Culture, Ethnicity and Diversity*

Klaus Desmet
SMU and CEPR

Ignacio Ortuño-Ortín
Universidad Carlos III

Romain Wacziarg
UCLA and NBER

March 2017

Abstract

We investigate the empirical relationship between ethnicity and culture, defined as a vector of traits reflecting norms, values and attitudes. Using survey data for 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. Although only a small portion of a country's overall cultural heterogeneity occurs between groups, we find that various political economy outcomes (such as civil conflict and public goods provision) worsen 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 Alberto Alesina, Georgy Egorov, James Fearon, Paola Giuliano, Wolfgang Keller, Keith Krehbiel, Giacomo Ponzetto, Enrico Spolaore, the editor, four anonymous referees and seminar participants at numerous universities for useful comments. We gratefully acknowledge financial support from the Spanish Ministry of Economics and Competitiveness (grants ECO2011-27014 and ECO2013-46091-P) and the UCLA Anderson Center for Global Management. 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 relationship 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 ethnicity and culture? Finally, is the relationship between ethnicity and culture important to understand salient political economy outcomes, such as civil conflict or public goods provision?

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 and to evaluate the statistical significance of ethnic identity. 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. We propose a new measure of cultural fractionalization, defined as 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 cultural and preference heterogeneity.

How can these seemingly contradictory results be reconciled? If most of cultural heterogeneity occurs within groups rather than between groups, then the correlation between ethnic diversity and cultural diversity will tend to be low. In spite of this, ethnicity could still carry some information about cultural values. This is indeed what we document. To do so, 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 from 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 variation in cultural norms is between groups. That is, the vast share of the variation is within groups, a result that mirrors well-known findings in population genetics. This explains the close-to-zero correlation between cultural heterogeneity and ethnic heterogeneity. The low share of between-group variation 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 variation in cultural values is about six 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 small in magnitude, matter for our understanding of political economy outcomes? To analyze whether the overlap between culture and ethnicity is relevant, we explore how ethnic heterogeneity, cultural heterogeneity and the overlap between culture and ethnicity affect civil conflict and public goods. We find empirically that both cultural and ethnic diversity have weak effects on civil conflict and public goods. If anything, higher cultural diversity reduces the probability of civil conflict and increases public goods. However, in countries where ethnicity is more strongly predictive of culture, as captured by a high χ^2, violent conflict is more likely, and public goods provision tends to be lower. Our interpretation of this empirical result is that in societies where individuals differ from each other in both ethnicity and culture, social antagonism is greater, and political economy outcomes are worse.

This paper is related to various strands of the literature on ethnolinguistic diversity. The first strand studies the relationship between ethnolinguistic diversity and political economy 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). 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 indices that improve on simple fractionalization. These measures take different forms: some account for distance between groups (Esteban and Ray, 1994, 2011, Duclos, Esteban and Ray, 2004, Bossert, d'Ambrosio and La Ferrara, 2011, Esteban, Mayoral and Ray, 2012); others look at income inequality between ethnic groups (Huber and Mayoral, 2013, Alesina, Michalopoulos and Papaioannou, 2016) or the historical depth of ethnic cleavages (Desmet, Ortuño-Ortín and Wacziarg, 2012); yet others consider heterogeneity between individuals rather than groups (Ashraf and Galor, 2013a, Arbatli, Ashraf and Galor, 2015). Our paper is related to this literature because we propose new indices of heterogeneity both between and within ethnic groups.

The third strand relates to the overlap of ethnicity with other dimensions: A political science literature on cross-cutting cleavages, starting with Rae and Taylor (1970), studies whether two dimensions of heterogeneity

might reinforce each other.[1] Of particular interest is the important recent paper by Gubler and Selway (2012) who also use a χ^2 index to look at how the overlap between ethnicity and other dimensions (income, geography and religious identity) affects civil war. Our work differs from theirs for four reasons: First, we focus on cultural values, and conduct a systematic analysis of how these values relate to ethnic identity, and how ethnic diversity and cultural diversity relate to each other. Second, we explicitly relate our measures to a simple model of social antagonism. Third, we develop new measures of cultural diversity and analyze their correlates. Fourth, we look at the effect of these indices on a broader range of political economy outcomes, beyond civil conflict.

Finally, a recent literature relates genetic differences - a measure associated with cultural differences - with political and economic outcomes. For instance, Spolaore and Wacziarg (2009, 2016) use an F_{ST} index of genetic distance between countries (rather than ethnic groups within countries) to capture barriers between populations. Ashraf and Galor (2013a) investigate the effect of genetic diversity on historical and contemporary economic performance. In Arbatli, Ashraf and Galor (2015), the same measure of genetic diversity is found to have a positive effect on the probability of civil conflict. The latter two papers were the first to consider measures of overall diversity between individuals within societies, something that previous measures of diversity (such as commonly used measure of ethnic group fractionalization) failed to do. Our approach also captures diversity between individuals, but rather than using genetic data, we measure cultural diversity using responses to surveys on norms, attitudes and preferences.[2]

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 is 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 than 5%.

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

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

[1] We discuss at length the relationship between our measurement framework and this literature on cross-cuttingness in Appendix A.3.

[2] Another related literature studies the socioeconomic effects of specific cultural traits, rather than cultural heterogeneity. Salient examples include Alesina, Giuliano and Nunn (2013); Giuliano (2007); Fernandez and Fogli (2009); Luttmer and Singhal (2011); Tabellini (2010); and Guiso, Sapienza and Zingales (2009).

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 each question in each country, and then examine the share of regressions 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. To capture the magnitude of the joint effect of ethnicity on culture, 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.

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, we 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 are grouped into question categories labeled from A to G).[3] In the end this left us 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.[4] 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 B1 provides the exact count by country). When combined across all waves, the average number of respondents across the countries in the sample, and across all questions, was $1,497$.

[3] 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).

[4] 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.

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 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 race rather than language usually defines ethnic identity. 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 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 existing classsifications that are widely used in the literature: we choose either ethnic identity or linguistic identity depending on which one gives us group shares that most resemble those in Alesina et al. (2003) and Fearon (2003). 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 common fractionalization measures. Indeed, our ethnic classification results in a fractionalization measure that is 74% correlated with the one from Alesina et al., and 73% correlated with the one from 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, sex, education and household income. We conduct extensive robustness tests on these controls, decribed below.

2.3 Results

Baseline results. Table 1 presents the overall share of regressions where ethnicity dummies are jointly significant at the 5% level, breaking down these results by region. Table 2 displays 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 B1 presents the results country by country.

Interesting findings emerge. First, the average share of questions for which ethnicity dummies are jointly significant, across all countries, is 43%. Thus, ethnic identity is an important determinant of responses to many questions.

Second, this average masks 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 (details and results appear in Appendix B.1 and Appendix Tables B2 and B3).

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.[5]

Fourth, the explanatory power of the regressions is quite low. Table 1 shows that 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. These results suggest that the extent to which ethnic identity can explain cultural attitudes is a small share of overall cultural variation, a theme to which we will return at length below.

Robustness and Extensions. We conduct a wide range of extensions and robustness tests on this exercise, reported in Appendix Tables B4 through B23. We first examine the comparative explanatory power of other sorts of cleavages: the respondents' subnational region, religion and city size. We replace ethnic dummies with dummies based on these dimensions of identity, to see if they have comparable explanatory power for culture. We find that regional identity has a larger explanatory power than ethnicity: dummies for respondent's region have joint significance in 75% of the regressions, with the R^2 rising from 3.6% without region dummies to 6.2% with them. In contrast, religious identity has on average smaller predictive power for culture, with religion dummies significant in 36% of the regressions and an average increase in the regression R^2 by only 1.3 percentage points. Finally, a set of dummies capturing the respondents' urban categories (by city size intervals) are jointly significant in 57% of the regressions, with an average R^2 increase of 1.6 percentage points. These results confirm that it is difficult to find respondent characteristics that explain a large share of the variation in responses to questions on cultural attitudes.

Second, we examine the robustness of our findings about ethnic identity to the inclusion of dummies for region, religion and urban categories. We find that the results are robust to these additional controls. The inclusion of regional dummies has the biggest impact, as the share of regressions where ethnicity dummies are jointly significant fall from 43% in the baseline to 31% when adding region effects. This is possibly

[5] 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.

due to the collinearity between ethnicity and region dummies - in many countries ethnic groups have a regional basis (Alesina and Zhuravskaya, 2011). For religion and urban categories, the effect on the share of significant ethnic dummies is less pronounced.

Third, we change the set of controls included in the regressions alongside ethnic dummies. In one robustness check, we remove education and income, which are possibly endogenous to values. In another, we add more controls: marital status, whether the respondent has children, and a set of dummies for respondents' occupational categories. Neither modification has a material effect on the baseline inferences drawn above, or on the breakdown of the results by region, question category or question type.

Finally, we examine whether the use of a linear probability model for binary questions has any effect on the results. Excluding the scale questions, we use a probit estimator instead. We conduct joint Wald tests on the ethnic dummies, and examine the increment in the pseudo-R^2. In Table 2 we showed that 42.7% of binary questions have significant F-tests using OLS. Now 41.5% of binary questions have significant Wald tests. The R^2 rose by 1.39 percentage points using OLS and now the pseudo-R^2 rises by 1.45 percentage points. In sum, the use of probit rather than OLS does not change our conclusions.

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 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 average 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 individuals with different cultures or different ethnicities. Antagonism could also stem from barriers that prevent fruitful interactions, 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. Of course, these sources of antagonism are not mutually exclusive. Later, we will calculate these indices and relate them to political economy outcomes, including them jointly to let the data tell us which source of antagonism is most relevant empirically.

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 traits are the answers to the q questions in the WVS (or any other survey of cultural attitudes). For the sake of simplicity, in this section we present our model of antagonism and the different indices of diversity for the case of just one question and r possible answers. Appendix A.1 provides the general case with $q \geq 1$. There, the level of antagonism and the indices of diversity are obtained by averaging the corresponding values over all

questions. Focusing on a given country, w_j is the share of the population that gives answer j and w_j^s is the share of individuals from ethnic group s that gives answer j.

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. When an individual gives answer j, his level of antagonism v_j is given by:

$$v_j = 1 - w_j \tag{2}$$

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_j measures the probability that a randomly chosen citizen disagrees with the answer j. Social antagonism v is the summation of all the individual levels of antagonism v_j, normalized by the population size n. Appendix A.1.1 shows that v can be rewritten as the following index of cultural fractionalization (CF):

$$CF = 1 - \sum_{j=1}^{r} w_j^2 \tag{3}$$

The cultural fractionalization (CF) index measures the probability that two randomly drawn individuals from a population give different answers to the question. Thus, in a society where antagonism is driven by differences in preferences and cultural values, 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. We postulate that in this society the level of antagonism of an individual from ethnic group s is the share of people who belong to a different ethnic group:

$$v_s = 1 - w^s \tag{4}$$

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 A.1.2 shows that v 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 by ethnic animosity or barriers between ethnic groups, the conventional index of ethnolinguistic fractionalization (ELF) should matter for political economy outcomes.

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. In addition, cultural differences between the members of the same ethnic group do not increase the level of antagonism. Ethnicity only matters if ethnic groups differ in their cultural values.

Take an agent from ethnic group s who gives answer j to the question. Suppose first that this agent only interacts with agents of her own ethnic group s. In this case, by definition the share of people within group s giving an answer identical to hers is w_j^s. This is the probability that a randomly chosen agent from the ethnic group s agrees with her. Now assume that this agent is equally likely to interact with anybody in society. In this case the probability she agrees with a randomly chosen individual in society is w_j. If the probability w_j is equal to w_j^s, she does not see any difference between her own ethnic group and society overall. However, if the proportion of people in society overall answering j is lower than the corresponding proportion within her own ethnic group, the agent experiences antagonism. In particular we assume that antagonism for an agent from group s who gives answer j, v_j^s, depends on the (relative) difference between these two shares:

$$v_j^s = \frac{w_j^s - w_j}{w_j} \tag{6}$$

Notice that if $w_j^s < w_j$ 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%.

If all ethnic groups are identical, i.e., the distribution of answers is independent of the distribution of ethnic groups, then $v_j^s = 0$. We average the individual levels of antagonism to obtain social antagonism v. 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 v as a measure that can be calculated from data, Appendix A.1.3 shows that it can be rewritten as:

$$\chi^2 = \sum_{s=1}^{S} \sum_{j=1}^{r} \frac{w^s (w_j - w_j^s)^2}{w_j} \tag{7}$$

Thus, if we believe that antagonism is driven by differences in culture across ethnic groups, we should observe a relationship between the χ^2 index of overlap and political economy outcomes.

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 compare the distribution of answers across groups. This is what the χ^2 index accomplishes.[6] Let n_j^s be the number of individuals who belong to ethnic group s and give answer j. We write $n^s = \{n_1^s, n_2^s, ..., n_r^s\}$. Under independence, the expected number of individuals that belong to ethnic group s and give answer j should be $w_j n^s$, while the observed frequency is n_j^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:

$$\overline{\chi}^2 = \sum_{s=1}^{S} \sum_{j=1}^{r} \frac{(n_j^s - w_j n^s)^2}{w_j n^s} \tag{8}$$

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

$$\chi^2 = \sum_{s=1}^{S} \sum_{j=1}^{r} \frac{n^s (w_j^s - w_j)^2}{n w_j} = \sum_{s=1}^{S} \sum_{j=1}^{r} \frac{w^s (w_j^s - w_j)^2}{w_j} \tag{9}$$

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

Thus, the χ^2 index depends on the average difference between the observed shares w_j^s and the expected shares w_j that would be observed if the distributions of ethnicity and 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 possible answers r.

Closely related to this index is Cramér's V, which is defined as $V = \sqrt{\chi^2/t}$, where t is the smaller of $S-1$ and $r-1$ (Cramér, 1946). This normalization ensures that V is always between zero and one. This is the index used by Selway (2011) to study how ethnicity overlaps with religious identity. Alesina and Zhuravskaya (2011) use an index of geographical ethnic segregation that is very much related to both χ^2 and to Cramér's V. Their index is the same as our χ^2 multiplied by the factor $1/(S-1)$, where S is the number of ethnic groups. We adopt the standard χ^2 index because there is no foundation in our model for the normalizations in Cramér's V or in the segregation index of Alesina and Zhuravskaya (2011).[7]

[6] For previous uses of this index, see Selway (2010) who examines the overlap between religion and ethnicity, Alesina and Zhuravskaya (2011) who analyze the overlap between geography and ethnicity as a way of measuring segregation, and Gubler and Selway (2012) who look at the overlap between ethnicity and religion, income and geography.

[7] Further, Reardon and Firebaugh (2002) analyze several indices of segregation and recommend the use of the Mutual Information Index. This index originates 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 uncertainty about how an individual answers the questions, resulting from knowing her 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).

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 that case, cultural cleavages and ethnolinguistic cleavages cross-cut. In contrast, if F_{ST} is equal to 1, knowing someone's ethnolinguistic identity allows a perfect prediction of their cultural attributes. In that case, cultural cleavages and ethnolinguistic cleavages would be reinforcing. F_{ST} is therefore a measure of overlap between cultural values and ethnolinguistic identity. F_{ST} 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.

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

$$CF^s = 1 - \sum_{j=1}^{r} \left(w_j^s\right)^2 \tag{10}$$

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

$$CF^W = \sum_{s=1}^{S} w^s CF^s = \sum_{s=1}^{S} w^s \left(1 - \sum_{j=1}^{r} \left(w_j^s\right)^2\right) \tag{11}$$

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

$$F_{ST} = \frac{CF - CF^W}{CF} \tag{12}$$

This is, for each question, the ratio of between-group cultural fractionalization divided by total fractionalization, i.e. Wright's fixation index F_{ST} (Wright, 1949, Nei, 1973).[9]

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.

[8] Another alternative is the more conventional index of cross-cuttingness used in the political science literature (Rae and Taylor, 1970, Selway, 2011). Appendix A.3 provides a detailed discussion of this index and its relationship with χ^2 and F_{ST}. We do not use this index of cross-cuttingness here as it is sensitive to changes in group sizes that are not associated with changes in the degree to which ethnicity is informative about a person's cultural attitudes, the concept we have sought to capture in this paper so far.

[9] In the general case, $q > 1$, we average across questions. There are of course many ways to do this. For instance, Cavalli-Sforza et al. (1994) separately average the numerator and the denominator of equation (12), and then take the ratio. We adopt the simpler method of averaging the question by question F_{ST}.

Appendix A.2 discusses the theoretical drawbacks of F_{ST} (Jost, 2008, Meirmans and Hedrick, 2011, and Jakobsson et al., 2013). We find empirically that χ^2 and F_{ST} are highly correlated (98%), so that these theoretical drawbacks are not consequential in our application.[10] Appendix A.2 also discusses the past uses of F_{ST} in the literature on the measurement of cultural heterogeneity, although in contexts different from ours (Bell et al., 2009, Ross et. al., 2013).

3.4 Distance-Based Measures and Polarization

We now briefly discuss two extensions to our diversity measures. The first extension takes into account distances between responses. If a question from the WVS has more than two answers ordered on a scale, then the distance between responses 1 and 4 is greater than the distance between responses 2 and 3. A society where half the people answers 1 and the other half answers 4 might be more culturally diverse than a society where half the people answers 2 and the other half answers 3. Fearon (2003) and Desmet, Ortuño-Ortín and Weber (2009) suggest that taking into account these distances is important when measuring diversity. Appendix A.4.1 develops indices of cultural fractionalization and the overlap between culture and ethnicity that incorporate distances between responses.

The second extension considers indices of polarization. In our model antagonism is captured by different fractionalization indices. Some authors have argued that polarization rather than fractionalization might be a better way to measure the underlying antagonism in a society (Reynal-Querol, 2002; Duclos, Esteban and Ray, 2004). In Appendix A.4.2 we propose alternative indices of CF, ELF and χ^2 based on polarization.

4 Ethnic Heterogeneity and Cultural Diversity

In this section we calculate the measures of heterogeneity described 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. Yet, Section 2 showed that ethnic identity has predictive power for cultural attitudes. To reconcile these seemingly contradictory results, we show that between-group heterogeneity in cultural attitudes is small compared to total heterogeneity, yet it is not zero. This explains both the low correlation between cultural diversity and ethnic diversity, and the fact that ethnic identity helps to predict cultural values. Moreover, the overlap between culture and ethnicity shows considerable variation across countries, consistent with the variation 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 using factor analysis and different question categories and types.

[10] In the case of a question with two possible answers, F_{ST} and χ^2 coincide exactly (see Workman and Niswander, 1970).

4.1 Cultural Diversity and Ethnolinguistic Fractionalization

Before describing the indices, some comments on the data are in order. We use the same WVS/EVS survey data as before, with two minor differences: First, we drop questions that are not asked in at least 50 countries, to ensure cross-country comparability of the indices.[11] Second, there is no longer any need to convert questions that admit multiple unordered answers into series of binary questions.

Appendix Figure B1 shows a world map with the values of cultural heterogeneity in the 76 countries in our sample, and Panel A in Table 3 displays simple summary statistics.[12] 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 Appendix Figure B2 with the one of CF in Appendix Figure B1, 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.

4.2 The Overlap Measures

Appendix Figure B3 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 most heterogeneity is within groups (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 significantly predicts responses to survey questions about values, norms and preference are the same regions where the overlap measures take on higher values.[13] 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

[11] The list of questions used to compute the heterogeneity measures, as well as their breakdown by question category and type, appears in Appendix Table B52. The final sample consists of 76 countries for which we can compute ELF, CF, χ^2 and F_{ST}.

[12] Appendix Table B51 presents the underlying values country by country.

[13] See Appendix Table B31. In fact, if one considers, 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}, one obtains 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.

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. 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). For instance, Lewontin (1972) famously showed that, of the 0.1% of human DNA that varies across individuals, only 6.3% of the variation was between racial groups, while 93.6% of the variation was within racial groups. Similarly, an average 87% of the human genetic variation occurs within continental groups, with the remaining variation occurring between populations from different continents (Jorde et al., 2000; Hinds et al., 2005).[14] Hence, by analogy with well-known results in population genetics, we should perhaps not be surprised that cultural diversity also occurs mostly between individuals within groups, rather than between groups.

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 cultural heterogeneity in the world is between countries, whereas only 1.2 percent is between ethnic groups within countries.[15] Even the country with the highest between-ethnic group F_{ST}, India, has a lower value than the one observed between countries. If, instead, we take regions within countries to be the groups, using the same definition of regions as in Section 2.3, we find an F_{ST} value of 3.1 percent, in between the value for ethnic groups and countries (Appendix Table B25).

Fourth, overlap measures bear 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). 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 a lack of a strong relationship. Morocco and Pakistan are culturally relatively homogeneous ($CF = 0.445$, rank 73), but the former has a low χ^2 (0.008, rank 67), whereas the latter has a high χ^2 (0.048, rank 11). As another example, compare Zambia and Chile. Both are in the top-10 of culturally most heterogeneous countries, but knowing someone's identity is more informative in Zambia ($\chi^2 = 0.083$, rank 4) than in Chile ($\chi^2 = 0.011$, rank 58).

Fifth, as expected, the degree of overlap between culture and ethnicity is higher in countries that are more ethnically diverse. The correlation between χ^2 and ethnolinguistic fractionalization is 0.620. For example, India ranks first both in terms of ELF (0.851) and χ^2 (0.128), whereas Argentina has both a low ELF

[14] One possible explanation for the large share of within-group diversity is outlined in Arbatli, Ashraf and Galor (2015): if initial diversity is positively associated with the formation of groups (Ashraf and Galor, 2013b), yet there are scale effects implying a lower bound on the size of groups, then societies with particularly diverse traits will display both a high level of overall diversity while also displaying high within-group diversity. Such a mechanism could partly explain why total diversity is much larger than between-group diversity - both when considering genetic and cultural diversity.

[15] 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}.

(0.131, rank 61) and a low χ^2 (0.009, rank 66). On the other hand, Nigeria has a very high ELF (0.767, rank 5) but a relatively low χ^2 (0.023, rank 36).

4.3 Factor Analysis, Breakdown by Question Groupings and Other Robustness

4.3.1 Factor Analysis

We rely on a large number of questions to calculate our diversity measures. Many are variations of each other and are likely to yield answers that are correlated. Although this need not be viewed as a problem - any possible repetition likely reflects the relative importance of the issue at hand - factor analysis provides a way to reduce the dimensionality of the question space. We use the iterative principal factor method, and determine that five is a reasonable number of factors.[16] These five factors explain just 34% of the variance in answers, suggesting that the question space cannot easily be reduced to just a few dimensions. A simple inspection of the factor loadings suggests that Factor 1 refers to confidence in public institutions, Factor 2 refers to attitudes towards moral issues, Factor 3 refers to the importance of god and religion, Factor 4 refers to attitudes towards cheating and bribery in the public sphere, and Factor 5 refers to different aspects of satisfaction and happiness.

We now treat the five factors as five questions, and the factor scores as survey respondents' answers to those questions. However, because we cannot use continuous variables to compute our diversity measures, we discretize each individual's factor scores by assigning a value of one to factor scores above the worldwide mean and zero otherwise. We then use these discretized factor scores to compute values of CF, F_{ST} and χ^2 for each of these five factors separately and also take the mean of these measures across factors.

Our main results are preserved with this alternative measurement approach. In particular, the correlation between CF and ELF ranges from -0.33 to 0.22, with the correlation between the average CF across factors and ELF being -0.04, very similar to the -0.03 value in the baseline (Panel F of Appendix Table B24). This confirms that there is no relation between cultural heterogeneity and ethnolinguistic fractionalization. The overlap measures also continue to be small. The average χ^2 across the five factors is 0.022, compared to 0.029 in the baseline (Panel C of Table 3). This confirms that knowing someone's ethnicity reveals very little information about her culture.[17]

4.3.2 Breakdown by Question Category and Type

We calculate our measures separately for each of the question categories identified by the WVS/EVS (labeled A through G), and across question types (binary, scale, and unordered response questions). The results are reported in Panel C of Table 3 supplemented with Appendix Table B24. We find, on average, a higher

[16] Beyond the fifth factor, additional eigenvalues drop substantially in magnitude (indicating they explain a smaller share of the variance) and the factor loadings tend to concentrate on only one or two questions (indicating that additional factors essentially refer to specific questions, thus no longer reducing the dimensionality).

[17] When using factor analysis, as soon as an individual does not answer one question, his factor score will be missing. This presents a tradeoff between the number of questions and the number of individuals. In our factor analysis we exclude countries that have fewer than 400 non-missing observations and we exclude questions that are not answered by at least 70% of the individuals. Doing so reduces the number of questions from 175 to 138 and the number of countries to from 76 to 51. Because of this we do not use indices obtained from factor analysis to conduct the analysis of conflict and public goods in Section 5.

degree of cultural heterogeneity (CF) for questions related to environmental policy and 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). Overall, there is a high degree of consistency in the magnitudes of our heterogeneity measures. For instance, F_{ST} is comprised in a tight band between 1% and 1.7% across categories.

We also examine 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). We 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 analyzed different question 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. Second, with binary questions the issue of the distance between answers does not arise, and polarization is perfectly correlated with fractionalization. Third, focusing exclusively on a subset of questions with the same number of possible answers prevents heterogeneity measures from varying simply because of differences in the number of answers.

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 broadest 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 consider nine questions relating to family values, child qualities, trust and beliefs.[18] 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 -0.22. In addition, the correlation between CF using the limited set of questions and CF using all questions is 0.49 and significant at the 1% level (the corresponding correlation for χ^2 is 0.83). This suggests that our results are not specific to our comprehensive approach.[19]

[18] 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): f053 (belief in hell).

[19] Additionally, we find no evidence that the results differ markedly for respondents located in large cities (of more than $500,000$ inhabitants) versus the rest (Table 3, Panel C).

4.4 Other Robustness

We finish this section by carrying out two more robustness checks. First, accounting for distances in responses may give a more accurate picture of cultural fractionalization and the overlap between ethnicity and culture. When using generalized CF and F_{ST} measures that incorporate distances between questions, the results do not change: the correlation between CF and ELF is 0.049 when taking into account distances, compared to -0.030 in the benchmark, and the average value of F_{ST} is 0.013 when incorporating distances, similar to 0.012 in the benchmark (Appendix Table B28).

Second, we use the Afrobarometer and the Latinobarómetro as alternative datasets and recompute our measures of diversity. In the case of Afrobarometer we find an F_{ST} value of 3.4%, higher than the 1.7% value for Sub-Saharan Africa in the WVS/EVS sample. Consistent with this, the correlation between CF and ELF is also higher (0.291), though it remains statistically insignificant at the 10% level. For Latinobarómetro, we also find that F_{ST} is very small (0.009), as it is for Latin America in the WVS/EVS dataset (Appendix Tables B29 and B30).[20] In sum, these alternative datasets confirm a higher degree of overlap between culture and ethnicity in Sub-Saharan Africa, and a lower level in Latin America, compared to the worldwide average.

4.5 The Correlates of Cultural Diversity

Table 4 displays simple correlations between our cultural diversity measures and various variables. The first panel considers existing measures of diversity. Chief among them is the measure of genetic diversity from Ashraf and Galor (2013a). There are clear conceptual links between cultural and genetic diversity. Both measures include the contribution of diversity across individuals within groups to overall societal diversity, something that most other measures of diversity (such as the commonly used measure of ethnic fractionalization) fail to do. Moreover, genes, like cultural traits, are transmitted intergenerationally with variation, and diversity in both genes and culture result from processes of historical change that affect different lines of descent differently (Boyd and Richerson, 1985, Richerson and Boyd, 2005). Finally, the functional forms of both measures are identical: genetic diversity is calculated as the probability that two randomly chosen individuals from a given society have different alleles of a randomly drawn gene locus.

Despite these conceptual commonalities, the correlation between genetic diversity and CF is small and statistically insignificant (-0.175). Differences between both measures of diversity stem from the fact that genetic diversity is based on genetic traits (genes) whereas cultural diversity is based on cultural traits (memes).[21] Initial conditions are different for genes and memes, and rates of cultural and genetic change differ from each other, in part because genetic traits can only be transmitted vertically, whereas cultural traits are also transmitted horizontally or obliquely (Cavalli-Sforza and Feldman, 1981, Boyd and Richerson, 1985). In addition, selective forces may occur at different speeds for genes and memes. Since CF and

[20] For the Latinobarómetro, one important difference is the strong positive correlation of 0.729 between CF and ELF. We do not put too much stock in this anomalous finding because group sampling in the Latinobarómetro is completely off: the correlation between ELF based on Latinobarómetro and ELF in either Alesina et al. (2003) or Fearon (2003) is around 0.05. In contrast, the correlation between the ELF based on the WVS/EVS and Alesina et al. (2003) and Fearon (2003) is above 0.7. Moreover, questions in Latinobarómetro focus nearly exclusively on politics, and hence do not capture the multi-dimensionality of culture. More details are available in Appendix Table B29 and B30.

[21] A scatter plot of CF against genetic diversity is included as Figure B5 in the Online Appendix.

genetic diversity capture heterogeneity along different dimensions (culture versus genes), they may bear distinct relationships with the political economy outcomes we study below, as we indeed find.

Turning to other measures of diversity, including linguistic fractionalization, ethnic polarization and ethnic inequality, we continue to see no relation with cultural fractionalization. The only measure of diversity to bear a positive correlation with CF is religious diversity, a notable fact because religion is a cultural trait akin to some of those included in the World Values Survey. With respect to correlations with χ^2, Table 4 reports positive but insignificant correlations with measures of ethnic, linguistic and religious segregation from Alesina and Zhuravskaya (2011) and ethnic income inequality from Alesina, Michalopoulos and Papaioannou (2016).

Table 4 also shows the correlation between cultural fractionalization and several other 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 0.598); and countries with a high per capita income exhibit a greater CF (the correlation is 0.373). The other correlations are mostly small and statistically insignificant.

Turning to the correlates of our overlap measures, Table 4 displays the quantitative magnitudes of the simple correlations of a set of country characteristics and χ^2. 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 economic development weakens 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 this is partly due to the fact that χ^2 is lower in Latin America than elsewhere.

5 Culture, Ethnicity and Political Economy Outcomes

In this section we examine the determinants of civil conflict and public goods. Our aim is to shed light on the mechanisms underlying the relation between ethnolinguistic diversity and these outcomes. Do differences in cultural traits matter? Does ethnic diversity *per se* affect outcomes? Are between-group cultural differences important?[22]

[22] Our analysis contributes to the extensive literature linking ethnic diversity to outcomes. Some authors have argued for broadly negative effects, particularly on public goods and civil conflict. Examples include La Porta et al. (1999), Alesina et al. (2003), Montalvo and Reynal-Querol (2005) and Desmet, Ortuño-Ortín and Wacziarg (2012). Other research focusing on economic performance and productivity has found either positive or hump-shaped effects. Examples include Ottaviano and Peri (2006) who found a positive effect of birthplace diversity on rents in U.S. cities, Alesina, Harnoss and Rapoport (2016) who uncovered a positive effect of birthplace diversity on economic performance, and Ashraf and Galor (2013a) who documented a hump-shaped relation between genetic diversity and economic development. A broader literature outside of economics has also found positive effects of diversity on productivity and problem-solving at the organizational level. A salient example is Hong and Page (2004), among many others.

5.1 Civil Conflict

Whether and how ethnolinguistic diversity affects civil conflict often depends on which measure of diversity is used. Fearon and Laitin (2003), using a measure of ethnic fractionalization, 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 that greater diversity reduces the probability of a civil war. In contrast, Montalvo and Reynal-Querol (2005), using a measure of ethnic polarization, find that it increases civil war incidence. Arbatli, Ashraf and Galor (2015), using a measure of genetic diversity, find a similar result of worsening conflict.

Some recent papers have explored not just the effect of diversity, but also that of between-group income inequality. Esteban, Mayoral and Ray (2012) analyze how civil conflict is related to ethnic fractionalization, ethnic polarization and between-group inequality. Likewise, Esteban and Ray (2011) and Huber and Mayoral (2013) examine the role of income inequality between and within ethnic groups as a determinant of civil conflict. Our paper shares with these various contributions a focus on between-group diversity. However, instead of focusing on income inequality between groups, we focus on cultural heterogeneity between groups. Another related paper is Gubler and Selway (2012) who use a χ^2 index to look at how the overlap between ethnicity and other dimensions (income, geography and religious identity) affects civil war. They do not examine the link between ethnicity and cultural values, which is our main focus here.

In the framework of Section 3, diversity may affect social antagonism through three non-mutually exclusive channels. Appendix A.5 provides a theoretical model based on a contest function approach that more explicitly links civil conflict to the same three channels. First, ethnolinguistically heterogeneous societies may have more diverse preferences and values, leading to increased overall antagonism and conflict (in this case, CF should predict conflict). Second, ethnolinguistic fractionalization may matter *per se* because of direct animosity, hatred or barriers between different ethnolinguistic groups (in this case, ELF should predict conflict). Third, civil conflict may arise more frequently when ethnic divisions and cultural differences reinforce each other (in this case, χ^2 should predict conflict).

5.1.1 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. We relate conflict incidence to our three sets of measures of diversity, CF, ELF and χ^2 (or F_{ST}):

$$C_{ct} = \beta_0 + \beta_1 CF_c + \beta_2 ELF_c + \beta_3 \chi^2_c + \boldsymbol{\beta}_4^{'} \mathbf{Z}_{ct} + \varepsilon_{ct} \tag{13}$$

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}.

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 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.

5.1.2 Results

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).[23] 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 0.87. That implies that a one standard deviation change in χ^2 (equal to 0.027) raises the probability of conflict by 2.349 percentage points. The baseline probability of being in a civil conflict is 14.2% in the sample for this regression, so the standardized effect of χ^2 amounts to about 17% 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.

Contrary to the idea that cultural differences *per se* should lead to conflict, Table 5 shows 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 0.038, 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.[24] Another possible interpretation is that within-group cohesion is important to mobilize a group for conflict, so that for a given level of between-group diversity, an increase in CF implies a larger degree of within-group diversity, making it difficult for a group to agree and mobilize to fight effectively.

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.[25]

Finally, Appendix B.2 outlines in detail endogeneity concerns that stem from the possible effect of civil

[23] Most regressions are based on 69 countries. Out of the 76 countries for which we are able to construct diversity measures from the WVS/EVS, we lose 4 countries due to missing data on civil conflict (e.g., Puerto Rico) and we lose another 3 because they do not have data on GDP per capita during the sample period (e.g., some of the countries of the former Yugoslavia). Note, furthermore, that the panel is unbalanced, as some countries, such as the former Soviet Republics, only enter in more recent years.

[24] Although our results indicate that both CF and χ^2 are statistically and economically significant, one may wonder whether they add much in terms of explaining observed variation in war incidence. Indeed, when comparing columns (1), (2) and (3) in Table 5, the relative increase in pseudo-R^2 when adding the three indices of diversity is quite small. However, this is due to the initial level of the pseudo-R^2 being high because of the inclusion of lagged war as a regressor. As shown in Appendix Table B40, once we drop lagged war, the relative increase in pseudo-R^2 when adding our diversity indices is around 30%.

[25] We also explore the possibility of nonlinearities in the effects of CF, ELF and χ^2. Appendix Table B45 reports our findings. We find little evidence of nonlinear effects.

conflict on cultural values, and discusses our extensive empirical work aimed at addressing them. Across a wide variety of approaches, we find that such reverse causality is unlikely to account for the patterns described here. Regarding omitted variables, we also control for a wide range of geographic factors that have been shown to affect ethnolinguistic diversity (Michalopoulos, 2012), finding that their inclusion does not change the results (Appendix Table B35).

5.1.3 Additional Robustness Checks

Alternative definitions of conflict. Some recent contributions to the conflict literature (for instance, Esteban, Mayoral and Ray, 2012, and Huber and Mayoral, 2013) use an alternative database on conflict from the Peace Research Institute Oslo (PRIO). 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." Appendix Table B37, column 1, reports the results for the PRIO dataset. With the 25 battle deaths threshold, we find a standardized effects of χ^2 and CF that are substantially larger than in our baseline (respectively 26% and -20%).

Rather than looking at the incidence of conflict, some papers have analyzed the determinants of onset of conflict. The results for onset, reported in Appendix Table B38, 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 1.78%. 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 increases the probability of civil war onset by 27%, 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).

Different subsets of questions. We now analyze whether our results are sensitive to the questions used to compute the different diversity measures. Appendix Table B36 reports the results by question category and type. In Panel A we see that the baseline results are quite robust across question categories: CF enters negatively in 6 of the 7 categories, although the overall results appear to have been driven mostly by questions on religion and morals, since CF based on other categories is not statistically significant at the 5% level. Results for χ^2 are stronger, with this variable entering with a significantly positive sign for 5 of the 7 categories. Since one of the categories relates to questions on national identity, one potential concern is whether our overall results are driven by that category. To exclude that possibility, we rerun our regression with all questions, except category G, and find essentially identical results to the baseline. In Panel B we also see robustness with respect to question type. The results are strongest for binary and scale questions, but χ^2 enters positively and significantly in all three cases.

Next, we expand 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 comes at the cost of greater heterogeneity across countries in the set of questions. The results appear in column 2 of Appendix Table B37. Reassuringly, nothing changes much:

the standardized effects of χ^2 and CF are largely unchanged at 17% and -8%, respectively.[26] Finally, we go the opposite way, focusing on nine questions often used in the literature on culture and economics (the questions are those listed in footnote 18). Again, the results are unchanged (column 3 of Appendix Table B37).

Different measures of diversity. We examine the robustness of our findings to the inclusion of different measures of diversity. The results are reported in Appendix Table B39. First, Arbatli, Ashraf and Galor (2015) found that genetic diversity is a strong predictor of civil conflict. We therefore run a robustness check where we add their measure of genetic diversity to our baseline regression (column 1). Consistent with Arbatli, Ashraf and Galor (2015), we find that genetic diversity indeed increases the incidence of civil conflict, but including this additional regressor does not change our results on CF, ELF and χ^2, neither qualitatively nor quantitatively. Why might CF and genetic diversity bear opposite signs? As we discussed in Section 4.5, these two measures capture distinct dimensions of diversity, based respectively on cultural traits and genes. Arbatli, Ashraf and Galor (2015) argue that genetic diversity may reflect low levels of trust and interpersonal cooperation. In contrast the pacifying effect of CF may be due to higher tolerance in societies where respondents feel free to express a wide range of values, attitudes and opinions. It is also possible, as suggested before, that a large degree of within-group cultural diversity makes it difficult for groups to agree, mobilize and fight.

Second, we estimate our baseline regression using F_{ST} rather than χ^2 as the measure of overlap between culture and ethnicity (column 2). 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%.

Third, Montalvo and Reynal-Querol (2005) have argued that polarization may be a better predictor of civil conflict than fractionalization. We therefore replace our standard CF and χ^2 measures with their polarization equivalents (column 3). While cultural polarization continues to reduce the incidence of conflict, the polarization overlap measure is only significant at the 6% level, though its standardized effect is 33%, substantially larger than that of χ^2.

Fourth, we use alternative measures of cultural fractionalization and overlap that incorporate distances between answers (column 4). The results continue to be very similar. In particular, the overlap measure remains statistically significant at the 1% level, and its standardized effect is 13%, similar to the baseline. It may not be surprising that the results do not change much, because accounting for distances only affects scale questions. Therefore, as a further robustness we consider diversity measures based exclusively on those scale questions (column 4). The results are similar, as the standardized effect of the distance overlap measure is 11%.

Ethnic and linguistic classification. As explained in Section 2, we characterize ethnic identity by relying on either ethnic or linguistic classifications. To further investigate whether one classification leads to different results from the other, we rerun the conflict regressions including a series of interaction terms

[26] 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.

between the diversity measures and a dummy equal to 1 if we relied on the ethnic classification for the corresponding country, and 0 if we relied on the linguistic classification. As can be seen in Appendix Table B42 (column 1), the results are unchanged: χ^2 continues to reduce the incidence of conflict, with a coefficient that is statistically significant at the 1% level, whereas the interaction terms are not statistically different from zero. As a further robustness check, we reran the conflict regressions, excluding either Latin America (where the classification is racial rather than linguistic) or sub-Saharan Africa (where the classification is linguistic rather than racial). The results are virtually unchanged (columns 2 and 3).

Alternative estimation methods. We also experiment with alternative estimation methods. Appendix Table B43 reports the results. Following most of the literature, in our baseline we used a dynamic panel specification. We now analyze whether our results change when we leave out the lagged dependent variable (column 1). As can be seen, the χ^2 continues to be statistically significant at the 1% level, but its standardized effect increases to more than 60%. Another option is to completely remove the time-dimension by retaining one observation per country. To do so, we define the dependent variable as the share of years a country is in a civil conflict (column 2). Again, the coefficient on our χ^2 index continues to be statistically significant, now at the 5% level. As further robustness checks, we also use a rare events logit estimator (column 3), a linear probability specification (column 4) and a dynamic panel based on decade averages rather than yearly data (column 5). As in all our previous regressions, an increase in χ^2 is associated with higher incidence of war, and CF bears a negative sign.

Geography and ethnicity. One relevant question is to what extent ethnic and cultural differences are related to regional differences. To investigate this question, we compute measures of regional fractionalization and of the overlap between regions and culture. That is, instead of using an individual's ethnicity as his identity, we use his region. Appendix Table B44 shows the results for two additional regressions. The first regression substitutes the χ^2 and the ELF measures based on ethnicities with the same measures based on regions (column 2). We find that the overlap between culture and regions has no statistically significant effect on the incidence of conflict. The second regression includes simultaneously the χ^2 and the ELF measures based on ethnicities and based on regions (column 3). We find that only the measures based on ethnicities are statistically significant. In particular, the overlap between culture and ethnicity continues to be associated with an increase in the incidence of conflict, and the effect is statistically significant at the 1% level.

5.2 Public Goods

Cultural and ethnic diversity are likely to also affect the provision of public goods. There is a large literature on the relationship between ethnic diversity and public goods provision. Alesina, Baqir and Easterly (1999) show that the share of spending on public goods across geographic units in the United States is inversely related to ethnic fractionalization. Miguel and Gugerty (2005) and Banerjee et al. (2005) find similar results for Kenya and India, respectively. La Porta et al. (1999) document this inverse relationship in a cross-section of countries. Alesina, Baqir and Hoxby (2004) discuss two different channels for why diversity may negatively

affect public goods: people from different ethnic groups may have different preferences over public goods, or even if they have the same preferences, they may dislike sharing the public goods with other groups. The first channel relates to cultural fractionalization, the second channel to ethnic fractionalization. Baldwin and Huber (2010) suggest a third channel, showing empirically that income inequality between ethnic groups is negatively associated with public goods provision in a cross-section of 46 countries. We ask instead whether cultural heterogeneity between groups matters for public goods provision.

5.2.1 Data and Specification

Taking a similar approach as in our conflict regressions, we now analyze how public goods provision is related to our three measures of diversity, CF, ELF and χ^2:

$$G_c = \beta_0 + \beta_1 CF_c + \beta_2 ELF_c + \beta_3 \chi^2_c + \boldsymbol{\beta}_4' \mathbf{Z}_c + \varepsilon_c \tag{14}$$

To capture public goods G_c, we use a variety of indicators. To tie our hands ex-ante on the choice of indicators, we start from the same 10 variables as in Desmet, Ortuño-Ortín and Wacziarg (2012). They cover a wide spectrum of public goods, including education, health and infrastructure. In our baseline analysis, we construct an index that is increasing in the amount and quality of public goods, by taking the first principal component of 8 of these underlying variables (we denote this index by $PG8$) and using it as our measure for G_c. In additional robustness checks, we consider each variable separately, as well as the first principal component of all 10 variables ($PG10$).[27] Our specification also controls for a wide and varying set of right-hand side variables, such as legal origin, regional dummies, per capita income, latitude and geographic factors.

5.2.2 Results

Table 6 reports results on the determinant of $PG8$. Consistent with our analysis of civil conflict, we find that an increase in χ^2 worsens public goods provision, with the relevant coefficient statistically significant at the 5% level in *all* specifications. Moreover, an increase in CF is usually associated with an improvement in public goods, whereas ELF does not have a significant effect. To assess the economic significance of the coefficients, we focus on the specification of column 1. The standardized beta on χ^2 (i.e. the effect of a one standard deviation increase in χ^2 as a fraction of a one standard deviation change in child mortality) is -43.4% and the corresponding standardized effect of CF is 38.7%. The remaining columns of Table 6 modify the specification, including fewer controls (column 2), adding log population and log income (column 3), substituting colonial origin dummies for legal origins (column 4), broadening the set of regional

[27] The 10 variables under consideration are: infant survival rate, measles immunization rate, hospital beds per capita, literacy rate, log school attainment, percentage access to improved sanitation, percentage access to improved water, a measure of infrastructure quality, road length per capita and railway length per capita. All these variables are from the World Bank's World Development Indicators, except for school attainment, which comes from the Barro-Lee database. The first principal component of 8 of these indicators (all but the literacy rate and railway length per capita) accounts for 66.4% of the variance in the data, and this index is available for 99 countries. The first principal component of all 10 indicators accounts for 61.2% of the variance in the data, and is available for 62 countries. Due to the availability of a larger sample for the first index, it is our focus for the bulk of the analysis.

effects (column 5) and adding a wide range of geographic and climatic controls (column 6). None of these modifications change the inferences drawn from the baseline specification.

We conducted further robustness tests on these public goods results, which broadly mirror those conducted for civil conflict. Appendix Table B46, column 1 uses the first principal component of all 10 measures of public goods. The sample falls from 59 to 42 countries, so while the magnitudes and signs of the coefficients are preserved, they become less significant (the p-value on χ^2 is slightly above 5%). The remaining columns revert to the original dependent variable, but consider measures of heterogeneity based on varying sets of questions - first expanding the set of questions, and then restricting questions to those from each WVS/EVS category. Expanding the set of questions has no effect on the results, and the coefficient on χ^2 is significant and negative for 5 of the 7 question categories. Appendix Table B47 further examines the robustness of the results to breaking down the underlying questions by type (binary, scale, unordered), considering distance and polarization measures, and splitting the sample between countries where ethnic identity comes from the WVS/EVS linguistic classification and those where it comes from the ethnicity classification. The overlap measure based on polarization is not statistically significant, indicating that our overlap measures based on fractionalization are better predictors of public goods provision. None of the other modifications have a material impact on our basic results: the overlap between culture and ethnicity is negatively related to public goods provision.

Appendix Tables B48 and B49 report the results using each of the 10 measures of public goods separately. The first table considers a parsimonious specification while the second adds many geographic controls. In both cases, χ^2 is a negative correlate of public goods at the 5% significance level for 6 out of the 10 public goods, and in a few other cases it is significant at the 10% level. We also find a general pattern of a positive effect of CF, and an insignificant effect of ELF. Thus, despite differences in the dependent variables, these regressions broadly confirm the pattern of correlations found with the summary measure of public goods based on the first principal component. In sum, our results on civil conflict broadly apply to public goods provision.[28]

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 WVS, 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

[28] We also examined the effect of our different diversity measures on income per capita. The results are presented in Appendix Table B50. They confirm our main findings: CF is positively and significantly related to income per capita, χ^2 is negatively and significantly related to income per capita, and ethnic fractionalization is not robustly correlated with income per capita. We do not wish, however, to make too much of these results, particularly because we believe that causality is likely to run in both directions. Of particular concern is the possibility that income affects prevalent cultural values over the long run. In contrast, for conflict and public goods, our results are robust to including per capita income on the right hand side.

several 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 and public goods. We found that ethnic fractionalization has no predictive power for these outcomes, but that cultural diversity, if anything, serves to reduce conflict and increase public goods provision. The measures of overlap between culture and ethnicity, χ^2 and F_{ST}, have a robust positive effect on the probability of civil conflict onset and incidence, and a negative impact on public goods. Thus, ethnic divisions matter for conflict and public goods when they are associated with cultural differences across ethnic groups.

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 him 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 for civil conflict and public goods.

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. 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 the χ^2 and F_{ST} indices tends to be small, indicating that the extent to which ethnicity is informative for culture 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. We found, moreover, that economic and political modernization may help to sever the link between culture and ethnicity. 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., R. Baqir and W. Easterly (1999), "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 114(4): 1243-1284.

Alesina, A. R. Baqir and C. Hoxby (2004), "Political Jurisdications In Heterogeneous Communities," *Journal of Political Economy*, 112(2): 348-396.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat and R. Wacziarg (2003), "Fractionalization," *Journal of Economic Growth*, 8(2): 155-194.

Alesina, A., J. Harnoss and H. Rapoport (2016), "Birthplace Diversity and Economic Prosperity," *Journal of Economic Growth*, 21(2): 101-138.

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*, 128(2): 469-530.

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

Alesina, A., S. Michalopoulos and E. Papaioannou (2016), "Ethnic Inequality," *Journal of Political Economy*, 124(2): 428-488.

Alesina, A. and E. Zhuravskaya (2011), "Segregation and the Quality of Government in a Cross Section of Countries," *American Economic Review*, 101(5): 1872-1911.

Algan, Y. and P. Cahuc (2010), "Inherited Trust and Growth," *American Economic Review*, 100(5): 2060-2092.

Arbatli, C. E., Q. Ashraf and O. Galor (2015), "The Nature of Civil Conflict," *working paper*, Brown University.

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

Ashraf, Q. and O. Galor (2013b), "Genetic Diversity and the Origins of Cultural Fragmentation," *American Economic Review Papers and Proceedings*, 103(3): 528-533.

Baldwin, K. and J. D. Huber (2010), "Economic versus Cultural Differences: Forms of Ethnic Diversity and Public Goods Provision," *American Political Science Review*, 104(4): 644-662.

Banerjee, A., Iyer, L. and Somanathan, R. (2005), "History, Social Divisions, and Public Goods in Rural India," *Journal of the European Economic Association,* 3(2-3): 639-647.

Barro, R. and R. McCleary (2003), "Religion and Economic Growth across Countries", *American Sociological Review,* 68(5): 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," *PNAS*, 106(2): 17671-17674.

Bossert, W., C. D'Ambrosio and E. La Ferrara (2011), "A Generalized Index of Fractionalization," *Economica*, 78(312): 723-750.

Boyd, R. and P. J. Richerson (1985), *Culture and The Evolutionary Process*, Chicago: The University of Chicago Press.

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

Cavalli-Sforza, L. L. and M. W. Feldman (1981), *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton: Princeton University Press.

Collier, P. and A. Hoeffler (2004), "Greed and Grievance in Civil War," *Oxford Economic Papers*, 56(4): 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*, 97(2): 322-338.

Desmet, K., I. Ortuño-Ortín and S. Weber (2009), "Linguistic Diversity and Redistribution," *Journal of the European Economic Association*, 7(6): 1291-1318.

Duclos, J. Y., J. M. Esteban and D. Ray (2004), "Polarization: Concepts, Measurement, Estimation," *Econometrica*, 72(6): 1737-1772.

Easterly, W. and R. Levine (1997), "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, 112(4): 1203-1250.

Edwards, A. W. F. (2003), "Human Genetic Diversity: Lewontin's Fallacy," *BioEssays*, 25(8): 798-801.

Esteban, J. M., and D. Ray (1994), "On the Measurement of Polarization," *Econometrica*, 62(4): 819-851.

Esteban, J. M., and D. Ray (2011), "A Model of Ethnic Conflict." *Journal of the European Economic Association*, 9(3): 496-521.

Esteban, J., L. Mayoral and D. Ray (2012), "Ethnicity and Conflict: An Empirical Study," *American Economic Review*, 102(4): 1310-1342.

Fearon, J. (2003), "Ethnic and Cultural Diversity by Country", *Journal of Economic Growth*, 8(2): 195-222.

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

Fernandez, R. and A. Fogli (2009), "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics,* 1(1): 146-177.

Fischer, R. and S. Schwartz (2010), "Whence Differences in Value Priorities? Individual, Cultural, or Artifactual Sources," *Journal of Cross-Cultural Psychology,* 20(10): 1-18.

Giuliano, P. (2007), "Living Arrangements in Western Europe: Does Cultural Origin Matter?" *Journal of the European Economic Association*, 5(5): 927-952.

Guiso, L., P. Sapienza and L. Zingales (2009), "Cultural Biases in Economic Exchange?" *Quarterly Journal of Economics*, 124(3): 1095-1131.

Gubler, J. R. and J. Selway (2012), "Horizontal Inequality, Crosscutting Cleavages, and Civil War", *Journal of Conflict Resolution*, 56(2): 206-232.

Hinds D. A. et. al. (2005), "Whole-Genome Patterns of Common DNA Variation in Three Human Populations," *Science* 307(5712): 1072-1079.

Hong, L. and S. Page (2004), "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers," *PNAS*, 101(46): 16385-16389.

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*, 193(2): 512-528.

Jorde, L.B. et al. (2000), "The Distribution of Human Genetic Diversity: A Comparison of Mitochondrial, Autosomal, and Y-Chromosome Data". *The American Journal of Human Genetics* 66(3): 979-88.

Jost, L. (2008), "GST and Its Relatives Do Not Measure Differentiation," *Molecular Ecology*, 17(18): 4015-4026.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R. Vishny (1999), "The Quality of Government," *Journal of Law, Economics and Organizations* 15(1): 222-279.

Lewontin, R. (1972), "The Apportionment of Human Diversity," *Evolutionary Biology*, 6: 381-398.

Luttmer, E. F. P., and M. Singhal (2011), "Culture, Context, and the Taste for Redistribution," *American Economic Journal: Economic Policy*, 3(1): 157-79.

Meirmans P. G. and P. W. Hedrick (2011), "Assessing Population Structure: F_{ST} and Related Measures," *Molecular Ecology Resources*, 11(1): 5-18.

Miguel, E. and Gugerty, M.K. (2005), "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics*, 89(11): 2325-2368.

Michalopoulos, S. (2012), "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 102(4): 1508-1539.

Montalvo, J. G. and M. Reynal-Querol (2005), "Ethnic Polarization, Potential Conflict and Civil War," *American Economic Review*, 95(3): 796-816.

Nei, M. (1973), "Analysis of Gene Diversity in Subdivided Populations," *PNAS*, 70(12): 3321-3323.

Ottaviano, G. and G. Peri (2006), "The Economic Value of Cultural Diversity: Evidence from US Cities." *Journal of Economic Geography*, 6(1): 9-44.

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

Reardon, S. and G. Firebaugh (2002), "Measures of Multi-Group Segregation," *Sociological Methodology*, 32: 33-67.

Reynal-Querol, M. (2002), "Ethnicity, Political Systems, and Civil Wars," *Journal of Conflict Resolution*, 46(1): 29-54.

Richerson, P. J. and R. Boyd (2005), *Not by Genes Alone*, Chicago: University of Chicago 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,* 280(1756).

Selway, J. (2010), "Cross-Cuttingness, Cleavage Structures and Civil War Onset," *British Journal of Political Science,* 41(1): 111-138.

Selway, J. (2011), "The Measurement of Cross-Cutting Cleavages and Other Multidimensional Cleavage Structures," *Political Analysis*, 19(1): 48-65.

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

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

Tabellini, G. (2010), "Culture and Institutions: Economic Development in the Regions of Europe," *Journal of the European Economic Association*, 8(4): 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*, 22(1): 24-49.

Wright, S. (1949), "The Genetical Structure of Populations," *Annals of Human Genetics*, 15(1): 323-354.

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	21,467	0.430	2.681	4.065	1.384
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	5,654	0.572	2.334	4.486	2.152
Of which: East and Southeast Asia	2,088	0.626	2.092	4.526	2.434
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	4,380	0.425	3.238	4.576	1.338
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	4,550	0.427	2.836	4.227	1.391
Binary from unordered response questions	7,029	0.362	1.616	2.707	1.091
Scale questions	9,888	0.479	3.367	4.956	1.589

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, F_{ST} and χ^2

Panel A: Summary Statistics (based on 76 observations)

	Mean	Standard Deviation	Minimum	Maximum
Cultural Fractionalization	0.529	0.037	0.427	0.602
ELF	0.390	0.258	0.000	0.852
F_{ST}	0.012	0.011	0.000	0.059
χ^2	0.029	0.025	0.000	0.128

Panel B: Correlations (based on 76 observations)

	Cultural Fractionalization	ELF	F_{ST}	χ^2
Cultural Fractionalization	1			
ELF	-0.030	1		
F_{ST}	0.179	0.620**	1	
χ^2	0.219	0.620**	0.981**	1

(** Significant at the 1% level)

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

	CF	F_{ST}	χ^2
Breakdown by Question Category			
SECTION A: PERCEPTIONS OF LIFE	0.414	0.012	0.020
SECTION B: ENVIRONMENT	0.596	0.011	0.026
SECTION C: WORK	0.549	0.011	0.026
SECTION D: FAMILY	0.516	0.012	0.028
SECTION E: POLITICS AND SOCIETY	0.612	0.011	0.031
SECTION F: RELIGION AND MORALE	0.525	0.014	0.043
SECTION G: NATIONAL IDENTITY	0.578	0.017	0.047
Breakdown by Question Type			
BINARY	0.350	0.013	0.013
UNORDERED RESPONSE QUESTIONS	0.574	0.012	0.026
SCALE	0.597	0.011	0.037
Restricted Set of 9 Questions Used in Literature			
9 QUESTIONS FROM LITERATURE	0.345	0.011	0.012
Breakdown by Size of Town			
MORE THAN 500,000 inhabitants	0.549	0.016	0.050
LESS THAN 500,000 inhabitants	0.544	0.012	0.035
Factor Analysis			
Factor 1	0.442	0.024	0.024
Factor 2	0.287	0.028	0.028
Factor 3	0.405	0.018	0.018
Factor 4	0.434	0.023	0.023
Factor 5	0.344	0.019	0.019
Average of 5 Factors	0.382	0.022	0.022

This panel displays a breakdown of the means of the various diversity measures by WVS question category, by question type (binary / unordered / scale), for a restricted set of 9 questions often used in the existing literature, by the size of the town where the interview was conducted and finally by factor, where factors were obtained from factor analysis over WVS questions (Section 4.3.1). Additional results and breakdowns are shown in Appendix Table B24.

Table 4 – Correlations of Different Variables with CF, ELF and χ^2

	Cultural Fractionalization	χ^2	ELF
Other Measures of Diversity			
Genetic Diversity	-0.175	0.186	0.010
Linguistic Diversity (Alesina et al., 2003)	0.018	0.579**	0.647**
Ethnic Diversity (Alesina et al., 2003)	-0.078	0.382**	0.736**
Religious Diversity (Alesina et al., 2003)	0.314**	0.086	0.109
RQ Ethnolinguistic Polarization	-0.066	0.188	0.503**
ER Ethnolinguistic Polarization	0.036	0.240*	0.206
Ethnic Diversity (Fearon, 2003)	0.014	0.432**	0.726**
Ethnic Greenberg (Fearon, 2003)	0.018	0.492**	0.572**
Ethnic Inequality GREG	-0.175	0.127	0.241*
Ethnic Inequality Ethnologue	-0.181	0.181	0.394**
Ethnic segregation	-0.250	0.207	0.446**
Linguistic segregation	-0.208	0.058	0.330*
Religious segregation	-0.214	0.187	0.264
Percentage Protestant	0.212	-0.044	-0.086
Percentage Catholic	0.258*	-0.264*	-0.034
Percentage Muslim	-0.597**	0.082	0.160
Geography			
Absolute Latitude	0.180	-0.355***	-0.529**
Area	0.097	-0.082	0.083
Roughness	0.025	-0.007	-0.060
Sub-Saharan Africa	0.083	0.234*	0.382**
Middle East & North Africa	-0.529**	-0.094	0.055
Europe & Central Asia	0.267*	-0.188	-0.396**
South Asia	-0.262*	0.263*	0.157
East Asia & Pacific	-0.026	0.219	-0.062
North America	0.126	-0.055	0.058
Latin America & Caribbean	0.112	-0.230	0.097
Economics & Development			
Log Population 1990-2010	-0.154	0.075	0.138
Log GDP per Capita 1990-2010	0.373***	-0.287**	-0.368**
Institutions			
Democracy	0.598***	-0.125	-0.312**
UK Legal Origin	0.080	0.443***	0.360**
French Legal Origin	-0.159	-0.226	-0.046
German Legal Origin	0.059	-0.158	-0.230
Scandinavian Legal Origin	0.101	-0.105	-0.224

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

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

	(1)	(2)	(3)	(4)	(5)	(6)
	CF only	**ELF Only**	**χ^2 only**	**Baseline**	**Add legal origins**	**Add GDP growth**
Cultural Fractionalization	**-0.187**			**-0.391****	**-0.472****	**-0.417****
	[-1.601]			**[-3.330]**	**[-3.490]**	**[-3.645]**
Ethnolinguistic Fractionalization		**0.019**		**-0.037**	**-0.041**	**-0.035**
		[1.092]		**[-1.769]**	**[-1.882]**	**[-1.779]**
Chi-square			**0.472****	**0.870****	**0.992****	**0.840****
			[2.738]	**[3.823]**	**[4.003]**	**[3.959]**
Lagged War	0.860**	0.868**	0.864**	0.840**	0.840**	0.833**
	[30.660]	[33.743]	[33.325]	[27.877]	[26.284]	[25.778]
Log Lagged GDP per capita	-0.005	-0.008	-0.007	0.003	0.005	0.003
	[-0.895]	[-1.736]	[-1.473]	[0.621]	[0.982]	[0.778]
Log Lagged Population	0.011**	0.012**	0.013**	0.012**	0.015**	0.011**
	[3.687]	[3.979]	[4.330]	[4.183]	[4.316]	[4.322]
Fraction Mountainous	0.041*	0.040*	0.041*	0.041**	0.034	0.032*
	[2.532]	[2.341]	[2.345]	[2.602]	[1.838]	[2.024]
Country with Non-Connected Territories	0.025	0.017	0.015	0.026	0.025	0.025
	[1.607]	[1.273]	[1.164]	[1.596]	[1.612]	[1.536]
Oil	0.020	0.021	0.025	0.022	0.014	0.015
	[1.561]	[1.497]	[1.797]	[1.618]	[1.343]	[1.330]
New State	0.200*	0.222*	0.240*	0.210*	0.255*	0.143*
	[2.144]	[2.287]	[2.358]	[2.143]	[2.157]	[2.109]
Instability	-0.009	-0.009	-0.009	-0.009	-0.010	-0.010
	[-1.001]	[-0.941]	[-0.997]	[-1.080]	[-1.034]	[-1.338]
Democracy Lagged (Polity 2, range -1 to 1)	0.006	0.004	0.001	0.003	0.006	0.003
	[0.902]	[0.561]	[0.205]	[0.490]	[0.746]	[0.473]
UK Legal Origin					0.009	
					[0.263]	
French Legal Origin					0.032	
					[0.965]	
Socialist Legal Origin					0.010	
					[0.296]	
GDP Growth						-0.119**
						[-3.497]
GDP Growth Lagged						-0.003
						[-0.080]
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

Robust z statistics in brackets, based on standard errors clustered at the country level. * significant at 5%; ** significant at 1%. Columns report marginal effects. All columns include dummies for Latin America and Caribbean; Sub-Saharan Africa; East and Southeast Asia.

This table reports estimates from a logit regression of a binary indicator of civil conflict on 3 diversity measures as well as a series of control variables. The data is a panel of most 69 countries from 1945 to 1999.

Table 6: Public Goods Provision and Diversity

	(1)	(2)	(3)	(4)	(5)	(6)
	Benchmark specification	**Removing legal origins**	**Income and population**	**Colonial origins**	**WB region dummies**	**Geography controls**
Cultural Fractionalization	**21.520** **[2.925]**	**22.386** **[3.493]**	**7.607** **[1.558]**	**21.847** **[3.674]**	**11.465*** **[2.239]**	**26.183** **[4.005]**
Ethnolinguistic Fractionalization	**1.478** **[1.720]**	**1.216** **[1.457]**	**0.893** **[1.354]**	**0.504** **[0.612]**	**0.975** **[1.485]**	**1.075** **[1.102]**
Chi-square	**-34.303**** **[-2.880]**	**-25.501*** **[-2.155]**	**-15.572*** **[-2.159]**	**-29.098*** **[-2.358]**	**-19.498**** **[-2.741]**	**-29.631*** **[-2.182]**
Absolute latitude	0.046* [2.406]	0.048* [2.563]	0.042** [3.250]	0.047* [2.432]	0.055** [3.236]	-0.035 [-0.711]
Log Population (1990 - 2010) average			-0.269** [-4.904]			
Log GDP per capita (1990 - 2010) average			0.647** [5.959]			
Constant	-11.106** [-3.025]	-12.430** [-4.315]	-5.364* [-2.364]	-11.862** [-4.313]	-6.367* [-2.391]	-5.967 [-1.420]
Observations	59	59	59	59	59	57
R-squared	0.774	0.727	0.882	0.794	0.882	0.849

Robust t-statistics in brackets; * significant at 5%; ** significant at 1%.

The dependent variable is the first principal component of 8 indicators of public goods quality and quantity: infant mortality rate, measles immunization rate, hospital beds per capita, log school attainment, percentage access to improved sanitation, percentage access to improved water, a measure of infrastructure quality and railway length per capita. These variables are averaged over 1990 to 2010. Estimation is by OLS.

All specifications except column (5) include regional dummies defined as follows: Latin America and Caribbean; Sub-Saharan Africa; East and Southeast Asia. For column (5), region dummies are from the World Bank regional nomenclature: Sub-Saharan Africa region; Middle East and North Africa region; Europe and Central Asia region; South Asia region; East Asia and Pacific region; Latin America and Caribbean region. North America is the omitted region.

All specifications except those of columns (2) and (4) include legal origins dummies: French legal origin, German legal origin, Scandinavian legal origin.

Column (4) includes colonial past dummies instead of legal origins dummies: Former French colony, former British colony, former Spanish colony and former Soviet colony.

Column (6) includes a set of additional geographic controls: Soil fertility, terrain ruggedness, mean elevation, island dummy, landlocked country dummy, log land area, temperature and precipitation.