agent of type k on question i.

Now assume that an agent of type k is equally likely to interact with anybody in society. In this case the probability an agent of type k agrees on question i with a randomly chosen individual in society is $w_{i,\xi(k,i)}$. If the probability $w_{i,\xi(k,i)}$ is equal to $w^{s(k)}_{i,\xi(k,i)}$, an agent of type k does not see any difference between her own ethnic group and society overall.[4] However, if the proportion of people in society overall answering $\xi(k,i)$ is lower than the corresponding proportion within her own ethnic group, the agent experiences antagonism. In particular we assume that antagonism for question i and an agent of type k, v_{ik} depends on the (relative) difference between these two shares:

$$v_{ik} = \frac{w^{s(k)}_{i,\xi(k,i)} - w_{i,\xi(k,i)}}{w_{i,\xi(k,i)}} \tag{6}$$

Notice that if $w^{s(k)}_{i,\xi(k,i)} < w_{i,\xi(k,i)}$ the individual experiences *negative antagonism*, i.e. she is happy to interact with people in society who give the same answer as she does in greater proportion than people in her own group. Suppose that I trust people, and that 50% of those in my ethnic group trust people. I feel antagonism toward the rest of society if the share of people in the rest of society that trust people is 20%, but I am quite happy if the share of people in the rest of society that trust people is 60%.

Averaging v_{ik} over all the q questions, giving the same weight to all of them, individual antagonism is:

$$v_k = \frac{1}{q}\sum_{i=1}^{q} \frac{w^{s(k)}_{i,\xi(k,i)} - w_{i,\xi(k,i)}}{w_{i,\xi(k,i)}} \tag{7}$$

[4] Here we assume that the agent takes all other ethnic groups as being a unique group.

Notice that if all ethnic groups are identical, i.e., if for each question the distribution of answers is independent of the distribution of ethnic groups, we have that $v_k = 0$. We add up the individual levels of antagonism across k, weighing by the population shares of each type k, to obtain social antagonism:

$$v = \frac{\sum_{k=1}^{K} v_k n_k}{n} = \frac{\sum_{k=1}^{K} \frac{1}{q} \sum_{i=1}^{q} \frac{w_{i,\xi(k,i)}^{s(k)} - w_{i,\xi(k,i)}}{w_{i,\xi(k,i)}} n_k}{n} \tag{8}$$

Again, if the distribution of answers within each group is the same as the distribution of answers in society overall, $v = 0$. If, on the contrary, culture and ethnicity overlap strongly, then v will be large.

To operationalize the v measure as a measure that can be calculated from data, Appendix 2C shows that it can be rewritten as:

$$\chi^2 = \frac{1}{q} \sum_{i=1}^{q} \sum_{s=1}^{S} \sum_{j=1}^{r(i)} \frac{w^s (w_{ij} - w_{ij}^s)^2}{w_{ij}} \tag{9}$$

Thus, if we believe that antagonism is driven purely by differences in culture across ethnic groups, we should observe a relationship between the χ^2 index of overlap and socioeconomic outcomes. Defining and using this index to measure the overlap between culture and ethnicity is an important contribution of this paper.

3.3.2 Heuristic Discussion of the χ^2 Index

To complement the discussion above, it is useful to give a heuristic sense of the meaning of the χ^2 index. χ^2 is based on comparing the distribution of average answers for a given group to the distribution of answers in the overall population. If the distribution of answers in a given ethnic group is exactly the same as in the entire population, then knowing a person's ethnic identity conveys no information about his cultural attributes. If instead the distributions are distinct, then there is overlap between ethnic identity and cultural attributes.

To measure the overlap between ethnolinguistic diversity and preference diversity we can compare the distribution of answers across groups. This is what the χ^2 accomplishes.[5] Let n_{ij}^s be the number of individuals who belong to ethnic group s and give answer j to question i. We write $n_i^s = \{n_{i1}^s, n_{i2}^s, ..., n_{ir(i)}^s\}$. Under independence, the expected number of individuals that belong to ethnic group s and give answer j to question i should be $w_{ij} n^s$, while the observed frequency is n_{ij}^s. The χ^2 index is based on the difference between the observed number of individuals of an ethnic group s that give answer j and the corresponding expected number of individuals under the assumption of independence between ethnicity and answers. So for each question i:

$$\chi_i^2 = \sum_{s=1}^{S} \sum_{j=1}^{r(i)} \frac{(w_{ij} n^s - n_{ij}^s)^2}{w_{ij} n^s} \tag{10}$$

[5] For previous uses of this index, see Selway (2010) who examines the overlap between religion and ethnicity and Alesina and Zhuravskaya (2011) who analyze the overlap between geography and ethnicity as a way of measuring segregation.

The value of χ_i^2 depends on the group sample sizes n^s. Since different countries have different sample sizes and we want to compare different values of χ_i^2 across countries, it is better to work from group shares than from the number of individuals in each group. Thus, we can divide the χ_i^2 index by n to obtain a normalized index:

$$\phi_i^2 = \sum_{s=1}^{S} \sum_{j=1}^{r(i)} \frac{n^s (w_{ij} - w_{ij}^s)^2}{n w_{ij}} = \sum_{s=1}^{S} \sum_{j=1}^{r(i)} \frac{w^s (w_{ij} - w_{ij}^s)^2}{w_{ij}} \tag{11}$$

where $w^s = n^s/n$.

If we combine the ϕ_i^2 from the different questions, we obtain the Chi-squared index derived above, χ^2:[6]

$$\chi^2 = \frac{1}{q} \sum_{i=1}^{q} \phi_i^2 = \frac{1}{q} \sum_{i=1}^{q} \sum_{s=1}^{S} \sum_{j=1}^{r(i)} \frac{w^s (w_{ij} - w_{ij}^s)^2}{w_{ij}} \tag{12}$$

Thus, the χ^2 index depends on the average difference between the observed shares w_{ij}^s and the expected shares w_{ij} that we would observe if the distribution of ethnicity and the distribution of culture were independent. This index has a minimum value of zero when there is no overlap. The maximum value depends on the number of ethnic groups, S, and the number of answers of each question, $r(i)$.[7]

3.3.3 An Alternative: The Fixation Index or F_{ST}

An alternative to the χ^2 index is F_{ST}, an index commonly used in population genetics to measure genetic differentiation or distance between groups (see Wright, 1949, and Nei, 1973).[8] In genetics, F_{ST} is a measure of relative heterogeneity: it is the ratio of between-group heterogeneity in genetic characteristics to total heterogeneity. Analogously, here we compute a cultural F_{ST} - the ratio of between-group cultural heterogeneity to total heterogeneity: when F_{ST} is 0, ethnic identity conveys no information about cultural attitudes, norms and values. In contrast, if F_{ST} is equal to 1, knowing someone's ethnolinguistic identity allows a perfect prediction of their cultural attributes. F_{ST} is therefore a measure of overlap between

[6] Strictly speaking, the index we use is ϕ^2 not χ^2. However, for simplicity we will use the term χ^2 in referring to the index based on population shares rather than number of individuals.

[7] An alternative to this index is Cramér's V, which is defined as $V = \sqrt{\frac{\chi^2}{n\ t}}$, where t is the smaller of the number of ethnic groups minus one and the number of answers minus one (Cramér, 1946). The index V is always between zero and one, and corrects for the different number of ethnic groups or answers. However, this index is hard to interpret as it does not derive from our model. Another alternative to χ^2 is the Mutual Information Index, originating from the concept of entropy in information theory. In our case this is a measure of the amount of information that ethnicity contains about values, i.e., the reduction in the uncertainty about how an individual answers the questions due to the knowledge of his ethnicity. It can be shown that χ^2 is up to an order of approximation equal to the Mutual Information Index (Cover and Thomas, 2006, p. 400).

[8] Another alternative is the more conventional index of cross-cuttingness used in the political science literature (Rae and Taylor, 1970, Selway, 2011). Appendix 3 provides a detailed discussion of this index and its relationship with χ^2 and F_{ST}.

cultural values and ethnolinguistic identity. An added advantage of F_{ST} is that it relates neatly to the already described measures of cultural diversity in terms of functional form - namely, it isolates the part of the variation in overall cultural diversity that occurs between groups.

Definition of F_{ST}. To define F_{ST}, we start from the probability that two randomly drawn individuals from ethnic group s give a different answer to question i (the within-group cultural diversity of group s):

$$CF_i^s = 1 - \sum_{j=1}^{r(i)} \left(w_{ij}^s\right)^2 \tag{13}$$

The population-weighted average of the within-group cultural fractionalization for question i can be written as:

$$CF_i^W = \sum_{s=1}^{S} w^s CF_i^s = \sum_{s=1}^{S} w^s \left(1 - \sum_{j=1}^{r(i)} \left(w_{ij}^s\right)^2\right) \tag{14}$$

The share of the total population's cultural fractionalization that is not due to within-group fractionalization for each question i is then:

$$F_{ST_i} = \frac{CF_i - CF_i^W}{CF_i} \tag{15}$$

This is, for each question, the ratio of between-group cultural fractionalization divided by total fractionalization. Averaging over all questions gives us Wright's fixation index, F_{ST} (Wright, 1949, Nei, 1973):[9]

$$F_{ST} = \frac{1}{q} \sum_{i=1}^{q} F_{ST_i} \tag{16}$$

The F_{ST} index measures the share of between-group heterogeneity in total cultural heterogeneity. If all ethnic groups were as heterogeneous as the total population, F_{ST} would be equal to 0, and there would be no between-group heterogeneity. In that case, cultural cleavages and ethnolinguistic cleavages cross-cut. Knowing someone's ethnolinguistic identity would give no information about his preferences or culture. Instead if all ethnic groups were to be homogeneous ($CF_i^s = 0$), F_{ST} would be equal to 1, and all heterogeneity would be between groups. In that case, cultural cleavages and ethnolinguistic cleavages would be reinforcing.

The advantage of F_{ST} is that it is well-known and captures intuitively a simple concept, as it represents how much one can predict answers to questions on norms, attitudes and preferences simply by knowing a respondent's ethnolinguistic identity. In the case of two ethnic groups and one question with only two possible answers this index ranges from 0 to 1. With two groups and more than two possible answers, or more generally when the number of answers exceeds the number of groups, there is always some within-group fractionalization and the index cannot reach 1.

[9] There are of course many ways to average across questions. For instance, Cavalli-Sforza et al. (1994) separately average the numerator and the denominator of equation (15), and then take the ratio. We adopt the simpler method of averaging the question by question F_{ST}^i.

Drawbacks of F_{ST}. While the F_{ST} index is very commonly used in population genetics, it does have some drawbacks, as explained for example in Jost (2008), Meirmans and Hedrick (2011) and Jakobsson et al. (2013). The most important drawback, outlined in Jost (2008), relates to the properties of the decomposition of within and between fractionalization in a context where these measures are bounded above by 1. To illustrate this potential problem, let us denote between-group fractionalization by D_i for question i. Such between-group fractionalization is defined by subtracting within-group fractionalization CF_i^W from total fractionalization CF_i, i.e., $D_i \equiv CF_i - CF_i^W$. Thus, this approach relies upon the additive decomposition of total fractionalization, but CF_i^W and D_i are not independent because we always have that $D_i + CF_i^W \leq 1$.[10] This constraint implies that D_i declines with within-group fractionalization CF_i^W regardless of the degree of cultural differentiation of ethnic groups.

A numerical example is useful to illustrate this drawback. Suppose that there is just one question and two ethnic groups of the same size. The question has four possible answers, a, b, c and d. The vector of answers for individuals from the first ethnic group is $\{0.1, 0.9, 0, 0\}$, i.e., 10% of them answer a, and 90% answer b. For the second ethnic group the vector of answers is $\{0.9, 0.1, 0, 0\}$. It is easy to check that in this society, $F_{ST} = 0.64$. Suppose a second society where those two vectors of answers are $\{0.5, 0.5, 0, 0\}$ and $\{0, 0, 0.5, 0.5\dot{}\}$. It is clear that in this society culture and ethnicity overlap more strongly than in the first society. However, in this case we have $F_{ST} = 0.33$.[11] The reason is that the second society displays a much higher degree of within-group heterogeneity than the first (a high CF^W), which drives down F_{ST} in spite of the higher degree of overlap between culture and ethnicity.

Our first overlap measure, χ^2, is not subject to this drawback, but as we will see empirically, it does not matter which index we use: while the χ^2 index comes out directly from our model of antagonism, empirically χ^2 and F_{ST} are almost perfectly correlated (in our sample the correlation is 97%, a result that also holds in simple simulations).[12]

Uses of F_{ST} in the past literature. We conclude this conceptual section by discussing a few papers that have used F_{ST} to measure between-group cultural heterogeneity, noting that their goals and methods are very different from ours. Bell et al. (2009) study inter-group competition and analyze whether there is more scope for selection based on cultural traits rather than on genetic traits. They use the WVS to compute a cultural F_{ST} measure between 150 pairs of neighboring countries. They show that this measure is an order of magnitude larger than an analogous measure of F_{ST} based on genetic data, suggesting a

[10] Jost (2008, pp 4018) provides a complete explanation of this constraint and its implication: "Additive partitioning of heterozygosity does not produce pure within-and between-subpopulation components; it is an incomplete partitioning". In our case cultural fractionalization is a parallel concept to heterozygosity in population genetics.

[11] A similar numerical example appears in Wright (1978).

[12] In the case of a question with two possible answers F_{ST} and χ^2 coincide exactly (see Workman and Niswander, 1970). Details of our simulations are available upon request.

greater scope for cultural rather than genetic selection. In contrast to our approach, they measure cultural heterogeneity between countries rather than between groups within countries. In another paper, Ross et al. (2013) compute a measure of cultural F_{ST} based on between-group variation in folktales across different European ethnic groups. Again they are interested in comparing patterns of cultural F_{ST} to those of genetic F_{ST}, finding some similarities between the two. In contrast to our work, neither of these papers is interested in using cultural F_{ST} to measure the degree of overlap between ethnicity and culture, or in understanding how cultural F_{ST} relates to overall cultural heterogeneity and overall ethnolinguistic diversity. Instead, their focus is on the importance of cultural F_{ST}, relative to genetic F_{ST}. Finally, Spolaore and Wacziarg (2009, 2013) use a genetic F_{ST} as a measure of intergenerational divergence in a wide range of human traits transmitted culturally or biologically, in order to estimate the effects of barriers between populations on political and economic outcomes. In contrast to our approach, their F_{ST} is based on neutral genes, not cultural attitudes, and it measures distance between countries rather than between ethnic groups within countries.

4 Ethnic Heterogeneity and Cultural Diversity

In this section we empirically calculate the measures of heterogeneity derived in Section 3, and describe their properties and correlates. We show that, contrary to the assumption of much of the past literature, measures of ethnic diversity and cultural diversity are uncorrelated with each other. At the same time, we know from Section 2 that ethnic identity does help to predict cultural attitudes. To reconcile these seemingly contradictory results, we show that although between-group heterogeneity in cultural attitudes is small compared to total heterogeneity, the overlap between culture and ethnicity shows considerable variation across countries, variation that mirrors that found in Section 2. We explore the correlates of these new measures, uncovering interesting patterns concerning characteristics of countries with a high degree of cultural diversity as well as those with a relatively high degree of overlap between culture and ethnicity. Finally, we examine the robustness of our results to calculating the various measures for different question categories and types.

4.1 Cultural Diversity and Ethnolinguistic Fractionalization

Before describing the indices, some comments on the data are in order. First, we use the same baseline set of questions from the integrated WVS/EVS dataset as in Section 2. Second, not all WVS/EVS questions are asked in all countries. In our benchmark analysis we drop questions that are not asked in at least 50 countries, to ensure cross-country comparability of the indices.[13] Third, since we are interested in

[13] The list of questions used to compute the heterogeneity measures, as well as their breakdown by question category and type, appears in Appendix Table A6.

relating cultural fractionalization to ethnolinguistic fractionalization, we focus exclusively on countries in the WVS/EVS for which we have ethnolinguistic information.[14] Taken together, this gives us information on 76 countries. Finally, there is no need here to convert questions that admit multiple unordered answers into series of binary questions to calculate our various indices.

Figure 1 shows a world map with the values of cultural heterogeneity in the 76 countries in our sample, and Panel A in Table 3 displays some simple summary statistics.[15] Darker-colored countries are more culturally diverse than lighter-colored ones. The most culturally diverse country is Zambia ($CF = 0.602$), and the least culturally diverse country is Jordan ($CF = 0.427$). Other interesting data points are France and India, with relatively high cultural heterogeneity, and Egypt, Indonesia and China, with relatively low numbers (a high degree of cultural conformism). The average value of CF across countries is 0.529.

The standard assumption in the literature is that cultural heterogeneity (CF) should be highly correlated with ethnolinguistic heterogeneity (ELF). Comparing the map of ELF in Figure 2 with the one of CF in Figure 1, it becomes immediately obvious that there are important differences. Countries such as Pakistan and Egypt have high levels of ethnolinguistic heterogeneity but low levels of cultural heterogeneity. At the other extreme are countries such as Germany and South Korea, which are ethnolinguistically fairly homogeneous but culturally diverse. The lack of a relationship between both types of heterogeneity is not limited to these few examples. The correlation between CF and ELF, displayed in Panel B of Table 3, is essentially zero, -0.030 to be exact. Ethnolinguistic diversity is therefore not associated with cultural diversity.

If cultural heterogeneity is not correlated with ethnolinguistic heterogeneity, what might it be related to? Table 4 shows the correlation between cultural fractionalization and several variables. Five correlates stand out: countries with a higher proportion of Muslims exhibit a lower CF (the correlation is -0.597); partly reflecting the previous correlation, countries located in North Africa and the Middle East show a similar negative correlation (-0.529); more religiously diverse countries are also culturally more diverse (correlation of 0.314); more democratic countries (measured by the Polity 4 index) have higher cultural fractionalization (the correlation is 0.586); and countries with a high per capita income exhibit a greater CF (the correlation is 0.382). The other correlations are mostly small and statistically insignificant. Countries in North America, Europe and Central Asia, and countries farther away from the equator, have, on average, greater cultural fractionalization.

[14] In principle we could of course use data on ethnolinguistic fractionalization from sources other than the WVS. We refrain from doing here so for two reasons. First, using the same data source makes the two indices, cultural and ethnolinguistic fractionalization, more easily comparable. Second, we later analyze the overlap between culture and ethnicity. For that we need to use ethnolinguistic identity and cultural values at the individual level, from the same source.

[15] Table A5 in the Appendix presents the underlying values country by country.

4.2 The Overlap Measures

Figure 3 shows a map of the χ^2 index for all countries in our database, and Table 3 Panel A reports summary statistics. Several observations are in order. First, χ^2 takes on low average values, indicating that the relative difference between the within-group heterogeneity and the overall heterogeneity is small (the mean value of χ^2 in our sample of 76 countries is 0.029). However, there is substantial variation in χ^2, with Asia (especially South Asia and Southeast Asia) and Sub-Saharan Africa displaying high values, while Europe, Russia and Latin America display relatively low values. Notable data points with high χ^2 values include India, Thailand and Zambia. Countries with low values include Japan, Russia, Poland and Italy. These patterns closely mirror those uncovered in Section 2. The regions where ethnicity could significantly predict responses to survey questions about values, norms and preference are the same regions where the overlap measures take on higher values.[16] These patterns help to explain why cultural diversity and ethnic diversity are uncorrelated, even though ethnic identity helps predict a large share of answers to questions on cultural attitudes: most of the heterogeneity is within groups.

Second, the ranking of countries is very similar across both the χ^2 and F_{ST} measures (Table 3, Panel B). The correlation between the two is 0.981 (and so is the Spearman rank correlation). This gives us confidence, despite very different functional forms, that these two measures capture common features of the data regarding the overlap between ethnicity and culture. In particular, the drawback of F_{ST} identified above does not seem very relevant empirically, since χ^2 and F_{ST} are very highly correlated, and χ^2 is not subject to the drawback. As with χ^2, the mean value of F_{ST} is low: the share of between-group variance in cultural attitudes relative to the overall variation is 0.012. A similar result is well-known in population genetics, where within-group variation in genetic characteristics swamps between-group variation (Cavalli-Sforza et al., 1994).

Third, although the overlap values are small, one could ask the question: "small relative to what?". To compare these values to a benchmark, we recompute χ^2 and F_{ST}, but now take the groups to be the different countries, rather than the different ethnicities within countries. We find values for χ^2 and F_{ST} of 0.162 and 0.074. These numbers are about six times larger than the ones we found before. Focusing on the F_{ST} measure, 7.4 percent of the cultural heterogeneity in the world is between countries, whereas only 1.2 percent is between ethnic groups within countries.[17] Even the country with the highest between-ethnic

[16] In fact, if you consider, for each country in our sample, the share of WVS/EVS questions for which ethnicity dummies are jointly significant predictors of individual responses, and correlate this share across countries with our measures of χ^2 and F_{ST}, you obtain correlations of 0.73 and 0.77, respectively. These high correlations exist despite the vastly different methodologies used in Section 2 and Section 4 to capture the degree of overlap between culture and ethnicity.

[17] The results are consistent with the average cultural F_{ST} across neighboring countries reported by Bell, Richerson and McElreath (2009). In fact, their reported F_{ST} (0.08) is remarkably close to ours (0.074). Our results are also in line with those in Fischer and Schwartz (2010), where the authors also use surveys of values to analyze the variability of answers both within and across nations using the Interclass Correlation Index, a measure closely related to F_{ST}.

group F_{ST}, India, has a lower value than the one observed between countries.

Fourth, reinforcing cleavages have a weak, positive correlation with cultural fractionalization. From Table 3 Panel B, the correlation between χ^2 and cultural fractionalization is 0.219 (this correlation is statistically significant at the 10% level). The corresponding number for F_{ST} is 0.179 (this correlation is not statistically significant at the 10% level). An example of this positive correlation is Malaysia, a country that is culturally heterogeneous ($CF = 0.563$) and where knowing someone's identity is relatively informative about that individual's culture ($\chi^2 = 0.092$). But other examples show the lack of a strong relationship. Germany is culturally heterogeneous ($CF = 0.576$), but has a low χ^2 ($\chi^2 = 0.009$).

Fifth, as expected, reinforcing cleavages are stronger in countries that are more ethnically diverse. The correlation between χ^2 and ethnolinguistic fractionalization is 0.620 (statistically significant at 1%); and the corresponding figure for F_{ST} is the same. This positive correlation can also be perceived by comparing Figure 3 with Figure 2. Ethnolinguistically diverse countries such as India, Philippines or Ethiopia also tend to have high levels of χ^2 or F_{ST}.

Finally, we investigate the correlates of our overlap measures. Table 4 displays the quantitative magnitudes of the simple correlations of a set of country characteristics and our two overlap measures. We find interesting descriptive patterns. Consistent with results in Section 2, χ^2 is higher in South Asia, East Asia and Sub-Saharan Africa, and it is lower in Latin America. χ^2 is also higher in countries with English legal origins, partly reflecting these spatial patterns. Per capita income is negatively associated with χ^2, as is latitude, indicating that poorer countries tend to display more overlap between culture and ethnicity. Thus, economic development severs the link between ethnicity and cultural values. Among variables describing the prevalence of various religions, only the percentage of Catholics is significantly (and negatively) correlated with χ^2, although it is hard to disentangle this correlation from the fact that χ^2 is lower in Latin America than elsewhere.[18]

4.3 Breakdown of the Measures by Question Category and Type

One concern with our exercise so far is that we are calculating our measures on a broad and diverse set of questions reflecting possibly different facets of culture. This was done to avoid arbitrarily picking questions from which to calculate our measures, instead letting the World Values Survey itself define the universe of values, preferences and norms under consideration. However, some questions may have to do with preferences over public policies, others with attitudes toward family, others with social trust, etc. The patterns identified above could be very different when looking at different categories of questions. For instance, the degree of correlation between CF and ELF might be higher for some question categories.

[18] One might think that our overlap measures may be picking up segregation: countries with higher levels of segregation may have a greater overlap between culture and ethnicity. Using the data on ethnic, linguistic and religious segregation from Alesina and Zhuravskaya (2011), we indeed find positive correlations, but they are not statistically significant.

Similarly, the ratio of between-group heterogeneity to overall heterogeneity could vary. We are also concerned about the degree to which each question category yields measures that are correlated across categories: for instance, is the degree of overlap between ethnicity and answers to questions pertaining to identity highly correlated with overlap measures calculated for other question categories? If this is the case, we can be more certain that results based on the broadest set of questions is representative of a general pattern. If not, we can learn interesting facts about various dimensions of heterogeneity across question categories.

We therefore calculated our measures separately for each of the question categories identified by the WVS/EVS (these are labeled A through G), and across question types (binary, scale, and unordered response questions). The results are reported in Panels C through G of Table 3. In panel C we find, on average, a higher degree of cultural heterogeneity (CF) for questions related to public policy such as relating to the environment and on politics and society (categories B and E), and lower heterogeneity for questions relating to perceptions of life (category A). Reflecting results in Section 2, we also find a higher degree of overlap between ethnicity and culture (χ^2) for questions relating to religion and morale (category F) and national identity (category G), relative to other categories. However, F_{ST} and χ^2 continue to take on small magnitudes; for instance F_{ST} is comprised in a tight band between 1% and 1.7% across categories. Despite these differences, there is a remarkable degree of consistency in the magnitudes of our heterogeneity measures across categories. Turning to correlations across categories, in Panel D we find that there is substantial heterogeneity in patterns of CF obtained from different question categories. The correlations are usually positive, except for question categories with few questions (such as category G) where the measures could be more volatile due to imperfect measurement. On the other hand, there is a much more substantial degree of correlation of χ^2 across categories, with correlations usually in the $0.7 - 0.9$ range. To summarize, CF displays patterns that differ across categories, χ^2 less so.

We also examined the correlations between CF, ELF, F_{ST} and χ^2 category by category. The correlation between ELF and CF, which was zero when CF was calculated across all questions, differs across categories. The correlation remains low in magnitude, varying between -0.26 (category F) and 0.30 (category A), yet we do find a weak positive correlation between ethnic and cultural heterogeneity measures for categories of questions reflecting perceptions of life and politics and society. All the other correlations are negative, contrary to the view that ethnic heterogeneity "captures" cultural heterogeneity. We also find that χ^2 and F_{ST} continue to be highly correlated with each other, category by category.

In addition to looking at different question categories, we also analyzed different question types - i.e. whether a question admits binary answers, answers on an ordered scale, or multiple unordered responses. Panel C reports some level differences in CF and χ^2 across these types. Cultural fractionalization is higher for unordered response questions and scale questions, compared to binary questions. This is not surprising: when given more possible answers, measured heterogeneity tends to increase. For the overlap measures, the F_{ST} are very similar across question types (on the order of 1%), whereas the χ^2 measures

are higher for scale and unordered response questions. Overall, focusing on binary questions has several advantages. First, when questions have only two answers, F_{ST} and χ^2 are identical.[19] Second, with scaled questions one could argue that the distance between possible answers should be taken into account when calculating heterogeneity metrics - no such concern exists for questions that have only two possible answers. Third, there is some advantage in focusing exclusively on a subset of questions with the same number of possible answers, since this avoids heterogeneity varying simply because of differences in the number of answers. Thankfully, our results are robust to focusing only on binary questions. For instance, Panels F and G of Table 3 show that the correlation between ELF and CF is close to zero no matter the question type. Moreover, the correlation between χ^2 and F_{ST} is above 0.98 for the scale and unordered response questions, so the high correlation between our two overlap measures is not driven simply by the inclusion of binary questions, for which they are identical.

Since we are interested in the general relation between culture and ethnicity, it is sensible not to cherry-pick questions. However, many papers on the cultural determinants of economic outcomes focus on a small set of questions that are deemed meaningful *a priori*, rather than taking a comprehensive approach by focusing on the broad set of questions, as we do. To assess whether our main findings would differ if we were to exclusively focus on a limited number of questions that have commonly been used in the literature, we take nine questions relating to family values, child qualities, trust and beliefs.[20] As can be seen in Panel C and Panel H of Table 3, this does not change any of the main results. In fact, it reinforces the finding that ethnolinguistic diversity is not a good proxy for cultural fractionalization: the correlation between ELF and CF is now negative, standing at -0.222. In addition, the correlation between CF, when using the limited set of questions, and CF, when using all questions, is positive and statistically significant at the 1% level. This suggests that our results are not specific to our comprehensive approach.

[19] When questions have more than two possible answers, the equivalence of F_{ST} and χ^2 breaks down (although empirically the two continue to be extremely highly correlated). In addition, when there are more than two possible answers, χ^2 is no longer bounded by 1. Normalizing χ^2 so it does not exceed 1 would require dividing it by the minimum of the number of answers and the number of ethnic groups. Since our theory provides no foundation for this normalization, we refrain from it.

[20] To be precise, the nine questions are: 1) Questions on family ties, used in Alesina and Giuliano (2013): a001 (family important in life), a025 (respect and love for parents), a026 (parents' responsibilities to their children). 2) Questions on important child qualities, used in Tabellini (2010): a030, a035, a038, a042 (hard work, tolerance and respect for other people, thrift, obedience). 3) Generalized trust, used in Algan and Cahuc (2010), among many others: a165 (most people can be trusted). 4) Belief in hell, used in Barro and McCleary (2003): f035 (belief in hell).

5 Culture, Ethnicity and Civil Conflict

In this section we examine the relationship between our various measures of diversity and civil conflict, to illustrate the possible uses of our measures. There remains a debate on whether ethnolinguistic diversity in fact affects civil conflict. Results seem to depend on which measures of diversity are used: using fractionalization, Fearon and Laitin (2003) find little evidence of an effect on conflict onset. Collier and Hoeffler (2004), using a measure of social fractionalization that combines ethnic and religious dimensions, find evidence that greater diversity reduces the probability of a civil war. In contrast, using a measure of ethnic polarization, Montalvo and Reynal-Querol (2005) find that it has a significantly positive effect on civil war incidence. Esteban, Mayoral and Ray (2012), using different measures of diversity jointly, found that they were significantly associated with civil conflict. Of particular note, a recent paper by Huber and Mayoral (2013) examines the role of income inequality between and within ethnic groups as a determinant of civil conflict, finding that within-group inequality affects conflict positively. Our paper shares with theirs a decomposition of overall differences between and within ethnic groups, and also uses survey data. However, the application is very different as their paper is interested in within- and between-ethnic group income inequality, as opposed to cultural differences.[21]

How might ethnolinguistic diversity affect civil conflict in our framework? First, ethnolinguistically heterogeneous societies may have more diverse preferences and values, leading to increased overall antagonism in society. Second, ethnolinguistic fractionalization may matter *per se* because of direct animosity, hatred or barriers between different ethnolinguistic groups. Third, civil conflict may arise more frequently when ethnic divisions and cultural differences reinforce each other. To evaluate which of these channels operates most strongly, we introduce measures of these three dimensions of heterogeneity simultaneously in regressions explaining civil conflict. If the first explanation is valid, then cultural fractionalization (CF) should affect civil conflict. If the second explanation is valid, then ethnolinguistic fractionalization (ELF) should matter. Finally, if the third explanation is valid, then the χ^2 (or F_{ST}) should matter.

5.1 Data and Specification

In our application, the aforementioned contributions to the study of civil conflict constitute the methodological starting point. Following the literature, we define a dummy variable C_{ct} equal to 1 if country c experiences a civil war in year t (to explore conflict incidence). We also define a separate dummy C^o_{ct} equal to 1 if in a given year a country experiences the onset of a new civil war (to study onset only). We relate these outcomes to our three sets of measures of diversity, CF, ELF and χ^2 (or F_{ST}):

$$C_{ct} = \beta_0 + \beta_1 CF_c + \beta_2 ELF_c + \beta_3 \chi^2_c + \boldsymbol{\beta}_4^{'} \mathbf{Z}_{ct} + \varepsilon_{ct} \tag{17}$$

[21] The functional forms are also different as they use a decomposition of the Gini index, not fractionalization.

where $\mathbf{Z}_{ct}$ is a vector of control variables commonly used in the literature. In particular, we use an expansive set of controls very close to the ones used in Fearon and Laitin (2003), Esteban, Mayoral and Ray (2012) and Desmet, Ortuño-Ortín and Wacziarg (2012). These include a variety of geographic variables, lagged per capita GDP, as well as lagged conflict, legal origins, and dummy variables for major geographic regions. These controls include most of the variables that were shown to be predictors of CF, χ^2 and F_{ST}. Hence, any effect of, for example, χ^2 on the probability of conflict is not proxying for the effects of these correlates.

The data on civil conflict and the control variables come from Fearon and Laitin (2003). In this database, a country is coded as being in a civil conflict when the conflict overall killed over $1,000$ people, with an average of at least 100 deaths a year and at least 100 deaths on both sides of the conflict. As an alternative source of data we use a database from the Peace Research Institute Oslo (PRIO), which has been used in recent contributions (for instance, Esteban, Mayoral and Ray, 2012, and Huber and Mayoral, 2013). There, a civil war is defined as "a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths." We also use a more stringent threshold of $1,000$ battle related deaths. When looking at conflict incidence, in our sample of countries, the correlation between PRIO25 and the Fearon-Laitin variable is 0.69, while the correlation between PRIO1000 and the Fearon-Laitin variable is 0.47 (the correlation between PRIO25 and PRIO1000 is 0.51).

5.2 Results

The results are reported in Tables 5 through 8. In Table 5 we examine the determinants of civil conflict incidence, introducing cultural fractionalization, ethnolinguistic fractionalization and χ^2, first individually and then jointly (columns 1 through 4). We find that, when introduced individually, χ^2 is a significant predictor of conflict incidence (at the 1% level), and this continues to be the case when all three measures are introduced jointly. In fact in that case the logit marginal effect of χ^2 becomes twice as large (column 4): it is equal to 0.870. That implies that a one standard deviation change in χ^2 (equal to 0.025) raises the probability of conflict by 2.175 percentage points. The baseline probability of being in a civil conflict is 14.173% in the sample for this regression, so the standardized effect of χ^2 amounts to about 15% of the probability of conflict. Our interpretation of this result is that ethnic divisions matter for civil conflict, but only when they overlap with cultural cleavages.

Another notable result from Table 5 is that cultural fractionalization (CF) tends to reduce the incidence of civil wars (even when controlling for several variables previously found to be correlates of CF, such as democracy, per capita income and region dummies). In column 4, the marginal effect of CF is -0.391. The standard deviation of CF is 0.037, which means that a one standard deviation increase in CF is associated with a 10% reduction in the probability of conflict. One interpretation of

this finding is that cultural diversity is the sign of a society that is tolerant of a multiplicity of values and preferences, and this tolerance reduces the incidence of civil conflict. A related interpretation is that cultural diversity is the sign of a society that embraces modernity more generally, and modernity is not fully captured by the included controls that correlate with CF. Ethnolinguistic fractionalization is insignificant, and has an unstable sign across specifications, although it tends to bear a negative sign when all measures of heterogeneity are entered together (columns 4-6). In columns (5) and (6) of Table 5 we add legal origins and GDP growth and its lag to the baseline specification, with little effect on the estimates on χ^2 and CF. None of these results change much when using probit rather than logit.[22]

Table 6 breaks down these results by question category and type. In Panel A we see that the baseline results are quite robust across categories: CF enters negatively in 6 of the 7 categories, although the overall results appear to have been driven mostly by questions on perceptions of life and religion and morale since CF based on other categories is not statistically significant. Results for χ^2 are stronger, with this variable entering with a significantly positive sign for 5 of the 7 categories (for one of the remaining categories, questions on the environment, χ^2 is based on only 4 questions, and may therefore be a noisy measure). Overall these results suggest considerable robustness in the pattern of coefficients across question category, and also allow us to determine what categories of questions are important to draw the overall inferences. In Panel B we also see robustness with respect to question type. As expected the results are strongest for binary and scale questions, but χ^2 enters positively and significantly in all three cases.

5.3 Endogeneity of CF and χ^2

As in most of the literature on civil conflict, we have so far treated our heterogeneity measures as exogenous to conflict. As long as we limit attention to ethnolinguistic fractionalization, and include a suitably expansive set of controls, this can be justified as ethnolinguistic fractionalization is very time-persistent and is unlikely to change very much as a result of conflict. The same cannot necessarily be said of cultural heterogeneity and the overlap of culture and ethnicity. The experience of civil wars can lead people to change their values and preferences, and respond differently to questions from the World Values Survey. This in turn can lead measures of cultural diversity such as CF to change as a result of conflict (though it is not clear *a priori* in what direction, hence the sign of the endogeneity bias is

[22] Moreover, Table A4 looks at the determinants of civil war *onset*. The results are very much in line with those for incidence. Of course, conflict onset is a much rarer event than conflict incidence, with the percentage of country-year observations featuring the onset of a civil war being equal to 1.780%. Given this fact, the marginal effect of χ^2 on conflict onset is found to be 0.175, implying that a one standard deviation increase in χ^2 reduces the probability of civil war onset by 25%, a sizeable effect. On the other hand, while it is still negative, the effect of cultural diversity on conflict onset is generally not significant at the 5% level. Finally, the effect of ethnolinguistic fractionalization continues to be insignificant, in line with results on civil war onset in Fearon and Laitin (2003).

not clear). Similarly, civil war, especially when there is an ethnic component, can change the salience of ethnic identity, leading ethnic groups to adopt values and attitudes that differ more than they did prior to the war, or on the contrary can lead the victor to impose their values and preferences on the vanquished, therefore affecting the degree of overlap between ethnicity and culture (again in an ambiguous direction).[23] This could be a problem particularly in our setting as the questions from the World Values Survey used to characterize preferences and values are from survey waves from 1981 to 2008, while our main civil war dataset (from Fearon and Laitin, 2003) covers 1945 to 1999. We already partly address this problem in the regressions shown so far as we adopt a dynamic specification for the incidence of civil wars, i.e., we include a term for lagged civil war on the right-hand side of the specification, in keeping with the usual practice in the literature (see in particular Fearon and Laitin, 2003, p. 84 and Esteban, Mayoral and Ray, 2012, p. 1318). Since civil war incidence is highly autocorrelated, this purges CF and χ^2 of much of their variation attributable to past wars.

To deal with any remaining endogeneity, we adopt a three-pronged approach. First, we focus on questions that display a high degree of persistence in cultural fractionalization across waves, and are thus less likely than other questions to respond endogenously to external events such as civil wars. For each question i, country c and WVS wave w, consider cultural fractionalization CF_{icw}. For each question i and each country c we compute the coefficient of variation of CF_{icw} across waves w, and average this coefficient of variation across countries for each question i. This gives a measure of persistence for each question i. We then remove from consideration every question with a coefficient of variation in excess of 0.1, which leaves us with about 60% of the questions used previously - the ones with the most time persistent value of CF_{icw}. We reran our baseline specification (the one in column 4 of Table 5) with CF and χ^2 computed from this restricted set of questions. The results are presented in column (1) of Table 7. The signs of our main effects remain the same, namely CF affects conflict negatively (albeit the effect is no longer statistically significant at the 5% level) and χ^2 affects conflict incidence positively and remains significant at the 5% level. The magnitude of the effect falls slightly in standardized terms, with a one standard deviation increase in χ^2 associated with a 10% increase in the probability of conflict.

Our second approach is to limit our sample to the post-1970 period. The idea is that if endogeneity were a strong concern, we should find different results in this subsample compared to the full sample. The argument could take various forms. On the one hand, if we limit attention to wars that occur closer to the date when we observe values, there is perhaps greater potential for recent wars to affect values, and then in turn cultural diversity and χ^2. On the other hand, if the lag with which war may affect values is substantial, by focusing on a recent sample, values may not yet have had time to change, and therefore CF and χ^2 may not yet have changed in response to civil war. In either case, if reverse causality were a concern we would observe different estimates of the effect of CF and χ^2 in the post-1970 sample and

[23] This is a prevalent theme in constructivist approaches to ethnicity.

in the whole sample. Column (3) of Table 7 presents the results, which are almost unchanged compared to the baseline regression of Column 4 of Table 5. Indeed, the standardized effect of χ^2 stands equal to the one previously calculated, at 15%. The standardized effect of CF on the probability of civil war also remains equal to -10%. These results have the added advantage of showing the stability and robustness of our estimated effects to the sample period under consideration.

Out third approach is a variation on the previous one. Here, we limit attention to respondents born before 1950 *and* to the post-1970 sample. The effects could once again go in a variety of directions, but the argument is again that the estimates would be *different* if endogeneity were a big concern. On the one hand, if respondents' values are formed in early adulthood and change little after that, since every respondent in the sample would be at least 20 years old in the event of a civil war, their cultural values may respond less than younger individuals to the event of a war. On the other hand, if one's view was that these individuals were the most likely to be affected by a civil war because they were most likely to be combatants or to be affected by the war in adult age, their values may be most likely to be affected by the war. Either way, war would affect values, and therefore potentially CF and χ^2 also (although, again, in unknown directions). Column (4) of Table 7 presents the results, but once again we find very little evidence of a different effect of cultural diversity and χ^2 on the probability of war: the standardized effects of these two variables are, respectively, -9% and 13%, close to those in column (4) of Table 5.

While we do not want to place too much weight on any one of these tests in isolation, taken together they do suggest that our main results are remarkably stable when looking only at the post-1970 sample, when looking only at respondents born before 1950, and when including only questions for which question-by-question cultural diversity CF_i is most stable across survey waves. Under reverse causality, if war had a strong causal effect on CF or χ^2, we would have expected estimates under each of these modifications to differ from the baseline. They do not differ materially.

5.4 Robustness Checks

In Table 8, we carry out a series of additional robustness tests. First, we estimate our baseline regression using F_{ST} rather than χ^2 as the measure of overlap between culture and ethnicity (column 1). The results do not change in any substantive way, as expected because F_{ST} is so highly correlated with χ^2. The standardized effect of F_{ST} is 16% while the effect of CF is -10%. Second, we change our data source for civil wars to data from PRIO. With the 25 battle deaths threshold (column 2), we find a standardized effects of χ^2 and CF that are much larger than in our baseline (respectively 27% and -20%). With the $1,000$ battle deaths threshold (not reported here) the corresponding effects stand at 14% and -10%, respectively, very similar to the baseline. We conclude that the signs and significance of the estimated effects are not sensitive to changes in the definition of a civil war.

In an additional robustness check, we focus on different sets of questions. We start by expanding

the number of questions included to calculate our measures of cultural diversity and overlap to those that were asked in at least 30 countries, rather than the more stringent criterion of 50 countries used previously. This expands the set of questions used in our calculations, at the cost of greater heterogeneity across countries in the set of questions. The results appear in column (3) of Table 8. Reassuringly, nothing changes much: the standardized effects of χ^2 and CF are largely unchanged at 17% and -9%, respectively.[24] We next go the opposite way, and focus on a small set of nine questions often used in the literature on culture and economics (the questions are those listed in footnote 20). Again, the results, reported in column (4), are unchanged. This greatly increases our confidence that our results do not depend on the set of questions used to calculate cultural diversity and overlap.

As a last robustness check, we add linguistic polarization in our baseline regression.[25] Montalvo and Reynal-Querol (2005) have argued that polarization is an important predictor of civil conflict. As can be seen in column (5), polarization is not statistically significant when all other measures of heterogeneity are also included. When comparing the coefficients on our three variables of interest to those of column (4) in Table 5, they are essentially identical. Taken together, the different robustness checks suggest that our main findings continue to stand: neither ethnolinguistic heterogeneity nor cultural heterogeneity increase the probability of conflict, but when culture and ethnicity reinforce each other, then conflict becomes more likely.

6 Conclusion

In this paper we studied the complex relationship between ethnicity and culture, defined as a vector of answers to a broad set of questions about norms, values and preferences. We uncovered novel results. First, ethnicity does serve to significantly predict cultural attitudes, to an extent that varies across geographic regions. Second, the share of variation in culture that is explained by ethnicity is very small. As a result, cultural diversity, defined as the average probability that two randomly chosen individuals respond differently to a question from the World Values Survey, is not correlated with ethnic diversity. Thus, ethnic fractionalization cannot readily be taken as a proxy for diversity in values, attitudes and preferences. Third, we derived and calculated several new indices measuring the extent of overlap between culture and ethnicity, stemming from a simple model of social antagonism. These measures display interesting geographic variation, with the degree of overlap being greatest in Sub-Saharan Africa and Asia, and smallest in Latin America. Fourth, as an application we used our new measures of cultural diversity and overlap to study the determinants of civil conflict, finding that ethnic fractionalization has no predictive power for civil conflict, but that cultural diversity has, if anything, a pacifying effect.

[24] Reassuringly, the correlation in our sample of 76 countries between CF using the 50-countries threshold and CF using the 30-countries threshold is very high - at 0.95. Similarly, the two versions of χ^2 bear a 0.99 correlation with each other.

[25] The linguistic polarization measure comes from Desmet, Ortuño-Ortín and Wacziarg (2012).

Our new measures of overlap between culture and ethnicity, χ^2 and F_{ST}, have a positive effect on the probability of civil conflict onset and incidence, indicating that ethnic divisions matter for conflict when they are associated with cultural differences across ethnic groups. Hence, we have identified the degree of overlap between culture and ethnicity as a new and robust determinant of civil conflict.

Our results parallel a famous debate in population genetics on within-group versus between-group genetic differentiation, going back to Lewontin (1972). Lewontin pointed out that between-race genetic variation was a very small part of overall variation, and that within-group diversity accounted for a much larger share of overall genetic variation. This led Lewontin to question the validity of the very concept of race. In a series of rejoinders, Edwards (2003), Dawkins (2005) and others argued that while between-group variation was small, it could still be a relevant part of the variation: humans share up to 99% of their DNA with some animals, yet the 1% that differs matters a lot to set the two groups apart. Lewontin's point on genetics mirrors our finding that between-ethnic group cultural variation is a small part of overall cultural variation, and that most of this variation occurs within-groups. Edwards' and Dawkins' argument also finds an echo in our work, since we argue that between-group variation, while a small share of the overall variation, matters greatly for civil conflict.

The question we posed here is also related to a continuing debate in the social sciences as to whether ethnic, linguistic and religious identities are "constructed" or reflect "primordial" differences between different groups of humans. Each of these traditions reflects a variety of viewpoints on the persistence of ethnic and cultural identities and a wide range of theories on the factors that gave rise to both ethnic and cultural differentiation. However, drawing a stark distinction between these two broad categories of views helps bring into focus a fundamental difference separating them: the primordialist view holds that ethnolinguistic divisions reflect deep differences between humans, the result of historical separation which allowed for cultural drift over centuries and millennia, so that the resulting ethnic divisions are associated with stark and persistent differences in culture, norms, values and preferences.[26] In contrast, constructivists view ethnic identities as the endogenous result of shifting patterns of power, some very recent, so that the association between ethnic identity and cultural differences, if there is one at all, would be context-dependent, malleable, and fleeting.

Our paper provides evidence consistent with a synthesis of both views: ethnicity is indeed associated with fundamental differences in values, attitudes and preferences, in line with a primordialist viewpoint. Moreover, to the extent that ethnic divisions matter for conflict, they only do so when they overlap with cultural cleavages, once again a result with primordialist connotations. However, there are many other sources of variation in culture, not associated with ethnic identity: the magnitude of our χ^2 and F_{ST} indices tends to be small, indicating that the extent to which ethnicity is informative for culture

[26] See the voluminous and growing literature on ethnic heterogeneity in economics, among which the recent paper by Michalopoulos (2012) constitutes an excellent illustration of the primordialist orientation of the economics approach to ethnicity.

is limited, a result that is more in line with the constructivist view. Moreover, some regions like Latin America feature a weak degree of association between culture and ethnicity, while others like Sub-Saharan Africa and Asia feature more overlap. The degree to which ethnic classifications reflect deep differences in cultural attitudes varies across regions, so the extent to which ethnic identities are primordially given or constructed varies across locations. Future work should continue to study the complex relationship between ethnicity and culture, a subject that had so far remained missing from the economics literature on ethnic heterogeneity.

References

Alesina, A. and E. La Ferrara (2005), "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, vol. 43, pp. 762-800.

Alesina, A., R. Baqir and W. Easterly (1999), "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, vol. 114, no. 4, pp. 1243-1284.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat and R. Wacziarg (2003), "Fractionalization," *Journal of Economic Growth*, vol. 8, no. 2, pp. 155-194.

Alesina, A. and P. Giuliano (2013), "Family Ties," in *Handbook of Economic Growth*, P. Aghion and S. Durlauf (eds.), North Holland, Elsevier.

Alesina, A., P. Giuliano and N. Nunn (2013), "On the Origins of Gender Roles: Women and the Plough," *Quarterly Journal of Economics*, vol. 128, no. 2, pp. 469-530.

Alesina, A., S. Michalopoulos and E. Papaioannou (2012), "Ethnic Inequality," *NBER Working Paper* #18512.

Alesina, A. and E. Zhuravskaya (2011), "Segregation and the Quality of Government in a Cross Section of Countries," *American Economic Review*, vol. 101, no. 5, pp. 1872-1911.

Algan, Y. and P. Cahuc (2010), "Inherited Trust and Growth," *American Economic Review*, vol. 100 no. 5, pp. 2060-92.

Arbatli, C. E., Q. Ashraf and O. Galor (2013), "The Nature of Civil Conflict," *working paper*, Brown University.

Ashraf, Q. and O. Galor (2013), "The Out of Africa Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *American Economic Review*, vol. 103, no. 1, pp. 1-46.

Barro, R. and R. McCleary (2003), "Religion and Economic Growth across Countries", *American Sociological Review,* vol. 68, no. 5, pp. 760-781.

Bell, A., P. J. Richerson and R. McElreath, (2009) "Culture Rather than Genes Provides Greater Scope for the Evolution of Large-Scale Human Prosociality," *Proceedings of the National Academy of Sciences*, vol. 106, no. 2, pp. 17671-17674.

Bossert, W., C. D'Ambrosio and E. La Ferrara (2011), "A Generalized Index of Fractionalization," *Economica*, vol. 78, no. 312, pp. 723-750.

Cavalli-Sforza, L. L., P. Menozzi and A. Piazza (1994), *The History and Geography of Human Genes*, Princeton: Princeton University Press.

Central Intelligence Agency (2009), *The World Factbook 2009*, Washington, DC: CIA, 2009.

Collier, P. and A. Hoeffler (2004), "Greed and Grievance in Civil War," *Oxford Economic Papers*, vol. 56, no. 4, pp. 563-595.

Cover, T. M. and J. A. Thomas (2006), *Elements of Information Theory*, Hoboken, New Jersey: John Wiley & Sons.

Cramér, H. (1946), *Mathematical Methods of Statistics*. Princeton: Princeton University Press.

Dawkins, R. (2005), *The Ancestor's Tale: A Pilgrimage to the Dawn of Evolution*, New York: Houghton Mifflin Harcourt.

Desmet, K., I. Ortuño-Ortín and R. Wacziarg (2012), "The Political Economy of Linguistic Diversity," *Journal of Development Economics*, vol. 97, no. 2, pp. 322-338.

Easterly, W. and R. Levine (1997), "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, vol. 112, no. 4, November, pp. 1203-1250.

Edwards, A. W. F. (2003), "Human Genetic Diversity: Lewontin's Fallacy," *BioEssays*, vol. 25, no. 8, pp. 798-801.

Esteban, J. M., and D. Ray (1994), "On the Measurement of Polarization," *Econometrica*, vol. 62, no. 4, pp. 819-851.

Esteban, J. M., and D. Ray (2004), "Polarization: Concepts, Measurement, Estimation," *Econometrica*, vol. 72, no. 6, pp. 1737-1772.

Esteban, J., L. Mayoral and D. Ray (2012), "Ethnicity and Conflict: An Empirical Study," *American Economic Review*, vol. 102, no. 4, pp. 1310-1342.

Fearon, J. and D. Laitin (2003), "Ethnicity, Insurgency, and Civil War," *American Political Science Review*, vol. 97, no. 1, February, pp. 75-90.

Fearon, J. (2003), "Ethnic and Cultural Diversity by Country", *Journal of Economic Growth*, vol. 8, no. 2, June, pp. 195-222.

Fernandez, R. and A. Fogli (2009), "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics, vol.* 1, no. 1, pp. 146–177.

Fischer, R. and S. Schwartz (2010), "Whence Differences in Value Priorities? Individual, Cultural, or Artifactual Sources," *Journal of Cross-Cultural Psychology,* vol. 20, no. 10, pp. 1–18.

Giuliano, P. (2007), "Living Arrangements in Western Europe: Does Cultural Origin Matter?" *Journal of the European Economic Association*, vol. 5, no. 5, pp. 927-952.

Guiso, L., P. Sapienza and L. Zingales (2009), "Cultural Biases in Economic Exchange?" *Quarterly Journal of Economics*, vol. 124, no. 3, pp. 1095-1131.

Huber, J. D. and L. Mayoral (2013), "Inequality, Ethnicity and Civil Conflict," *Barcelona GSE Working Paper*, no. 744.

Jakobsson, M., M. D. Edge and N. A. Rosenberg, (2013), "The Relationship Between F_{ST} and the Frequency of the Most Frequent Allele," *Genetics*, vol. 193, pp. 512-528.

Jost, L. (2008), "GST and Its Relatives Do Not Measure Differentiation," *Molecular Ecology*, vol. 17, no. 18, pp. 4015-4026, September.

Lewontin, R. (1972), "The Apportionment of Human Diversity," *Evolutionary Biology*, vol. 5, pp. 381–398

Luttmer, E. F. P., and M. Singhal (2011), "Culture, Context, and the Taste for Redistribution," *American Economic Journal: Economic Policy*, vol. 3, no. 1 pp. 157-79.

Meirmans P. G. and P. W. Hedrick (2011) "Assessing Population Structure: FST and Related Measures," *Molecular Ecology Resources*, vol. 11, pp. 5-18.

Michalopoulos, S. (2012), "The Origins of Ethnolinguistic Diversity," *American Economic Review*, vol. 102, no. 4, pp. 1508–1539.

Montalvo, J. G. and M. Reynal-Querol (2005), "Ethnic Polarization, Potential Conflict and Civil War," *American Economic Review*, vol. 95, no. 3, June, pp. 796-816.

Nei, M (1973) "Analysis of Gene Diversity in Subdivided Populations," *Proceedings of the National Academic of Sciences, USA*, vol. 70, pp. 3321-3323.

Rae, D. W. and M. Taylor (1970), *The Analysis of Political Cleavages*, New Haven: Yale University Press.

Ross R. M., S. J. Greenhill and Q. D. Atkinson (2013), "Population Structure and Cultural Geography of a Folktale in Europe," *Proceedings of the Royal Society B,* vol. 280.

Selway, J. (2010) "Cross-Cuttingness, Cleavage Structures and Civil War Onset," *British Journal of Political Science,* vol. 41, pp. 111–138.

Selway, J. (2011) "The Measurement of Cross-Cutting Cleavages and Other Multidimensional Cleavage Structures," *Political Analysis*, vol. 19, pp. 48-65.

Spolaore E. and R. Wacziarg (2009), "The Diffusion of Development," *Quarterly Journal of Economics*, vol. 124, no. 2, May, pp. 469-529.

Spolaore E. and R. Wacziarg (2013), "War and Relatedness," *working paper, UCLA and Tufts University.*

Tabellini, G. (2010), "Culture and Institutions: Economic Development in the Regions of Europe," *Journal of the European Economic Association*, vol. 8, no. 4, pp. 677-716.

Workman, P. L. and J. D. Niswander (1970), "Population Studies on Southwestern Indian Tribes. II. Local Genetic Differentiation in the Papago," *American Journal of Human Genetics*, vol. 22, no. 1, pp. 24–49.

Wright, S. (1949). "The Genetical Structure of Populations," *Annals of Human Genetics*, vol. 15, pp. 323–354.

Wright, S. (1978), *Evolution and the Genetics of Populations, Volume 4: Variability Within and Among Natural Populations.* Chicago: University of Chicago Press.

Table 1 – Joint Significance of Ethnolinguistic Dummies in Questions from the World Values / European Values Integrated Surveys, by Region

	# of regres-sions	Share of regressions w/ jointly significant ethnic dummies	R^2 without ethnic dummies	R^2 with ethnic dummies	ΔR^2
Whole Sample	**21,469**	**0.430**	**2.688**	**4.074**	**1.386**
Africa	**3,623**	**0.548**	**2.468**	**4.064**	**1.597**
Of which: Sub-Saharan Africa	2,724	0.616	2.369	4.274	1.905
Of which: North Africa	899	0.344	2.766	3.430	0.663
Europe	**7,769**	**0.373**	**3.045**	**4.144**	**1.099**
Of which: Western and Southern Europe	2,369	0.313	3.567	4.399	0.832
Of which: Eastern and Central Europe	5,400	0.399	2.816	4.032	1.215
Asia	**5,656**	**0.571**	**2.360**	**4.519**	**2.159**
Of which: East and Southeast Asia	2,090	0.626	2.161	4.614	2.452
Of which: South Asia	852	0.667	2.899	6.363	3.463
Of which: Southwestern and Central Asia	1,511	0.479	2.223	3.391	1.168
Of which: Middle East	1,203	0.525	2.494	4.464	1.971
America	**3,749**	**0.235**	**2.480**	**3.188**	**0.708**
Of which: North America	741	0.513	3.157	4.075	0.918
Of which: Latin America and Caribbean	3,008	0.166	2.313	2.970	0.656
Oceania	**672**	**0.342**	**3.669**	**4.509**	**0.840**

Note: North America is defined here as Canada and the US. Mexico is included with Latin America and the Caribbean. R^2 is expressed in % terms.

Table 2 - Joint Significance of Ethnolinguistic Dummies in Questions from the World Values / European Values Integrated Surveys, by Question Category and Question Type

	# of regres-sions	Share of regressions with jointly significant ethnic dummies	R^2 without ethnic dummies	R^2 with ethnic dummies	ΔR^2
Breakdown by Question Category					
A: Perceptions of Life	4,382	0.425	3.270	4.618	1.347
B: Environment	971	0.427	2.185	3.640	1.454
C: Work	2,409	0.398	2.404	3.546	1.143
D: Family	1,319	0.445	3.240	4.599	1.359
E: Politics and Society	9,046	0.409	2.407	3.717	1.310
F: Religion and Morals	2,316	0.516	3.268	5.043	1.775
G: National Identity	1,026	0.495	1.801	3.682	1.881
Breakdown by Question Type					
Binary questions	4,551	0.426	2.854	4.248	1.394
Binary from unordered response questions	7,029	0.362	1.616	2.707	1.091
Scale questions	9,889	0.479	3.373	4.965	1.592

Note: This result does not change if you break it down by continent: there is little difference in shares of questions with significant ethnolinguistic dummies when the breakdown by category is done continent by continent. R^2 is expressed in % terms.

Table 3 – Summary Statistics for the Main Indices of Ethnic Heterogeneity, Cultural Diversity, FST and χ^2.

Panel A: Summary Statistics

	Mean	Standard Deviation	Minimum	Maximum
Cultural Fractionalization	0.5291	0.0366	0.4273	0.6024
ELF	0.3896	0.2584	0.0000	0.8517
F_{ST}	0.0116	0.0110	0.0000	0.0588
χ^2	0.0291	0.0254	0.0000	0.1281

(Summary statistics based on 76 observations)

Panel B: Correlations

	Cultural Fractionalization	ELF	F_{ST}	χ^2
Cultural Fractionalization	1			
ELF	-0.0303	1		
F_{ST}	0.1787	0.6198**	1	
χ^2	0.2191	0.6203**	0.9813**	1

(** Significant at the 1% level; correlations based on 76 observations)

Panel C: Means of CF, F_{ST} and χ^2 by Question Category and Type

	Cultural Fractionalization	F_{ST}	χ^2
Breakdown by Question Category			
SECTION A: PERCEPTIONS OF LIFE	0.4137	0.0117	0.0196
SECTION B: ENVIRONMENT	0.5955	0.0106	0.0261
SECTION C: WORK	0.5493	0.0111	0.0264
SECTION D: FAMILY	0.5156	0.012	0.0275
SECTION E: POLITICS AND SOCIETY	0.6122	0.0106	0.0311
SECTION F: RELIGION AND MORALE	0.5254	0.0143	0.0434
SECTION G: NATIONAL IDENTITY	0.5777	0.0172	0.0468
Breakdown by Question Type			
BINARY	0.3548	0.0125	0.0125
UNORDERED RESPONSE QUESTIONS	0.5740	0.0120	0.0262
SCALE	0.5970	0.0108	0.0368
Restricted Set of 9 Questions Used in Literature			
9 QUESTIONS FROM LITERATURE	0.3454	0.0109	0.0125

Panel D: Correlations among Cultural Diversity and χ^2 by Question Category

	CF Overall	CF A	CF B	CF C	CF D	CF E	CF F
CF Category A	0.4320	1					
CF Category B	0.3858	0.1400	1				
CF Category C	0.7321	0.1847	-0.0397	1			
CF Category D	0.5957	0.1219	0.3717	0.2825	1		
CF Category E	0.7568	0.3015	0.2767	0.5652	0.4767	1	
CF Category F	0.8559	0.0771	0.2535	0.6178	0.4917	0.4817	1
CF Category G	-0.0091	-0.1324	0.2337	-0.3535	0.4105	-0.1039	0.0740

(Based on 76 observations)

	χ^2 Overall	χ^2 A	χ^2 B	χ^2 C	χ^2 D	χ^2 E	χ^2 F
χ^2 Category A	0.9671	1					
χ^2 Category B	0.7654	0.7163	1				
χ^2 Category C	0.7978	0.7376	0.9228	1			
χ^2 Category D	0.9206	0.9152	0.6849	0.6605	1		
χ^2 Category E	0.9775	0.9183	0.7764	0.8075	0.8943	1	
χ^2 Category F	0.9413	0.9031	0.6434	0.6812	0.8251	0.8755	1
χ^2 Category G	0.6317	0.5552	0.4493	0.4195	0.6637	0.6336	0.5535

(Based on 76 observations)

Question categories are defined as follows: A: Perceptions of Life (42 questions); B: Environment (4 questions); C: Work (25 questions); D: Family (12 questions); E: Politics and Society (59 questions); F: Religion and Morale (30 questions); G: National Identity (3 questions).

Panel E: Correlations among Cultural Diversity and χ^2 by Question Type

	CF All	CF Binary	CF Scale
CF Binary	0.5988	1	
CF Scale	0.9184	0.5588	1
CF Unordered	0.6032	0.4672	0.507

	χ^2 All	χ^2 Binary	χ^2 Scale
χ^2 Binary	0.9442	1	
χ^2 Scale	0.9958	0.9304	1
χ^2 Unordered	0.9535	0.9299	0.9372

Binary: 38 questions; unordered response questions: 26 questions; scale: 100 questions

Panel F: Correlations between Diversity Measures by Question Category

	Correlations with		
	Cultural Fractionalization	**ELF**	**F_{ST}**
Question Category A			
ELF	0.2958	1	
F_{ST}	0.4646	0.5643	1
χ^2	0.4570	0.5769	0.9859
Question Category B			
ELF	-0.1176	1	
F_{ST}	-0.2786	0.5561	1
χ^2	-0.2277	0.5611	0.9891
Question Category C			
ELF	-0.0333	1	
F_{ST}	0.1331	0.6232	1
χ^2	0.3019	0.5557	0.8898
Question Category D			
ELF	-0.0715	1	
F_{ST}	0.0726	0.6291	1
χ^2	0.0863	0.6429	0.9769
Question Category E			
ELF	0.1650	1	
F_{ST}	0.1981	0.6196	1
χ^2	0.2308	0.6367	0.9834
Question Category F			
ELF	-0.2552	1	
F_{ST}	-0.0135	0.4767	1
χ^2	0.1193	0.4764	0.8922
Question Category G			
ELF	-0.2210	1	
F_{ST}	-0.0126	0.4298	1
χ^2	0.1368	0.3587	0.9126

(Based on 76 observations)

Question categories are defined as follows: A: Perceptions of Life (42 questions); B: Environment (4 questions); C: Work (25 questions); D: Family (12 questions); E: Politics and Society (59 questions); F: Religion and Morale (30 questions); G: National Identity (3 questions).

Panel G: Correlations between Diversity Measures by Question Type

	Cultural Fractionali-zation	ELF	FST
Binary Response Questions			
ELF	0.1345	1	
F_{ST}	0.2455	0.5506	1
χ^2	0.2556	0.5556	1
Unordered Response Questions			
ELF	-0.2816	1	
F_{ST}	-0.0222	0.5422	1
χ^2	-0.0359	0.5671	0.9818
Scale Response Questions			
ELF	0.1107	1	
F_{ST}	0.2674	0.6313	1
χ^2	0.2869	0.6260	0.9860

Binary: 38 questions; unordered response questions: 26 questions; scale: 100 questions

Panel H: Correlations between Diversity Measures for Restricted Set of 9 Questions

	Cultural Fractionali-zation	ELF	FST
Restricted Set of 9 Questions Used in Literature			
ELF	-0.2215	1	
F_{ST}	0.0222	0.5508	1
χ^2	-0.0225	0.5856	0.9688

Table 4 – Correlations of Different Variables with CF, F_{ST} and χ^2

	Cultural Fractionalization	F_{ST}	χ^2
ELF	-0.0303	0.6198**	0.6203**
Linguistic Diversity (Alesina et al., 2003)	0.0184	0.5609**	0.5785**
Ethnic Diversity (Alesina et al., 2003)	-0.0781	0.3611*	0.3824**
Religious Diversity (Alesina et al., 2003)	0.3143**	0.0815	0.0864
Percentage Protestant	0.2122	-0.0573	-0.0436
Percentage Catholic	0.2578**	-0.2847*	-0.2643*
Percentage Muslim	-0.5968**	0.0997	0.0821
Absolute Latitude	0.1804	-0.3178*	-0.3547**
Area	0.0971	-0.0787	-0.0823
Roughness	0.0254	0.0145	-0.0067
Log Population 1990-2010	-0.1437	0.0581	0.0751
Log GDP per Capita 1990-2010	0.3819**	-0.2644*	-0.2869*
Sub-Saharan Africa	0.0831	0.1664	0.2342*
Middle East & North Africa	-0.5286**	-0.0682	-0.0937
Europe & Central Asia	0.2670*	-0.1577	-0.1880
South Asia	-0.2624*	0.2919*	0.2630*
East Asia & Pacific	-0.0264	0.2022	0.2193*
North America	0.1261	-0.0390	-0.0549
Latin America & Caribbean	0.1118	-0.2363*	-0.2296*
Democracy	0.5862**	-0.1414	-0.1491
UK Legal Origin	0.0796	0.4231*	0.4430**
French Legal Origin	-0.1589	-0.2190	-0.2258
German Legal Origin	0.0592	-0.1351	-0.1582
Scandinavian Legal Origin	0.1005	-0.1245	-0.1050
Ethnic segregation	-0.2503	0.1884	0.2065
Linguistic segregation	-0.2076	0.0668	0.0581
Religious segregation	-0.2137	0.2183	0.1869

* significant at 5%; ** significant at 1%

Table 5 - Incidence of Civil Conflict and Diversity
(Dependent Variable: Incidence of Civil Conflict)

	(1) Incidence	(2) Incidence	(3) Incidence	(4) Incidence	(5) Incidence	(6) Incidence
Cultural Fractionalization	**-0.187**			**-0.391****	**-0.472****	**-0.417****
	[-1.60]			**[-3.33]**	**[-3.49]**	**[-3.65]**
Ethnolinguistic Fractionalization		**0.019**		**-0.037**	**-0.041**	**-0.035**
		[1.09]		**[-1.77]**	**[-1.88]**	**[-1.78]**
Chi Square			**0.472****	**0.870****	**0.992****	**0.840****
			[2.74]	**[3.82]**	**[4.00]**	**[3.96]**
Lagged War	0.860**	0.868**	0.864**	0.840**	0.840**	0.833**
	[30.66]	[33.74]	[33.33]	[27.88]	[26.28]	[25.78]
Log Lagged GDP per capita	-0.005	-0.008	-0.007	0.003	0.005	0.003
	[-0.90]	[-1.74]	[-1.47]	[0.62]	[0.98]	[0.78]
Log Lagged Population	0.011**	0.012**	0.013**	0.012**	0.015**	0.011**
	[3.69]	[3.98]	[4.33]	[4.18]	[4.32]	[4.32]
% Mountainous Terrain	0.000*	0.000*	0.000*	0.000**	0.000	0.000*
	[2.53]	[2.34]	[2.34]	[2.60]	[1.84]	[2.02]
Non Contiguous	0.025	0.017	0.015	0.026	0.025	0.025
	[1.61]	[1.27]	[1.16]	[1.60]	[1.61]	[1.54]
Oil	0.020	0.021	0.025	0.022	0.014	0.015
	[1.56]	[1.50]	[1.80]	[1.62]	[1.34]	[1.33]
New State	0.200*	0.222*	0.240*	0.210*	0.255*	0.143*
	[2.14]	[2.29]	[2.36]	[2.14]	[2.16]	[2.11]
Instability	-0.009	-0.009	-0.009	-0.009	-0.010	-0.010
	[-1.00]	[-0.94]	[-1.00]	[-1.08]	[-1.03]	[-1.34]
Democracy Lagged (Polity 2)	0.001	0.000	0.000	0.000	0.001	0.000
	[0.90]	[0.56]	[0.20]	[0.49]	[0.75]	[0.47]
Latin America and Caribbean	0.017	0.007	0.018	0.058	0.042	0.066*
	[0.98]	[0.48]	[0.99]	[1.92]	[1.54]	[1.99]
Sub-Saharan Africa	0.022	0.001	0.000	0.051	0.064	0.060
	[0.90]	[0.05]	[0.02]	[1.53]	[1.53]	[1.65]
East and Southeast Asia	-0.008	-0.010	-0.017*	-0.016	-0.017	-0.015
	[-0.68]	[-1.12]	[-2.25]	[-1.86]	[-1.84]	[-1.86]
UK Legal Origin					0.009	
					[0.26]	
French Legal Origin					0.032	
					[0.97]	
Socialist Legal Origin					0.010	
					[0.30]	
GDP Growth						-0.119**
						[-3.50]
GDP Growth Lagged						-0.003
						[-0.08]
Observations	2,921	2,921	2,921	2,921	2,705	2,850
Pseudo R-squared	0.752	0.752	0.754	0.758	0.754	0.771

Logit estimation, based on at most 69 countries from 1945 to 1999, standard errors clustered at country level. The columns report marginal effects.
Robust z statistics in brackets. * significant at 5%; ** significant at 1%

Table 6 - Incidence of Civil Conflict and Diversity, by Question Category and Type
(Dependent Variable: Incidence of Civil Conflict)

Panel A – By Question Category

	A (42q)	**B (4q)**	**C (25q)**	**D (12q)**	**E (59q)**	**F (30q)**	**G (3q)**
Cultural Fractionalization	-0.200 [-1.88]	0.025 [0.40]	-0.062 [-1.06]	-0.123 [-1.11]	-0.135 [-0.91]	-0.119** [-3.35]	-0.035 [-1.38]
Ethnolinguistic Fractionalization	-0.016 [-0.96]	0.021 [1.22]	0.011 [0.49]	-0.034 [-1.60]	-0.028 [-1.38]	-0.007 [-0.44]	-0.000 [-0.01]
Chi Square	0.756** [3.02]	-0.009 [-0.07]	0.075 [0.60]	0.843** [3.29]	0.687** [3.26]	0.394** [3.19]	0.202* [2.47]
Observations	2,971	2,925	2,971	2,921	2,971	2,971	2,916
Pseudo R-squared	0.757	0.760	0.753	0.756	0.756	0.759	0.755

Robust z statistics in brackets. * significant at 5%; ** significant at 1%
Logit estimation, based on 68 countries from 1945 to 1999, standard errors clustered at the country level. The columns report marginal effects.
All columns include controls for lagged war, log lagged GDP per capita, log lagged population, % mountainous terrain, non-contiguous country dummy, oil dummy, new state dummy, instability dummy, democracy lagged (Polity 2), Latin America and Caribbean dummy, Sub-Saharan Africa dummy, East and Southeast Asia dummy.
Question categories are defined as follows: A: Perceptions of Life (42 questions); B: Environment (4 questions); C: Work (25 questions); D: Family (12 questions); E: Politics and Society (59 questions); F: Religion and Morale (30 questions); G: National Identity (3 questions).

Panel B – By Question Type

	Binary (38q)	**Scale (100q)**	**Unordered (26q)**
Cultural Fractionalization	-0.249** [-2.73]	-0.257* [-2.01]	-0.058 [-0.58]
Ethnolinguistic Fractionalization	-0.011 [-0.63]	-0.021 [-1.04]	-0.017 [-0.90]
Chi Square	1.257** [4.59]	0.562** [2.95]	0.585* [2.57]
Observations	2,971	2,971	2,925
Pseudo R-squared	0.759	0.757	0.762

Robust z statistics in brackets. * significant at 5%; ** significant at 1%
Logit estimation, based on 68 countries from 1945 to 1999, standard errors clustered at the country level. The columns report marginal effects.
All columns include controls for lagged war, log lagged GDP per capita, log lagged population, % mountainous terrain, non-contiguous territory dummy, oil dummy, new state dummy, instability dummy, democracy lagged (Polity 2), Latin America and Caribbean dummy, Sub-Saharan Africa dummy, East and Southeast Asia dummy.

Table 7 – Endogeneity of Chi-Square and Cultural Fractionalization
(Dependent Variable: Incidence of Civil Conflict)

	(1) Persistent Questions	(2) Non-Persistent Questions	(3) Post-1970	(4) Post-1970 Cohort < 1950
Cultural Fractionalization	-0.222	-0.241**	-0.504**	-0.358**
	[-1.49]	[-3.87]	[-2.64]	[-2.74]
Ethnolinguistic Fractionalization	-0.014	-0.022	-0.022	-0.017
	[-0.76]	[-1.26]	[-0.64]	[-0.66]
Chi Square	0.458*	0.697**	1.071**	0.370**
	[2.21]	[3.65]	[2.60]	[2.65]
Observations	2,677	3,078	1,785	1,636
Pseudo R-squared	0.758	0.758	0.793	0.796

Logit estimation, standard errors clustered at country level. The columns report marginal effects. Robust z statistics in brackets. * significant at 5%; ** significant at 1%.
All columns include controls for lagged war, log lagged GDP per capita, log lagged population, % mountainous terrain, non-contiguous territory dummy, oil dummy, new state dummy, instability dummy, democracy lagged (Polity 2), Latin America and Caribbean dummy, Sub-Saharan Africa dummy, East and Southeast Asia dummy.

Table 8 – Robustness Tests
(Dependent Variable: Incidence of Civil Conflict)

	(1) FST	(2) PRIO25	(3) Broader Set Questions	(4) 9 Questions	(5) Polarization
Cultural Fractionalization	-0.361**	-0.894**	-0.323**	-0.217**	-0.387**
	[-3.20]	[-3.27]	[-2.80]	[-2.79]	[-3.13]
Ethnolinguistic Fractionalization	-0.038	-0.058	-0.035	-0.030	-0.037
	[-1.85]	[-1.34]	[-1.66]	[-1.63]	[-1.80]
Overlap measure[a]	2.021**	1.683**	0.855**	1.293**	0.864**
	[4.34]	[4.41]	[3.71]	[3.45]	[3.61]
Polarization					0.003
					[0.13]
Observations	2,921	2,834	2,921	3,078	2,921
Pseudo R-squared	0.758	0.591	0.757	0.756	0.758

Logit estimation, based on at most 69 countries from 1945 to 1999, standard errors clustered at country level. The columns report marginal effects. Robust z statistics in brackets. * significant at 5%; ** significant at 1%
a: FST in column (1), Chi-square in all other columns.
All columns include controls for lagged war, log lagged GDP per capita, log lagged population, % mountainous terrain, non-contiguous territory dummy, oil dummy, new state dummy, instability dummy, democracy lagged (Polity 2), Latin America and Caribbean dummy, Sub-Saharan Africa dummy, East and Southeast Asia dummy.

Figure 1: Cultural Fractionalization (CF)

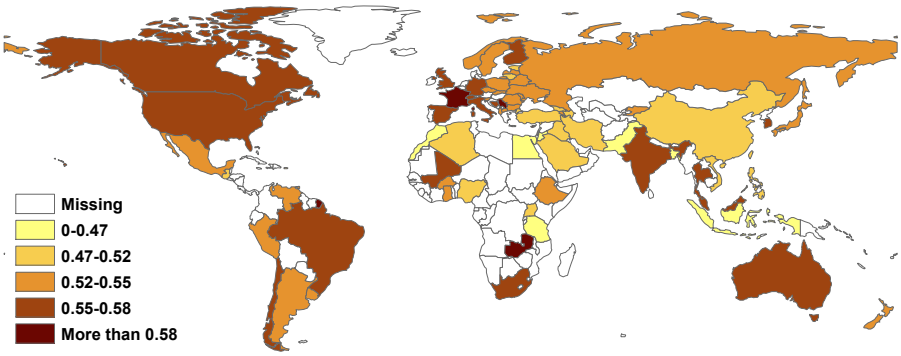

Figure 2: ELF

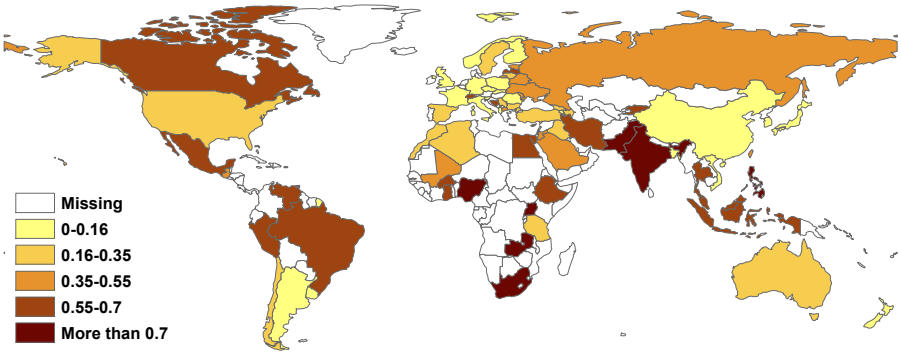

Figure 3: Chi-Square