CESifo WORKING PAPERS

6989 2018

April 2018

Diversity and Conflict

Cemal Eren Arbatlı, Quamrul H. Ashraf, Oded Galor, Marc Klemp

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6989*
Category 6: Fiscal Policy, Macroeconomics and Growth

Diversity and Conflict

Abstract

This research advances the hypothesis and establishes empirically that interpersonal population diversity has contributed significantly to the emergence, prevalence, recurrence, and severity of intrasocietal conflicts. Exploiting an exogenous source of variations in population diversity across nations and ethnic groups, it demonstrates that population diversity, as determined predominantly during the exodus of humans from Africa tens of thousands of years ago, has contributed significantly to the risk and intensity of historical and contemporary internal conflicts, accounting for the confounding effects of geographical, institutional, and cultural characteristics, as well as for the level of economic development. These findings arguably reflect the adverse effect of population diversity on interpersonal trust, its contribution to divergence in preferences for public goods and redistributive policies, and its impact on the degree of fractionalization and polarization across ethnic, linguistic, and religious groups.

JEL-Codes: D740, N300, N400, O110, O430, Z130.

Keywords: social conflict, population diversity, ethnic fractionalization, ethnic polarization, interpersonal trust, political preferences.

Cemal Eren Arbatlı
Faculty of Economic Sciences
National Research University
Higher School of Economics
Moscow / Russia

Quamrul H. Ashraf
Department of Economics
Williams College
USA – Williamstown, MA 01267
Quamrul.H.Ashraf@williams.edu

Oded Galor
Department of Economics
Brown University
USA – Providence, RI 02912
Oded.Galor@brown.edu

Marc Klemp
Department of Economics
University of Copenhagen
Denmark – 1353 Copenhagen K
marc.klemp@econ.ku.dk

March 2018

We thank the editor, four anonymous referees, Ran Abramitzky, Alberto Alesina, Yann Algan, Sascha Becker, Moshe Buchinsky, Matteo Cervellati, Carl-Johan Dalgaard, David de la Croix, Emilio Depetris-Chauvin, Paul Dower, Joan Esteban, James Fenske, Raquel Fernández, Boris Gershman, Avner Greif, Pauline Grosjean, Elhanan Helpman, Murat Iyigun, Noel Johnson, Garett Jones, Mark Koyama, Stelios Michalopoulos, Steven Nafziger, Nathan Nunn, John Nye, Ömer Özak, Elias Papaioannou, Sergey Popov, Stephen Smith, Enrico Spolaore, Uwe Sunde, Mathias Thoenig, Nico Voigtländer, Joachim Voth, Romain Wacziarg, Fabian Waldinger, David Weil, Ludger Wößmann, Noam Yuchtman, Alexei Zakharov, and seminar participants at George Mason University, George Washington University, HSE/NES Moscow, the AEA Annual Meeting, the conference on "Deep Determinants of International Comparative Development" at Brown University, the workshop on "Income Distribution and Macroeconomics" at the NBER Summer Institute, the conference on "Culture, Diversity, and Development" at HSE/NES Moscow, the conference on "The Long Shadow of History: Mechanisms of Persistence in Economics and the Social Sciences" at LMU Munich, the fall meeting of the NBER Political Economy Program, the session on "Economic Growth" at the AEA Continuing Education Program, the workshop on "Biology and Behavior in Political Economy" at HSE Moscow, and the Economic Workshop at IDC Herzliya for valuable comments. Gregory Casey, Ashwin Narayan, Daniel Prinz, and Jeffrey Wang provided excellent research assistance. Arbatlı acknowledges research support from the Center for Advanced Studies (CAS) at the National Research University Higher School of Economics (HSE). Ashraf acknowledges research support from the NSF (SES-1338738), the Hellman Fellows Program, and the Oakley Center for Humanities and Social Sciences at Williams College. Galor acknowledges research support from the NSF (SES-1338426) and the Population Studies and Training Center (PSTC) at Brown University. The PSTC receives core support from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (5R24HD041020). Klemp acknowledges research support from the Carlsberg Foundation, the Danish Research Council (grant numbers 1329-00093 and 1327-00245), and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Global Fellowship grant agreement number 753615).

1 Introduction

Over the course of the 20th century, in the period following World War II, civil conflicts have been responsible for more than 16 million casualties worldwide, well surpassing the cumulative loss of human life associated with international conflicts. Nations plagued by civil conflict have experienced significant fatalities from violence, substantial loss of productive resources, and considerable declines in their standards of living. While the number of countries experiencing conflict has declined from its peak in the early 1990s, as many as 35 nations have been afflicted by the prevalence of civil conflict since 2010, and more than a quarter of all nations encountered the incidence of civil conflict for at least a decade during the 1960–2013 time horizon.[1]

This research explores the origins of the prevailing variation in the emergence, prevalence, recurrence, and severity of intrasocietal conflicts across countries, regions, and ethnic groups, highlighting one of their deepest roots, molded during the dawn of the dispersion of anatomically modern humans across the globe and by its differential impact on the level of population diversity across regions. Exploiting an exogenous source of variations in population diversity across nations and ethnic groups, the study advances the hypothesis and establishes empirically that interpersonal population diversity, as determined predominantly during the exodus of humans from Africa tens of thousands of years ago, has contributed significantly to the risk and intensity of historical as well as contemporary conflicts. It argues that the adverse effect of interpersonal population diversity on interpersonal trust and cooperation, its contribution to divergence in preferences for public goods and redistributive policies, and its impact on the degree of fractionalization and polarization across ethnic, linguistic, and religious groups have fostered social, political, and economic disorder and have, thus, magnified the vulnerability of society to internal conflicts.

Population diversity at the national or subnational level may contribute to intergroup as well as intragroup conflicts through several mechanisms. First, population diversity may have an adverse effect on the prevalence of mutual trust and cooperation, and excessive diversity could therefore depress the level of social capital below a threshold that could have averted the emergence of social, political, and economic grievances or prevented the culmination of such grievances to violent hostilities. Second, to the extent that population diversity captures interpersonal divergence in preferences for public goods and redistributive policies, overly diverse societies may find it difficult to reconcile such differences through collective action, thereby intensifying their susceptibility to internal antagonisms. Third, insofar as population diversity reflects interpersonal heterogeneity in traits that are differentially rewarded by the geographical, institutional, and technological environment, it can potentially cultivate resentments that are rooted in economic inequality, thereby magnifying society's vulnerability to internal belligerence.

Moreover, the prehistorical variation in the level population diversity across regions and its potential role in facilitating the formation of ethnic groups may have contributed to the emergence of social conflicts. In particular, following the "out of Africa" migration of humans, the initial endowment of population diversity in each region may have catalyzed the process of group formation, reflecting the trade-off associated with the size of each group. While a larger group may benefit from economies of scale, its productivity tends to be affected adversely by its incohesiveness. Thus, in light of the detrimental impact of population diversity on social cohesiveness, a larger initial endowment of population diversity could have plausibly led to the emergence of a larger number of groups, and due to the forces of "cultural drift" and "biased transmission" of cultural markers (e.g., language dialects, customs, traditions, and social norms), intergroup cultural divergence

[1]These estimates are based on the UCDP/PRIO Armed Conflict Dataset, Version 4-2014a (Gleditsch et al., 2002; Themnér and Wallensteen, 2014), compiled by the Uppsala Conflict Data Program (UCDP) and the Peace Research Institute Oslo (PRIO).

would have become more pronounced, thus leading to the formation of distinct collective identities along ethnic lines. The emergent fragmentation, in turn, could have plausibly facilitated intergroup conflict, fueling excessive intergroup competition and dissension, and creating fertile grounds for the use of a divide-and-rule strategy by political elites.

The empirical exploration of the pivotal contribution of interpersonal population diversity to conflicts within nations and ethnic groups necessitates a measure that encompasses the various dimensions of population diversity. While the level of population diversity at the national level can be partly captured by ethnolinguistic fractionalization and polarization, these measures predominantly reflect only the proportional representation of ethnic groups in the population (and, rarely, pairwise distances amongst ethnic groups), abstracting from the importance of the degree of interpersonal diversity *within* each ethnic group to the overall level of diversity in the national population. The underlying assumption in the construction of these measures is that ethnic groups are homogenous, despite overwhelming evidence for the existence of substantial diversity within rather than across ethnic groups. These measures are therefore unable to capture the contribution of the various layers of population diversity to conflict within a national population. Moreover, by construction, they cannot shed light on the role of population diversity within an ethnic group on intragroup conflict.

In contrast, the degree of population diversity within each ethnic group, as determined in the course of the exodus of anatomically modern humans from Africa tens of thousands of years ago, could shed light on the effect of interpersonal population diversity on not only intragroup conflicts but conflicts across ethnic groups as well. This measure captures the probability that two individuals, selected at random from the relevant population, are genetically different from one another with respect to a given spectrum of traits. Moreover, measures of predicted genetic diversity for national populations permits the exploration of the role of the various layers of population diversity on social conflict, reflecting all three dimensions of heterogeneity at the country level – the proportional representation of each ethnic group, the pairwise genetic distances that exist amongst these groups, and the degree of interpersonal genetic diversity within each group. Thus, the analysis at the ethnic-homeland level in this research is conducted based on a novel georeferenced data set of ethnic groups for which genetic diversity is either observed or can be predicted, whereas the national-level analysis is conducted based on ancestry-adjusted predicted genetic diversity, as constructed by Ashraf and Galor (2013a).

Exploiting variations across either countries or ethnic homelands, the analysis demonstrates that interpersonal population diversity – determined primarily over the course of the "out of Africa" expansion of anatomically modern humans tens of thousands of years ago – has contributed significantly to the risk and intensity of various forms of contemporary as well as historical intrasocietal conflicts. Furthermore, it uncovers evidence suggesting that the reduced-form influence of population diversity may indeed operate through some of the aforementioned hypothesized mechanisms. In particular, the country-level analysis documents that the contribution of population diversity to intrastate conflicts has plausibly operated via the number of ethnic groups in the population, the prevalence of generalized interpersonal trust, and the degree of dispersion in political preferences.

The dual analysis at the national and at the ethnic-homeland levels has several virtues. First, the focus on nations as well as ethnic groups permits the exploration of the impact of population diversity on the emergence of conflicts in societies of different scales, suggesting that population diversity reduces social cohesion and increases the likelihood of social conflicts within national as well as *subnational* populations. Second, since the boundaries of ethnic homelands largely predate the formation of modern nation states, the ethnic-homeland level analysis mitigates potential concerns regarding the endogeneity of contemporary national boarders to population diversity and internal conflicts (Alesina and Spolaore, 2003). Third, the focus on ethnic groups as well as on national populations permits the analysis to disentangle the impact of population diversity within an ethnic group, from the impact of ethnic diversity across groups, in the emergence of conflicts. Fourth, because the population within an ethnic homeland has been largely native to that location, at least since the precolonial era, the analysis at the ethnicity level circumvents potential concerns about the effect of conflicts on migrations across countries and, thus, about their potential impact on the global distribution of national population diversity.

The research exploits several empirical strategies to mitigate concerns about the potential role of reverse causality, omitted cultural, geographical, and human characteristics, as well as sorting in the observed association between population diversity and intrasocietal conflicts. In particular, the positive association between the extent of *observed* population diversity within an ethnic group and intragroup conflict may reflect reverse causality from conflict to population diversity. It is not inconceivable that in the course of human history, conflicts within ethnic groups have operated towards the homogenization of these populations, reducing their observed levels of diversity. Hence, in order to address this concern, the analysis exploits *predicted* population diversity, rather than observed diversity, to explore the impact of diversity on internal conflicts at the ethnicity level . In particular, since the observed population diversity of a geographically indigenous contemporary ethnic group decreases with distance along ancient migratory paths from East Africa due to the serial founder effect (e.g., Harpending and Rogers, 2000; Ramachandran et al., 2005; Prugnolle, Manica and Balloux, 2005; Ashraf and Galor, 2013a), migratory distance from Africa is employed to predict population diversity for a globally representative sample of more than 1,200 ethnic groups from the *Ethnographic Atlas*.[2]

Furthermore, the association between population diversity and internal conflicts at the ethnicity level may be governed or biased by omitted cultural, geographical, and human characteristics. Thus, in order to mitigate this concern, the empirical analysis implements two related strategies. In light of the serial founder effect, the analysis exploits the migratory distance from Africa to the location of each ethnic group as an instrumental variable for its observed level of population diversity, and as a predictor of population diversity in an extended sample that additionally includes groups for which diversity is unobserved. Nevertheless, there are several plausible scenarios that could a priori weaken the credibility of this methodology. First, selective migration out of Africa, or natural selection along the migratory paths, could have affected human traits and, therefore, conflict independently of the impact of migratory distance from Africa on the degree of *diversity* in human traits. Second, migratory distance from Africa could be correlated with distances from focal historical locations (e.g., technological frontiers) and could, therefore, capture the effect of these other distances on the process of development and the emergence of conflicts, rather than the effect of these migratory distances via population diversity.

[2]Predicted diversity is based on two widely accepted theories from the field of population genetics – the "out of Africa" hypothesis of human origins, and the existence of a serial founder effect associated with the subsequent demic expansion of humans to the rest of the globe. According to the well-established "out of Africa" hypothesis, the human species, having evolved to its anatomically modern form in East Africa more than 200,000 years ago, embarked on populating the entire globe in a stepwise migration process beginning around 90,000–60,000 BP. In addition, the contemporary worldwide distribution of genetic diversity across prehistorically indigenous ethnic groups overwhelmingly reflects a serial founder effect – i.e., a chain of ancient population bottlenecks – originating in East Africa. In particular, because the spatial diffusion of humans to the rest of the world occurred in a series of discrete steps, where in each step, a subgroup of individuals left their parental colony to establish a new settlement farther away, carrying with them only a subset of the genetic diversity of their parental colony, the genetic diversity of a prehistorically indigenous ethnic group as observed today decreases with increasing distance along ancient human migratory paths from East Africa.

These potential concerns are mitigated, however, by the following observations. First, while migratory distance from Africa has a significant negative association with the degree of diversity in traits, it has no association with the *mean* level of traits in a population, such as height, weight, skin reflectance, and IQ (Ashraf and Galor, 2013a), conditional on distance from the equator. Second, conditional on migratory distance from East Africa, migratory distances from historical technological frontiers in the years 1, 1000, and 1500 do not qualitatively alter the impact of predicted diversity on internal conflicts, reinforcing the justification for reliance on the "out of Africa" hypothesis and the serial founder effect for identifying the influence of population diversity on intrasocietal conflicts.

Moreover, a rather unlikely threat to identification would emerge if the actual migratory paths from Africa would have been correlated with geographical characteristics that are directly conducive to conflicts (e.g., soil quality, ruggedness, climatic conditions, and propensity to trade). This would have implausibly necessitated that the conduciveness of these geographical characteristics to conflicts would be aligned not only along the main root of the migratory path out of Africa but also along each of the main forks that emerge from this primary path. The further a location is from Africa, the lower the conduciveness of these geographical factors to conflicts would need to be. In particular, in several important forks of this migration process (e.g., the Fertile Crescent and the associated eastward migration into Asia versus the westward migration into Europe), the geographical characteristics that are conducive to conflicts would have to diminish symmetrically along divergent secondary migratory paths. Nevertheless, in order to further mitigate this implausible concern, the analysis establishes that the results are qualitatively unaffected when it accounts for the potentially confounding influence of a wide range of geographical characteristics of the homeland of each ethnic group. In addition, the analysis accounts for spatial dependence across observations (in both internal conflicts and unobserved heterogeneity) as well as regional fixed effects, capturing time-invariant unobserved heterogeneity in each region and, hence, identifying the association between interpersonal population diversity and internal conflicts across societies from the same geographical region.[3]

The observed association between population diversity and internal conflicts at the ethnic-homeland level may further reflect the sorting of less diverse populations into geographical niches that are less conducive to conflicts. While this unlikely sorting mechanism would not affect the existence of a positive association between population diversity and conflicts, it could render the proposed interpretation of the association less credible. Nevertheless, in light of the serial founder effect and the tight negative association between migratory distance from Africa and population diversity, such sorting would require that the spatial distribution of ex-ante conflict risk would have to be negatively correlated with migratory distance from Africa. As argued above, however, this implausibly necessitates that the conduciveness of geographical characteristics to conflicts would have to be negatively aligned with the primary migratory path out of Africa and with each of the main subsequent forks, diminishing symmetrically along divergent secondary migratory paths. Despite the implausibility of this scenario, concerns regarding the sorting mechanism are further mitigated by the fact that the analysis accounts for heterogeneity in a wide range of geographical characteristics across ethnic homelands, spatial autocorrelation, and regional fixed effects.

Further, to the extent that interregional migration flows in the post-1500 era, and thus the proportional representation of ethnic groups within each national population, may have been spurred by historically persistent spatial patterns of conflict risk (and, possibly, by unobserved or noisy correlates of the propensity for conflict), the ancestry-adjusted measure of national population

[3] The analysis also establishes that, based on inference from selection on observables, bias in the estimated association between diversity and conflict, arising from selection on unobservables, is not a concern.

diversity may be endogenous to contemporary intrastate conflicts. Thus, to mitigate concerns of endogeneity in the country-level analysis, two alternative empirical strategies are developed, yielding remarkably similar results.

The first strategy confines the analysis to exploiting variations in a sample of countries that only belong to the Old World (i.e., Africa, Europe, and Asia), where the population diversity of contemporary national populations predominantly reflects the population diversity of indigenous populations that became native to their current locations well before the colonial era. This strategy rests on the observation that post-1500 population movements within the Old World did not result in the significant admixture of populations that are genetically very distant from one another. The second strategy exploits variations in a globally representative sample of countries using a two-stage estimator, in which the migratory distance of a country's prehistorically native population from East Africa is employed as an excluded instrument for the ancestry-adjusted diversity of its contemporary national population. This strategy utilizes the extraordinarily strong and negative first-stage impact of migratory distance from the cradle of humankind on the contemporary worldwide distribution of population diversity across prehistorically indigenous ethnic groups. It rests on the identifying assumption that the migratory distance of a country's prehistorically native population from East Africa is exogenous to the risk of intrastate conflict faced by the country's overall population in the last half-century.

The empirical analysis at the country level establishes that, accounting for the potentially confounding effects of geographical characteristics, institutional factors, ethnolinguistic fragmentation, outcomes of economic development, and continent fixed effects, an increase in national population diversity that corresponds to the movement from the 10th to the 90th percentile of its global cross-country distribution (i.e., a movement from the diversity level of the Republic of Korea to that of the Democratic Republic of Congo) is associated with 2.5 new civil conflict outbreaks during the 1960–2008 time horizon (relative to a sample mean of 1 and a standard deviation of 1.5 new civil conflict outbreaks). In addition, this increase in diversity is also associated with (i) an increase in the likelihood of observing the incidence of civil conflict in any given 5-year interval during the 1960–2008 period from 19 percent to 35 percent; (ii) an increase in the likelihood of observing the onset of a new civil conflict in any given year during the 1960–2008 time horizon from 0.8 percent to 4 percent; (iii) an increase in the likelihood of observing the incidence of one or more intragroup factional conflict events at any point in the 10-year interval between 1990 and 1999 from 15 percent to 73 percent; and (iv) an increase in the intensity of social unrest by either 30 percent or 51 percent of a standard deviation of the observed distribution of intrastate conflict severity across countries during the 1960–2008 time period (depending on the employed measure of intrastate conflict severity).

Similarly, the analysis at the ethnic-homeland level establishes that, accounting for the potentially confounding influence of a wide range of geographical and historical factors, outcomes of economic development, and regional fixed effects, an increase in the population diversity of an ethnic group from the 10th percentile (e.g., the Huave people of Mesoamerica) to the 90th percentile (e.g., the Nuer people of Central Africa) of its global distribution is associated with an increase in the spatiotemporal prevalence of conflict by 1.2 percentage points *per year* and *per unit area* of the *ethnic homeland* during the 1989–2008 time period (relative to a sample mean of 0.4 and a standard deviation of 0.8 percentage points). Further, this change in ethnic population diversity is also associated with an increase of about 260 in the total number of conflict events and an increase of over 6,400 in the total number of deaths across all conflict events during the same time horizon.

The remainder of the paper is organized as follows: Section 2 highlights the added value of this research to the related literature; Section 3 discusses the country-level analysis, starting with the data and empirical framework employed for identifying the impact of national population diversity on various outcomes associated with intrastate conflict over the past half-century, and then revealing the baseline findings for each of several conflict outcomes; Section 4 analogously discusses the ethnicity-level analysis, beginning with the data and empirical framework employed, and then presenting the baseline findings for the impact of ethnic population diversity on conflict within ethnic homelands; Section 5 presents the findings from various complementary analyses of some of the mechanisms that can potentially mediate the influence of population diversity on conflicts at both national and subnational levels; and finally, Section 6 concludes. In the interest of conserving space, the results from all robustness checks are reported in the appendix.

2 Advancements with Respect to the Related Literature

This study is the first to empirically establish the importance of interpersonal population diversity as a significant trigger of civil discordance. The analysis relates to several well-established lines of inquiry. First, the paper contributes to the vast literature on the empirical determinants of civil conflict. The origins of civil conflict have been the focus of intensive research over the past two decades (Sambanis, 2002; Collier and Hoeffler, 2007; Blattman and Miguel, 2010). One of the major unresolved debates in this literature – originating from seminal studies by Collier and Hoeffler (1998, 2004) and Fearon and Laitin (2003) – concerns the role of social, political, and economic grievances as determinants of the risk of civil conflict, relative to other factors such as the capability of the state to subdue armed opposition groups, the conduciveness of geographical characteristics towards rebel insurgencies, or the opportunity cost of engaging in rebellions. The present study advances an understanding of the nature of grievance-related mechanisms in civil conflict, highlighting the role of interpersonal population diversity and its deep determinants on the emergence of intragroup as well as intergroup social divisions.

In particular, motivated by the conventional wisdom that intergroup competition over ownership of productive resources or exclusive political power, along with conflicting preferences for public goods and redistributive policies, are more difficult to reconcile in societies that are fragmented ethnolinguistically, the role of *fractionalization* was initially at the forefront of empirical analyses of the underlying determinants of civil conflict. Nevertheless, early evidence regarding the influence of ethnic, linguistic, and religious fractionalization on the risk of civil conflict in society had been largely inconclusive (Collier and Hoeffler, 1998, 2004; Fearon and Laitin, 2003), arguably due in part to conceptual limitations associated with fractionalization indices. The subsequent introduction of *polarization* indices to empirical analyses of civil conflict has led to more affirmative findings (Montalvo and Reynal-Querol, 2005; Esteban, Mayoral and Ray, 2012), demonstrating that intergroup grievances are indeed contributors to the risk of civil conflict in society.

An important shortcoming of existing measures of ethnolinguistic fragmentation (based on either intergroup fractionalization or polarization indices) is that they are unable to account for the potentially critical role of intragroup heterogeneity in augmenting the risk of conflict in society at large. For instance, theoretical models of conflict (e.g., Esteban and Ray, 2011a) have established a positive link between intragroup heterogeneity (in individual income or wealth) and the risk of intergroup conflict due to complementarities between human and material inputs; a link that cannot be directly tested in a cross-country framework using conventional measures of ethnolinguistic fragmentation.[4] As such, a central virtue of the measure of population diversity advanced in the current study is that it captures, amongst other elements, the diversity across individuals *within*

[4] A similar point is made by Blattman and Miguel (2010). While one may empirically examine the link between intragroup inequality (rather than population diversity) and intergroup conflict, the endogeneity of income inequality would preclude a causal interpretation of the observed relationship.

ethnic groups – an advantage that a-priori permits population diversity to retain both economically and statistically significant explanatory power for the potential of civil conflict in society, even after accounting for the influence of conventional measures of ethnolinguistic fragmentation.

As discussed earlier, the diversity of a national population, proxied by its genetic diversity, captures interpersonal heterogeneity due to variations within ethnic groups as well as differences in ethnic markers across groups. Thus, the measure partly reflects the impact of intragroup cleavages that conventional measures of ethnolinguistic fragmentation are unable to capture. Nevertheless, even as a proxy for *interethnic* divisions, the employment of genetic diversity in empirical analyses of conflict brings about substantial insights relative to existing proxies that are based on fractionalization and polarization indices. Specifically, notwithstanding some notable exceptions (e.g., Fearon, 2003; Desmet, Ortuño-Ortín and Weber, 2009; Esteban, Mayoral and Ray, 2012), the commonly used measures of ethnolinguistic fragmentation typically do not exploit information beyond the proportional representations of ethnolinguistically differentiated groups in the national population – namely, they implicitly assume that these ethnic groups are internally homogenous and culturally "equidistant" from one another. In contrast, the measure of national population diversity incorporates information on pairwise intergroup genetic distances, as well as the genetic diversity within each ethnic group, as determined predominantly over the course of the "out of Africa" demic diffusion of humans to the rest of the globe tens of thousands of years ago.[5]

Moreover, conventional measures of ethnolinguistic fragmentation are potentially tainted by both measurement error and endogeneity in empirical analyses of civil conflict. In particular, the individual shares of different ethnolinguistic groups in a national population may be noisily observed in general and may even be systematically mismeasured in more conflict-prone societies, due to (i) the incoherent and often inconsistent nature of ethnic categories in the data from national censuses (Fearon, 2003); (ii) the endogenous "political economy" of national census categorizations of subnational groups; and (iii) the endogenous constructivism of individual self-identification with an ethnic group (Eifert, Miguel and Posner, 2010; Caselli and Coleman, 2013; Besley and Reynal-Querol, 2014). In addition, due to atrocities and voluntary or forced migrations associated with historical conflict events, to the extent that temporal persistence in the potential for conflict in society is driven by factors other than interethnic cleavages, the ethnolinguistic configuration of a national population cannot be considered exogenous to the contemporaneous risk of civil conflict (Fletcher and Iyigun, 2010). Although the national-level measure of population diversity exploits information on the population shares of subnational groups possessing ethnically differentiated ancestries, the fact that the endowment of population diversity in a given location was overwhelmingly determined during the prehistoric "out of Africa" expansion of humans permits the analysis to exploit a plausibly exogenous source of the contemporary cross-country variation in this measure, thereby mitigating the biases associated with measurement and endogeneity issues that plague the widely used proxies of ethnolinguistic fragmentation.

[5]In this respect, the more sophisticated measures of ethnolinguistic fragmentation – such as (i) the Greenberg index of "cultural diversity," as measured by Fearon (2003) and Desmet, Ortuño-Ortín and Weber (2009), or (ii) the ethnolinguistic polarization index of Esteban and Ray (1994), as measured by Desmet, Ortuño-Ortín and Weber (2009) and by Esteban, Mayoral and Ray (2012) – incorporate information on pairwise linguistic distances, wherein pairwise linguistic proximity monotonically increases in the number of shared branches between any two languages in a hierarchical linguistic tree, relative to the maximum possible number of branches, which is 15. This information, however, is still somewhat constrained by the nature of a hierarchical linguistic tree, in the sense that the languages residing at the same level of branching of the tree are all necessarily equidistant from one another. On the other hand, the genetic distance between any two ethnic groups in a contemporary national population predominantly reflects the prehistoric migratory distance between their respective ancestral populations (from the precolonial era), and as follows from the smooth continuity of geographical distances, the advanced population diversity measure ends up incorporating more *continuous* information on intergroup distances.

Second, considering the fact that the diversity of a national population reflects interpersonal differences that are associated with heterogeneity amongst subnational groups in ethnic markers, the current study contributes to a vast literature that establishes the adverse influence of the ethnolinguistic fragmentation of a national population on various societal outcomes, including the rate of economic growth, the quality of national institutions, the extent of financial development, efficiency in the provision of public goods, the level of social capital, and the potential for civil conflict (Easterly and Levine, 1997; Alesina, Baqir and Easterly, 1999; Alesina and La Ferrara, 2005). However, since genetic diversity at the national level additionally captures the degree of heterogeneity within each ethnic group as well as the pairwise distances amongst all such groups, the current analysis is uniquely positioned to capture the contribution of these additional dimensions of diversity to social dissonance and aggregate inefficiency.

Third, in light of the fact that the contemporary variation in population diversity predominantly reflects the influence of ancient population bottlenecks that occurred during the "out of Africa" demic expansion of humans to the rest of the world tens of thousands of years ago, the paper contributes to the exploration of the role of deeply rooted geographical, sociocultural, institutional, and human characteristics to comparative economic development (see, e.g., surveys by Galor, 2011; Spolaore and Wacziarg, 2013; Nunn, 2014). In particular, the analysis highlights the importance of the prehistorically determined macrogenetic structure of human populations (e.g., Spolaore and Wacziarg, 2009, 2013, 2014; Ashraf and Galor, 2013a,b, forthcoming). In contrast to the country-level analysis in the current study, which focuses on the link between national population diversity (as proxied by its genetic diversity) and the potential of social conflict within the population, Spolaore and Wacziarg (2016) examines the impact of genetic distance between any two national populations on their proclivity to engage in war, documenting a negative relationship between genetic distance and interstate warfare. Their finding is broadly consistent with the view that if genetic relatedness proxies for unobserved similarity in preferences over rival and excludable goods and, thus, the resources necessary for producing them, then violent contentions over the ownership and control of such resources would be more likely to arise between national populations that are genetically closer to one another.[6]

Last but not least, the current study relates to various perspectives on the formation of ethnic identities, the manifestation of grievances across ethnic boundaries, and the culmination of such grievances to intergroup conflict in society. Social theory features at least three approaches. In particular, the primordialist or essentialist perspective (e.g., Shils, 1957; Geertz, 1973; Brewer, 1979, 1991, 1997; Van den Berghe, 1981, 1995; Horowitz, 1985, 1999; Connor, 1994) harbors the view that because ethnicity is ultimately rooted in perennial notions of kinship and group-belonging, interethnic relations in society can be charged with the potential for conflict, especially when "groupthink" is conditioned by deep sources of enmity against other groups or the desire to dominate them. On the other hand, the instrumentalist-constructivist approach (e.g., Barth, 1969; Bates, 1983; Horowitz, 1985, 1999; Hardin, 1995; Brass, 1997; Brubaker, 2004) argues that although ethnic identities can be conditioned by past conflict events, conflict in society may simply manifest itself along endogenous interethnic boundaries for pragmatic reasons, including but not limited to the mobilization of ethnic networks by "ethnic entrepreneurs" as devices for effective monitoring, enforcement against free-riding, and easier access to financing. Finally, advocates of the modernist viewpoint (e.g., Bates, 1983; Gellner, 1983; Wimmer, 2002) stress that interethnic conflict arises from increased competition over scarce resources, especially when previously marginalized groups

[6]It may be noted that because dissensions in the context of intrastate conflict often arise from grievances associated with incompatibilities in preferences over public (rather than private) goods, to the extent that the diversity of a national population is associated with divergence in such preferences across subnational groups in the population, the empirical findings of Spolaore and Wacziarg (2016) are not necessarily inconsistent with those from our analysis.

that were excluded from the nation-building process experience socioeconomic modernization and, thus, begin to challenge the status quo.

The current research is consistent with both primordialist and instrumentalist perspectives. Since the initial endowment of interpersonal population diversity at a given location – as governed by the "out of Africa" demic diffusion process – may have facilitated the endogenous formation of coalitional groups at that location in prehistory, with collective identities thereafter diverging over time under the forces of "cultural drift," a reduced-form link between the prehistorically determined diversity of a national population and its contemporary risk of interethnic conflict may well be apparent in the data, regardless of whether these groups today are mobilized into conflict by "ethnic entrepreneurs" that aim to reinforce ethnic identities for their private interests or whether such identities entirely reflect primordial notions of kinship and group-belonging, with conflict between groups being driven by deeply rooted grievances.[7]

3 Population Diversity and Conflict at the Country Level

3.1 Data and Empirical Framework

This section outlines the order in which we present our results, describe our identification strategies and the key variables we employ in the country-level analyses. We first focus on contemporary conflicts. Here, our regressions comprise different specifications, exploiting variations in either cross-country or repeated cross-country data. We investigate the explanatory power of interpersonal population diversity for observed variations in three different dimensions of civil conflict. First, we run cross-country regressions using the average frequency of new conflict outbreaks over the whole sample period. Second, we account for persistence of conflicts, focusing on the likelihood of conflict prevalence (incidence) and using 5-year intervals in a country as our unit of analysis. Third, we the study the likelihood of conflict outbreak (onset) on an annually repeated cross-country sample.

Next, we investigate the relation between population diversity and two less well-explored, but nevertheless important, dimensions of social conflict. Namely, we examine the impact of population diversity on the intensive margin of conflict in society. We exploit cross-country variations over time in the *severity* of social unrest, reflected by two alternative measures of the intensity of intrastate conflict.

Our measure of diversity, by virtue of capturing both intergroup and intragroup personal differences, may not only deepen intergroup incompatibilities in a society but also diminish social cohesion within subnational groups. Therefore, next we set out to investigate this possibility by exploiting cross-country variations in population diversity and the incidence of *intragroup* factional conflicts in a national population. We conclude our country-level analysis by exploring whether deep-rooted interpersonal diversity in genetic traits also predicts conflicts of the more distant past.

[7] The argument that population diversity in the distant past may have contributed to ethnic heterogeneity as observed in the modern era (Ashraf and Galor, 2013b) is consistent with the sociobiological perspective of ethnic origins (e.g., Van den Berghe, 1981, 1995), rooted mainly in the dual-inheritance theory of gene-culture coevolution from the field of evolutionary anthropology (e.g., Durham, 1991; Cavalli-Sforza, Menozzi and Piazza, 1994). Accordingly, like most other mammals, human beings exhibit nepotistic behavior, including greater loyalty to their immediate kin, extended family, or clan, because such behavior can ultimately serve to maximize the likelihood of passing on one's traits successfully to future generations. As such, the formation of collectives in prehistory is partly viewed as a manifestation of "extended nepotism," with subsequent intergroup differentiation of collective identities occurring over a long expanse of time through the forces of "cultural drift" (Cavalli-Sforza and Feldman, 1981) and "biased transmission" (Boyd and Richerson, 1985; Henrich and McElreath, 2003) of cultural markers – e.g., language dialects, customs and traditions, and norms of social conduct – that partly serve to distinguish the "outsiders" from the "insiders" of a group.

Following the norm in the empirical literature on civil conflict, we confine our contemporary analysis to the post-1960 time period. Most of the previous European colonies in Sub-Saharan Africa, the Middle East, and South and Southeast Asia had become independent nation states by 1960. Thus, this time horizon permits an assessment of the correlates of civil conflict at the national level, independently of their interactions with the presumably distorting contemporaneous influence of the colonial powers on the potential for internal conflict. Our baseline sample for contemporary analysis of conflict contains information on 143 countries for the period 1960–2008. 119 of these countries are in the Old World.

In the following sub-sections, we describe our main outcome variables, explain how we measure interpersonal population diversity, and briefly introduce the covariates included in our baseline specifications. We describe all additional outcome and control variables in Section 3.2, where we present our empirical findings.

3.1.1 Main Outcome Variables: Frequency, Incidence, and Onset of Civil Conflict

The core dependent variables in our analysis reflect various outcomes related to either *overall* civil conflict (i.e., of the type that includes also those conflicts not related to ethnic divisions) or *ethnic* civil conflict. The main data source that we rely on for conflict events is the UCDP/PRIO Armed Conflict Dataset, Version 4-2012.[8] By definition, a civil conflict refers to an internal armed conflict between the government of a state and one or more internal opposition groups, fighting over a given incompatibility. For our main analysis, we employ the most comprehensive conflict coding PRIO25 which encompasses all conflict events that resulted in 25 or more battle-related deaths in a given year.

Recent evidence uncovered by Ashraf and Galor (2013b) supports the notion that following the prehistoric "out of Africa" migration of humans, the population diversity of an indigenous settlement may have served as a domain over which endogenous group selection and subsequent intergroup cultural or ethnic differentiation had taken place at that location. As already mentioned, this deeper mechanism may indeed be one of the primary channels through which population diversity influences the potential for intergroup conflict in contemporary national populations. One way to assess the validity of this argument is to investigate the influence of population diversity on the potential for those types of civil conflict in which interethnic divisions are presumably a more germane issue. As such, our analysis also focuses on outcomes associated with *ethnic* civil conflict, as defined by Wimmer, Cederman and Min (2009) (henceforth referred to as WCM09). Using the UCDP/PRIO Armed Conflict Dataset as their primary data source for civil conflict events, WCM09 additionally apply an "ethnic" categorization to the data. They identify those conflict events wherein the opposition group(s) (i) either explicitly pursue ethnonationalist aims (e.g., attempt to secure ethnonational self-determination, ethnoregional autonomy, or language and other cultural rights) or are motivated by ethnic concerns (e.g., ethnic balance of power in the government or ethnic discrimination), and (ii) they recruit fighters and forge political alliances on the basis of ethnic affiliations. Since WCM09 employ an earlier version (Version 3-2005b) of the UCDP/PRIO Armed Conflict Dataset, the sample we employ for ethnic conflict regressions contains only 141 countries and limited to the 1960–2005 period.

Depending on the unit of analysis, our outcome variables capture different dimensions of either overall or ethnic civil conflict. In our cross-country regressions, for instance, the outcome variables record the annual frequency (i.e., the average number per year) of "new" civil conflict outbreaks —involving a new issue of incompatibility and/or a new set of nonstate actors fighting

[8]For more information about the data, see Gleditsch et al. (2002) and Themnér and Wallensteen (2012).

against the government— over the relevant sample period. These outcome variables therefore reflect the number of distinct incompatibilities between state actors and armed opposition groups that have, on average, escalated to a full-blown conflict —as defined by the battle-related death threshold— on a yearly basis.

Many civil conflicts, however, span several years and may even comprise multiple conflict episodes that are separated by one or more years of inactivity —i.e., years of actual peace or in which the annual battle-related death toll is below the specified threshold. Our regressions based on repeated cross-country data exploit this temporal dimension of civil conflict. Specifically, in our regressions explaining the incidence (prevalence) of civil conflict, the outcome variable is an indicator, coded 1 for each country-period (a period typically being a 5-year time interval) in which there is at least one active conflict-year observed, and 0 otherwise. For example, A PRIO25 civil conflict is said to be active in a given year if the clashes between the conflicting parties resulted in more than 25 battle-related fatalities in that year. Finally, we employ annually repeated cross-country data to examine the predictive power of population diversity for the *onset* of a new conflict episode in a given country-year. The onset year of a PRIO25 civil conflict corresponds to either the first year of a new PRIO25 conflict or the first year of a recurring conflict after more than one year of inactivity.

3.1.2 Population Diversity and the Empirical Strategy

Our proxy for interpersonal population diversity in the country-level analyses is a measure of contemporary genetic diversity at the country level. In this section we first describe the construction of this measure. Then we discuss potential problems regarding identification and explain how we address them.

Genetic diversity measurements at the level of ethnic groups constitute the building blocks of our proxy variable for country-level population diversity. Observed genetic diversity at the ethnic group level is measured by an index referred to by population geneticists as expected heterozygosity. This index conceptually reflects the probability that two individuals, selected at random from the relevant population, are genetically different from one another with respect to a given spectrum of genetic traits. The index itself is constructed by population geneticists using data on allelic frequencies —i.e., the frequency with which a gene variant or allele (e.g., the brown versus blue variant of the eye-color genetic trait) occurs in a given population.[9] Using information on the allelic frequencies in a given population for a particular gene or DNA locus, it is possible to compute a gene-specific heterozygosity statistic (i.e., the probability that two randomly selected individuals will differ with respect to the gene in question), which when averaged across multiple, say m, genes, yields an aggregate expected heterozygosity for the population of the following form:

$$H_{exp} = 1 - \frac{1}{m} \sum_{l=1}^{m} \sum_{i=1}^{k_l} p_i^2, \qquad (1)$$

where l denotes a single gene or DNA locus with k_l observed variants or alleles in the population, and p_i denotes the frequency of occurrence of the ith allele.

[9] In molecular genetics, an allele is defined as any one of a number of viable DNA codings (formally, a sequence of nucleotides) that occupy a given locus (or position) in a chromosome. Chromosomes themselves are packages for carrying strands of DNA molecules in cells, and they comprise multiple loci that typically correspond to some of the observed discrete units of heredity (or genes) in living organisms. For additional details on basic concepts and definitions from molecular and population genetics, the interested reader is referred to Griffiths et al. (2000) and Hartl and Clark (2007).

Like standard measures of ethnolinguistic fragmentation, based on fractionalization or polarization indices, *observed* genetic diversity might be endogenous in an empirical model of civil conflict, since it could be tainted by genetic admixtures resulting from the movement of populations across space, triggered by cross-regional differences in patterns of historical conflict potential, the nature of political institutions, and levels of economic prosperity. To circumvent this problem, we employ the measure of contemporary genetic diversity introduced to the literature by Ashraf and Galor (2013a). This measure captures —amongst other dimensions of contemporary genetic diversity at the national level, as explained below— the component of observed interpersonal diversity within a country's indigenous ethnic groups that is *predicted* by migratory distance from Addis Ababa, Ethiopia to the country's modern-day capital city, along prehistoric land-connected human migration routes.[10]

Exploiting the explanatory power of a serial founder effect associated with the "out of Africa" migration process, the predicted population diversity of a country's prehistorically indigenous population is generated by applying the coefficients obtained from an ethnic-group-level regression (e.g., Ramachandran et al., 2005; Prugnolle, Manica and Balloux, 2005) of expected heterozygosity on migratory distance from Addis Ababa, Ethiopia, in a sample comprising 53 globally representative ethnic groups from the Human Genome Diversity Cell Line Panel.[11] According to population geneticists, these ethnic groups have not only been prehistorically native to their current geographical locations, but they have also been largely isolated from genetic flows from other ethnic groups.[12]

The measure of expected heterozygosity in the sample of 53 HGDP-CEPH ethnic groups is constructed using data on allelic frequencies for a particular class of DNA loci called microsattelites, residing in non-protein-coding or "neutral" regions of the human genome —i.e., regions that do not directly result in phenotypic expression. Therefore, this measure of observed genetic diversity has the advantage of not being tainted by the differential forces of natural selection that may have operated on these populations since their prehistoric exodus from Africa. Importantly, however, we expect that the observed socioeconomic influence of expected heterozygosity in microsattelites should reflect the unobserved impact of diversity in phenotypically and cognitively expressed genomic material, in light of mounting evidence from the fields of physical and cognitive anthropology.[13]

In the absence of systematic and large-scale population movements across geographically (and, thus, genetically) distant regions, as had been largely true during the precolonial era, the interpersonal diversity of the prehistorically native population in a given location serves as a good proxy for the contemporary population diversity of that location. While this continues to remain true to a large extent for nations in the Old World (i.e., Africa, Europe, and Asia), post-1500

[10]These routes incorporate five obligatory intermediate waypoints, namely Cairo, Egypt; Istanbul, Turkey; Phnom Penh, Cambodia; Anadyr, Russia; and Prince Rupert, Canada. In contrast to a measure of direct geodesic distance from East Africa, the use of these intermediate waypoints ensures that the measure of migratory distance more accurately reflects the fact that humans did not cross large bodies of water in the course of their prehistoric exodus from Africa.

[11]The Human Genome Diversity Cell Line is compiled by the Human Genome Diversity Project (HGDP) in collaboration with the Centre d'Etudes du Polymorphisme Humain (CEPH).

[12]For a more detailed description of the HGDP-CEPH Human Genome Diversity Cell Line Panel data set, the interested reader is referred to Cann et al. (2002). A broad overview of the Human Genome Diversity Project is provided by Cavalli-Sforza (2005).

[13]This body of evidence establishes serial founder effects —associated with the prehistoric "out of Africa" migration process— on worldwide spatial patterns in various forms of intragroup phenotypic and cognitive diversity, including phonemic diversity (Atkinson, 2011) as well as interpersonal diversity in skeletal features pertaining to cranial characteristics (Manica et al., 2007; von Cramon-Taubadel and Lycett, 2008; Betti et al., 2009), dental attributes (Hanihara, 2008), and pelvic traits (Betti et al., 2013).

population flows from the Old World to the New World have had a considerable impact on the ethnic composition and, thus, the contemporary interpersonal diversity of national populations in the Americas and Oceania. Thus, instead of employing the interpersonal diversity of prehistorically native populations (i.e., precolonial diversity) at the expense of limiting our entire analysis to the Old World, we adopt the measure of ancestry-adjusted genetic diversity from Ashraf and Galor (2013a) as our main proxy for contemporary population diversity. Using the shares of different groups in a country's modern-day population, this measure accounts for (i) the diversity within the ethnic groups that can trace own ancestry around year 1500 to their current homelands, (ii) the diversity of those descended from immigrant settlers over the past half-millennium, and (iii) the additional component of population diversity at the national level that arises from the pairwise genetic distances amongst these different subnational groups.[14]

Yet, ancestry-adjusted population diversity may still be afflicted by endogeneity bias because it accounts for the impact of cross-country migrations in the post-1500 era on the diversity of contemporary national populations. These migrations may have been spurred by historically persistent spatial patterns of conflict. We employ two alternative strategies to address this issue. The first strategy is to exploit variations across countries that only belong to the Old World, where as discussed previously, the interpersonal diversity of contemporary national populations overwhelmingly reflects the diversity within populations that have been native to their current locations since well before the colonial era.[15] This strategy exploits the fact that the great human migrations of the post-1500 era had systematically differential impacts on the genetic composition of national populations in the Old World versus the New World. Specifically, although post-1500 population flows had a dramatic effect on the interpersonal diversity of national populations in the Americas and Oceania, the diversity of resident populations in Africa, Europe, and Asia remained largely unaltered, primarily because native populations in the Old World were not subjected to substantial inflows of migrant settlers that were descended from genetically distant ancestral populations. By confining our analysis to the Old World, we effectively exploit the spatial variation in contemporary population diversity that largely coincides with the variation in diversity of prehistorically indigenous populations. This prehistoric diversity was overwhelmingly determined by an ancient serial founder effect associated with the "out of Africa" migration process.

The second strategy is to exploit variations in our global sample of countries with instrumental variables (IV) estimators, i.e. 2SLS or IV Probit. We employ the migratory distance of the prehistorically native populations in each country from East Africa as an instrument for the country's contemporary interpersonal diversity. This strategy utilizes the fact that the mark of ancient population bottlenecks that occurred during the prehistoric "out of Africa" demic diffusion of humans across the globe continues to be seen on average in the worldwide pattern of genetic diversity across contemporary national populations —a fact reflected by the sizable correlation of 0.750 between our proxies for precolonial and contemporary population diversity in our global sample of countries. In addition, this strategy rests on the identifying assumption that the migratory distance of a country's prehistorically indigenous population from East Africa is plausibly excludable from an empirical model of the risk of civil conflict faced by its modern

[14]The data on the population shares of these different subnational groups at the country level are obtained from the World Migration Matrix, 1500–2000 of Putterman and Weil (2010), who compile for each country in their data set, the share of the country's population in 2000 that is descended from the population of every other country in 1500. For an in-depth discussion of the methodology underlying the construction of the ancestry-adjusted measure of genetic diversity, the reader is referred to the data appendix of Ashraf and Galor (2013a).

[15]This is a pattern that primarily arises from the fact that historical cross-country migrations in the Old World did not result in the admixture of populations that are genetically distant from one another.

national population, conditional on our large set of controls for the geographical and institutional determinants of conflict as well as the correlates of economic development.

3.1.3 Control Variables: Geography, Institutions, Ethnolinguistic Fragmentation, and Development Outcomes

The vast empirical literature on the determinants of civil conflict has emphasized a large number of potentially contributing factors. Drawing on this literature, we include a large set of control variables in our baseline specifications. All other control variables that we use in robustness checks are discussed in corresponding Appendices.

Geographical Characteristics Given that the predicted intragroup component of our ancestry-adjusted measure of population diversity varies linearly with prehistoric migratory distance from East Africa, we control for a wide range of geographical attributes that may be correlated with migratory distance and that can also reasonably impart a reduced-form influence on conflict risk through channels unrelated to population diversity. Absolute latitude and distance to the nearest waterway, for instance, can exert an influence on economic development and, thus, on conflict potential through climatological, institutional, and trade-related mechanisms. We also control for total land area of a country. Larger territories may mechanically experience more conflict while also hosting a bigger and more diverse national population.[16] Rugged terrains can provide safe havens for rebels and enable them to sustain continued resistance by protecting them from numerically and militarily superior government forces (Fearon and Laitin, 2003). Moreover, in regions with rough terrains, subgroups of a regional population may be geographically more isolated. Such isolation may strengthen the forces of "cultural drift" and ethnic differentiation among these groups (Michalopoulos, 2012). This in turn may increase the potential for intergroup conflict. Finally, in light of evidence that conditional on their respective country-level means, greater intracountry dispersion in agricultural land suitability and elevation can contribute to ethnolinguistic diversity (Michalopoulos, 2012), these natural attributes could also impart an indirect influence on conflict propensity through the ethnolinguistic fragmentation of the population.[17] To account for these factors, we also include in our baseline set of covariates terrain ruggedness, the mean and range of both agricultural land suitability and elevation [18] Finally, our baseline specifications additionally account for a complete set of continent fixed effects to ensure that the estimated reduced-form impact of population diversity on conflict potential is not simply reflecting the latent influence of unobserved time-invariant cultural, institutional, and geographical factors at the continent level.[19]

[16] Land area can also account for any bias that might arise if our measure of population diversity, by virtue of being based on migratory distance from East Africa to the modern-day capital city of a country, is less comparable across countries of different geographical size.

[17] Although we directly control for measures of ethnolinguistic fragmentation in our full empirical model, those measures are afflicted by endogeneity bias, and beyond that, their exogenous geographical determinants may still explain some unobserved component of intrapopulation heterogeneity in ethnic and cultural traits, thereby exerting some residual influence on the potential for conflict in society.

[18] The data for absolute latitude, total land area, and distance to the nearest waterway are obtained from the Central Intelligence Agency (2006), the World Bank (2006), and Gallup, Sachs and Mellinger (1999), respectively. Nordhaus (2006) provides disaggregated geospatial data at a 1-arc-minute resolution on surface undulation and elevation, from which we derive our country-level aggregate measures of terrain ruggedness and the mean and range of elevation. Finally, we obtain the country-level aggregate measures of the mean and range of agricultural land suitability directly from the data set of Michalopoulos (2012). See the data appendices of Ashraf and Galor (2013a,b) for further details.

[19] In addition to "soaking up" the possibility of omitted-variable bias from unobserved time-invariant characteristics at the continent level, the need to account for continent fixed effects is perhaps even more binding for observed non-geographical factors, given the potential for systematic measurement error at the continent level in covariates reflecting cultural and institutional characteristics.

Institutional Factors Colonial legacies may have significantly shaped the political economy of interethnic cleavages in newly independent states (Posner, 2003). More generally, the heritage of colonial rule and the identity of the former colonizers may have important ramifications for the nature and stability of contemporary political institutions at the national level, thereby influencing the potential for conflict in society.We consider two different sets of covariates in our baseline specifications to control for the impact of colonial legacies. Depending on the unit of analysis, the first set comprises either binary indicators for the historical prevalence of colonial rule (as is the case in our cross-country regressions) or time-varying measures of the lagged prevalence of colonial rule (as is the case in our regressions using repeated cross-country data). In either case, we distinguish between colonial rule by the U.K., France, and any other major colonizing power.

The second set of covariates comprises time-invariant binary indicators for British and French legal origins, included to account for any latent influence of legal codes and institutions that may not necessarily be captured by colonial experience.[20] Our baseline specifications additionally include three control variables, all based on yearly data at the country level from the Polity IV Project (Marshall, Gurr and Jaggers, 2009), in order to account for the direct influence of contemporary political institutions on the risk of civil conflict. The first variable is based on an ordinal index that reflects the degree of executive constraints in any given year, whereas the other two variables are based on binary indicators for the type of political regime, reflecting the prevalence of either democracy (when the polity score is above 5) or autocracy (when the polity score is below -5) in a given year.[21]

Ethnolinguistic Fragmentation Previous empirical findings regarding the role of ethnic fragmentation have generally been somewhat mixed, exhibiting substantial sensitivity to model specifications and conflict codings (Fearon and Laitin, 2003; Collier and Hoeffler, 2004; Hegre and Sambanis, 2006). Moreover, theoretical work on the link between the ethnic composition of a society and the risk of civil conflict suggests that ethnic fractionalization by itself may be insufficient to fully capture the conflict potential that can be attributed to broader ethnolinguistic configurations of the population (Esteban and Ray, 2011b). In light of their well-grounded structural foundations, indices of polarization have gained popularity as a substitute for —or in addition to— the fractionalization measures commonly considered by empirical analyses of civil conflict.[22] Indeed, many empirical studies find that ethnic polarization is a stronger predictor of the likelihood of civil conflict (e.g., Montalvo and Reynal-Querol, 2005; Esteban, Mayoral and Ray, 2012).

We include two time-invariant controls in our baseline model to capture the influence of the ethnolinguistic composition of national populations on the potential for civil conflict. Our first proxy is the well-known ethnic fractionalization index of Alesina et al. (2003), reflecting the probability that two individuals, randomly selected from a country's population, will belong to different ethnic groups. Our second proxy for this channel is an index of ethnolinguistic polarization, obtained from the data set of Desmet, Ortuño-Ortín and Wacziarg (2012). The authors provide measures of several such polarization indices, constructed at different levels of aggregation of linguistic groups in a country's population (based on hierarchical linguistic trees). The specific polarization measure

[20] The country-level indicators for British and French legal origins are obtained directly from the data set of La Porta et al. (1999). The measures of historical and contemporary colonial rule, on the other hand, are constructed using a number of secondary sources, and we refer the reader to the data appendix of Ashraf and Galor (2013b) for further details.

[21] The prevalence of anocracy, occurring when the polity score is between -5 and 5, therefore serves as the omitted political regime category.

[22] Esteban and Ray (1994) provide the first serious attempt to measure polarization, derive its theoretical properties, and highlight its role in contributing to the potential for social conflict.

we employ corresponds to the most disaggregated level of the linguistic tree and reflects the extent of polarization across subnational groups classified according to modern-day languages.[23]

Development Outcomes Ashraf and Galor (2013a) have shown that interpersonal population diversity, as proxied by genetic diversity, confers a hump-shaped influence on productivity at the country level. At the same time, scholars of civil conflict have linked various factors, codetermined with the level of development, such as natural resource revenues, the size of a country's population, and income per capita to conflict risk. For example, natural resources can foster the risk of civil conflict by weakening political institutions and facilitating state capture, easing the financial constraints on rebel organizations (e.g., Fearon and Laitin, 2003; Dube and Vargas, 2013; Collier and Hoeffler, 2004; Angrist and Kugler, 2008), increasing vulnerability of political elite to terms-of-trade shocks (e.g., Humphreys, 2005) or raising the return to regional secession (e.g., Ross, 2006).

Population size is also a standard covariate in the empirical literature. One reason is that operational definitions of civil conflict typically impose a death threshold, and violence-related casualties may be mechanically related to the size of population. In addition, a larger population may imply a larger recruitment pool for rebels (Fearon and Laitin, 2003). Therefore to the extent more populous countries exhibit greater intrapopulation heterogeneity, they could also harbor stronger motives for secessionist conflicts (Collier and Hoeffler, 1998; Alesina and Spolaore, 2003; Desmet et al., 2011).

Average living standards can influence civil conflict potential in a country through several ways. One argument, due to Grossman (1991) and Hirshleifer (1995), is that higher per-capita incomes raise the opportunity cost for potential rebels to engage in insurrections, thus predicting an inverse relationship between the level or growth rate of income, on the one hand, and the risk of civil conflict, on the other (Collier and Hoeffler, 1998, 2004; Miguel, Satyanath and Sergenti, 2004). Another argument, due to Hirshleifer (1991) and Grossman (1999), is that by raising the return to predation, higher per-capita incomes can contribute to the risk of rapacious activities over society's resources, consistently with empirical findings from some of the aforementioned studies on the link between income from natural resources and conflict potential. Furthermore, to the extent that income per capita serves as a proxy for state capabilities (Fearon and Laitin, 2003), a higher level of per-capita income can reflect the notion of a state that is better able to prevent or defend itself against rebel insurgencies; an idea that has also found some recent empirical support (e.g., Bazzi and Blattman, 2014).

In the light of the preceding discussion, we add to our baseline specifications controls for per-capita value of oil production borrowed from Ross (2013) as well as total population size and GDP per capita borrowed from Maddison (2010). We apply a log transformation to each of these variables before including them as covariates in our regressions.

Yet, we expect many of the aforementioned controls for institutional quality, ethnolinguistic fragmentation, and the correlates of economic development to be endogenous in an empirical model of civil conflict, and as such, their estimated coefficients in our regressions do not permit a causal interpretation. Nonetheless, controlling for these factors is essential to minimize specification errors and assess the extent to which the reduced-form influence of interpersonal diversity on conflict potential can be attributed to more conventional explanations in the literature.

Table A.23 in Appendix A presents the summary statistics of all the variables in the baseline samples exploited by our cross-country analyses of overall civil conflict frequency.

[23] We prefer employing Desmet, Ortuño-Ortín and Wacziarg (2012) as the data source for ethnolinguistic polarization, primarily due to the more comprehensive geographical coverage of their data set, relative to other potential data sources such as Montalvo and Reynal-Querol (2005) and Esteban, Mayoral and Ray (2012).

3.2 Empirical Results

This section presents our main findings from several country-level analyses. First, we establish the highly significant and robust reduced-form causal influence of population diversity on various intrastate conflict outcomes over the past half-century. We commence with the results of our baseline cross-country regressions that explain the average frequency of both overall civil conflict and ethnic conflict outbreaks in the post-1960 time period. We next discuss the results of our conflict incidence and onset regressions that exploit variations in repeated cross-country data. Third, we demonstrate that, besides frequency and persistence, our measure of diversity also explains severity of conflict and other forms of social unrest. Fourth, we present evidence that interpersonal population diversity is a significant predictor of conflict within ethnic groups. We conclude this section with an empirical analysis of conflicts over the period 1400-1799 to show that our deep-rooted measure of diversity had a very long-term influence on conflict potential that persisted across centuries. For our analysis of each conflict outcome, we conduct several robustness checks, the results from which are collected and discussed in Appendix A.

3.2.1 Analysis of Civil Conflict Frequency in Cross-Country Data

Our cross-country regressions attempt to explain the variation across countries in the annual frequency of new civil conflict onsets —i.e., the average number of new civil conflict eruptions per year— over the sample period. Specifically, the baseline empirical model for our cross-country analysis is as follows.

$$CF_i = \beta_0 + \beta_1 \widehat{GD}_i + \beta_2' GEO_i + \beta_3' INS_i + \beta_4' ETH_i + \beta_5' DEV_i + \varepsilon_i, \quad (2)$$

where CF_i is the (log transformed) average number of new civil conflict outbreaks per year in country i; $\widehat{GD}_i$ is the ancestry-adjusted population diversity of the national population; GEO_i, INS_i, ETH_i, and DEV_i are the respective vectors of control variables for geographical characteristics (including continent fixed effects), institutional factors, ethnolinguistic fragmentation, and the correlates of economic development, as described in Section 3.1; and finally, ε_i is a country-specific disturbance term. All the time-varying controls for institutional factors and development outcomes enter the model as their respective temporal means over the relevant sample period – namely, 1960–2008 in the regressions explaining the annual frequency of new PRIO25 civil conflict outbreaks, and 1960–2005 in the regressions explaining the same outcome for WCM09 ethnic civil conflict onsets.

Before moving to results from our baseline analysis of conflict frequency, we show two sets of regressions. The bivariate regressions in Table 1 show how the unconditioned influence of population diversity compares with the influence of other well-known diversity measures that capture the degree of ethnolinguistic fragmentation of a national population. In Table 2 we run "horse race" regressions between population diversity, on the one hand, and various combinations of the measures of ethnolinguistic fragmentation, on the other.[24] As is evident from Table 1, population diversity appears as a positive and statistically significant correlate of civil conflict frequency. Specifically, the estimated coefficient suggests that a move from the 10th to the 90th percentile of the cross-country population diversity distribution is associated with an increase in conflict frequency by 0.014 new PRIO25 civil conflict outbreaks per year, a relationship that is statistically significant at the 1 percent level. Bearing in mind that the sample mean of the dependent variable is about 0.021

[24]The sample employed in both tables is larger than our baseline sample of 143 countries due to the fact that the latter is conditioned on the availability of data on our baseline controls for geographical characteristics, institutional factors, and development outcomes.

Table 1: Population Diversity vs. Other Diversity Measures in Explaining the Frequency of Civil Conflict Onset across Countries – Bivariate Regressions

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
	Log number of new PRIO25 civil conflict onsets per year, 1960–2008					
Population diversity (ancestry adjusted)	0.212***					
	(0.076)					
Ethnic fractionalization (Alesina et al., 2003)		0.024***				
		(0.008)				
Linguistic fractionalization (Alesina et al., 2003)			0.032***			
			(0.009)			
Religious fractionalization (Alesina et al., 2003)				0.006		
				(0.009)		
Ethnolinguistic fractionalization (Desmet et al., 2012)					0.026***	
					(0.008)	
Ethnolinguistic polarization (Desmet et al., 2012)						0.007
						(0.009)
Observations	154	154	154	154	154	154
R^2	0.032	0.037	0.081	0.002	0.061	0.003
Effect of 10th–90th %ile move in diversity measure	0.014***	0.017***	0.025***	0.004	0.022***	0.005
	(0.005)	(0.005)	(0.007)	(0.006)	(0.007)	(0.006)

Notes: This table employs bivariate regressions to assess the unconditional importance of contemporary population diversity versus other well-known diversity measures for explaining the cross-country variation in the annual frequency of new PRIO25 civil conflict onsets during the 1960–2008 time period. In each regression, the estimated effect associated with increasing the corresponding diversity measure from the tenth to the ninetieth percentile of the measure's cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

outbreaks per year, this association is also of sizable economic significance, reflecting 40.3 percent of a standard deviation across countries in the temporal frequency of new civil conflict onsets. In terms of the other diversity variables, the different measures of ethnic and linguistic fractionalization enter their respective bivariate regressions with positive and statistically significant coefficients, although in the absence of conditioning covariates, neither religious fractionalization nor ethnolinguistic polarization appear to be significantly associated with the temporal frequency of new civil conflict outbreaks.

The results in Table 2 indicate that the positive and statistically significant association of population diversity with conflict frequency does not vanish even after its potential influence through the degree of ethnic fragmentation is accounted for. Notably, the coefficient on population diversity remains qualitatively robust throughout this analysis, although its statistical significance drops to the 5 percent level —primarily due to a modest decrease in its point estimate— when conditioned on the measures of ethnic and linguistic fractionalization. This provides suggestive evidence that certain dimensions of ethnolinguistic fragmentation could be capturing an important —though not the only— proximate channel that potentially mediates the deeper influence of interpersonal population diversity on the propensity for conflict in society.[25].

Table 3 presents the results from our baseline cross-country analysis. In Panel A, the outcome variable is the average annual frequency of new PRIO25 civil conflict onsets during the 1960–2008 time horizon. In Panel B, the outcome is annual frequency of WCM09 ethnic civil conflict onsets during the 1960–2005 period. To provide an appropriate benchmark for the subsequent

[25]The results from a more systematic investigation of this and two other potential mechanisms are presented in Table 10 and will be discussed in Section 5.1

Table 2: Interpersonal Population Diversity vs. Other Diversity Measures in Explaining the Frequency of Civil Conflict Onset across Countries – "Horse race" Regressions

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS
	Log number of new PRIO25 civil conflict onsets per year, 1960–2008									
Population diversity (ancestry adjusted)	0.212***	0.178**	0.146**	0.209***	0.157**	0.207***	0.183**	0.158**	0.158**	0.160**
	(0.076)	(0.073)	(0.072)	(0.072)	(0.074)	(0.076)	(0.071)	(0.069)	(0.073)	(0.068)
Ethnic fractionalization (Alesina et al., 2003)		0.021***					0.021***			−0.004
		(0.007)					(0.007)			(0.014)
Linguistic fractionalization (Alesina et al., 2003)			0.029***					0.031***		0.029*
			(0.009)					(0.009)		(0.016)
Religious fractionalization (Alesina et al., 2003)				0.001			−0.003	−0.009		−0.010
				(0.009)			(0.009)	(0.009)		(0.010)
Ethnolinguistic fractionalization (Desmet et al., 2012)					0.023***				0.026**	0.007
					(0.008)				(0.011)	(0.023)
Ethnolinguistic polarization (Desmet et al., 2012)						0.005			−0.008	−0.006
						(0.008)			(0.012)	(0.015)
Observations	154	154	154	154	154	154	154	154	154	154
Adjusted R^2	0.026	0.046	0.084	0.019	0.066	0.021	0.041	0.082	0.063	0.066
Effect of 10th–90th %ile move in diversity	0.014***	0.011**	0.009**	0.013***	0.010**	0.013***	0.012***	0.010**	0.010**	0.010**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)

Notes: This table employs regressions that run "horse races" between contemporary population diversity and other well-known diversity measures to assess their relative importance for explaining the cross-country variation in the annual frequency of new PRIO25 civil conflict onsets during the 1960–2008 time period, establishing the robustness of population diversity over other diversity measures as a predictor of conflict frequency. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

regressions, In Column 1, we rerun the bivariate regression in Table 1 on our baseline sample of 143 countries. Beginning with Column 2, we progressively include an expanding set of covariates to the specification. We first incorporate exogenous geographical conditions and then additionally account for semi-endogenous institutional factors, before including the more endogenous outcomes of economic development until our full empirical model in Column 8 is attained.

After accounting for the potentially confounding influence of geographical conditions (Column 2), population diversity continues to remain statistically significant at the 1 percent level, but now, its coefficient is more than twice as large as the unconditional estimate in Column 1. This is true both for overall and ethnic civil conflicts. This increase appears to be largely driven by the inclusion of absolute latitude and the range of agricultural land suitability as covariates to the model, as both variables enter the regression significantly and with expected signs.[26] The top panels in Figure 1 depict the positive and statistically significant cross-country relationship in Column 2, between population diversity and the annual frequency of new PRIO25 civil conflict onsets, both in our full sample of countries (Subfigure A) and in a sample that omits apparently influential outliers (Subfigure B). The corresponding scatter plots for ethnic conflicts are presented at the bottom two panels.

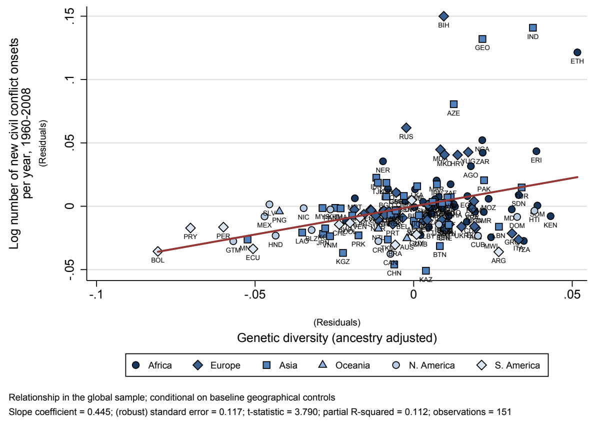
The point estimate for interpersonal population diversity becomes somewhat diminished once we condition the specification to only exploit intracontinental variations (Column 3). However, even after including a complete set of continent dummies, our coefficients of interest in both panels remain notably larger than the unconditioned estimates from Column 1. A move from the 10th to the 90th percentile of the cross-country population diversity distribution is associated with an increase in conflict frequency by 0.021 overall civil conflict outbreaks per year —or, equivalently,

[26] Specifically, countries located farther from the Equator have seen fewer conflict outbreaks on average, while those with greater dispersion in their respective land endowments have experienced such outbreaks more frequently, a result that plausibly reflects the conflict-promoting role of ethnolinguistic fragmentation, following the rationale provided by the findings of Michalopoulos (2012).

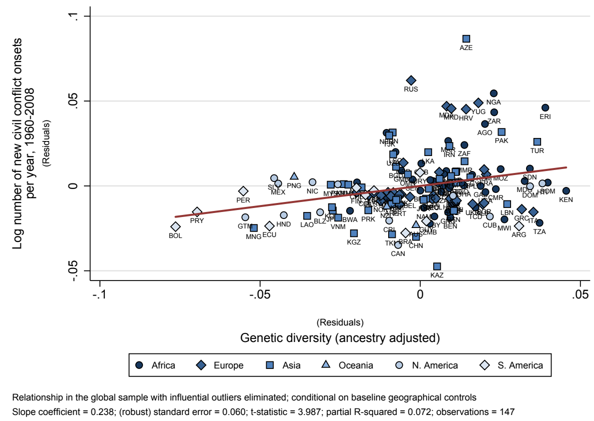
Table 3: Population Diversity and the Frequency of Civil Conflict Onset across Countries – The Baseline Analysis

Cross-country sample:	Global								Old World		Global	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS	(11) 2SLS	(12) 2SLS
Panel A	Log number of new PRIO25 civil conflict onsets per year, 1960–2008											
Population diversity (ancestry adjusted)	0.182**	0.422***	0.322**	0.366**	0.350**	0.390**	0.377**	0.398**	0.639**	0.855**	0.599***	0.805***
	(0.077)	(0.123)	(0.150)	(0.171)	(0.168)	(0.179)	(0.184)	(0.183)	(0.257)	(0.333)	(0.231)	(0.275)
Ethnic fractionalization					0.011		0.006	0.007		0.012		−0.002
					(0.012)		(0.012)	(0.012)		(0.014)		(0.013)
Ethnolinguistic polarization						0.013	0.010	0.010		0.007		0.019
						(0.013)	(0.014)	(0.013)		(0.015)		(0.013)
Absolute latitude		−0.404***	−0.440*	−0.331	−0.225	−0.356	−0.292	0.149	−0.333	0.255	−0.529**	−0.116
		(0.119)	(0.255)	(0.263)	(0.320)	(0.256)	(0.307)	(0.287)	(0.301)	(0.356)	(0.243)	(0.296)
Land area		0.765	1.825	1.709	1.972	1.719	1.862	1.586	4.177	4.114	1.626	1.311
		(2.119)	(2.287)	(2.358)	(2.382)	(2.403)	(2.436)	(2.675)	(2.797)	(2.743)	(2.247)	(2.631)
Ruggedness		0.038	0.028	0.030	0.036	0.032	0.035	0.056	0.041	0.080	0.034	0.054
		(0.038)	(0.044)	(0.044)	(0.047)	(0.045)	(0.047)	(0.047)	(0.053)	(0.054)	(0.043)	(0.042)
Mean elevation		−0.016*	−0.015	−0.017*	−0.018*	−0.018*	−0.018*	−0.020**	−0.019	−0.025**	−0.016*	−0.023**
		(0.009)	(0.009)	(0.010)	(0.010)	(0.011)	(0.011)	(0.010)	(0.012)	(0.012)	(0.009)	(0.009)
Range of elevation		0.009**	0.009**	0.009**	0.008*	0.008*	0.008*	0.004	0.009*	0.003	0.010**	0.005
		(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.006)	(0.005)	(0.004)	(0.004)
Mean land suitability		0.013	0.018	0.016	0.019	0.019	0.020	0.006	0.018	0.006	0.018	0.003
		(0.012)	(0.013)	(0.015)	(0.014)	(0.015)	(0.014)	(0.016)	(0.015)	(0.020)	(0.013)	(0.015)
Range of land suitability		0.013	0.014	0.012	0.011	0.014	0.013	0.010	0.019	0.007	0.017	0.015
		(0.008)	(0.011)	(0.012)	(0.012)	(0.014)	(0.014)	(0.015)	(0.013)	(0.016)	(0.012)	(0.014)
Distance to nearest waterway		0.008	0.005	0.007	0.006	0.007	0.007	0.002	0.000	−0.001	0.004	0.001
		(0.009)	(0.010)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.013)	(0.009)	(0.010)
Executive constraints, 1960–2008 average				0.004	0.004	0.004	0.004	0.006*		0.005		0.008**
				(0.003)	(0.003)	(0.003)	(0.003)	(0.003)		(0.004)		(0.004)
Fraction of years under democracy, 1960–2008				−0.015	−0.014	−0.015	−0.014	−0.012		−0.002		−0.017
				(0.019)	(0.019)	(0.019)	(0.019)	(0.018)		(0.019)		(0.017)
Fraction of years under autocracy, 1960–2008				−0.006	−0.005	−0.005	−0.005	−0.008		−0.009		−0.007
				(0.017)	(0.016)	(0.017)	(0.016)	(0.016)		(0.017)		(0.015)
Log oil production per capita, 1960–2008 average								0.002**		0.002*		0.002*
								(0.001)		(0.001)		(0.001)
Log population, 1960–2008 average								0.003		0.004		0.003
								(0.003)		(0.003)		(0.003)
Log GDP per capita, 1960–2008 average								−0.015***		−0.016***		−0.016***
								(0.005)		(0.005)		(0.004)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Legal origin dummies				×	×	×	×	×		×		×
Colonial history dummies				×	×	×	×	×		×		×
Observations	143	143	143	143	143	143	143	143	119	119	143	143
Partial R^2 of population diversity		0.117	0.046	0.056	0.051	0.063	0.056	0.066	0.094	0.141		
Partial R^2 sum of other diversity measures					0.006	0.010	0.007	0.007		0.009		
First-stage adjusted R^2											0.755	0.759
First-stage partial R^2 of migratory distance											0.478	0.438
First-stage F statistic											211.910	103.087
Adjusted R^2	0.019	0.190	0.197	0.192	0.190	0.193	0.188	0.235	0.256	0.308		
Effect of 10th–90th %ile move in diversity	0.012**	0.027***	0.021**	0.024**	0.023**	0.025**	0.025**	0.026**	0.032**	0.042**	0.039***	0.052***
	(0.005)	(0.008)	(0.010)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.013)	(0.016)	(0.015)	(0.018)
Panel B	Log number of new WCM09 ethnic civil conflict onsets per year, 1960–2005											
Population diversity (ancestry adjusted)	0.217***	0.418***	0.375**	0.385**	0.352**	0.408**	0.366*	0.391**	0.780***	0.904**	0.707***	0.795***
	(0.082)	(0.121)	(0.152)	(0.183)	(0.176)	(0.189)	(0.192)	(0.191)	(0.277)	(0.356)	(0.254)	(0.297)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for ethnic diversity					×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	141	141	141	141	141	141	141	141	117	117	141	141
Partial R^2 of population diversity		0.092	0.050	0.049	0.042	0.054	0.042	0.052	0.108	0.123		
Partial R^2 sum of other diversity measures					0.017	0.007	0.011	0.008		0.015		
First-stage adjusted R^2											0.753	0.760
First-stage partial R^2 of migratory distance											0.475	0.438
First-stage F statistic											206.014	97.246
Adjusted R^2	0.024	0.127	0.165	0.139	0.146	0.138	0.140	0.207	0.194	0.233		
Effect of 10th–90th %ile move in diversity	0.014***	0.027***	0.024**	0.025**	0.023**	0.027**	0.024*	0.025**	0.038***	0.045**	0.046***	0.052***
	(0.005)	(0.008)	(0.010)	(0.012)	(0.011)	(0.012)	(0.012)	(0.012)	(0.014)	(0.018)	(0.016)	(0.019)

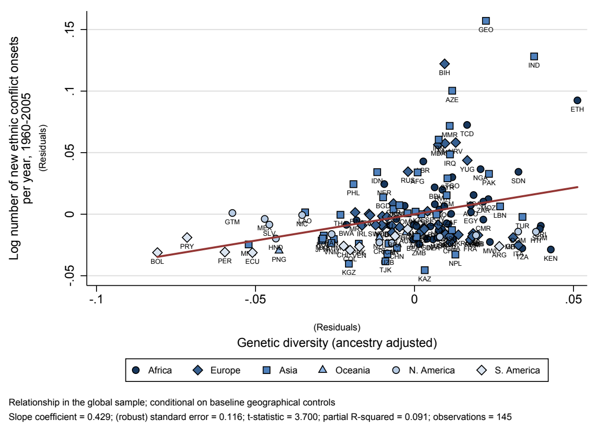
Notes: This table exploits cross-country variations to establish a significant positive reduced-form impact of contemporary population diversity on the annual frequency of (i) new PRIO25 civil conflict onsets during the 1960–2008 time period (Panel A); and (ii) new WCM09 ethnic civil conflict onsets during the 1960–2005 time period (Panel B), conditional on other well-known diversity measures as well as the proximate geographical, institutional, and development-related correlates of conflict. All time-dependent covariates shown in Panel A correspond to their averages or shares for the 1960–2005 time period in Panel B. For regressions based on the global sample, the set of continent dummies includes five indicators for Africa, Asia, North America, South America, and Oceania, whereas for regressions based on the Old-World sample, the set includes two indicators for Africa and Asia. The set of legal origin dummies includes two indicators for British and French legal origins, and the set of colonial history dummies includes three indicators for experience as a colony of the U.K., France, and any other major colonizing power. The 2SLS regressions exploit prehistoric migratory distance from East Africa to the indigenous (precolonial) population of a country as an excluded instrument for the country's contemporary population diversity. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.



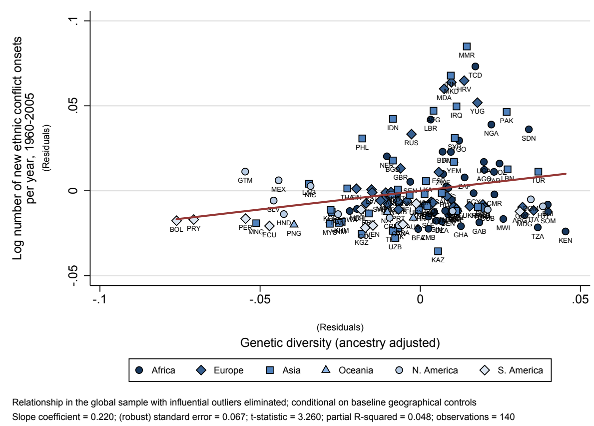
(a) Civil conflicts (full sample)



(b) Civil conflicts (outliers omitted)



(c) Ethnic civil conflicts (full sample)



(d) Ethnic civil conflicts (outliers omitted)

Figure 1: Population Diversity and the Frequency of Civil Conflict Onset across Countries

Notes: This figure depicts the global cross-country relationship between contemporary population diversity and the annual frequency of either (i) new PRIO25 civil conflict onsets during the 1960–2008 time period (Panels (a) and (b)); or (ii) new WCM09 ethnic civil conflict onsets during the 1960–2005 time period (Panels (c) and (d)), conditional on the baseline geographical correlates of conflict, as considered by the specifications in Column 2 of Panels A and B of Table 3. For each type of conflict, the relationship is depicted for either an unrestricted sample of countries (left panel) or a sample without apparently influential outliers (right panel). Each of the four panels presents an added-variable plot with a partial regression line. Given that the unrestricted samples employed by the left panels are not constrained by the availability of data on the additional covariates considered by the analysis in Table 3, the regression coefficients reported in these panels are marginally different from those presented in Column 2 of Panels A and B of Table 3. The set of influential outliers omitted from the sample in the top-right panel includes Bosnia and Herzegovina (BIH), Ethiopia (ETH), Georgia (GEO), and India (IND), whereas in the bottom-right panel, this set additionally includes Azerbaijan (AZE).

65.6 percent of a standard deviation in the cross-country conflict frequency distribution. This relationship is statistically significant at the 5 percent level. The effect size implied by a similar move is slightly larger for ethnic conflict frequency (0.024 conflicts per year).

Controlling for the potentially confounding impact of colonial legacy and more contemporary institutional factors does not significantly affect the stability and significance of the coefficient on population diversity (Column 4) relative to the estimate in the previous column. This finding is especially reassuring in light of the fact that the coefficients associated with some of the institutional covariates are likely to be afflicted by endogeneity bias.

The regressions in Columns 5–7 indicate that when additionally subjected to controls for ethnic fractionalization and ethnolinguistic polarization, either individually or jointly, the point estimate of the coefficient on population diversity continues to remain largely stable in both magnitude and statistical precision.[27] In contrast, neither ethnic fractionalization nor ethnolinguistic polarization appear to possess any significant explanatory power for the cross-country variation in the temporal frequency of civil conflict outbreaks, conditional on population diversity and our baseline set of geographical and institutional covariates. Specifically, the partial R^2 statistics associated with the regression in Column 7 suggest that while the residual cross-country variation in population diversity can explain 5.6 percent of the residual cross-country variation in conflict frequency, only 0.7 percent of such variation is explained by the residual variations in the two measures of ethnolinguistic fragmentation.

The regression in Column 8 reassuringly indicates that the inclusion of controls for oil production per capita, population size, and GDP per capita to our specification hardly sways the point estimate of the coefficient on population diversity, which remains remarkably stable in both magnitude and statistical significance when compared to the estimates from previous columns. In particular, our coefficient of interest from this regression suggests that conditional on our complete set of controls for geographical characteristics, institutional factors, ethnolinguistic fragmentation, and the correlates of economic development, a move from the 10th to the 90th percentile of the cross-country population diversity distribution is associated with an increase in conflict frequency by 0.026 new PRIO25 civil conflict outbreaks per year (or 81.2 percent of a standard deviation in the cross-country conflict frequency distribution, which is comparable to a move from the 50th to the 90th percentile of this distribution). Moreover, the adjusted R^2 statistic of the regression suggests that our baseline empirical model explains 23.5 percent of the cross-country variation in conflict frequency, whereas the partial R^2 statistic associated with population diversity indicates that the residual cross-country variation in population diversity can explain 6.6 percent of the residual cross-country variation in conflict frequency.

Addressing Endogeneity Our results thus far demonstrate a significant and robust cross-country association between population diversity and the temporal frequency of civil conflict onsets over the last half-century, even after conditioning the analysis on a sizable set of controls for geographical characteristics, institutional factors, ethnolinguistic fragmentation, and development outcomes. Nevertheless, this association could be marred by endogeneity bias, in light of the possibility that the large-scale human migrations of the post-1500 era —incorporated into our ancestry-adjusted measure of population diversity for contemporary national populations— and the spatial pattern of conflicts in the modern era could be codetermined by common unobserved forces (e.g., the spatial pattern of *historical* conflicts) that may not be fully captured by our control variables. Coefficient stability analyses we present in Appendix A.1 suggest that selection on unobservables should be unreasonably strong to fully drive our results. Nonetheless we cannot rely on the OLS point estimates to evaluate the substantive importance of population diversity in explaining conflict. As discussed previously in Section 3.1, we use two alternative identification strategies to address this issue. In Columns 9 and 10, we implement our first approach to causal identification by simply restricting the OLS estimator to exploit variations in a subsample of countries that only belong to the Old World. Then in Columns 11 and 12, we run 2SLS regressions using the migratory distance of the prehistorically native population in each country from East

[27]By restricting both fractionalization and polarization measures to enter our regressions linearly, our baseline approach follows Esteban, Mayoral and Ray (2012), but we nevertheless checked the robustness of our main finding to employing alternative specifications that allow for both a linear and a quadratic term in ethnic fractionalization, and we found qualitatively similar results (not reported).

Africa as an instrument for the country's contemporary population diversity. The identifying assumption is that the migratory distance from East Africa is exogenous to the risk of civil conflict in the post-1960 time period. Our instrument plausibly satisfies the exclusion restriction, conditional on our rich set of geographical controls, institutional factors and other correlates of economic development.

Comparing the specifications with the same set of covariates, the two alternative identification strategies yield remarkably similar results, with the point estimate of the coefficient on population diversity being noticeably larger in magnitude, relative to its less well-identified counterpart (in either Column 3 or Column 8), based on an OLS regression in our global sample of countries. The Old World regressions in Columns 9 and 10 suggest that a move from the 10th to the 90th percentile of the cross-country population diversity distribution in the Old World leads to an increase in conflict frequency by 0.032 to 0.042 new PRIO25 civil conflict outbreaks per year. These correspond to 91.9 to 123 percent of a standard deviation in the cross-country conflict frequency distribution in the Old World. Both estimates are statistically significant at the 5 percent level. 2SLS estimates on the global sample suggest that a move from the 10th to the 90th percentile of the global cross-country population diversity distribution leads to an increase in conflict frequency by 0.039 to 0.052 new PRIO25 civil conflict outbreaks per year. These estimates are equivalent to 122 and 164 percent of a standard deviation in the global cross-country conflict frequency distribution respectively and both are statistically significant at the 1 percent level. The corresponding estimates from ethnic conflict regressions in Panel B are also significant and similar in magnitude to those obtained for overall civil conflicts.

In our view, there are three distinct rationales —perhaps operating in tandem— for why our better-identified point estimates of the coefficient on population diversity are larger than their less well-identified counterparts. First, the spatial pattern of social conflict may exhibit long-term persistence, for reasons other than population diversity. If persistent conflict spurred emigrations and atrocities that gradually led to systematically more homogenous populations (Fletcher and Iyigun, 2010), there should be a downward bias in the estimated coefficient on population diversity in an OLS regression that explains the global variation in civil conflict potential in the modern era.

A second plausible explanation is that the pattern of conflict risk in the modern era, especially across populations in the New World that experienced a substantial increase in diversity from migrations in the post-1500 era, has been influenced not so much by the higher population diversity of the immigrants but more so by the unobserved (or observed but noisily measured) human capital that European settlers brought with them, the colonization strategies that they pursued, and the sociopolitical institutions that they established. To the extent that these unobserved factors associated with European settlers in the New World served, in one way or another, to reduce the risk of social conflict in the modern national populations of the Americas and Oceania, they could also introduce a negative bias in the OLS-estimated relationship between population diversity and conflict risk in a global sample of countries.

A third possible rationale is that in the end, population diversity explains the conflict propensity of a population mostly through its prehistorically determined component. This component may have contributed to the formation and ethnic differentiation of *native* groups in a given location and, thus, to more deeply rooted interethnic divisions amongst these groups. As such, conditional on continent fixed effects that absorb any systematic differences in the pattern of post-1500 population flows into locations in the Old World versus the New World, our ancestry-adjusted measure of interpersonal diversity —that incorporates the diversity of both native and nonnative groups in a contemporary national population— might be a noisy proxy for the "true" measure of prehistorically determined population diversity. Due to this "measurement error", the influence of the ancestry-adjusted measure of population diversity might be attenuated in an OLS regression that exploits worldwide variations.

Given that both of our identification strategies ultimately exploit the variation in population diversity across populations that have been prehistorically indigenous to their current locations, either by omitting the modern national populations of the New World from the estimation sample or by instrumenting contemporary population diversity in a globally representative sample of countries with the prehistoric migratory distance of a country's geographical location from East Africa, our better-identified estimates mitigate all the aforementioned sources of negative bias.

Robustness Checks In Section A.1 of Appendix A, we present several robustness checks for our cross-country analysis of the influence of population diversity on the temporal frequency of either overall or ethnic civil conflict outbreaks in the post-1960 time horizon. We demonstrate that our main findings are qualitatively robust to **(1)** additionally considering the influence of average intergroup genetic distance in the national population as well as the interaction between our measure of population diversity and intergroup genetic distance (Table A.1); **(2)** accounting for the potentially confounding influence of various ecological and climatic covariates, including the temporal means and volatilities of annual temperature and precipitation over the relevant sample period as well as ecological fractionalization and polarization (Table A.2); **(3)** accounting for timing of the Neolithic Revolution, state antiquity, duration of human settlement and distance from regional frontier in 1500 (Table A.3); **(4)** accounting for inequality across ethnic homelands as well as overall spatial inequality in nighttime lights within a country (Table A.4); **(5)** accounting for linguistic rather than ethnic fractionalization as a covariate in the baseline specifications (Table A.5); **(6)** controlling for values of time-varying baseline covariates measured in the initial year instead of their temporal means over the sample period (Table A.6); **(7)** explaining the total count rather than the annual frequency of new conflict onsets over the relevant time horizon (Table A.7); **(8)** accounting for spatial dependence across observations (Tables A.8 and A.9); **(9)** computing bootstrapped standard errors (Table A.10); **(10)** examining the temporal frequency of overall civil conflict outbreaks in the sample of countries for which data on ethnic civil conflict events are available (Table A.11); and **(11)** eliminating a priori statistically influential world regions from the estimation sample (Table A.12).

3.2.2 Analysis of Civil Conflict Incidence in Repeated Cross-Country Data

The second dimension of civil conflict that we examine is its temporal prevalence. Specifically, exploiting the time structure of quinquennially repeated cross-country data, we investigate the predictive power of population diversity for the likelihood of observing the incidence of one or more active conflict episodes in a given 5-year interval during the post-1960 time horizon. We estimate the following probit model using maximum-likelihood estimation.

$$\begin{aligned} CP^*_{i,t} = \gamma_0 + \gamma_1 \widehat{GD}_i + \gamma_2' GEO_i + \gamma_3' INS_{i,t-1} + \gamma_4' ETH_i + \gamma_5' DEV_{i,t-1} + \gamma_6 C_{i,t-1} \\ + \gamma_7' \delta_t + \eta_{i,t} \equiv \gamma' Z_{i,t} + \eta_{i,t}; \end{aligned} \tag{3}$$

$$C_{i,t} = \mathbf{1}\left[CP^*_{i,t} \geq D^*\right]; \tag{4}$$

$$Pr\left(C_{i,t} = 1 | Z_{i,t}\right) = Pr\left(CP^*_{i,t} \geq D^* | Z_{i,t}\right) = \Phi\left(\gamma' Z_{i,t} - D^*\right), \tag{5}$$

where $CP^*_{i,t}$ is a latent variable measuring the potential for an active conflict episode in country i during any given 5-year interval, t, and it is modeled as a linear function of explanatory variables. Further, the time-invariant explanatory variables $\widehat{GD}_i$, GEO_i, and ETH_i are all as previously defined, but now, the time-varying covariates included in $INS_{i,t-1}$ and $DEV_{i,t-1}$ enter as their respective temporal means over the previous 5-year interval. δ_t is a vector of time-interval (5-year period) dummies, and $\eta_{i,t}$ is a country-period-specific disturbance term.[28] By specifying each of our time-varying controls to enter the model with a one-period lag, we aim to mitigate the concern that the use of contemporaneous measures of these covariates may exacerbate reverse-causality bias in their estimated coefficients.[29] Finally, we assume that contemporary conflict potential additionally depends on the lagged incidence of civil conflict, $C_{i,t-1}$, which accounts for the possibility that countries with a conflict experience in the immediate past may exhibit a higher conflict potential in the current period, mainly because of the intertemporal spillovers that are common to most conflict processes —e.g., the self-reinforcing nature of past casualties on either side of a conflict.[30] Because the continuous variable reflecting conflict potential, $CP^*_{i,t}$, is unobserved, its level can only be inferred from the binary incidence variable, $C_{i,t}$, indicating whether the latent conflict potential was sufficiently intense for the annual battle-related death threshold of a civil conflict episode to have been surpassed during a given 5-year interval. As is evident from equations (4)-(5), D^* is the corresponding threshold for unobserved conflict potential, and it appears as an intercept in $\Phi(.)$, the cumulative distribution function for the disturbance term, $\eta_{i,t}$.

We present our results for the temporal prevalence (or incidence) of civil conflict in Panel A of Table 4. The first four columns report our findings from regressions explaining the incidence of PRIO25 civil conflict episodes, whereas the remaining four collect our results from regressions explaining the incidence of WCM09 ethnic civil conflict episodes. To keep the exposition concise, we only report better-identified point estimates, namely the Probit regressions in the Old World sample and the IV Probit regressions in the global sample.[31] For each outcome variable and for each identification strategy, we estimate two distinct specifications; one that partials out the influence of only exogenous geographical covariates (including continent fixed effects), and the other that conditions the analysis on the full set of controls in our baseline empirical model of conflict incidence.

The IV probit regression presented in Column 4 suggests that conditional on our complete set of baseline controls, a 1 percentage point increase in population diversity leads to an increase in the quinquennial likelihood of a PRIO25 civil conflict incidence by 2.49 percentage points. This average marginal effect is statistically significant at the 1 percent level. The corresponding marginal effect on the quinquennial likelihood of a WCM09 ethnic civil conflict incidence is 2.11 percentage points —an estimated average marginal effect that is also statistically significant at the 1 percent level (Column 8).

The plots presented in the top row of Figure 2 illustrate precisely how the *predicted* likelihoods —associated with the incidence of either PRIO25 civil conflicts (Subfigure A) or WCM09

[28]We confirm the robustness of our analysis of conflict incidence to exploiting variations in annually (rather than quinquennially) repeated cross-country data. Naturally, in those regressions, the time-dependent covariates enter as their lagged annual values (instead of their lagged 5-year temporal means) and time fixed effects are captured by a set of year dummies.

[29]An alternative method to address the reverse-causality problem, in the context of quinquennially repeated cross-country data, is to control for time-dependent covariates as measured in the initial year of each 5-year interval. Although this method would retain the first period-observation for each country, which is dropped under the current specification, it leaves open the possibility that the presence or absence of an active conflict in the first year of each period may still exert a direct influence on the time-varying controls.

[30]In adopting this strategy, our analysis of conflict incidence follows Esteban, Mayoral and Ray (2012). We also note here that because our measure of population diversity is time-invariant (as is indeed the case with all known measures of ethnolinguistic fragmentation, based on fractionalization or polarization indices), we are unable to account for country fixed effects in our model or exploit dynamic panel estimation methods, despite the time dimension in our repeated cross-country data. In all our regressions exploiting such data, however, the robust standard errors of the estimated coefficients are always clustered at the country level.

[31]Probit results for the global sample are available upon request.

Table 4: Interpersonal Population Diversity and the Incidence or Onset of Civil Conflict in Repeated Cross-Country Data

Cross-country sample:	Old World		Global		Old World		Global	
	(1) Probit	(2) Probit	(3) IV Probit	(4) IV Probit	(5) Probit	(6) Probit	(7) IV Probit	(8) IV Probit
Panel A	Quinquennial PRIO25 civil conflict incidence, 1960–2008				Quinquennial WCM09 ethnic civil conflict incidence, 1960–2005			
Population diversity (ancestry adjusted)	11.883***	12.043**	12.540***	12.833***	20.171***	21.488***	15.379***	15.732***
	(4.502)	(4.686)	(4.215)	(4.807)	(5.830)	(6.174)	(5.198)	(5.897)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions		×		×		×		×
Controls for oil, population, and income		×		×		×		×
Controls for ethnic diversity		×		×		×		×
Observations	944	944	1,154	1,154	927	927	1,039	1,039
Countries	119	119	141	141	117	117	129	129
Pseudo R^2	0.423	0.457			0.516	0.549		
Marginal effect of diversity	2.231***	2.139***	2.538***	2.491***	2.600***	2.596***	2.170***	2.107***
	(0.810)	(0.816)	(0.848)	(0.947)	(0.740)	(0.755)	(0.725)	(0.783)
Panel B	Annual PRIO25 civil conflict onset, 1960–2008				Annual WCM09 ethnic civil conflict onset, 1960–2005			
Population diversity (ancestry adjusted)	9.088***	10.302***	10.053***	11.561***	13.251***	14.710***	10.120***	10.885***
	(2.931)	(3.132)	(2.700)	(3.304)	(3.532)	(3.801)	(3.236)	(4.077)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Peace duration cubic splines	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions		×		×		×		×
Controls for oil, population, and income		×		×		×		×
Controls for ethnic diversity		×		×		×		×
Observations	3,849	3,828	4,896	4,874	3,607	3,585	4,038	4,016
Countries	119	119	141	141	117	117	129	129
Pseudo R^2	0.120	0.141			0.146	0.180		
Marginal effect of diversity	0.455***	0.508***	0.492***	0.560***	0.620***	0.669***	0.563***	0.677**
	(0.153)	(0.163)	(0.157)	(0.191)	(0.181)	(0.187)	(0.210)	(0.278)

Notes: This table exploits variations in repeated cross-country data to establish a significant positive reduced-form impact of contemporary population diversity on the likelihood of observing (i) the incidence of a PRIO25 civil conflict in any given 5-year interval during the 1960–2008 time period (Panel A, Columns 1–4); (ii) the incidence of a WCM09 ethnic civil conflict in any given 5-year interval during the 1960–2005 time period (Panel A, Columns 5–8); (iii) the onset of a new PRIO25 civil conflict in any given year during the 1960–2008 time period (Panel B, Columns 1–4); and (iv) the onset of a new WCM09 ethnic civil conflict in any given year during the 1960–2005 time period (Panel B, Columns 5–8), conditional on other well-known diversity measures as well as the proximate geographical, institutional, and development-related correlates of conflict. The controls for geography include absolute latitude, land area, ruggedness, distance to the nearest waterway, the mean and range of agricultural suitability, and the mean and range of elevation. The controls for ethnic diversity include ethnic fractionalization and polarization. The controls for institutions include a set of legal origin dummies, comprising two indicators for British and French legal origins, as well as six time-dependent covariates that capture the lagged values of executive constraints, two indicators for the type of political regime (democracy and autocracy), and three indicators for experience as a colony of the U.K., France, and any other major colonizing power. The controls for oil, population, and income include three time-dependent covariates that capture the log-transformed lagged values of oil production per capita, total population, and GDP per capita. In Panel A, all aforementioned time-dependent covariates assume their average annual values over the previous 5-year interval, whereas in Panel B, they assume their annual values from the previous year. To account for duration and temporal dependence in conflict outcomes, all regressions control for the lagged incidence of conflict, and all regressions in Panel B additionally control for a set of cubic splines of the number of peace years. For regressions based on the global sample, the set of continent dummies includes five indicators for Africa, Asia, North America, South America, and Oceania, whereas for regressions based on the Old-World sample, the set includes two indicators for Africa and Asia. The IV probit regressions exploit prehistoric migratory distance from East Africa to the indigenous (precolonial) population of a country as an excluded instrument for the country's contemporary population diversity. The estimated marginal effect of a 1 percentage point increase in population diversity is the average marginal effect across the entire cross-section of observed diversity values, and it reflects the increase in either the quinquennial likelihood of a conflict incidence (Panel A) or the annual likelihood of a conflict onset (Panel B), both expressed in percentage points. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

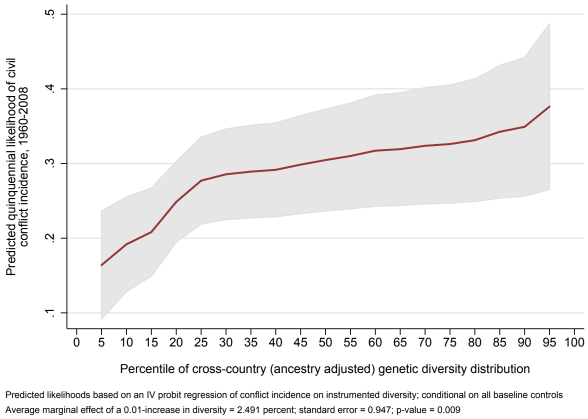
ethnic civil conflicts (Subfigure B)— vary as one moves along the global cross-country diversity distribution, based on the IV probit regressions from Columns 4 and 8 in Panel A of Table 4.[32] The economically significant influence of population diversity is evident in these plots, which indicate that a move from the 10th to the 90th percentile of the cross-country diversity distribution in the relevant estimation sample leads to an increase in the predicted quinquennial likelihood of civil conflict incidence from 19.2 percent to 34.9 percent for PRIO25 civil conflicts, and from 12.2 percent to 23.8 percent for WCM09 ethnic civil conflicts.

Robustness Checks In Section A.2 of Appendix A, we establish that our baseline estimates of the impact of population diversity on civil conflict incidence are qualitatively insensitive to **(1)** accounting for ecological diversity and polarization as well as the influence of time-varying climatic variables measured in the previous 5-year interval (Panel A in Table A.13); **(2)** accounting for the deep-rooted legacy of various correlates of historical development such as the history of statehood, time elapsed since the Neolithic Revolution and the duration of human settlement (Panel A in Table A.14); **(3)** controlling for ethnic and spatial inequality in nighttime lights within each country (Panel A in Table A.15); **(4)** accounting for alternative distributional indices of intergroup diversity (Alesina et al., 2003; Fearon, 2003; Esteban, Mayoral and Ray, 2012) and for additional time-invariant geographical and historical correlates of conflict potential, including the percentage of mountainous terrain, the presence of any noncontiguous subnational territories, the intensity of the disease environment (Table A.16); **(5)** considering alternative definitions and types of intrastate conflict as the outcome variable, such as the prevalence of large-scale civil conflicts —i.e., "civil wars"— as well as intrastate conflicts involving only nonstate actors (Table A.17); **(6)** exploiting variations in annually rather than quinquennially repeated cross-country data (Table A.18); **(7)** empirically modeling conflict prevalence using either classical logit or "rare events" logit (King and Zeng, 2001) estimators, in lieu of the standard probit estimator (Panel A in Table A.19); and **(8)** assessing statistical significance of our estimates under two-way cluster-robust standard errors that allow arbitrary within-period correlation across observations in addition to arbitrary within-country correlation (Panel A in Table A.20).

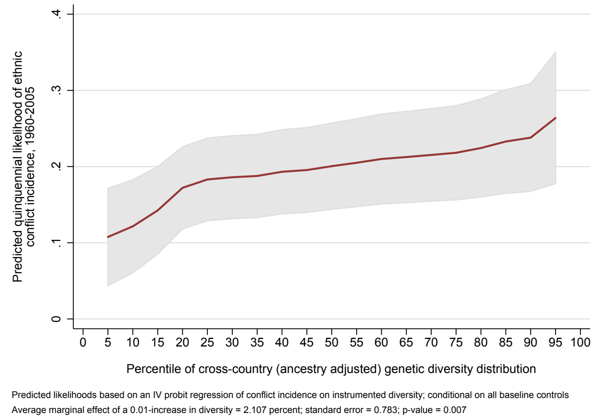
3.2.3 Analysis of Civil Conflict Onset in Repeated Cross-Country Data

The third dimension of conflict examined by our analysis is the onset of civil conflict. Unlike the model of conflict incidence, the onset model focuses solely on explaining the outbreak of conflict events, classifying the subsequent years into which a given conflict persists as nonevent years (akin to civil peace), unless they coincide with the eruption of another conflict. Conceptually, this model assesses the extent to which population diversity at the national level influences sociopolitical instability by *triggering* conflicts, rather than only contributing to their perpetuation over time. The probit model for our analysis of conflict onset is similar to the model of conflict incidence, described by equations (3)-(5), except that now, following the convention in the literature, (i) we exploit variations in annually repeated cross-country data, with our binary outcome variable assuming a value of 1 if a country-year observation coincides with the first year of a "new period" of conflict (as discussed below), and 0 otherwise; and (ii) a set of cubic splines of the number of preceding years of uninterrupted peace is included as a control, along with year dummies, in order to account for temporal or duration dependence (Beck, Katz and Tucker, 1998). Further, to mitigate issues

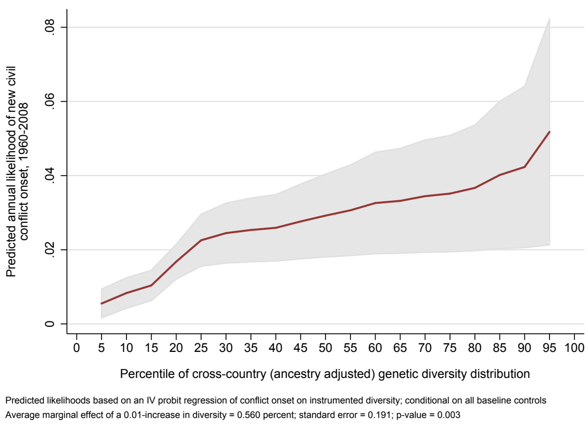
[32] Similar to Figure 2, the two panels of Figure A.1 in Appendix A.3 depict the manner in which the predicted quinquennial likelihoods —associated with the incidence of either PRIO25 civil conflicts (Subfigure A) or WCM09 ethnic civil conflicts (Subfigure B)— respond as one moves along the cross-country diversity distribution in the Old World, based on the probit regressions from Columns 2 and 6.



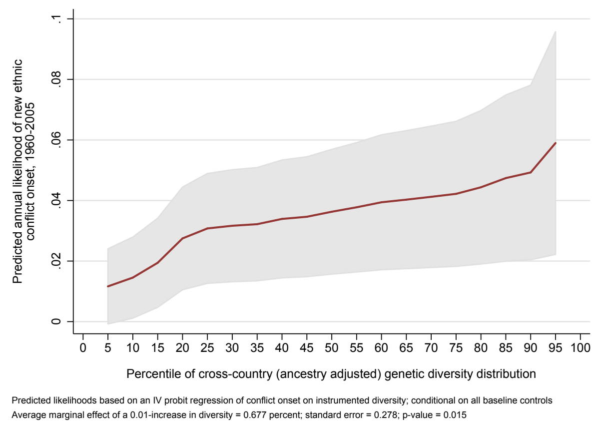
(a) Civil conflict incidence



(b) Ethnic civil conflict incidence



(c) Civil conflict onset



(d) Ethnic civil conflict onset

Figure 2: Population Diversity and the Incidence or Onset of Civil Conflict

Notes: This figure depicts the influence of contemporary population diversity at the country level on the *predicted* likelihood of observing (i) the incidence of a PRIO25 civil conflict in any given 5-year interval during the 1960–2008 time period (Panel (a)); (ii) the incidence of a WCM09 ethnic civil conflict in any given 5-year interval during the 1960–2005 time period (Panel (b)); (iii) the onset of a new PRIO25 civil conflict in any given year during the 1960–2008 time period (Panel (c)); and (iv) the onset of a new WCM09 ethnic civil conflict in any given year during the 1960–2005 time period (Panel (d)), conditional on other well-known diversity measures, the proximate geographical, institutional, and development-related correlates of conflict, controls for temporal dependence in conflict outcomes, and continent and time dummies, as considered by the specifications in Columns 4 and 8 of Panels A and B of Table 4 for the global sample of countries. In each panel, the predicted likelihood of the conflict outcome is illustrated as a function of the percentile of the cross-country diversity distribution, and the shaded area reflects the 95-percent confidence-interval region of the depicted relationship.

of causal identification of the influence of population diversity on conflict onset, we implement the same two strategies followed by our analyses of conflict frequency and conflict incidence.

The results for the onsets of overall and ethnic civil conflicts are presented in Panel B of Table 4. Irrespective of the identification strategy employed, or the set of covariates considered by the specification, population diversity confers a highly statistically significant and qualitatively robust positive influence on the annual likelihood of civil conflict outbreaks. To elucidate the economic significance of this impact in a globally representative sample of countries, the estimated average marginal effects associated with the IV probit regressions in Columns 4 and 8 suggest that, accounting for the influence of geographical conditions, institutional factors, ethnolinguistic fragmentation, and development outcomes, a 1 percentage point increase in population diversity leads to an increase in the annual likelihood of a PRIO25 civil conflict outbreak by 0.560 percentage points, and it leads to an increase in the annual likelihood of a new WCM09 ethnic civil conflict eruption by 0.677 percentage points.

The plots in the bottom row of Figure 2 show the predicted likelihoods of conflict onset as we move along the cross-country distribution of population diversity. In response to a move from the 10th to the 90th percentile of the cross-country diversity distribution, the predicted annual likelihood of a PRIO-NC onset event rises from 0.834 percent to 4.23 percent (Subfigure C), and that of a new WCM09 ethnic civil conflict outbreak rises from 1.45 percent to 4.93 percent (Subfigure D).

Robustness Checks In Section A.2 of Appendix A, we demonstrate that our baseline findings regarding the impact of population diversity on civil conflict onset are qualitatively unaltered after **(1)** accounting for ecological diversity and polarization as well as the influence of time-varying climatic variables measured in the previous 5-year interval (Panel B in Table A.13); **(2)** accounting for the deep-rooted legacy of various correlates of historical development such as the history of statehood, time elapsed since the Neolithic Revolution and the duration of human settlement (Panel B in Table A.14); **(3)** controlling for ethnic and spatial inequality in nighttime lights within each country (Panel B in Table A.15); **(4)** empirically modeling conflict prevalence using either classical logit or "rare events" logit (King and Zeng, 2001) estimators, in lieu of the standard probit estimator (Panel B in Table A.19); **(5)** assessing statistical significance of our estimates under two-way cluster-robust standard errors that allow arbitrary within-year correlation across observations in addition to arbitrary within-country correlation (Panel B in Table A.20); **(6)** accounting for the influence of additional correlates of the propensity for conflict outbreaks, including the time-invariant "ethnic dominance" indicator of Collier and Hoeffler (2004) and the time-varying "political instability" and "new state" indicators of Fearon and Laitin (2003) (Table A.21); and **(7)** accounting for contemporaneous and lagged impact of annual export price shocks to various commodities studied in Bazzi and Blattman (2014) (Table A.22).

3.2.4 Analysis of Intrastate Conflict Severity in Repeated Cross-Country Data

Our findings thus far establish that the interpersonal diversity of a contemporary national population is a robust and significant reduced-form contributor to the risk of civil conflict in society, as manifested by the frequency, prevalence, and emergence of both overall and ethnic civil conflict events over the past half-century. The outcome variables employed in these analyses have all been based on binary measures that are subject to a predefined threshold of annual battle-related casualties being surpassed for the identification of civil conflict events. Therefore our results so far reflect, broadly speaking, the influence of interpersonal population diversity on the *extensive* margin of conflict. In this section, we explore the influence of population diversity on the *intensive* margin of conflict. In particular, we employ both ordinal and continuous measures that capture the "severity" of intrastate conflicts and social strife including but not limited to armed conflict.

The first measure of conflict intensity that we examine exploits information on the apparent "magnitude scores" associated with "major episodes" of intrastate armed conflict, as reported by the Major Episodes of Political Violence (MEPV) data set (Marshall, 2010).[33] According to this data

[33]The specific version of the MEPV data set that we employ provides annual information for a total of 175 countries over the 1946–2008 time period. See http://www.systemicpeace.org/warlist.htm for further details on our measure of conflict intensity from the MEPV data set.

set, a "major episode" of armed conflict involves both (i) a minimum of 500 directly related fatalities in total; and (ii) systematic violence at a sustained rate of at least 100 directly related casualties per year. Importantly, for each such episode of conflict, the MEPV data set provides a "magnitude score" —namely, an ordinal measure on a scale of 1 to 10 of the episode's destructive impact on the directly affected society, incorporating information on multiple dimensions of conflict severity, including the capabilities of the state, the interactive intensity (means and goals) of the oppositional actors, the area and scope of death and destruction, the extent of population displacement, and the duration of the episode. The specific outcome variable from the MEPV data set that we employ reflects the aggregated magnitude score across all conflict episodes that are classified as one of four types of intrastate conflict —namely, civil war, civil violence, ethnic war, and ethnic violence.[34] In particular, this variable is reported by the MEPV data set at the country-year level, with nonevent years for a country being coded as 0.

Our second measure of conflict intensity is based on annual time-series data on a continuous index of social conflict at the country level, as reported by the Cross-National Time-Series (CNTS) Data Archive (Banks, 2010). Rather than adopting an ad hoc fatality-related threshold for the identification of conflict events, this index provides an aggregate summary of the general level of social discordance in any given country-year, by way of presenting a weighted average, following the methodology of Rummel (1963), across all observed occurrences of eight different types of sociopolitical unrest, including assassinations, general strikes, guerrilla warfare, major government crises, political purges, riots, revolutions, and anti-government demonstrations.[35]

Given that our analysis of conflict severity follows Esteban, Mayoral and Ray (2012) in terms of exploiting variations in quinquennially repeated cross-country data, for each country in our sample, we collapse the annual data on both measures of conflict intensity to a quinquennial time series, by assigning to any given 5-year interval in our post-1960 sample period, the maximum level of conflict intensity reflected by that measure across all years in that 5-year interval. In all regressions, we account for temporal dependence in conflict severity by allowing both the lagged observation of the outcome variable and a full set of time-interval (5-year period) dummies to enter the specification. As before, whenever time-varying covariates are allowed to enter the specification, they do so with a one-period lag. Since the units, in which either of our proxies for conflict intensity are measured in the data, have no natural interpretation, we use standardized outcome variables in the regressions.

Table 5 presents the results from our analysis of the influence of interpersonal diversity on intrastate conflict severity —as reflected by either the MEPV aggregate magnitude score of conflict intensity (Columns 1–4) or the CNTS index of social conflict (Columns 5–8).[36] Notwithstanding the

[34]Specifically, all episodes of intrastate conflict in the MEPV data set are categorized along two dimensions. With respect to the first dimension, an episode may be considered either (i) one of "civil" conflict, involving rival political groups; or (ii) one of "ethnic" conflict, involving the state agent and a distinct ethnic group. In terms of the second dimension, however, an episode may be either (i) one of "violence," involving the use of instrumental force, without necessarily possessing any exclusive goals; or (ii) one of "war," involving violent activities between distinct groups, with the intent to impose a unilateral result to the contention.

[35]The specific weights (reported in parentheses) assigned to the different types of sociopolitical unrest considered by the index are as follows: assassinations (25), general strikes (20), guerrilla warfare (100), major government crises (20), political purges (20), riots (25), revolutions (150), and anti-government demonstrations (10). For further details, the reader is referred to the codebook of the CNTS data archive, available at http://www.databanksinternational.com/32.html.

[36]Despite the fact that our measure of conflict intensity from the MEPV data set is ordinal rather than continuous in nature, we choose to pursue least-squares (as opposed to maximum-likelihood) estimation methods when examining this particular outcome variable, primarily because this permits us to conveniently exploit both of our identification strategies. Specifically, although we are able to qualitatively replicate our key findings from Columns 1–2 using ordered probit rather than OLS regressions (results not shown), the absence (to our knowledge) of a readily available

Table 5: Population Diversity and the Severity of Civil Conflict in Repeated Cross-Country Data

Cross-country sample:	Old World		Global		Old World		Global	
	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS	(5) OLS	(6) OLS	(7) 2SLS	(8) 2SLS
	Quinquennial MEPV civil conflict severity, 1960–2008				Quinquennial CNTS social conflict index, 1960–2008			
Population diversity (ancestry adjusted)	3.897**	5.102**	3.741**	4.516**	6.182**	9.126***	5.644**	7.611***
	(1.854)	(2.037)	(1.769)	(2.019)	(2.912)	(2.815)	(2.760)	(2.912)
Lagged conflict severity	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions		×		×		×		×
Controls for oil, population, and income		×		×		×		×
Controls for ethnic diversity		×		×		×		×
Observations	944	944	1,173	1,173	942	942	1,171	1,171
Countries	119	119	143	143	119	119	143	143
Partial R^2 of population diversity	0.005	0.008			0.007	0.012		
Partial R^2 sum of other diversity measures		0.001				0.001		
First-stage adjusted R^2			0.769	0.791			0.769	0.791
First-stage partial R^2 of migratory distance			0.449	0.416			0.446	0.413
First-stage F statistic			180.390	113.647			170.303	105.199
Adjusted R^2	0.597	0.598			0.217	0.233		
Effect of 10th–90th %ile move in diversity	0.209**	0.274**	0.252**	0.304**	0.332**	0.490***	0.379**	0.512***
	(0.100)	(0.109)	(0.119)	(0.136)	(0.156)	(0.151)	(0.186)	(0.196)

Notes: This table exploits variations in repeated cross-country data to establish a significant positive reduced-form impact of contemporary population diversity on the severity of conflict, as reflected by (i) the maximum value of an annual ordinal index of conflict intensity (from the MEPV data set) across all years in any given 5-year interval during the 1960-2008 time period; and (ii) the maximum value of an annual continuous index of the degree of social unrest (from the CNTS data set) across all years in any given 5-year interval during the 1960-2008 time period, conditional on other well-known diversity measures as well as the proximate geographical, institutional, and development-related correlates of conflict. Given that both measures of conflict severity are expressed in units that have no natural interpretation, their intertemporal cross-country distributions are standardized prior to conducting the regression analysis. The controls for geography include absolute latitude, land area, ruggedness, distance to the nearest waterway, the mean and range of agricultural suitability, and the mean and range of elevation. The controls for ethnic diversity include ethnic fractionalization and polarization. The controls for institutions include a set of legal origin dummies, comprising two indicators for British and French legal origins, as well as six time-dependent covariates that capture the average annual values over the previous 5-year interval of executive constraints, two indicators for the type of political regime (democracy and autocracy), and three indicators for experience as a colony of the U.K., France, and any other major colonizing power. The controls for oil, population, and income include three time-dependent covariates that capture the log-transformed average annual values over the previous 5-year interval of oil production per capita, total population, and GDP per capita. To account for temporal dependence in conflict outcomes, all regressions control for the severity of conflict in the previous 5-year interval. For regressions based on the global sample, the set of continent dummies includes five indicators for Africa, Asia, North America, South America, and Oceania, whereas for regressions based on the Old-World sample, the set includes two indicators for Africa and Asia. The 2SLS regressions exploit prehistoric migratory distance from East Africa to the indigenous (precolonial) population of a country as an excluded instrument for the country's contemporary population diversity. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of standard deviations of the intertemporal cross-country distribution of conflict severity. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

measure for conflict intensity examined, the identification strategy exploited, or the set of covariates considered by the specification, the results from our analysis of conflict severity in Table 5 establish population diversity as a qualitatively robust and statistically significant reduced-form contributor to the intensive margin of intrastate conflict. A move from the 10th to the 90th percentile of the cross-country diversity distribution in the relevant sample leads to an increase in conflict severity by between 27.4 percent and 30.4 percent of a standard deviation from the observed distribution of the MEPV magnitude score of conflict intensity. The same move in population diversity leads to

IV counterpart of the ordered probit regression model precludes conducting a similar robustness check on our key findings from Columns 3–4.

an increase in CNTS index of social conflict by 49 percent to 51.2 percent of a standard deviation of its observed distribution.

3.2.5 Analysis of Intragroup Factional Conflict Incidence in Cross-Country Data

One crucial dimension in which our measure of population diversity at the national level adds value beyond all known indices of ethnolinguistic fragmentation is that the index we employ as a proxy for population diversity, incorporates information on interpersonal heterogeneity not only across group boundaries but within such boundaries as well. As such, from a conceptual viewpoint alone, and in contrast to measures that capture the degree of ethnolinguistic fragmentation of a national population, to the extent that interpersonal heterogeneity can be expected to give rise to social, political, and economic grievances that culminate to violent contentions even across ethnically or linguistically homogenous subgroups, our measure is naturally better-suited to empirically link intrapopulation diversity with the incidence of such forms of conflict in society. Our analysis in this section elucidates precisely this virtue of our measure, by exploiting cross-country variations in the likelihood of observing the incidence of one or more *intragroup* factional conflict events during the 1990–1999 time period.

The primary source of our data on the incidence of intragroup factional conflict events across the globe is the Minorities at Risk (MAR), Phase IV data set (Minorities at Risk Project, 2009), which provides for each country with a national population of at least half a million, information on each subnational (i.e., nonstate communal) group that is considered a "minority at risk" —namely, an ethnopolitical group that (i) collectively suffers or benefits from systematic discriminatory treatment vis-à-vis other groups in the national population; and/or (ii) collectively mobilizes resources in defense or promotion of its self-defined interests. Specifically, for each such subnational group, the MAR data set furnishes an indicator for whether the group experienced any intragroup factional conflict event during the 1990–1999 time horizon. For our purposes, we simply aggregate this information to the country level, by coding a binary variable that reflects whether any of the MAR groups within a given country had an experience with intragroup factional conflict over this time span.[37]

The results from our cross-country analysis of intragroup factional conflict events are collected in Table 6. Depending on the identification strategy we employ, we use either a Probit or IV Probit estimator. For each of our two identification strategies, we present the results from estimating three alternative specifications. The first two of these specifications follow from our expositional methodology in previous sections, in that one conditions the analysis only on exogenous geographical covariates (including continent fixed effects), whereas the other partials out the influence of our full set of baseline controls including institutional factors, ethnolinguistic fragmentation, and development outcomes. Because the MAR groups in a given country may not be representative of all of its subnational groups, if higher interpersonal diversity in a national population happens to be associated with a higher prevalence of MAR groups, and if MAR groups also happen to face a higher risk (relative to non-MAR groups) of intragroup factional conflict, then any observed positive influence of population diversity on the incidence of such conflict could be spurious. Therefore our third specification augments the full baseline model with additional controls for the total number and total share of all MAR groups in the national population. Finally, given that our analysis of intragroup factional conflict incidence exploits a standard cross-country framework, wherever relevant, our time-varying controls for institutional factors and development

[37] To be sure, because the MAR data set does not provide information on the specific timing of intragroup factional conflict events, beyond the fact that they occurred at some point in the 1990–1999 time interval, we are restricted by the data to conduct our analysis in a cross-country framework, rather than in a repeated cross-country sample.

TABLE 6: Interpersonal Population Diversity and the Incidence of Intragroup Factional Conflict across Countries

Cross-country sample:	Old World			Global		
	(1) Probit	(2) Probit	(3) Probit	(4) IV Probit	(5) IV Probit	(6) IV Probit
	MAR intragroup conflict incidence, 1990–1999					
Population diversity (ancestry adjusted)	25.761**	35.605**	48.599**	26.992***	41.653***	48.924***
	(11.897)	(16.944)	(19.653)	(9.055)	(12.430)	(12.089)
Continent dummies	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×
Controls for institutions		×	×		×	×
Controls for oil, population, and income		×	×		×	×
Controls for ethnic diversity		×	×		×	×
Controls for number/share of MAR groups			×			×
Observations	84	84	84	103	103	103
Pseudo R^2	0.226	0.397	0.429			
Marginal effect of diversity	7.853**	8.368**	10.910***	8.111***	10.876***	12.320***
	(3.322)	(3.744)	(4.091)	(2.517)	(3.271)	(3.007)

Notes: This table exploits cross-country variations to establish a significant positive reduced-form impact of contemporary population diversity on the likelihood of observing one or more factional conflicts *within* the "minorities at risk" (MAR) groups of a country's population in the 1990–1999 time period, conditional on other well-known diversity measures, the proximate geographical, institutional, and development-related correlates of conflict, and measures capturing the distribution of MAR groups in the national population. The controls for geography include absolute latitude, land area, ruggedness, distance to the nearest waterway, the mean and range of agricultural suitability, and the mean and range of elevation. The controls for ethnic diversity include ethnic fractionalization and polarization. The controls for institutions include a set of legal origin dummies, comprising two indicators for British and French legal origins, a set of colonial history dummies, comprising three indicators for experience as a colony of the U.K., France, and any other major colonizing power, and three time-dependent covariates that capture the average annual values over the 1990–1999 time period of executive constraints and two indicators for the type of political regime (democracy and autocracy). The controls for oil, population, and income include three time-dependent covariates that capture the log-transformed average annual values over the 1990–1999 time period of oil production per capita, total population, and GDP per capita. The controls for the distribution of MAR groups include two covariates that capture the number of MAR groups and their total share in the national population. For regressions based on the global sample, the set of continent dummies includes five indicators for Africa, Asia, North America, South America, and Oceania, whereas for regressions based on the Old-World sample, the set includes two indicators for Africa and Asia. The IV probit regressions exploit prehistoric migratory distance from East Africa to the indigenous (precolonial) population of a country as an excluded instrument for the country's contemporary population diversity. The estimated marginal effect of a 1 percentage point increase in population diversity is the average marginal effect across the entire cross-section of observed diversity values, and it reflects the increase in the likelihood of an intragroup factional conflict incidence in the 10-year interval, 1990–1999, expressed in percentage points. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

outcomes enter the specification as their respective temporal means over the 1990–1999 time interval.

Turning to our findings in Table 6, the results obtained across all specifications and identification strategies invariably indicate that population diversity contributes substantially to the risk of intragroup factional conflict events in society. This influence is not only highly statistically significant but considerable in terms of economic significance as well. For instance, exploiting variations in our globally representative sample of countries, the IV probit regression presented in Column 5 suggests that, conditional on our complete set of baseline controls, a 1 percentage point increase in population diversity leads to an increase in the likelihood of observing the incidence of one or more intragroup factional conflict events in the 10-year interval between 1990 and 1999 by almost 10.9 percentage points. Figure 3 depicts the predicted likelihood of observing one or more intragroup factional conflicts as a function of the percentile of the cross-country diversity

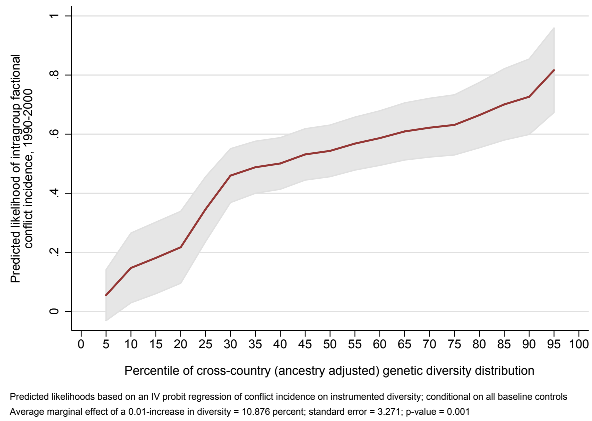


FIGURE 3: Interpersonal Population Diversity and the Incidence of Intragroup Factional Conflict

Notes: This figure depicts the influence of contemporary population diversity at the country level on the *predicted* likelihood of observing one or more factional conflicts *within* the "minorities at risk" (MAR) groups of a country's population during the 1990–1999 time period, conditional on other well-known diversity measures, the proximate geographical, institutional, and development-related correlates of conflict, and continent dummies, as considered by the specification in Column 5 of Table 6 for the global sample of countries. The predicted likelihood of observing one or more intragroup factional conflicts is illustrated as a function of the percentile of the cross-country diversity distribution, and the shaded area reflects the 95-percent confidence-interval region of the depicted relationship.

distribution. A move from the tenth to the ninetieth percentile in this distribution is predicted to raise the likelihood of conflict from about 15 percent to 70 percent.[38]

3.2.6 Analysis of Historical Conflict Outcomes in Cross-Country Data

So far, our analysis was confined to intra-state conflicts of the last half-century. This was a natural choice as it allowed us to focus on post-independence period in former European colonies, ensure a better quality for conflict codings and employ various standard controls used in civil conflict regressions. Having said that, there is a priori no reason why our hypothesis about the role of population diversity would not extend to conflicts that occurred in the distant past.

In this section we investigate whether population diversity predicts historical conflict events in a cross-section of countries. We use information on the locations of violent conflict events over the period 1400-1799, as compiled by Brecke (1999) and geocoded in Dincecco, Fenske and Onorato (2015). We map these events to countries according to modern borders.[39] The sample period excludes 19th century colonial wars related to the Scramble for Africa. This conforms with our goal to examine the deep-rooted legacy of population diversity. After all, most of these wars were the consequence of local resistance to colonial powers or driven by conflicting interests of the colonizers. Thus, we do not expect them to be related to population diversity in a meaningful way.

Brecke adopts a definition of violent conflict that is based on Cioffi-Revilla (1996): "An occurrence of purposive and lethal violence among 2+ social groups pursuing conflicting political goals that results in fatalities, with at least one belligerent group organized under the command of authoritative leadership. The state does not have to be an actor. Data can include massacres

[38]Figure A.2 in Appendix A.3 depicts the predicted likelihood of intragroup factional conflict as considered by the specification in Column 2 of Table 6 for the Old-World sample of countries. The estimated effect is somewhat smaller within the Old World than in the Global sample, but it is still fairly sizable.

[39]Brecke's data has been used by several studies including Iyigun (2008), Zhang et al. (2011), Besley and Reynal-Querol (2014), Michalopoulos and Papaioannou (2016) and Lagerlöf (2014) among others.

of unarmed civilians or territorial conflicts between warlords." Brecke's list includes those conflicts that resulted in at least 32 conflict-related deaths.[40] While it does not systematically distinguish between intrastate and inter-state conflicts, the latter appear to form the basis of the list. As is well known, the recorded conflicts do not represent the whole universe of conflicts. Yet, the conflict list should contain almost all major conflicts documented by historians.

In contrast to our modern conflict regressions, our variable of interest in the current analysis is predicted population diversity that is *not adjusted* to account for the impact of post-1500 population flows. Clearly, the effects of such migration as reflected in the ancestral composition of contemporary populations would not be relevant over the time frame we study. Also, since our variable of interest reflects intra-population diversity, it is not clear at the outset if we should expect any significant relationship between population diversity in a given location and inter-state —as opposed to intra-state— conflicts in that same location. For a conflict location (a modern country in our sample) that is sufficiently close to home territories of both warring parties, our measure can be a reasonable driver of conflict. Since we study conflicts during a time period when long-distance campaigns were quite rare due to the constraints imposed by transportation and warfare technologies, population diversity can in principle explain a considerable part of the cross-country variation in inter-state conflicts. However, even over our sample period of 1400-1799, we expect technological improvements to dilute the predictive power of population diversity for these conflict events.

In Table 7, we summarize our results in ten columns. All regressions include the baseline geographic controls we have employed in the modern conflict analysis. It is worth noting that the Brecke list likely suffers from a regional bias in coverage due to differences in the quality of primary sources as well as in the nature and scale of the events.[41] We try to overcome this potential problem by including region dummies in all our regressions.

In the first five columns, the outcome variable measures the total number of distinct conflict events (in natural logarithm) over different time periods.[42] In Column 1, we use all conflicts for which we have information on locations, i.e. conflicts during the whole sample period of 1400-1799. Next, in Columns 2 and 3, we focus on the two early centuries where we expect the data to be least contaminated by inter-state wars fought by warring parties whose combined population diversity is not representative of the population diversity of the location in which the war was fought. In Columns 4-5, we focus on the last two centuries of our sample period. In Columns 6-10, we investigate the explanatory power of population diversity for conflict incidence at the extensive margin. To do so, we repeat the same analysis as in the first five columns except that now we estimate a probit model where the outcome is a conflict onset dummy that records whether there was any conflict event over the specified time interval.

Our results indicate that pre-colonial population diversity confers a highly statistically significant influence on both the number and the incidence of conflicts that took place several centuries ago. This is true both for conflicts that occurred over the century prior to the discovery of the New World as well as the century that followed. In line with the prior that the influence of population diversity on conflicts ought to dissipate in periods marred by mostly international or interregional conflicts —particularly, between genetically disparate groups like colonial powers versus pre-colonially native populations—, the association between population diversity and conflicts is weaker in the latter centuries of the time period that we study (Columns 4-5 and Columns 9-10).

[40] This fatality level corresponds to a magnitude of 1.5 or higher on Richardsons (1960) base-10 log conflict scale.

[41] For example, primary sources on historical warfare in Sub-Saharan Africa are relatively scarce (Reid, 2014), and unlike the large-scale campaigns common in European warfare, historical conflicts in Africa more often took the form raiding wars.

[42] Before log-transformation, we add one to the number of conflicts to retain countries without any recorded conflict.

Table 7: Precolonial Population Diversity and the Occurrence of Historical Conflicts across Countries

Historical period:	1400–1799	1400–1499	1500–1599	1600–1699	1700–1799	1400–1799	1400–1499	1500–1599	1600–1699	1700–1799
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) Probit	(7) Probit	(8) Probit	(9) Probit	(10) Probit
	Number of conflict onsets in historical period					Onset of any conflict in historical period				
Population diversity	24.511***	19.745***	21.186***	8.452*	9.484*	32.436***	57.868***	53.954***	16.554*	15.518*
	(6.452)	(4.006)	(4.054)	(4.676)	(4.895)	(9.599)	(11.827)	(11.524)	(9.038)	(9.337)
Region dummies	×	×	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×	×	×
Observations	155	155	155	155	155	155	155	155	155	155
Partial R^2 of population diversity	0.089	0.111	0.124	0.019	0.032					
Adjusted R^2	0.415	0.431	0.448	0.354	0.300					
Pseudo R^2						0.258	0.395	0.350	0.250	0.202
Effect of 10th–90th %ile move in diversity	47.602***	12.161***	14.751***	5.058*	4.151*	0.670***	0.658***	0.657***	0.533***	0.491***
	(12.531)	(2.468)	(2.822)	(2.798)	(2.143)	(0.058)	(0.040)	(0.043)	(0.153)	(0.149)

Notes: This table exploits cross-country variations to establish a significant positive reduced-form impact of indigenous (precolonial) population diversity on (i) the number of conflict onsets (Columns 1–5); and (ii) the likelihood of observing one or more conflict onsets (Columns 6–10), either during the entire 1400–1799 time period (Columns 1 and 6) or in each century therein (Columns 2–5 and 7–10), conditional on the baseline geographical correlates of conflict. The controls for geography include absolute latitude, land area, ruggedness, distance to the nearest waterway, the mean and range of agricultural suitability, and the mean and range of elevation. The set of region dummies includes five indicators for the Sub-Saharan Africa, Middle East and North Africa, Europe and Central Asia, East Asia and Pacific, and Latin America and the Caribbean regions. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of either the number of conflict onsets (Columns 1–5) or the percentage-point increase in the likelihood of a conflict onset (Columns 6–10) during the time period examined by the regression. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

The size of the estimated effects are fairly large. OLS estimates imply that, depending on the time period, a move from the 10th to the 90th percentile in the cross-country distribution of population diversity leads to 4.2 to 14.7 more conflicts per century. These magnitudes are significantly larger than those implied by comparable specifications for modern civil conflicts.[43] This difference might reflect the waning, albeit significant, influence of population diversity as we get closer to recent time periods. But it could also be a rather mechanical consequence of measurement problems and the fact that our historical analysis lumps all intra-state and inter-state conflicts together, whereas the modern analysis is confined to civil conflicts only. Probit regressions imply that a move from the 10th to the 90th percentile in the cross-country diversity distribution, on average, leads to an increase in the likelihood of observing a conflict over the period 1400-1799 by 0.67 (or 1.5 standard deviations). Corresponding magnitudes are significantly larger for the two early centuries than for the two later centuries over this period, possibly reflecting the relatively lower predictive power of population diversity in explaining conflicts in later centuries.

These findings attest to the deep-rooted influence population diversity exerts on the conflict potential. They also suggest that the reduced-form causal effects we find in the modern analysis are unlikely to be driven by the political and institutional environment of the post-1960 period. Instead, they likely reflect a more persistent effect.

[43] In Column 3 of Table 3, the estimated impact of the same move in the cross-country distribution of population diversity was 0.02 additional civil conflict outbreaks per year, i.e. two conflicts per century. The 2SLS estimates for the same specification, reported in Column 11 of the same table, were larger (4 conflicts per century) but still significantly fewer than the average magnitude implied by our historical conflict analysis over the entire period of 1400-1799.

4 Population Diversity and Conflict at the Ethnicity Level

This section explores the origins of the prevailing variation in the prevalence and severity of intra-societal conflicts within ethnic groups. This ethnic level analysis mitigates potential concerns regarding the endogeneity of contemporary national borders to population diversity and the intensity of conflict. Moreover, since populations within ethnic homelands have been largely native to these locations at least since the pre-colonial era, the ethnic level analysis further circumvents potential concerns about the effect of conflict on migrations across countries and its potential impact on population diversity. In addition, the exploration of the effect of population diversity on conflict within ethnic groups, as well as within national populations, permits the analysis to disentangle the impact of population diversity within an ethnic group, from the impact of ethnic diversity across groups, and to establish that the effect of population diversity on the risk of conflict is independent of the population scale.

4.1 Data

The ethnic level analysis is conducted, based on a novel geo-referenced dataset consisting of ethnic groups, for which genetic diversity is either observed, or can be predicted. In particular, the geo-referenced dataset exploits the recently assembled data on observed genetic diversity within 232 indigenous ethnic groups across the globe that have been largely isolated and shielded from genetic admixture (Pemberton, DeGiorgio and Rosenberg, 2013).[44] The distribution of these ethnic groups across the globe is depicted in Figure 4 and the summary statistics of this measure of genetic diversity, as documented in Table B.5, establishes that observed diversity ranges from 0.77 among ethnic groups in Africa to 0.58 among those in South America. The geo-referenced dataset maps the genetic diversity of each ethnic group to the geographical characteristics of its ethnic homeland. The data consists of 230 ethnic groups for which genetic diversity is observed.[45] In addition the geo-referenced data consists of the entire ethnic groups in the *Ethnographic Atlas* for which genetic diversity is predicted.[46]

As to the measures of conflicts at the ethnic-group level, in line with the country-level analysis, the main measure is based on the UCDP/PRIO Armed Conflict Dataset (Gleditsch et al., 2002). In particular, the analysis focuses on the average yearly share of the area of each ethnic homeland that was within the boundaries of internal armed conflict events occurring between the government of a state and internal opposition groups over the period 1989–2008. This measure is calculated using the gridded PRIO data (PRIO-GRID version 1.01) as reported by Tollefsen, Strand and Buhaug (2012) based on the UCDP/PRIO Armed Conflict Dataset. This variable is denoted as the *spatio-temporal prevalence of UCDP/PRIO conflicts*. Furthermore, a secondary measure is based on the number of events recorded within each ethnic homeland in the UCDP geo-referenced Event Dataset (Sundberg, Eck and Kreutz, 2012; Croicu and Sundberg, 2015).

[44] This dataset combines eight human genetic diversity datasets based on the 645 loci that they share, including the HGDP-CEPH Human Genome Diversity Cell Line Panel used by Ashraf and Galor (2013a).

[45] The analysis includes all observations on ethnic groups in Pemberton, DeGiorgio and Rosenberg (2013), excluding two ethnicities (the Surui and the Ache of South America) that are largely viewed by population geneticists as extreme outliers in terms of genetic diversity (e.g. Wang et al., 2007). In particular, Ramachandran et al. (2005) omit the Surui, as "an extreme outlier in a variety of previous analyses", and do not include the Ache either. Furthermore, these ethnicities have the lowest levels of genetic diversity in the sample and the largest residuals of an OLS regression of genetic diversity on migratory distance from Addis Ababa. Including these observations, nevertheless, does not affect the qualitative analysis.

[46] Further details on the construction of the data set is presented in Section B.1 in the appendix.

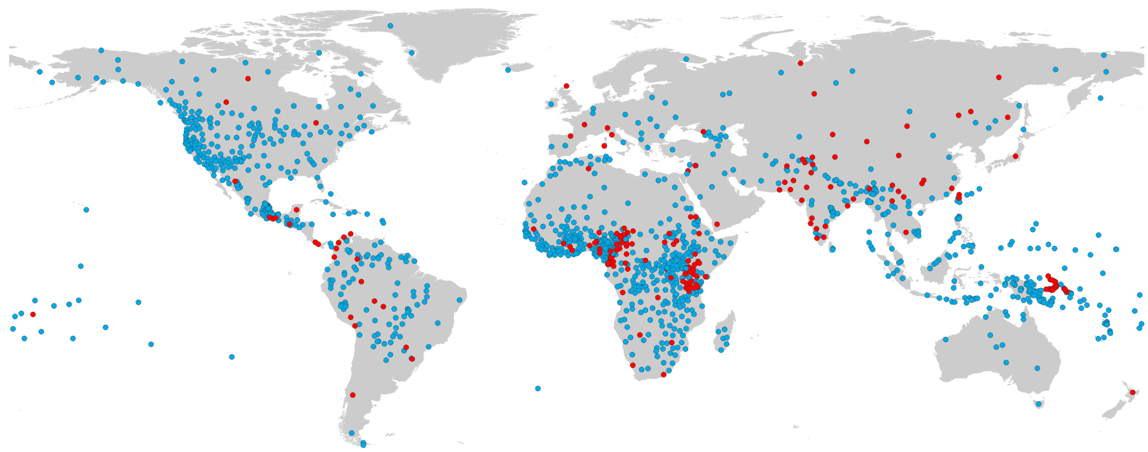


FIGURE 4: The Spatial Distribution of Ethnic Homelands

Notes: This map depicts the global spatial distribution of ethnic homelands. Each point represents the geodesic centroid of the historical homeland of an ethnic group from the *Ethnographic Atlas*. Points in red depict groups for which information on both observed and predicted population diversity are available. On the other hand, points in blue depict groups for which information on only predicted population diversity is available.

4.2 Empirical Strategy

The analysis implements several empirical strategies to mitigate concerns about the potential role of reverse causality, omitted cultural, geographical and human characteristics, as well as sorting in the observed association between population diversity and civil conflicts within ethnic groups. In particular, the positive associations between the extent of the observed population diversity within an ethnic group and civil conflict may reflect reverse causality from conflict to population diversity. It is not inconceivable that in the course of human history conflicts within ethnic groups have operated towards a homogenization of the population, reducing its observed levels of diversity. Hence, in order to remove concerns about reverse causality, as well as concerns about sample limitations, the ethnic level analysis exploits *predicted* population diversity rather than observed diversity to explore the effect of diversity of civil conflict in the ethnic level analysis. In particular, since observed population diversity within a geographically indigenous contemporary ethnic group decreases with distance along ancient migratory paths from East Africa, as established by the serial founder effect (e.g., Harpending and Rogers, 2000; Ramachandran et al., 2005; Prugnolle, Manica and Balloux, 2005; Ashraf and Galor, 2013a), and as depicted for the extended Pemberton sample in Figure 5, migratory distance from Africa is exploited to predict population diversity for all ethnic groups in the *Ethnographic Atlas*.

Furthermore, the associations between ethnic level population diversity and civil conflicts may be governed or biased by omitted cultural, geographical and human characteristics. Thus, in order to mitigate these concerns, the empirical analysis exploits two related strategies. In light of the serial founder effect, the analysis exploits the migratory distance from Africa to each ethnic group as an instrumental variable for the observed level of population diversity, and as a predictor for its level of diversity. Nevertheless, there are several plausible scenarios that would weaken this identification strategy. First, selective migration out of Africa, or natural selection operating in different ways along the migratory paths, could have affected human traits and therefore conflict independently of the effect of migratory distance from Africa on the degree of *diversity* in human traits. Second, migratory distance from Africa could be correlated with distances from focal historical locations (e.g., technological frontiers) and could therefore capture the effect of these distances on the process

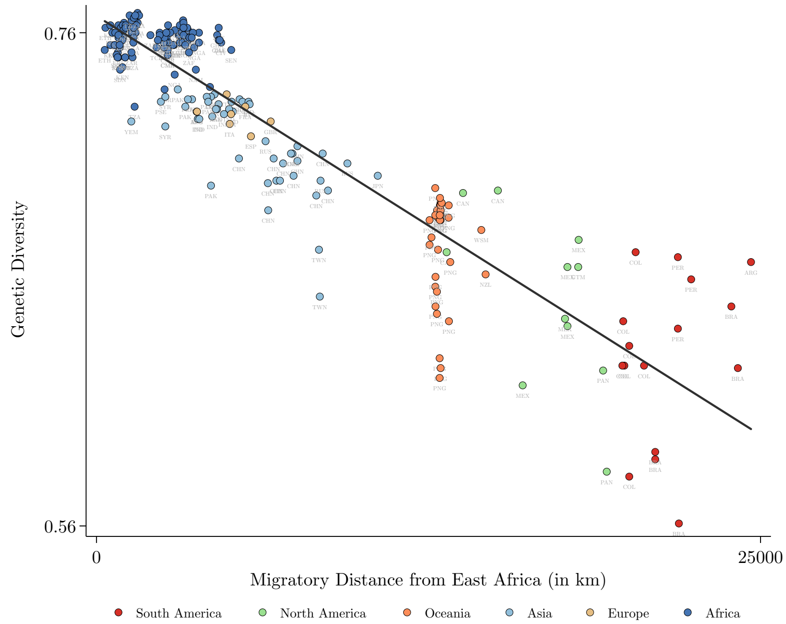


Figure 5: Migratory Distance from East Africa and Observed Population Diversity across Ethnic Groups

Notes: This figure depicts the relationship between prehistoric migratory distance from East Africa and observed population diversity in a sample of 230 ethnic groups (Pemberton et al., 2013). The negative relationship reflects the operation of a serial founder effect originating from East Africa during the demic expansion of humans from the cradle of humankind to the rest of the globe.

of development and the emergence of conflicts, rather than the effect of these migratory distances via population diversity.

These potential concerns are mitigated, however, by the following observations. First, while migratory distance from Africa has a significant negative association with the degree of genetic diversity, it has no association with the *mean* level of human traits, such as height, weight, skin reflectiveness, and IQ (Ashraf and Galor, 2013a), conditional on the distance from the equator, Second, conditional on migratory distance from East Africa, migratory distances from historical technological frontiers in the years 1, 1000, and 1500 do not affect the impact of population diversity on conflict, reinforcing the justification for the reliance on the out of Africa hypothesis and the serial founder effect.

Moreover, a highly implausible threat to the identification strategy would emerge if the actual migration path out of Africa would have been correlated with geographical characteristics that are directly conducive to conflicts (e.g., soil quality, ruggedness, climatic conditions, and propensity to trade). This, however, would have implausibly necessitated that the conduciveness of these geographical characteristics to conflict would be, on average, aligned along the main root of the migratory path out of Africa, as well as along each of the main forks that emerge from this primary path. The further the location is from Africa the lower the conduciveness of these geographical factors to conflict would need to be. In particular, in several important forks in the course of this migration process (e.g., the fertile crescent and the associated eastward migration towards east Asia and western migration towards Europe), the geographical characteristics that are conducive to conflicts would have to diminish symmetrically along these diverging migratory roots. Nevertheless, in order to further mitigate this highly implausible concern, the analysis establishes that the results are unaffected qualitatively if it accounts for the potentially confounding effects of a wide range of geographical factors in the homeland of each ethnic group. In addition, in order to further mitigate concerns regarding the role of omitted variables, the analysis accounts for spatial auto-correlation as well as regional fixed effects, capturing time-invariant unobserved heterogeneity in each region and hence identifying the association between interpersonal diversity and conflict within a geographical region rather than across regions. Furthermore, it establishes that selection on unobservables is not a concern.

The observed associations between population diversity and the extent of conflicts may further reflect the sorting of less diverse populations into geographical niches characterized by lower conflict. While this implausible sorting would not affect the existence of a positive association between population diversity and the extent of conflict, it could weaken the proposed mechanism. However, in view of the serial founder effect and the tight negative association between migratory distance from Africa and population diversity, sorting would necessitate that the ex-ante spatial distribution of conflict would have to be negatively correlated with migratory distance from Africa. As argued above, this would have implausibly necessitated that the conduciveness of geographical characteristics to conflict would be negatively aligned with the primary migratory path out of Africa, as well as with each of its diverging forks, and diminishing symmetrically along these diverging migratory roots. Nevertheless, to further mitigate this highly implausible scenario, the empirical analysis accounts for the potentially confounding effects of a wide range of geographical characteristics, as well as regional fixed effects.

4.3 Empirical Results

This subsection establishes a highly significant and robust reduced-form impact of observed and predicted diversity within an ethnic group on intra-societal conflicts within the ethnic homeland. The analyzes explores the effect of population diversity within ethnic groups on the spatio-temporal prevalence of conflicts, as well as on the extensive and the intensive margins of conflicts at the ethnic-level. The empirical specifications in the ethnic level analysis follows rather closely the specifications in the country-level analysis, assuring the comparability of the findings.[47]

Table 8 presents the results of the baseline analysis of the influence of population diversity within an ethnic group on the log spatio-temporal prevalence of UCDP/PRIO conflicts during 1989–2008. Column 1 establishes a highly significant relation between observed diversity across the 230 ethnic groups and the conflict measure, conditional only on the world-region fixed effects. Column 2, demonstrates that the association remains highly significant and stable if one accounts for the potentially confounding effects of some exogenous geographical factors. Column 3 establishes that — as depicted in Figure 6 — accounting for additional exogenous climatic variables which have been shown to be relevant for conflict, the association between observed diversity and conflict remains significant. The coefficient estimate suggests that an increase in the population diversity from the 10th percentile of the observed level of diversity (e.g., the Guaraní people of South America) to the 90th percentile (e.g., the Bulu people of South Africa) corresponds to an increase in the prevalence of spatio-temporal conflict by 0.33 percentage points (compared to a sample mean of 0.13 and

[47] The set of covariates in the ethnic-level analysis corresponds to the one considered in the country-level analysis. In particular, latitude, ruggedness, the spatial means and standard deviations of elevation and land suitability, and the distance to the nearest waterway are all part of the baseline set of geographical covariates in the country-level analyses. In addition, the climatological covariates — the temporal means and volatilities of temperature and precipitation —- are included in the robustness checks conducted in the appendix. Finally, as reported in the appendix, other control variables such as the time since initial settlement, malaria endemicity, luminosity, and population density have their counterparts in the country-level analysis, in the form of baseline measures of either economic development or the disease environment.

Table 8: Population Diversity and the Spatiotemporal Prevalence of Conflict across Ethnic Homelands

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) 2SLS	(8) 2SLS
	Log spatio-temporal prevalence of UCDP/PRIO conflicts, 1989–2008							
Observed diversity	31.166***	30.763***	21.953**	26.130**			94.424***	88.930***
	[9.094]	[9.253]	[10.347]	[11.129]			[35.527]	[29.949]
Predicted diversity					74.368***	75.000***		
					[5.951]	[5.999]		
Latitude		-0.090***	-0.113***	-0.104**	-0.012	0.012	-0.112***	-0.091**
		[0.018]	[0.040]	[0.045]	[0.017]	[0.017]	[0.040]	[0.042]
Ruggedness		0.238	0.091	0.143	0.163**	0.143*	0.456*	0.380*
		[0.146]	[0.176]	[0.197]	[0.077]	[0.079]	[0.245]	[0.210]
Elevation		-0.846***	0.000	0.067	0.062	0.080	-0.256	-0.182
		[0.294]	[0.456]	[0.468]	[0.191]	[0.193]	[0.508]	[0.508]
S.D. of elevation		-1.460**	-1.111	-0.675	-0.962**	-0.916*	-2.004*	-1.134
		[0.734]	[0.921]	[1.031]	[0.476]	[0.477]	[1.161]	[1.137]
Mean land suitability (native crops)		-0.159*	-0.145*	-0.125	0.016	0.038	-0.240**	-0.182**
		[0.088]	[0.085]	[0.088]	[0.022]	[0.024]	[0.098]	[0.086]
S.D. of land suitability (native crops)		0.535***	0.405*	0.310	0.062	0.112	0.381	0.276
		[0.194]	[0.240]	[0.232]	[0.093]	[0.096]	[0.251]	[0.241]
Change in land suitability (Columbian Exchange)		0.258***	0.010	0.069	-0.240***	-0.241***	0.053	0.081
		[0.083]	[0.091]	[0.099]	[0.037]	[0.037]	[0.093]	[0.094]
Share desert		2.414**	-2.856**	-2.921**	-1.667***	-1.738***	-2.930**	-3.020**
		[0.961]	[1.147]	[1.204]	[0.403]	[0.413]	[1.214]	[1.224]
Nearest waterway		-0.061	-0.051	-0.092	0.040	0.016	-0.054	-0.093
		[0.079]	[0.078]	[0.081]	[0.043]	[0.043]	[0.070]	[0.074]
Average temperature			0.153**	0.159**	0.078***	0.117***	0.174***	0.199***
			[0.063]	[0.070]	[0.025]	[0.025]	[0.067]	[0.070]
Volatility of temperature			7.303***	7.433***	1.570**	1.353**	8.288***	8.233***
			[1.829]	[1.962]	[0.660]	[0.632]	[1.825]	[1.853]
Average precipitation			0.017***	0.016**	0.002	0.001	0.021***	0.019**
			[0.006]	[0.007]	[0.002]	[0.002]	[0.008]	[0.008]
Volatility of precipitation			-0.129***	-0.121***	0.012	0.024*	-0.139***	-0.130***
			[0.039]	[0.044]	[0.012]	[0.012]	[0.048]	[0.049]
Time since initial settlement				0.507		0.537***		1.024*
				[0.495]		[0.203]		[0.596]
Malaria endemicity				0.961		-0.655		0.002
				[1.351]		[0.550]		[1.370]
Log Luminosity				-0.179*		-0.202***		-0.371***
				[0.105]		[0.037]		[0.143]
Regional dummies	×	×	×	×	×	×	×	×
Additional climatic covariates			×	×	×	×	×	×
Decile of population density dummies				×		×		×
Sample	Observed	Observed	Observed	Observed	Extended	Extended	Observed	Observed
Observations	230	230	230	230	1251	1251	230	230
Bootstrapped standard error					[11.493]***	[11.563]***		
Effect of increasing genetic diversity	0.461***	0.455***	0.325**	0.386**	1.215***	1.225***	1.397***	1.315***
from the 10th to the 90th percentile	[0.135]	[0.137]	[0.153]	[0.165]	[0.097]	[0.098]	[0.525]	[0.443]
Adjusted R^2	0.159	0.312	0.482	0.498	0.444	0.461		
First-stage F statistic							10.645	17.311
Beta		30.515	17.722	23.885	65.069	68.517		

Notes: This table exploits cross-ethnicity variations to establish a significant positive reduced-form impact of contemporary population diversity on the log spatio-temporal prevalence of UCDP/PRIO conflicts during the 1989–2008 period, conditional on the confounding effects of geographical and development-related characteristics. For regressions based on the global sample, the set of continental and regional dummies includes indicators for Europe, Asia, North America, South America, Oceania, North Africa, and Sub-Saharan Africa. Additional climatic covariates refer to the average diurnal temperature range, average cloud cover, and average temperature range in the homeland. The 2SLS regressions exploit prehistoric migratory distance from East Africa to each ethnic homeland as an excluded instrument for the observed population diversity of the ethnic group. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the change in the average yearly share of the area of each ethnic homeland that was within the boundaries of internal armed conflict over the period 1989–2008. The "Beta" statistic is the estimated effect of population diversity, if the proportion of selection on observables and unobservables is equal, and the maximal R^2 equal to 1.3 times the observed R^2 (Oster, forthcoming). Robust standard errors are reported in square brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent.

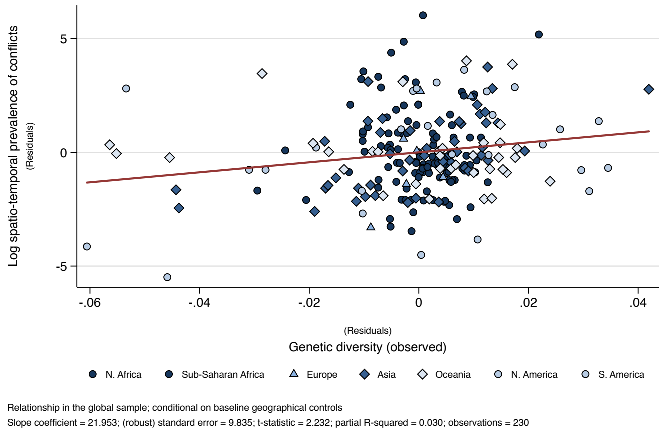


FIGURE 6: Observed Population Diversity and the Spatiotemporal Prevalence of Conflict across Ethnic Homelands

Notes: This figure depicts the global cross-ethnic homeland relationship between contemporary observed population diversity and the log spatio-temporal prevalence of UCDP/PRIO conflicts, 1989–2008, conditional on the baseline geographical correlates of conflict, as considered by the specification in Column 3 of Table 8.

a standard deviation of 0.25).[48] Finally, Column 4, demonstrates that the association remains significant and stable if one accounts for a set of potentially endogenous confounders (luminosity, malaria endemicity, time since settlement, and population density dummies).

In light of the potential endogeneity of observed population diversity, Columns 5 and 6 examine the reduced-form impact of predicted population diversity, based on migratory distance from East Africa, on the prevalence of spatio-temporal conflicts in a sample of 1,251 ethnic groups.[49] They establish that the reduced-form impact of predicted diversity on the prevalence of spatio-temporal conflicts is positive, highly significant and stable as one accounts for geographical and subsequently climatic characteristics.

In particular, as depicted in Figure 7, the estimated coefficient of interest in Column 5 suggests that an increase in the predicted diversity from the 10th percentile (e.g., the Huave people of Mesoamerica) to the 90th percentile (e.g., the Nuer people of Central Africa) corresponds to an increase in the prevalence of spatio-temporal conflict by roughly 1.21 percentage points (compared to a sample mean of 0.15 and a standard deviation of 0.29). The larger coefficient on predicted diversity, relative to that on observed diversity —which also holds in the observed diversity sample— suggests that indeed the use of predicted diversity mitigates the endogenous reduction in the diversity of ethnic populations due to conflict. Furthermore, as established in Column 6, the estimate of the coefficient of interest remains stable when one accounts for the set of potentially endogenous confounders.

Finally, using the migratory distance from Africa as an instrumental variable for observed population diversity, the 2SLS regression analysis reported in Columns 7 and 8, suggests that there exists a highly significant reduced-form impact of population diversity on conflict, accounting for the potentially confounding effects of geographical and climatic characteristics, regional fixed-effects,

[48] See Table B.5 in the appendix.

[49] Climatic variables are not defined for 16 ethnic groups out of the complete set of 1,267 observations in the *Ethnographic Atlas*. The inclusion of these groups in regressions not accounting for climatic variables does not affect the qualitative conclusions.

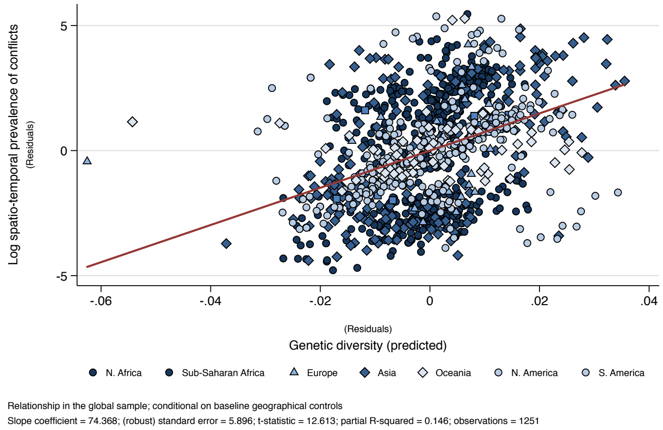


FIGURE 7: Predicted Population Diversity and the Spatiotemporal Prevalence of Conflict across Ethnic Homelands

Notes: This figure depicts the global cross-ethnic homeland relationship between contemporary predicted population diversity and the log spatio-temporal prevalence of UCDP/PRIO conflicts, 1989–2008, conditional on the baseline geographical correlates of conflict, as considered by the specification in Column 5 of Table 8.

as well as development outcomes.[50] In line with the results based on predicted diversity, once the potential reduction in the diversity of ethnic groups due to conflict is accounted for, the estimated coefficient of interest in Column 7 suggests that an increase in population diversity from the 10th percentile of the observed level of diversity (e.g., the Guaraní people of South America) to the 90th percentile (e.g., the Bulu people of South Africa) corresponds to an increase in the prevalence of spatio-temporal conflict by 1.40 percentage points (compared to a sample mean of 0.13 and a standard deviation of 0.25).

As established in the appendix, the results are robust towards a large number of alternative confounders and specifications. In particular, Table B.1 establishes the robustness of the findings to accounting for migratory distances from historical technological frontiers; Table B.2 demonstrates the robustness of the findings to controlling for ecological diversity and ecological polarization; Table B.3 establishes that the results are unaffected qualitatively by spatial dependence in the errors using SAR models; and Table B.4 shows that the results are robust to the use of Poisson regression. Furthermore, the results are robust to the bootstrapping of the standard errors, accounting for the use of generated regressors, and for selection on unobservables (see the notes for Table 8).

Table 9 explores the effect of predicted population diversity on the extensive and intensive margins of conflict. Column 1 establishes that observed diversity is positively correlated with the number of conflicts. Column 2 shows that this association is robust to controlling for the full set of exogenous control variables. Column 3, establishes the effect of predicted diversity on the extensive margin of conflicts, as depicted in Figure B.1. Columns 4–6 establish that predicted diversity has a highly significantly positive effect on the extensive margin of deaths, as depicted in Figure B.2, and Column 7 establishes that predicted diversity is a highly significant predictor for the number of deaths per conflict in the extended sample.

The findings suggest that population diversity is qualitatively as important a contributor to conflicts within ethnic homelands as it is for national populations. Furthermore, they also suggest

[50]The first-stage F-statistic indicates that the migratory distance is not a weak instrument.

Table 9: Population Diversity and Alternative Conflict Outcomes across Ethnic Homelands

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS
	Extensive margin of conflicts			Extensive margin of deaths			Intensive margin of deaths
Observed diversity	29.509***	19.293***		36.554***	20.955**		
	[6.889]	[6.360]		[9.162]	[9.267]		
Predicted diversity			34.883***			49.175***	14.292***
			[4.212]			[5.915]	[2.278]
Latitude		-0.059	0.027**		-0.109**	0.018	-0.010
		[0.036]	[0.011]		[0.053]	[0.016]	[0.007]
Ruggedness		-0.187	-0.105*		-0.151	-0.051	0.054
		[0.148]	[0.058]		[0.248]	[0.084]	[0.036]
Elevation		-0.093	0.076		-0.519	-0.119	-0.195**
		[0.414]	[0.126]		[0.568]	[0.181]	[0.077]
S.D. of elevation		0.132	-0.569*		-0.675	-1.237***	-0.667***
		[0.836]	[0.302]		[1.112]	[0.425]	[0.184]
Mean land suitability (native crops)		-0.059	-0.024		-0.083	-0.034	-0.010
		[0.069]	[0.016]		[0.105]	[0.025]	[0.013]
S.D. of land suitability (native crops)		0.491***	0.236***		0.637**	0.304***	0.068
		[0.182]	[0.075]		[0.266]	[0.109]	[0.044]
Change in land suitability (Columbian Exchange)		0.029	-0.107***		0.056	-0.179***	-0.073***
		[0.087]	[0.030]		[0.133]	[0.046]	[0.023]
Share desert		-1.093	-1.267***		-2.012	-1.592***	-0.325
		[1.178]	[0.349]		[1.720]	[0.525]	[0.275]
Nearest waterway		-0.074	0.109**		-0.039	0.136**	0.026
		[0.085]	[0.044]		[0.136]	[0.062]	[0.023]
Average temperature		0.146***	0.116***		0.193**	0.141***	0.025**
		[0.052]	[0.017]		[0.077]	[0.026]	[0.012]
Volatility of temperature		4.789***	0.761		7.674***	1.969***	1.209***
		[1.388]	[0.463]		[1.929]	[0.718]	[0.325]
Average precipitation		0.006	-0.001		0.009	0.000	0.001
		[0.004]	[0.001]		[0.005]	[0.002]	[0.001]
Volatility of precipitation		-0.043*	0.006		-0.061*	0.010	0.004
		[0.025]	[0.007]		[0.035]	[0.011]	[0.005]
Regional dummies	×	×	×	×	×	×	×
Additional climatic covariates		×	×		×	×	×
Sample	Observed	Observed	Extended	Observed	Observed	Observed	Extended
Observations	230	230	1251	230	230	1251	1251
Effect of increasing genetic diversity	403.685***	263.922***	262.542***		7406.506**	6410.692***	16.540***
from the 10th to the 90th percentile	[94.238]	[87.006]	[31.702]		[3275.258]	[771.047]	[2.636]
Adjusted R^2	0.272	0.506	0.394	0.261	0.532	0.423	0.380

Notes: This table exploits cross-ethnicity variations to establish a significant positive impact of contemporary population diversity, predicted by prehistoric migratory distance from East Africa on the log number of UCDP/GED conflicts, the log number of UCDP/GED deaths, and the log number of UCDP/GED deaths per conflict, during the 1989–2008 period, accounting for geographical and development-related correlates of conflict. The set of continent and regional dummies includes indicators for Europe, Asia, North America, South America, Oceania, North Africa, and Sub-Saharan Africa. Additional climatic covariates refer to the average diurnal temperature range, average cloud cover, and average temperature range in the homeland. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the prevalence of spatio-temporal of conflict. Robust standard errors are reported in square brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent.

that the estimated relationship of interest is conceptually robust to addressing the endogeneity of modern national borders to both national population diversity and civil conflicts.

5 Potential Mediating Channels

What are the proximate factors that can explain the adverse reduced-form influence of interpersonal population diversity on different forms and dimensions of social conflict? The following section, explores the potential mediating channels at the national as well as the subnational level.

5.1 Ethnic Diversity, Interpersonal Trust, and Dispersion in Political Preferences at the Country Level

This subsection examines some of our hypothesized proximate mechanisms that can potentially mediate the positive reduced-form cross-country relationship between population diversity and the risk of intrastate conflict, as reflected by the annual frequency of new PRIO25 civil conflict outbreaks during the 1960–2008 time period. Specifically, we provide evidence that our main cross-country empirical finding may partly be an expression of (i) the contribution of interpersonal population diversity to the degree of ethnolinguistic fragmentation at the country level, measured by the total number of ethnic groups in a national population Fearon (2003);[51] (ii) the adverse influence of population diversity on social capital, based on data from the World Values Survey (2006, 2009) (henceforth referred to as WVS) on the prevalence of generalized interpersonal trust in a country's population;[52] and (iii) the association between population diversity and heterogeneity in preferences for public goods and redistributive policies at the national level, as captured by the intracountry dispersion in self-reported individual political positions on a politically "left"–"right" categorical scale, based on data from the WVS.[53]

Table 10 reports the findings from our empirical examination of the aforementioned three potential mechanisms through which population diversity can partly contribute to the risk of intrastate conflict in society. For each posited channel, we present the results from estimating three different OLS regressions, exploiting worldwide variations in a common sample of countries, conditioned primarily by the availability of data on the mediating variable in question. In addition, throughout our analysis, we restrict our specifications to partialling out the influence of only our baseline set of geographical covariates (including continent or regional fixed effects). We do not include potentially endogenous control variables, many of which (like GDP per capita) may well be

[51] Unlike measures of ethnolinguistic fragmentation that are based on fractionalization or polarization indices, the number of ethnic groups in the national population is potentially less endogenous in an empirical model of the risk of civil conflict, in light of the fact that this measure is not additionally tainted by the incorporation of information on the endogenous shares of the different subnational groups.

[52] In particular, this well-known measure of social capital reflects the proportion in a given country of all respondents (from across five different waves of the WVS, conducted over the 1981–2009 time horizon) that opted for the answer "Most people can be trusted" (as opposed to "Can't be too careful") when responding to the survey question "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?"

[53] Specifically, this country-level measure of heterogeneity in political attitudes reflects the intracountry standard deviation across all respondents (sampled over five different waves of the WVS during the 1981–2009 time horizon) of their self-reported positions on a categorical scale from 1 (politically "left") to 10 (politically "right") when answering the survey question "In political matters, people talk of 'the left' and 'the right.' How would you place your views on this scale, generally speaking?" Given that the unit of measurement of this particular variable does not possess any natural interpretation, we standardize the cross-country distribution of this variable prior to conducting our regressions.

Table 10: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – Mediating Channels

Mediating channel:	Cultural fragmentation			Interpersonal trust			Preference heterogeneity		
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS
	Log number of ethnic groups	Frequency of new PRIO25 civil conflict onsets, 1960–2008		Prevalence of interpersonal trust	Frequency of new PRIO25 civil conflict onsets, 1960–2008		Variation in political attitudes	Frequency of new PRIO25 civil conflict onsets, 1960–2008	
Population diversity (ancestry adjusted)	5.431***	0.326**	0.259**	−2.151***	0.643**	0.583**	16.963***	0.548**	0.488*
	(1.884)	(0.141)	(0.124)	(0.807)	(0.258)	(0.277)	(6.371)	(0.246)	(0.290)
Log number of ethnic groups			0.012**						
			(0.006)						
Prevalence of interpersonal trust						−0.028			
						(0.033)			
Variation in political attitudes									0.004
									(0.007)
Continent/region dummies	×	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×	×
Observations	147	147	147	84	84	84	81	81	81
Partial R^2 of population diversity	0.054	0.041	0.025	0.105	0.084	0.063	0.111	0.059	0.042
Adjusted R^2	0.350	0.136	0.157	0.432	0.167	0.161	0.398	0.190	0.183
Effect of 10th–90th %ile move in diversity	2.236***	0.022**	0.018**	−0.124***	0.038**	0.034**	0.974***	0.032**	0.029*
	(0.776)	(0.010)	(0.008)	(0.046)	(0.015)	(0.016)	(0.366)	(0.014)	(0.017)

Notes: This table exploits cross-country variations to demonstrate that the significant positive reduced-form influence of contemporary population diversity on the annual frequency of new PRIO25 civil conflict onsets during the 1960–2008 time period, conditional on the baseline geographical correlates of conflict, is at least partly mediated by each of three potentially conflict-augmenting proximate channels that capture the contribution of population diversity to (i) the degree of cultural fragmentation, as reflected by the number of ethnic groups in the national population (Columns 1–3); (ii) the diminished prevalence of generalized interpersonal trust at the country level (Columns 4–6); and (iii) the extent of heterogeneity in preferences for redistribution and public-goods provision, as reflected by the intracountry dispersion in individual political attitudes on a politically "left"–"right" categorical scale (Columns 7–9). For each of the three mediating channels examined, the first regression documents the impact of population diversity on the proximate variable in the channel, the second presents the reduced-form influence of population diversity on conflict, and the third runs a "horse race" between population diversity and the proximate variable to establish reductions in the magnitude and explanatory power of the reduced-form influence of population diversity on conflict. All three regressions for each channel are conducted using a common cross-country sample, conditioned by the availability of data on the relevant variables employed by the analysis of the channel in question. The controls for geography include absolute latitude, land area, ruggedness, distance to the nearest waterway, the mean and range of agricultural suitability, and the mean and range of elevation. The regressions for the "cultural fragmentation" channel control for the full set of continent dummies (i.e., five indicators for Africa, Asia, North America, South America, and Oceania), whereas for the "trust" and "preference heterogeneity" channels, given the smaller degrees of freedom afforded by the more limited sample of countries, the regressions control for a more modest set of region dummies, including two indicators for Sub-Saharan Africa and Latin America and the Caribbean. Given that the unit of measurement for the variable reflecting the degree of intracountry dispersion in political attitudes has no natural interpretation, its cross-country distribution is standardized prior to conducting the relevant regressions. The estimated effect associated with increasing diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of (i) the actual number of ethnic groups in the national population in Column 1; (ii) the fraction of individuals in a country who "think that most people can be trusted" in Column 4; (iii) the number of standard deviations of the cross-country distribution of the national-level dispersion in political attitudes in Column 7; and (iv) the number of new conflict onsets per year in all the remaining columns. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

afflicted by reverse causality from the temporal frequency of civil conflict onsets and may also be determined in part by both population diversity and the mediating variable.

For our analysis of each mechanism, we proceed by first regressing the mediating variable on population diversity. These regressions are presented in Columns 1, 4, and 7. All coefficients on the mediating variables are statistically significant at the 1 percent level. They suggest that conditional on exogenous geographical factors, a move from the 10th to the 90th percentile of the cross-country diversity distribution in the relevant sample is associated with (i) an increase by 2.236 in the total number of ethnic groups in a national population; (ii) a decrease in the prevalence of generalized interpersonal trust at the country level by 12.4 percent; and (iii) an increase in the intracountry dispersion in individual political attitudes by 97.4 percent of a standard deviation from the cross-country distribution of this particular measure.[54]

The latter two regressions in our analysis of each hypothesized channel establish that the quantitative importance of population diversity as a predictor of the risk of civil conflict becomes diminished in both magnitude and explanatory power once the reduced-form influence of population diversity on the temporal frequency of civil conflict outbreaks is conditioned on the mediating variable of interest. Specifically, a comparison of the regressions in Columns 2 versus 3 indicates that when conditioned on the total number of ethnic groups in the national population, the influence of population diversity on conflict frequency, in terms of the response associated with a move from the 10th to the 90th percentile of the cross-country diversity distribution, is reduced in magnitude by 20.5 percent (from 0.022 to 0.018 new PRIO25 civil conflict onsets per year). The explanatory power of population diversity for conflict frequency, as reflected by the partial R^2 statistic, diminishes by 37.3 percent. The corresponding results obtained for each of the other two posited mechanisms are qualitatively similar but somewhat more muted, possibly due to greater measurement error in the relevant mediating variable. In particular, when conditioned on either the prevalence of generalized interpersonal trust in the national population or the intracountry dispersion in political attitudes, the magnitude of the response in conflict frequency that is associated with a move from the 10th to the 90th percentile of the cross-country diversity distribution decreases by either 9.25 percent (Columns 5 versus 6) or 11.1 percent (Columns 8 versus 9). Unlike the relation between the number of ethnic groups and frequency of civil conflicts, neither the prevalence of interpersonal trust nor the dispersion in political attitudes has a statistically significant relation with conflict frequency, conditional on population diversity and exogenous geographical factors. This finding, however, is consistent with attenuation bias afflicting the coefficients associated with the latter two mediating variables, in line with the aforementioned assertion regarding a potentially larger amount of white noise in the measurement of these variables.

One important caveat regarding the interpretation of our findings in Table 10 is that the mediating variables considered here may themselves be endogenous in a model of conflict risk. As corroborated by evidence from recent studies (e.g., Fletcher and Iyigun, 2010; Rohner, Thoenig and Zilibotti, 2013; Besley and Reynal-Querol, 2014), the unobserved historical cross-regional pattern of conflict risk may have partly contributed to the contemporary variations observed across countries in the degree of ethnolinguistic fragmentation, the prevalence of interpersonal trust, and the intracountry dispersion in revealed political preferences. Past conflicts possibly triggered movements of ethnic groups across space and reinforced extant inter-ethnic cleavages along with the social, political, and economic grievances associated with such divisions. Thus, we do not interpret our findings as being definitely reflective of the actual role of these factors as mediators. In order to assess our hypothesized mechanisms more conclusively, one would clearly need to exploit an independent exogenous source of variation for each of these proximate factors. This is a task that we leave open for future exploration.

5.2 Interpersonal Trust at the Individual Level

The proposed hypothesis suggests that interpersonal population diversity is conducive to conflict partly due to its adverse effect on trust and social cohesiveness. This section examines this channel,

[54] The three scatter plots presented in Figure A.3 in Appendix A.3 depict these statistically significant cross-country relationships, conditional on our baseline set of geographical covariates (including continent or regional fixed effects). The plots show the relationship between population diversity and (i) the total number of ethnic groups in a national population (Panel A); (ii) the prevalence of generalized interpersonal trust at the country level (Panel B); and (iii) the intracountry dispersion in political attitudes (Panel C).

using individual data to explore the relationship between interpersonal population diversity and interpersonal trust.

The analysis establishes that, in line with the proposed hypothesis, a higher degree of population diversity is indeed associated with a lower level of interpersonal trust, suggesting that the impact of diversity on the prevalence of conflict could plausibly be related to the adverse effect of diversity on trust.

5.2.1 Ethnic-Homeland Population Diversity and Individual Trust in Africa

The first layer of the analysis explores the relationship between observed population diversity in ethic homelands in Africa and the level of interpersonal trust of individuals (surveyed by the Afrobarometer) who are originated in these homelands and are either residing in their ethnic homelands or in other regions of Africa. The analysis accounts for time-invariant unobserved heterogeneity in the host country (e.g., geographical, cultural and institutional characteristics), mitigating possible concerns about the confounding effect of host country-specific characteristics.

Table 11 presents the regression analysis of various measures of trust across individuals in Africa on interpersonal population diversity in their ethnic homelands.[55] Panel A explores the association between observed population diversity in the individual's ethnic homeland and intra-group trust. Column 1 establishes that observed diversity in the individual ethnic homeland is significantly negatively correlated with the intra-group trust, accounting for the host country (i.e., country of residence) fixed effects. The estimated association remains negative and significant if one accounts for the potentially confounding effects of the age and the sex of the individual respondent (Column 2) and the historical levels of slave exports (Column 3), and it becomes highly significant if one accounts for the affluence of the respondent's town or village as captured by the presence of a school, electricity, piped water, sewage, and a health clinic (Column 4), and for whether the location of the respondent is urban (Column 5). The estimated association remains negative and highly significant if fixed effects for living conditions (Column 6), education (Column 7), and religion (Column 8) are accounted for. Finally, the coefficient estimate is smaller in absolute size, but remains negative and highly significant if one accounts for home-country fixed effects and thus identifying the effect of observed diversity on trust based on individual variation across ethnicities originating in the same home country who are residing in the same host country. The coefficient suggests that an increase in the observed population diversity from the 10th percentile of the observed level of diversity (e.g., individuals belonging to the Mandinka people) to the 90th percentile (e.g., individuals belonging to the Turu people) corresponds to a 0.14 points decrease in intra-group trust (compared to a sample mean of 1.47, a minimum of 0, a maximum of 3, and a standard deviation of 1.00).

Panel B explores the association between observed population diversity and trust in relatives, following the same specification in each column as reported in Panel A. The analysis suggests that observed population diversity is negatively associated with trust in relatives. In particular, as reported in Column 9, accounting for home-country fixed effects, this negative association is highly significant. Finally, Panel C explores the association between observed population diversity and trust in neighbors, following the same specification in each column as reported in Panel A. The analysis suggests that observed population diversity is negatively associated with trust in neighbors. In particular, as reported in Column 9, accounting for home-country fixed effects, this negative association is highly significant.

[55] The classification of individuals and their association with various ethnic homelands is based Nunn and Wantchekon (2011). Moreover, for the sake of comparability, the empirical specifications are parallel as well.

Table 11: Ethnic-Homeland Population Diversity and Individual-Level Trust in Africa

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS
Panel A					Intra-group trust				
Observed diversity	-23.010**	-21.851**	-28.775**	-26.399***	-26.380***	-26.105***	-25.436***	-18.258***	-13.843***
	(10.472)	(10.148)	(11.959)	(9.402)	(8.992)	(8.025)	(7.588)	(6.221)	(4.780)
Age		0.006***	0.007***	0.006***	0.006***	0.006***	0.004***	0.004***	0.004***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Male		-0.036*	-0.038**	-0.036*	-0.036*	-0.037*	-0.018	-0.029	-0.027
		(0.019)	(0.019)	(0.020)	(0.020)	(0.021)	(0.028)	(0.036)	(0.037)
Slave exports (Atlantic and Indian)			-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
School present				-0.136***	-0.136***	-0.141***	-0.125***	-0.124***	-0.121***
				(0.039)	(0.040)	(0.042)	(0.037)	(0.022)	(0.022)
Electricity present				-0.238***	-0.239***	-0.237***	-0.232***	-0.238***	-0.242***
				(0.063)	(0.048)	(0.046)	(0.040)	(0.041)	(0.042)
Piped water present				-0.040	-0.040	-0.038	-0.034	-0.028	-0.028
				(0.038)	(0.048)	(0.047)	(0.048)	(0.047)	(0.048)
Sewage present				0.028	0.028	0.028	0.034	0.059	0.059
				(0.045)	(0.044)	(0.049)	(0.053)	(0.071)	(0.071)
Health clinic present				0.022*	0.022*	0.023*	0.028***	0.039***	0.034***
				(0.012)	(0.013)	(0.012)	(0.010)	(0.013)	(0.008)
Living in an urban area					0.003	-0.000	0.012	-0.005	-0.007
					(0.067)	(0.069)	(0.060)	(0.045)	(0.044)
Host country FE	×	×	×	×	×	×	×	×	×
Living conditions FE						×	×	×	×
Education FE							×	×	×
Religion FE								×	×
Home country FE									×
Observations	3448	3448	3448	3448	3448	3448	3448	3448	3448
Adjusted R^2	0.220	0.227	0.236	0.248	0.248	0.249	0.253	0.261	0.264
Effect of increasing genetic diversity	-0.230**	-0.219**	-0.288**	-0.264***	-0.264***	-0.261***	-0.254***	-0.183***	-0.138***
from the 10th to the 90th percentile	[0.105]	[0.101]	[0.120]	[0.094]	[0.090]	[0.080]	[0.076]	[0.062]	[0.048]
Panel B					Trust in relatives				
Observed diversity	-35.552**	-34.369**	-33.129**	-31.452**	-32.089**	-32.690***	-31.891**	-22.995*	-11.804***
	(13.839)	(13.492)	(14.179)	(12.885)	(12.735)	(12.467)	(13.221)	(12.832)	(4.430)
Controls as in same column of Panel A	×	×	×	×	×	×	×	×	×
Observations	3460	3460	3460	3460	3460	3460	3460	3460	3460
Adjusted R^2	0.062	0.070	0.070	0.090	0.091	0.093	0.099	0.106	0.107
Effect of increasing genetic diversity	-0.356**	-0.344**	-0.331**	-0.315**	-0.321**	-0.327***	-0.319**	-0.230*	-0.118***
from the 10th to the 90th percentile	[0.138]	[0.135]	[0.142]	[0.129]	[0.127]	[0.125]	[0.132]	[0.128]	[0.044]
Panel C					Trust in neighbours				
Observed diversity	-27.106**	-25.428**	-28.361**	-26.249***	-26.537***	-27.161***	-27.117***	-20.518***	-17.375***
	(11.228)	(10.655)	(12.161)	(10.083)	(9.622)	(8.279)	(8.498)	(6.922)	(6.153)
Controls as in same column of Panel A	×	×	×	×	×	×	×	×	×
Observations	3452	3452	3452	3452	3452	3452	3452	3452	3452
Adjusted R^2	0.131	0.145	0.147	0.162	0.162	0.163	0.170	0.177	0.177
Effect of increasing genetic diversity	-0.271**	-0.254**	-0.284**	-0.262***	-0.265***	-0.272***	-0.271***	-0.205***	-0.174***
from the 10th to the 90th percentile	[0.112]	[0.107]	[0.122]	[0.101]	[0.096]	[0.083]	[0.085]	[0.069]	[0.062]

Notes: This table presents the results of an individual-level OLS regression analysis of interpersonal trust towards individuals of the same ethnicity (as reported recorded in Nunn and Wantchekon (2011)) on observed population diversity in the ancestral ethnicity of these individuals, controlling for a range of individual characteristics (i.e., age, gender, living conditions, education, religion), the presence of a school, electricity, piped water, sewage, a health clinic, in the local area, whether the local area is urban, and the intensity of Atlantic and Indian slave exports. In addition, the analysis accounts for host country fixed effects as well as fixed effects associated with the country in which the homeland of the individual's ethnicity is located.

Table 12: Country-of-Origin Population Diversity and Individual-Level Trust among Second-Generation U.S. Immigrants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Trust			
Ancestry-adjusted genetic diversity	-7.008***	-8.318***	-7.750***	-7.810***	-7.476***	-8.045***	-8.605***
	(0.011)	(0.712)	(0.577)	(1.007)	(2.126)	(2.325)	(2.599)
Parental continent FE	×	×	×	×	×	×	×
Year FE		×	×	×	×	×	×
Age FE			×	×	×	×	×
Sex FE			×	×	×	×	×
Income FE				×	×	×	×
Education FE						×	×
Religion FE					×	×	×
Region in the USA FE							×
Observations	1149	1149	1149	906	906	906	906
Adjusted R^2	0.062	0.090	0.168	0.186	0.194	0.216	0.231
Effect of increasing genetic diversity	-0.493***	-0.585***	-0.545***	-0.549***	-0.526***	-0.566***	-0.605***
from the 10th to the 90th percentile	[0.001]	[0.050]	[0.041]	[0.071]	[0.150]	[0.164]	[0.183]

Notes: This table presents the results of an individual-level OLS regression analysis of interpersonal trust among second-generation migrants in the US on population diversity in their parental country of origin (as captured by ancestry-adjusted predicted diversity Ashraf and Galor (2013a)), accounting for a range of individual-level socioeconomic characteristics (i.e., age, gender, income, religion, education), as well as time period fixed effects, parental region fixed effects, and the host region fixed-effect in the US. The trust variable takes the value 1 if the respondent replies that "people can people be trusted"; 2 if the respondent replies that "it depends", and 3 if the respondent replies that "people can be trusted". The measure of population diversity is the same as that used in the country-level analysis.

5.2.2 Country-of-Origin Population Diversity and Individual Trust amongst Second-Generation Immigrants to the U.S.

This subsection explores the effect of the level of population diversity in the parental country of origin of second-generation migrants in the United States on their level of trust (as reported in the General Social Survey, GSS). The analysis accounts for time-invariant unobserved heterogeneity in the host country (e.g., geographical, cultural, and institutional characteristics), mitigating possible concerns about the confounding effect of host country-specific characteristics. Moreover, since population diversity in the parental country of origin is distinct from population diversity in the country of residence, the estimated effect of population diversity in the country of origin captures the inter-generationally transmitted effect of population diversity on trust.

Table 12 present the association between the trust of second-generation migrants and the degree of population diversity in their parental country of origin.[56] Column 1 establishes a negative and highly significant association between population diversity in the parental country of origin and trust of second-generation migrants. This highly significant negative association remains largely stable if one accounts for interview-year fixed effects (Column 2), the respondent's age-specific and sex-specific fixed effects (Column 3), relative income of the family at the age of 16 years (Column 4), religion (Column 5), and educational attainment (Column 6). Moreover, the results are robust to controlling for region-specific fixed effects within the United States. The estimate of the coefficient of interest in Column 7 suggests that an increase in population diversity in the parental country of origin from the 10th percentile of the predicted contemporary level of diversity (e.g., individuals of Mexican descent) to the 90th percentile (e.g., individuals of Hungarian descent) corresponds to a

[56] The specifications follow the standards in the "epidemiological approach" to examining intergenerational persistence in individual cultural attitudes and values (Fernández, 2011; Algan and Cahuc, 2013; Galor and Özak, 2016).

decrease in trust by 0.61 points (compared to a sample mean of 1.87, a minimum of 0, a maximum of 3, and a standard deviation of 0.97).

6 Concluding Remarks

This research advances the hypothesis and establishes empirically that interpersonal population diversity, as determined predominantly during the exodus of humans from Africa tens of thousands of years ago, has contributed significantly to the emergence, prevalence, recurrence, and severity of intrasocietal conflicts. Exploiting an exogenous source of variations in population diversity across nations and ethnic groups, it demonstrates that interpersonal population diversity has contributed significantly to the risk and intensity of historical and contemporary internal conflicts, accounting for the confounding effects of geographical, institutional, and cultural characteristics, as well as for the level of economic development. These findings arguably reflect the adverse effect of population diversity on interpersonal trust, its contribution to divergence in preferences for public goods and redistributive policies, and its impact on the degree of fractionalization and polarization across ethnic, linguistic, and religious groups.

References

Ahlerup, Pelle, and Ola Olsson. 2012. "The Roots of Ethnic Diversity." *Journal of Economic Growth*, 17(2): 71–102.

Alesina, Alberto, and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature*, 43(3): 762–800.

Alesina, Alberto, and Enrico Spolaore. 2003. *The Size of Nations.* Cambridge, MA:MIT Press.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth*, 8(2): 155–194.

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics*, 114(4): 1243–1284.

Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou. 2016. "Ethnic Inequality." *Journal of Political Economy*, 124(2): 428–488.

Algan, Yann, and Pierre Cahuc. 2013. "Trust and Growth." *Annual Review of Economics*, 5: 521–549.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113(1): 151–184.

Angrist, Joshua D., and Adriana D. Kugler. 2008. "Rural Windfall or a New Resource Curse? Coca, Income, and Civil Conflict in Colombia." *Review of Economics and Statistics*, 90(2): 191–215.

Ashraf, Quamrul, and Oded Galor. 2013a. "The "Out of Africa" Hypothesis, Human Genetic Diversity, and Comparative Economic Development." *American Economic Review*, 103(1): 1–48.

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

Ashraf, Quamrul H., and Oded Galor. forthcoming. "The Macrogenoeconomics of Comparative Development." *Journal of Economic Literature.*

Atkinson, Quentin D. 2011. "Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa." *Science*, 332(6027): 346–349.

Banks, Arthur S. 2010. "Cross-National Time-Series Data Archive." Databanks International, Jerusalem, Israel. http://www.databanksinternational.com/ (accessed March 23, 2012).

Barth, Fredrik, ed. 1969. *Ethnic Groups and Boundaries: The Social Organization of Culture Difference.* Olso, Norway:Universitetsforlaget.

Bates, Robert H. 1983. "Modernization, Ethnic Competition, and the Rationality of Politics in Contemporary Africa." In *State versus Ethnic Claims: African Policy Dilemmas.* , ed. Donald Rothchild and Victor A. Olorunsola, 152–171. Boulder, CO:Westview Press.

Bazzi, Samuel, and Christopher Blattman. 2014. "Economic Shocks and Conflict: Evidence from Commodity Prices." *American Economic Journal: Macroeconomics*, 6(4): 1–38.

Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-Series–Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science*, 42(4): 1260–1288.

Bellows, John, and Edward Miguel. 2009. "War and Local Collective Action in Sierra Leone." *Journal of Public Economics*, 93(11-12): 1144–1157.

Besley, Timothy, and Marta Reynal-Querol. 2014. "The Legacy of Historical Conflict: Evidence from Africa." *American Political Science Review*, 108(2): 319–336.

Betti, Lia, François Balloux, William Amos, Tsunehiko Hanihara, and Andrea Manica. 2009. "Distance from Africa, Not Climate, Explains Within-Population Phenotypic Diversity in Humans." *Proceedings of the Royal Society B: Biological Sciences*, 276(1658): 809–814.

Betti, Lia, Noreen von Cramon-Taubadel, Andrea Manica, and Stephen J. Lycett. 2013. "Global Geometric Morphometric Analyses of the Human Pelvis Reveal Substantial Neutral Population History Effects, Even across Sexes." *PLoS ONE*, 8(2): e55909.

Blattman, Christopher, and Edward Miguel. 2010. "Civil War." *Journal of Economic Literature*, 48(1): 3–57.

Bockstette, Valerie, Areendam Chanda, and Louis Putterman. 2002. "States and Markets: The Advantage of an Early Start." *Journal of Economic Growth*, 7(4): 347–369.

Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process.* Chicago, IL:University of Chicago Press.

Brass, Paul R. 1997. *Theft of an Idol: Text and Context in the Representation of Collective Violence.* Princeton, NJ:Princeton University Press.

Brecke, Peter. 1999. "Violent Conflicts 1400 A.D. to the Present in Different Regions of the World." Paper presented at the 1999 Annual Meeting of the Peace Science Society, October 8–10.

Brewer, Marilynn B. 1979. "In-Group Bias in the Minimal Intergroup Situation: A Cognitive-Motivational Analysis." *Psychological Bulletin*, 86(2): 307–324.

Brewer, Marilynn B. 1991. "The Social Self: On Being the Same and Different at the Same Time." *Personality and Social Psychology Bulletin*, 17(5): 475–482.

Brewer, Marilynn B. 1997. "The Social Psychology of Intergroup Relations: Can Research Inform Practice?" *Journal of Social Issues*, 53(1): 197–211.

Brubaker, Rogers. 2004. *Ethnicity without Groups.* Cambridge, MA:Harvard University Press.

Burke, Marshall B., Edward Miguel, Shanker Satyanath, John A. Dykema, and David B. Lobell. 2009. "Warming Increases the Risk of Civil War in Africa." *Proceedings of the National Academy of Sciences*, 106(49): 20670–20674.

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel. 2015. "Climate and Conflict." *Annual Review of Economics*, 7: 577–617.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business & Economic Statistics*, 29(2): 238–249.

Cann, Howard M., Claudia de Toma, Lucien Cazes, Marie-Fernande Legrand, Valerie Morel, Laurence Piouffre, Julia Bodmer, et al. 2002. "A Human Genome Diversity Cell Line Panel." *Science*, 296(5566): 261–262.

Caselli, Francesco, and Wilbur John Coleman, II. 2013. "On the Theory of Ethnic Conflict." *Journal of the European Economic Association*, 11(S1): 161–192.

Cavalli-Sforza, L. Luca. 2005. "The Human Genome Diversity Project: Past, Present and Future." *Nature Reviews Genetics*, 6(4): 333–340.

Cavalli-Sforza, L. Luca, and Marcus W. Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach.* Princeton, NJ:Princeton University Press.

Cavalli-Sforza, L. Luca, Paolo Menozzi, and Alberto Piazza. 1994. *The History and Geography of Human Genes.* Princeton, NJ:Princeton University Press.

Central Intelligence Agency, The. 2006. "The World Factbook 2006." The Central Intelligence Agency, Washington, DC. https://www.cia.gov/library/publications/the-world-factbook/ (accessed March 18, 2007).

Cervellati, Matteo, Uwe Sunde, and Simona Valmori. 2017. "Pathogens, Weather Shocks and Civil Conflicts." *Economic Journal*, 127(607): 2581–2616.

Cioffi-Revilla, Claudio. 1996. "Origins and Evolution of War and Politics." *International Studies Quarterly*, 40(1): 1–22.

Collier, Paul, and Anke Hoeffler. 1998. "On Economic Causes of Civil War." *Oxford Economic Papers*, 50(4): 563–573.

Collier, Paul, and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers*, 56(4): 563–595.

Collier, Paul, and Anke Hoeffler. 2007. "Civil War." In *Handbook of Defense Economics, Vol. 2: Defense in a Globalized World.* , ed. Todd Sandler and Keith Hartley, 711–740. Amsterdam, The Netherlands:Elsevier, North-Holland.

Conley, T. G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1–45.

Connor, Walker. 1994. *Ethnonationalism: The Quest for Understanding.* Princeton, NJ:Princeton University Press.

Croicu, Mihai, and Ralph Sundberg. 2015. "UCDP Georeferenced Event Dataset Codebook version 4.0." Department of Peace and Conflict Research, Uppsala University. http://ucdp.uu.se/downloads/ged/ucdp-ged-40-codebook.pdf.

Desmet, Klaus, Ignacio Ortuño-Ortín, and Romain Wacziarg. 2012. "The Political Economy of Linguistic Cleavages." *Journal of Development Economics*, 97(2): 322–338.

Desmet, Klaus, Ignacio Ortuño-Ortín, and Shlomo Weber. 2009. "Linguistic Diversity and Redistribution." *Journal of the European Economic Association*, 7(6): 1291–1318.

Desmet, Klaus, Michel Le Breton, Ignacio Ortuño-Ortín, and Shlomo Weber. 2011. "The Stability and Breakup of Nations: A Quantitative Analysis." *Journal of Economic Growth*, 16(3): 183–213.

Dincecco, Mark, James Fenske, and Massimiliano Gaetano Onorato. 2015. "Is Africa Different? Historical Conflict and State Development." IMT Lucca EIC Working Paper No. 08/2015, IMT Institute for Advance Studies Lucca.

Drukker, David M., Ingmar R. Prucha, and Rafal Raciborski. 2013. "Maximum Likelihood and Generalized Spatial Two-Stage Least-Squares Estimators for a Spatial-Autoregressive Model with Spatial-Autoregressive Disturbances." *Stata Journal*, 13(2): 221–241.

Dube, Oeindrila, and Juan F. Vargas. 2013. "Commodity Price Shocks and Civil Conflict: Evidence from Colombia." *Review of Economic Studies*, 80(4): 1384–1421.

Durham, William H. 1991. *Coevolution: Genes, Culture, and Human Diversity.* Stanford, CA:Stanford University Press.

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

Eifert, Benn, Edward Miguel, and Daniel N. Posner. 2010. "Political Competition and Ethnic Identification in Africa." *American Journal of Political Science*, 54(2): 494–510.

Esteban, Joan, and Debraj Ray. 2011a. "A Model of Ethnic Conflict." *Journal of the European Economic Association*, 9(3): 496–521.

Esteban, Joan, and Debraj Ray. 2011b. "Linking Conflict to Inequality and Polarization." *American Economic Review*, 101(4): 1345–1374.

Esteban, Joan, Laura Mayoral, and Debraj Ray. 2012. "Ethnicity and Conflict: An Empirical Study." *American Economic Review*, 102(4): 1310–1342.

Esteban, Joan-María, and Debraj Ray. 1994. "On the Measurement of Polarization." *Econometrica*, 62(4): 819–851.

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

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

Fenske, James. 2013. "Does Land Abundance Explain African Institutions?" *Economic Journal*, 123(573): 1363–1390.

Fenske, James. 2014. "Ecology, Trade and States in Pre-Colonial Africa." *Journal of the European Economic Association*, 12(3): 612–640.

Fernández, Raquel. 2011. "Does Culture Matter?" In *Handbook of Social Economics, Vol. 1.* , ed. Jess Benhabib, Alberto Bisin and Matthew O. Jackson, 481–510. Amsterdam, The Netherlands:Elsevier, North-Holland.

Fletcher, Erin, and Murat Iyigun. 2010. "The Clash of Civilizations: A Cliometric Investigation." http://www.colorado.edu/economics/courses/iyigun/fractionalization013109.pdf.

Gallup, John Luke, Jeffrey D. Sachs, and Andrew D. Mellinger. 1999. "Geography and Economic Development." *International Regional Science Review*, 22(2): 179–232.

Galor, Oded. 2011. *Unified Growth Theory.* Princeton, NJ:Princeton University Press.

Galor, Oded, and Ömer Özak. 2016. "The Agricultural Origins of Time Preference." *American Economic Review*, 106(10): 3064–3103.

Geertz, Clifford. 1973. *The Interpretation of Cultures.* New York, NY:Basic Books.

Gellner, Ernest. 1983. *Nations and Nationalism.* Ithaca, NY:Cornell University Press.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. 2002. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research*, 39(5): 615–637.

Griffiths, Anthony J. F., Jeffrey H. Miller, David T. Suzuki, Richard C. Lewontin, and William M. Gelbart. 2000. *An Introduction to Genetic Analysis.* . 7th ed., New York, NY:W. H. Freeman & Co.

Grossman, Herschel I. 1991. "A General Equilibrium Model of Insurrections." *American Economic Review*, 81(4): 912–921.

Grossman, Herschel I. 1999. "Kleptocracy and Revolutions." *Oxford Economic Papers*, 51(2): 267–283.

Hanihara, Tsunehiko. 2008. "Morphological Variation of Major Human Populations Based on Nonmetric Dental Traits." *American Journal of Physical Anthropology*, 136(2): 169–182.

Hardin, Russell. 1995. *One for All: The Logic of Group Conflict.* Princeton, NJ:Princeton University Press.

Harpending, Henry, and Alan Rogers. 2000. "Genetic Perspectives on Human Origins and Differentiation." *Annual Review of Genomics and Human Genetics*, 1: 361–385.

Hartl, Daniel L., and Andrew G. Clark. 2007. *Principles of Population Genetics.* . 4th ed., Sunderland, MA:Sinauer Associates, Inc.

Hegre, Håvard, and Nicholas Sambanis. 2006. "Sensitivity Analysis of Empirical Results on Civil War Onset." *Journal of Conflict Resolution*, 50(4): 508–535.

Heizer, Robert F. 1978. *Handbook of North American Indians, Vol. 8: California.* Washington, DC:Smithsonian Institution.

Henrich, Joseph, and Richard McElreath. 2003. "The Evolution of Cultural Evolution." *Evolutionary Anthropology*, 12(3): 123–135.

Hirshleifer, Jack. 1991. "The Technology of Conflict as an Economic Activity." *American Economic Review: Papers & Proceedings*, 81(2): 130–134.

Hirshleifer, Jack. 1995. "Anarchy and its Breakdown." *Journal of Political Economy*, 103(1): 26–52.

Horowitz, Donald L. 1985. *Ethnic Groups in Conflict.* Berkeley, CA:University of California Press.

Horowitz, Donald L. 1999. "Structure and Strategy in Ethnic Conflict: A Few Steps Toward Synthesis." In *Annual World Bank Conference on Development Economics 1998.* , ed. Boris Pleskovic and Joseph E. Stiglitz, 345–370. Washington, DC:The World Bank.

Hsiang, Solomon M., Marshall Burke, and Edward Miguel. 2013. "Quantifying the Influence of Climate on Human Conflict." *Science*, 341(6151): 1235367/1–14.

Humphreys, Macartan. 2005. "Natural Resources, Conflict, and Conflict Resolution: Uncovering the Mechanisms." *Journal of Conflict Resolution*, 49(4): 508–537.

Iyigun, Murat. 2008. "Luther and Suleyman." *Quarterly Journal of Economics*, 123(4): 1465–1494.

King, Gary, and Langche Zeng. 2001. "Logistic Regression in Rare Events Data." *Political Analysis*, 9(2): 137–163.

König, Michael D., Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti. 2017. "Networks in Conflict: Theory and Evidence From the Great War of Africa." *Econometrica*, 85(4): 1093–1132.

Lagerlöf, Nils-Petter. 2014. "Population, Technology and Fragmentation: The European Miracle Revisited." *Journal of Development Economics*, 108: 87–105.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny. 1999. "The Quality of Government." *Journal of Law, Economics, and Organization*, 15(1): 222–279.

Maddison, Angus. 2010. "Statistics on World Population, GDP and Per Capita GDP, 1–2008 AD." Groningen Growth and Development Centre, Groningen, The Netherlands. http://www.ggdc.net/maddison/oriindex.htm (accessed January 2, 2011).

Manica, Andrea, William Amos, François Balloux, and Tsunehiko Hanihara. 2007. "The Effect of Ancient Population Bottlenecks on Human Phenotypic Variation." *Nature*, 448(7151): 346–348.

Marshall, Monty G. 2010. "Major Episodes of Political Violence (MEPV) and Conflict Regions, 1946–2008." Center for Systemic Peace, Vienna, VA. http://www.systemicpeace.org/inscrdata.html (accessed February 10, 2011).

Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2009. "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2009." Center for Systemic Peace, Vienna, VA. http://www.systemicpeace.org/inscrdata.html (accessed January 2, 2011).

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

Michalopoulos, Stelios, and Elias Papaioannou. 2016. "The Long-Run Effects of the Scramble for Africa." *American Economic Review*, 106(7): 1802–1848.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy*, 112(4): 725–753.

Minorities at Risk Project, The. 2009. "Minorities at Risk Dataset." Center for International Development and Conflict Management, College Park, MD. http://www.cidcm.umd.edu/mar/mar_data.asp (accessed August 24, 2012).

Montalvo, José G., and Marta Reynal-Querol. 2005. "Ethnic Polarization, Potential Conflict, and Civil Wars." *American Economic Review*, 95(3): 796–816.

Murdock, George Peter. 1959. *Africa: Its Peoples and Their Culture History.* New York, NY:McGraw-Hill Book Co., Inc.

Nordhaus, William D. 2006. "Geography and Macroeconomics: New Data and New Findings." *Proceedings of the National Academy of Sciences*, 103(10): 3510–3517.

Nunn, Nathan. 2014. "Historical Development." In *Handbook of Economic Growth, Vol. 2A.* , ed. Philippe Aghion and Steven N. Durlauf, 347–402. Amsterdam, The Netherlands:Elsevier, North-Holland.

Nunn, Nathan, and Leonard Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review*, 101(7): 3221–3252.

Oster, Emily. forthcoming. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics*. https://doi.org/10.1080/07350015.2016.1227711.

Pemberton, Trevor J., Michael DeGiorgio, and Noah A. Rosenberg. 2013. "Population Structure in a Comprehensive Genomic Data Set on Human Microsatellite Variation." *G3: Genes, Genomes, and Genetics*, 3(5): 891–907.

Posner, Daniel N. 2003. "The Colonial Origins of Ethnic Cleavages: The Case of Linguistic Divisions in Zambia." *Comparative Politics*, 35(2): 127–146.

Prugnolle, Franck, Andrea Manica, and François Balloux. 2005. "Geography Predicts Neutral Genetic Diversity of Human Populations." *Current Biology*, 15(5): R159–R160.

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

Ramachandran, Sohini, Omkar Deshpande, Charles C. Roseman, Noah A. Rosenberg, Marcus W. Feldman, and L. Luca Cavalli-Sforza. 2005. "Support from the Relationship of Genetic and Geographic Distance in Human Populations for a Serial Founder Effect Originating in Africa." *Proceedings of the National Academy of Sciences*, 102(44): 15942–15947.

Reid, Richard. 2014. "The Fragile Revolution: Rethinking War and Development in Africa's Violent Nineteenth Century." In *Africa's Development in Historical Perspective.* , ed. Emmanuel Akyeampong, Robert H. Bates, Nathan Nunn and James A. Robinson, 393–423. Cambridge, UK:Cambridge University Press.

Rohner, Dominic, Mathias Thoenig, and Fabrizio Zilibotti. 2013. "Seeds of Distrust: Conflict in Uganda." *Journal of Economic Growth*, 18(3): 217–252.

Ross, Michael. 2006. "A Closer Look at Oil, Diamonds, and Civil War." *Annual Review of Political Science*, 9: 265–300.

Ross, Michael L. 2013. "Oil and Gas Data, 1932–2011." http://dvn.iq.harvard.edu/dvn/dv/mlross/faces/study/StudyPage.xhtml?studyId=87422&versionNumber=2 (accessed December 26, 2013).

Rummel, Rudolph J. 1963. "Dimensions of Conflict Behavior Within and Between Nations." *General Systems Yearbook*, 8: 1–50.

Sambanis, Nicholas. 2002. "A Review of Recent Advances and Future Directions in the Quantitative Literature on Civil War." *Defence and Peace Economics*, 13(3): 215–243.

Shils, Edward. 1957. "Primordial, Personal, Sacred and Civil Ties: Some Particular Observations on the Relationships of Sociological Research and Theory." *British Journal of Sociology*, 8(2): 130–145.

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

Spolaore, Enrico, and Romain Wacziarg. 2013. "How Deep Are the Roots of Economic Development?" *Journal of Economic Literature*, 51(2): 325–369.

Spolaore, Enrico, and Romain Wacziarg. 2014. "Long-Term Barriers to Economic Development." In *Handbook of Economic Growth, Vol. 2A.* , ed. Philippe Aghion and Steven N. Durlauf, 121–176. Amsterdam, The Netherlands:Elsevier, North-Holland.

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

Sundberg, Ralph, Kristine Eck, and Joakim Kreutz. 2012. "Introducing the UCDP Non-State Conflict Dataset." *Journal of Peace Research*, 49(2): 351–362.

Themnér, Lotta, and Peter Wallensteen. 2012. "Armed Conflicts, 1946–2011." *Journal of Peace Research*, 49(4): 565–575.

Themnér, Lotta, and Peter Wallensteen. 2014. "Armed Conflicts, 1946–2013." *Journal of Peace Research*, 51(4): 541–554.

Tollefsen, Andreas Forø, Håvard Strand, and Halvard Buhaug. 2012. "PRIO-GRID: A Unified Spatial Data Structure." *Journal of Peace Research*, 49(2): 363–374.

Van den Berghe, Pierre L. 1981. *The Ethnic Phenomenon.* New York, NY:Elsevier, North-Holland.

Van den Berghe, Pierre L. 1995. "Does Race Matter?" *Nations and Nationalism*, 1(3): 357–368.

von Cramon-Taubadel, Noreen, and Stephen J. Lycett. 2008. "Brief Communication: Human Cranial Variation Fits Iterative Founder Effect Model with African Origin." *American Journal of Physical Anthropology*, 136(1): 108–113.

Wang, Sijia, Cecil M. Lewis Jr., Mattias Jakobsson, Sohini Ramachandran, Nicolas Ray, Gabriel Bedoya, Winston Rojas, Maria V. Parra, Julio A Molina, Carla Gallo, et al. 2007. "Genetic Variation and Population Structure in Native Americans." *PLoS Genetics*, 3(11): 2049–2067.

Weidmann, Nils B., Jan Ketil Rød, and Lars-Erik Cederman. 2010. "Representing Ethnic Groups in Space: A New Dataset." *Journal of Peace Research*, 47(4): 491–499.

Wimmer, Andreas. 2002. *Nationalist Exclusion and Ethnic Conflict: Shadows of Modernity.* Cambridge, UK:Cambridge University Press.

Wimmer, Andreas, Lars-Erik Cederman, and Brian Min. 2009. "Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Data Set." *American Sociological Review*, 74(2): 316–337.

World Bank, The. 2006. "The World Development Indicators 2006." The World Bank, Washington, DC. http://data.worldbank.org/data-catalog/world-development-indicators (accessed February 15, 2007).

World Values Survey, The. 2006. "European and World Values Surveys, Four-Wave Integrated Data File, 1981–2004, version 20060423." The World Values Survey Association, Stockholm, Sweden. http://www.worldvaluessurvey.org/ (accessed August 24, 2010).

World Values Survey, The. 2009. "World Values Survey, 1981–2008 Official Aggregate, version 20090914." The World Values Survey Association, Stockholm, Sweden. http://www.worldvaluessurvey.org/ (accessed August 24, 2010).

Zhang, David D., Harry F. Lee, Cong Wang, Baosheng Li, Qing Pei, Jane Zhang, and Yulun An. 2011. "The Causality Analysis of Climate Change and Large-Scale Human Crisis." *Proceedings of the National Academy of Sciences*, 108(42): 17296–17301.

Appendix A Appendix for the Country-Level Analyses

A.1 Robustness Checks for the Analysis of Civil Conflict Frequency in Cross-Country Data

In this appendix section, we present several robustness checks for our cross-country analysis of the influence of the contemporary population diversity —as proxied by the measure of ancestry-adjusted genetic diversity— on the temporal frequency of either overall or ethnic civil conflict outbreaks in the post-1960 time horizon.

Intergroup Genetic Distance and Civil Conflict In Table A.1 we establish that even after controlling for the influence of average intergroup genetic distance in the national population —as well as the interaction between population diversity and intergroup genetic distance.—, our variable of interest (ancestry-adjusted population diversity) remains a significant predictor of overall and ethnic civil conflict propensity regardless of the regression sample and specification. The additional controls for distance and the interaction term enter the regression significantly when considering the global sample of countries. In addition, the reduced-form coefficients associated with these terms in these regressions are interesting in that they are consistent with a convex combination of previous models of conflict. For further details, see the corresponding table notes.

Robustness to Ecological/Climatic Covariates A nascent interdisciplinary literature (e.g., Burke et al., 2009; Hsiang, Burke and Miguel, 2013; Burke, Hsiang and Miguel, 2015) has emphasized the role of climatic factors, like temperature and precipitation, as important correlates of the risk of civil conflict. Fenske (2014) shows that ecological diversity facilitated state centralization in pre-colonial African states. To keep our main specifications from becoming too unwieldy, we chose to exclude the aforementioned climatic and ecological variables from our baseline set of covariates, especially because this set already included a sizable vector of geographical factors that are known to be correlated with the former. In Table A.2, however, we establish that population diversity remains a significant predictor of overall and ethnic civil conflicts when we augment our baseline set of covariates in Table 3 with controls for (i) time-invariant fractionalization and polarization measures of the ecological diversity of land (e.g., Fenske, 2014); and (ii) the temporal mean and volatility of climatic experience (e.g., Burke, Hsiang and Miguel, 2015) with respect to annual temperature and annual precipitation over the relevant post-1960 time period. For further details, see the corresponding table notes.

Robustness to Accounting for Deep-Rooted Determinants of Economic Development In Table A.3, we establish robustness of our baseline cross-country analyses of civil conflict to *additionally* accounting for the potentially confounding influence of other deep-rooted determinants of comparative economic development. We augment the analyses in Table 3 with controls for (i) the time elapsed since the onset of the Neolithic Revolution (e.g., Ashraf and Galor, 2013a); (ii) an index of experience with institutionalized statehood since antiquity (e.g., Bockstette, Chanda and Putterman, 2002); (iii) the time elapsed since initial human settlement in prehistory (e.g., Ahlerup and Olsson, 2012); and (iv) the great-circle distance to the closest regional technological frontier in the year 1500 (e.g., Ashraf and Galor, 2013a). Regardless of the sample and specification, contemporary population diversity remains a significant predictor of annual frequency of overall and ethnic civil conflict onsets.

Robustness to Accounting for Ethnic and Spatial Inequality In Table A.4 we conduct robustness checks to *additionally* account for intrastate economic inequality (e.g., Alesina, Michalopoulos and Papaioannou, 2016), as captured by the subnational spatial distribution of per-capita adjusted nighttime luminosity in the year 2000 across either (i) the georeferenced homelands of ethnic groups (ethnic inequality); or (ii) 2.5×2.5-degree geospatial grid cells (spatial inequality). The specifications examined in this table are otherwise identical to those reported in Table 3. Ethnic —but not spatial— inequality enters some of the specifications with a positive coefficient that is statistically significant at the 10 percent level. However, our results establish that the positive and significant influence of ancestry-adjusted population diversity on the annual frequeny of overall and ethnic civil conflicts cannot be attributed to the confounding influence of these factors.

Robustness to Accounting for Linguistic Fractionalization As is evident from the results of our bivariate and "horse race" regressions that examined the influence of various diversity measures on civil conflict frequency, the linguistic fractionalization index of Alesina et al. (2003) entered some of these regressions (i.e., in Tables 1 and 2) with a statistically significant coefficient, much like the closely related ethnic fractionalization index from the same study. Due to the sizable cross-country correlation between these two fractionalization measures, however, rather than exploiting both variables simultaneously, we chose to employ the more widely used of the two indices – namely, ethnic fractionalization – as one of the many covariates in our baseline analysis of the influence of population diversity on either overall or ethnic civil conflict frequency. In Table A.5, we therefore examine the sensitivity of our baseline findings from Table 3 to employing the *linguistic* fractionalization index of Alesina et al. (2003) in lieu of our baseline control for the *ethnic* fractionalization index from the same source. Reassuringly, the results verify that all our baseline findings regarding the significant influence of population diversity on the temporal frequency of either overall of ethnic civil conflict onsets remain qualitatively intact under these alternative specifications.

Robustness to Using Initial Values of Time-Varying Covariates In Table A.6 we use the initial or year-1960 values of the time-dependent baseline controls employed in Table 3 (i.e., the degree of executive constraints and indicators for democracy and autocracy, oil production per capita, total population, and GDP per capita), rather than their respective temporal averages over the relevant post-1960 time period. To motivation behind this robustness check is to show that our baseline estimates of the influence of population diversity in Table 3 are not driven by any ex ante bias that may arise if temporal averages of the time-varying controls computed over the entire sample period are endogenous to the frequency of civil conflict onsets over the same period. Reassuringly, population diversity continues to be a significant predictor of conflict frequency, regardless of the estimation strategy (OLS or 2SLS) or the sample (global or the Old World countries).

Interpersonal Population Diversity and the Count of Civil Conflict Onset across Countries Given that our baseline cross-country regressions employ least-squares estimation, we apply a log transformation to each of our outcome variables in order to partly address the issue that their cross-country distributions are positively skewed with excess zeros, arising from the fact that new civil conflict onsets are generally rare events in cross-sectional data. An alternative approach to this issue, however, is to employ an estimation method that is tailored to the analysis of over-dispersed count data. In Table A.7, we consider the *total count* rather than the annual frequency of civil conflict onsets over the relevant post-1960 time period as the outcome variable. The regressions in Columns 1-10 are estimated using the negative binomial as opposed to a least squares estimator to account for over-dispersion.[1] Given the absence of a negative binomial estimator that permits instrumentation, in lieu of implementing the instrument-based identification strategy in the global

[1] In light of the over-dispersed nature of the number of overall and ethnic civil conflicts, both of which possess cross-country distributions with coefficients of variation larger than unity, an analysis that is based on the negative-binomial model is indeed most appropriate.

sample of countries, Columns 11 and 12 examine robustness to employing the Poisson rather than the negative binomial estimator. Our estimates in Columns 9-12 suggest that an increase in population diversity from the tenth to the ninetieth percentile of its cross-country distribution translates into 1.17 to 1.30 additional civil conflicts (0.80 to 0.89 standard deviation) in the Old World sample and 0.88 to 0.93 additional civil conflicts (0.65 to 0.68 standard deviation) in the Global sample over the 1960-2008 time period.

Robustness to Accounting for Spatial Dependence As with any empirical analysis that exploits spatial variations in cross-sectional data, autocorrelation in disturbance terms across observations could be biasing our estimates of the standard errors in our baseline cross-country analyses of conflict frequency. Table A.8 reports, for all regressions from Table 3, standard errors that are corrected for cross-sectional spatial dependence, using the methodology proposed by Conley (1999). To perform this robustness check, the spatial distribution of observations is specified on the Euclidean plane using the full set of pairwise geodesic distances between country centroids, and the spatial autoregressive process across residuals is modeled as varying inversely with distance from each observation up to a maximum threshold of 25,000 kilometers, thus admitting the possibility of spatial dependence at a global scale. The GMM specifications in this table correspond to the 2SLS specifications from Table 3. Reassuringly, depending on the specification examined, the corrected standard errors of the estimated coefficient on population diversity are either similar in magnitude or noticeably smaller when compared to their heteroskedasticity robust counterparts in our baseline analyses.

As an alternative way of accounting for spatial dependence across observations, in Table A.9 we repeat the baseline analysis in Table 3 by employing spatial-autoregressive models —with spatial-autoregressive disturbances (SARAR(1,1))— which are estimated via the generalized spatial two-stage least squares (GS2SLS) estimator (e.g., Drukker, Prucha and Raciborski, 2013). To perform this robustness check, involving the estimation of the AR(1) coefficients, λ and ρ, respectively associated with the spatial lags in the outcome variable and the error term, the estimator exploits an inverse-distance spatial weighting matrix for the regression sample, based on the great-circle distances between the geodesic centroids of country pairs. Again, all our baseline results remain qualitatively intact. Therefore, as far as our analysis of the influence of population diversity on conflict frequency is concerned, issues concerning spatial dependence do not pose a threat to identification.

Robustness to Accounting for Interpersonal Population Diversity as a Generated Regressor Our proxy for contemporary population diversity is a generated regressor in the empirical specifications, because it is projected from an implicit zeroth-stage relationship between prehistoric migratory distance from East Africa and expected heterozygosity in the HGDP-CEPH sample of 53 ethnic groups. Table A.10 establishes robustness of the standard-error estimates to accounting for a potential bias due to the use of generated regressors. To perform this robustness check, the current analysis adopts the two-step bootstrapping technique implemented by Ashraf and Galor (2013a) for computing the standard-error estimates for all models estimated in Table 3. The reader is referred to that work for additional details on the technique. Reassuringly, all of the boot-strapped standard errors are slightly lower than their robust counterparts reported in Table 3.

Robustness to Examining the Frequency of Overall Civil Conflict Onsets in the Ethnic Civil Conflict Sample Our baseline cross-country analysis for *overall* civil conflict frequency (Panel A of Table 3) uses a different sample than for our baseline cross-country analysis of *ethnic* civil conflict frequency (Panel B of Table 3). Since the parameter estimates from these two samples are not comparable, they are not very informative about the relative influence of population diversity on *overall* versus *ethnic* civil conflicts. In order to provide an appropriate benchmark for making *quantitative* comparisons between our findings with respect to the influence of population diversity on ethnic versus overall civil conflict frequency, Table A.11 replicates our baseline cross-country analysis of the temporal frequency of overall civil conflict onsets in Panel A of Table 3, this time using an outcome variable that reflects the PRIO25 civil conflict coding of WCM09 – i.e., our data source for ethnic civil conflict events that is based on an earlier version of the UCDP/PRIO Armed Conflict Dataset. As is apparent from comparing the results presented in Panel B of Table 3 versus Table A.11, the reduced-form impact of population diversity is indeed markedly stronger on the temporal frequency of overall (rather than only ethnic) civil conflict outbreaks, a finding that is consistent with our priors that the influence of interpersonal diversity on manifestations of intrastate conflict —more broadly defined— operates through mechanisms associated with social divisions that go well beyond merely ethnopolitical incompatibilities.

Robustness to the Elimination of Regions from the Estimation Sample Following the norm in cross-country empirical studies of civil conflict, we investigate whether our baseline findings —specifically, with respect to the influence of population diversity on either overall or ethnic civil conflict frequency in the relevant globally representative sample of countries— are driven by potentially influential observations in a given world region. Table A.12 conducts a robustness check on the results associated with the fully specified empirical models in the baseline cross-country analysis, as shown in Columns 8 and 12 of Table 3. In the table, we eliminate from our global sample one-at-a-time the following world regions: Sub-Saharan Africa (SSA), Middle East and North Africa (MENA), East Asia and Pacific (EAP), and Latin America and the Caribbean (LAC). Due to the lower degrees of freedom afforded by the regression samples with eliminated regions, the current analysis omits continent dummies from the empirical models in order to preserve as much of the cross-country variation in conflict as possible. The findings obtained under this robustness check reassuringly reveals that our baseline findings are not qualitatively sensitive to the exclusion of any potentially influential world region from our full estimation samples, in the sense that population diversity retains its significant explanatory power for the temporal frequency of either overall or ethnic civil conflict onsets in all restricted samples.

Selection on Observables and Unobservables Following the method first developed by Altonji, Elder and Taber (2005), we exploit the idea that the amount of selection on the unobserved variables in a model can be inferred from the amount of selection on the observed explanatory variables, thus permitting an assessment of how much larger the selectivity bias from unobserved heterogeneity needs to be, relative to the bias from selection on observables, in order to fully explain away the coefficient on our explanatory variable of interest.[2] Specifically, we compare the estimated coefficient, $\hat{\beta}_1^R$, on population diversity from a restricted model with its estimated coefficient, $\hat{\beta}_1^F$, from an augmented model with the full set of controls. We examine the Altonji, Elder and Taber (2005) ratio, $AET = \hat{\beta}_1^F/(\hat{\beta}_1^R - \hat{\beta}_1^F)$. Intuitively, a higher absolute value for AET indicates that the additional control variables included in the augmented model, relative to the restricted one, are not sufficient to explain away the estimated coefficient on population diversity in the full specification, and as such, this coefficient cannot be completely attributed to omitted-variable bias unless the amount of selection on unobservables is much larger than that on observables. We also compute two alternative statistics, δ and β^* suggested by Oster (forthcoming). δ measures how strongly

[2] Altonji, Elder and Taber (2005) develop this method for the case where the explanatory variable of interest is binary in nature, while Bellows and Miguel (2009) consider the case of a continuous explanatory variable. Roughly speaking, the assumption underlying assessments of this type is that the covariation of the outcome variable with observables, on the one hand, and its covariation with unobservables, on the other, are identically related to the explanatory variable of interest. Altonji, Elder and Taber (2005) provide some sufficient conditions for such an assumption to hold.

correlated any unobservables need to be with population diversity relative to observables in order to account for the full size of the coefficient on population diversity. But it differs from AET by accounting for the relevance of observable controls in explaining the variation in the outcome variable. The idea behind Oster's adjustment is that adding controls which do not move R^2 much would leave more room for unobservables that are correlated with the variable of interest. Oster's β^* statistic gives the estimated value of the coefficient on population diversity, if unobservables where as correlated as the observables. Oster (forthcoming) shows that if zero does not belong to the interval created by the estimated coefficient on population diversity and β^*, then we can reject the hypothesis that our coefficient of interest is exclusively driven by unobservables.

We consider Column 3 in Table 3 as our restricted model. This specification includes, besides population diversity, our baseline geographical controls and continent fixed effects. We assess coefficient stability as we augment this specification by adding the remaining controls to obtain our full baseline model in Column 8. The resulting AET ratio is -5.2 and it suggests that selection on unobservables would have to be at least five times larger than the selection on observables to account for the full size of the estimated coefficient of population diversity.[3] Oster's δ statistic on the other hand is -9.2, indicating that unobservables should be even more strongly correlated with population diversity (more than nine times the correlation with observables) to drive our estimate down to zero. Assuming that the unobservables are equally correlated with population diversity as the observables and that these correlations had the same sign, the estimated coefficient for diversity would be $\beta^* = 2.56$ if we were able to control for all these unobservables. Alternatively, if we assume the degree of selection on observables and unobservables to be of equal strength but work in opposite directions, the resulting β^* would be 0.15. Thus, in either scenario, the interval between our actual coefficient estimate from our full specification (0.398) and β^* excludes zero.[4] Therefore we conclude that it is improbable that omitted variables generate our results.

[3]The negative sign indicates that selection on unobservables need to move our coefficient estimate in the opposite direction compared to the selection on observables.

[4]All reported Oster's statistics are computed under the most conservative assumption that $R^2_{max} = 1$, i.e. the entire variation in conflict frequency would be explained with our model if we could add all unobservables correlated with population diversity in the model.

Table A.1: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – The Interaction of Diversity and Intergroup Distance

Cross-country sample:	Global		Old World		Global		Old World	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS
	Log number of new conflict onsets per year, based on:							
	PRIO25 civil conflicts, 1960–2008				WCM09 ethnic civil conflicts, 1960–2005			
Population diversity (ancestry adjusted)	0.661***	0.866***	0.700**	0.836**	0.786***	0.889***	0.782***	0.823**
	(0.240)	(0.299)	(0.282)	(0.357)	(0.250)	(0.321)	(0.294)	(0.366)
Average intergroup genetic distance	10.864***	13.341***	20.835	−10.126	12.331***	13.709***	−1.244	−26.307
	(3.729)	(4.527)	(20.279)	(25.143)	(3.757)	(4.930)	(21.753)	(27.677)
Diversity × distance	−15.408***	−18.967***	−29.236	12.314	−18.068***	−20.033***	−0.516	33.504
	(5.296)	(6.502)	(27.097)	(33.758)	(5.427)	(7.151)	(29.624)	(37.363)
Continent dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions		×		×		×		×
Controls for ethnic diversity		×		×		×		×
Controls for oil, population, and income		×		×		×		×
Observations	140	140	117	117	139	139	116	116
Partial R^2 of population diversity	0.104	0.160	0.100	0.115	0.117	0.136	0.098	0.087
Partial R^2 of genetic distance	0.066	0.098	0.004	0.001	0.067	0.079	0.000	0.005
Partial R^2 of interation term	0.065	0.096	0.005	0.001	0.070	0.081	0.000	0.004
Adjusted R^2	0.234	0.299	0.249	0.307	0.211	0.265	0.202	0.247
Effect of 10th–90th %ile move in diversity	0.044***	0.058***	0.035**	0.041**	0.053***	0.060***	0.039***	0.041**
	(0.016)	(0.020)	(0.014)	(0.018)	(0.017)	(0.022)	(0.015)	(0.018)

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of civil conflict onsets, as shown in Columns 7–10 of Table 3. Specifically, it establishes robustness to *additionally* considering the influence of average intergroup genetic distance in the national population as well as the interaction between population diversity and intergroup genetic distance. The measure of average intergroup genetic distance for a country is based on (i) the contemporary population shares of ethnic groups, differentiated by their precolonial ancestry or "source" populations as of the year 1500 (e.g., Putterman and Weil, 2010); and (ii) the pairwise interpopulation genetic distances amongst these ancestral populations, as predicted by their pairwise interpopulation migratory distances, exploiting the relationship between bilateral migratory distance and bilateral genetic distance across ethnic group pairs in the HGDP-CEPH sample of 53 ethnic groups. The results indicate that both population diversity and intergroup genetic distance contribute to the propensity of the population for civil conflict, conditional on one another, but the reduced-form impact of either variable is mitigated by the level of the other. This finding is consistent with a model of conflict in which the depth of intergroup cleavages increases the propensity of intergroup conflict over public goods but greater intragroup heterogeneity partly moderates this effect by increasing the coordination costs for any group to enter into conflict. It is also consistent with a model in which intragroup heterogeneity – by concentrating complementary material and labor inputs for conflict into different segments of a group's population – increases intergroup conflict propensity (e.g., Esteban and Ray, 2011a) but greater intergroup divergence in preferences over private goods mitigates this effect by reducing the space of issues over which intergroup contention may arise (e.g., Spolaore and Wacziarg, 2016). Further, in network-based models of conflict involving multiple groups (e.g., König et al., 2017), greater intergroup divergence could mitigate conflict propensity by reducing the strength of intergroup network alliances within one side or another of such conflicts. The specifications examined in this table are otherwise identical to those reported in Columns 7–10 of Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.2: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – Robustness to Accounting for Ecological/Climatic Covariates

Cross-country sample:	Global								Old World		Global	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS	(11) 2SLS	(12) 2SLS
Panel A	Log number of new PRIO25 civil conflict onsets per year, 1960–2008											
Population diversity (ancestry adjusted)	0.182**	0.410***	0.309**	0.332**	0.321*	0.350*	0.338*	0.345*	0.730**	0.891**	0.745***	0.882***
	(0.077)	(0.123)	(0.155)	(0.168)	(0.165)	(0.177)	(0.181)	(0.179)	(0.314)	(0.365)	(0.282)	(0.306)
Ecological fractionalization		0.003	0.008	0.009	0.008	0.008	0.007	0.010	0.012	0.018	0.001	0.002
		(0.021)	(0.020)	(0.024)	(0.025)	(0.024)	(0.024)	(0.025)	(0.024)	(0.030)	(0.020)	(0.023)
Ecological polarization		0.015	0.018	0.015	0.015	0.014	0.015	0.003	0.020	−0.002	0.021	0.003
		(0.020)	(0.019)	(0.021)	(0.022)	(0.021)	(0.022)	(0.024)	(0.021)	(0.028)	(0.018)	(0.021)
Annual temperature, 1960–2008 average		0.001	0.000	0.001	0.001	0.001	0.001	0.000	0.000	0.000	0.001	0.000
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Annual precipitation, 1960–2008 average		0.003	0.001	0.002	0.001	0.001	0.001	−0.002	0.011	0.004	0.006	0.003
		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.008)	(0.007)	(0.006)
Volatility of annual temperature, 1960–2008		−0.004	−0.012	−0.005	−0.007	−0.006	−0.007	−0.013	−0.028	−0.034	−0.011	−0.016
		(0.017)	(0.017)	(0.021)	(0.022)	(0.022)	(0.022)	(0.022)	(0.019)	(0.025)	(0.018)	(0.021)
Volatility of annual precipitation, 1960–2008		−0.031	−0.025	−0.047	−0.039	−0.046	−0.041	−0.032	−0.091	−0.032	0.008	0.010
		(0.057)	(0.056)	(0.057)	(0.058)	(0.057)	(0.058)	(0.059)	(0.095)	(0.097)	(0.060)	(0.063)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Control for ethnic fractionalization					×		×	×		×		×
Control for ethnolinguistic polarization						×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	143	143	143	143	143	143	143	143	119	119	143	143
Partial R^2 of Population diversity		0.085	0.035	0.040	0.037	0.043	0.040	0.044	0.087	0.109		
Partial R^2 sum of other diversity measures					0.006	0.007	0.005	0.006		0.010		
First-stage adjusted R^2											0.760	0.759
First-stage partial R^2 of migratory distance											0.397	0.375
First-stage F statistic											106.392	61.287
Adjusted R^2	0.019	0.162	0.172	0.170	0.167	0.168	0.162	0.203	0.241	0.278		
Effect of 10th–90th %ile move in diversity	0.012**	0.027***	0.020**	0.022**	0.021*	0.023**	0.022*	0.022*	0.036**	0.044**	0.049***	0.057***
	(0.005)	(0.008)	(0.010)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.016)	(0.018)	(0.018)	(0.020)
Panel B	Log number of new WCM09 ethnic civil conflict onsets per year, 1960–2005											
Population diversity (ancestry adjusted)	0.217***	0.457***	0.402**	0.406**	0.378**	0.417**	0.369*	0.356*	0.971***	1.046**	0.886***	0.884***
	(0.082)	(0.127)	(0.156)	(0.181)	(0.174)	(0.192)	(0.195)	(0.192)	(0.343)	(0.400)	(0.311)	(0.337)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography and climate		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Control for ethnic fractionalization					×		×	×		×		×
Control for ethnolinguistic polarization						×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	141	141	141	141	141	141	141	141	117	117	141	141
Partial R^2 of Population diversity		0.085	0.049	0.048	0.042	0.050	0.038	0.038	0.118	0.117		
Partial R^2 sum of other diversity measures					0.018	0.002	0.017	0.012		0.035		
First-stage adjusted R^2											0.758	0.759
First-stage partial R^2 of migratory distance											0.394	0.368
First-stage F statistic											101.734	56.699
Adjusted R^2	0.024	0.137	0.163	0.150	0.158	0.144	0.150	0.192	0.204	0.231		
Effect of 10th–90th %ile move in diversity	0.014***	0.030***	0.026***	0.026**	0.025**	0.027**	0.024*	0.023*	0.048***	0.052***	0.058***	0.057***
	(0.005)	(0.008)	(0.010)	(0.012)	(0.011)	(0.012)	(0.013)	(0.012)	(0.017)	(0.020)	(0.020)	(0.022)

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of civil conflict onsets, as shown in Table 3. Specifically, it establishes robustness to *additionally* accounting for the potentially confounding influence of (i) time-invariant fractionalization and polarization measures of the ecological diversity of land (e.g., Fenske, 2014); and (ii) the temporal mean and volatility of climatic experience (e.g., Burke, Hsiang and Miguel, 2015) with respect to annual temperature and annual precipitation over the relevant post-1960 time period. The specifications examined in this table are otherwise identical to those reported in Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the 2SLS regressions. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.3: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – Robustness to Accounting for Deep-Rooted Determinants of Economic Development

Cross-country sample:	Global								Old World		Global	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS	(11) 2SLS	(12) 2SLS
Panel A	Log number of new PRIO25 civil conflict onsets per year, 1960–2008											
Population diversity (ancestry adjusted)	0.213***	0.416***	0.345**	0.418**	0.408**	0.439**	0.424**	0.445**	0.894**	1.129***	0.697***	0.895***
	(0.080)	(0.122)	(0.149)	(0.176)	(0.174)	(0.186)	(0.188)	(0.189)	(0.348)	(0.391)	(0.263)	(0.284)
Ethnic fractionalization					0.012		0.008	0.003		0.005		−0.007
					(0.014)		(0.015)	(0.014)		(0.016)		(0.014)
Ethnolinguistic polarization						0.010	0.006	0.006		0.005		0.015
						(0.013)	(0.014)	(0.013)		(0.016)		(0.012)
Log years since Neolithic Revolution		0.005	0.007	0.012	0.011	0.011	0.011	0.012	−0.009	−0.007	0.003	0.009
		(0.005)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.012)	(0.012)	(0.008)	(0.007)
Log index of state antiquity		0.008**	0.007	0.007*	0.008*	0.007	0.008*	0.006	0.008*	0.006	0.008**	0.008*
		(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)
Log duration of human settlement		0.002	−0.000	0.001	0.001	0.001	0.001	0.003	0.002	0.007*	−0.001	0.003
		(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
Log distance from regional frontier in 1500		0.002	0.002	0.002	0.001	0.002	0.001	0.001	0.003	0.001	0.001	0.000
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	133	133	133	133	133	133	133	133	109	109	133	133
Partial R^2 of population diversity		0.098	0.049	0.070	0.067	0.075	0.068	0.080	0.117	0.175		
Partial R^2 sum of other diversity measures					0.006	0.006	0.004	0.002		0.002		
First-stage adjusted R^2											0.759	0.770
First-stage partial R^2 of migratory distance											0.450	0.441
First-stage F statistic											75.018	59.152
Adjusted R^2	0.028	0.223	0.203	0.223	0.220	0.220	0.214	0.266	0.282	0.351		
Effect of 10th–90th %ile move in diversity	0.014***	0.028***	0.023**	0.028**	0.028**	0.030**	0.029**	0.030**	0.044**	0.056***	0.047***	0.061***
	(0.005)	(0.008)	(0.010)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.017)	(0.019)	(0.018)	(0.019)
Panel B	Log number of new WCM09 ethnic civil conflict onsets per year, 1960–2005											
Population diversity (ancestry adjusted)	0.234***	0.361***	0.346**	0.368*	0.343*	0.388*	0.340*	0.367*	0.954**	1.103**	0.710**	0.795**
	(0.085)	(0.112)	(0.153)	(0.195)	(0.189)	(0.203)	(0.202)	(0.204)	(0.408)	(0.452)	(0.310)	(0.327)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography and deep determinants		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for ethnic diversity					×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	133	133	133	133	133	133	133	133	109	109	133	133
Partial R^2 of population diversity		0.060	0.039	0.042	0.037	0.045	0.035	0.045	0.102	0.130		
Partial R^2 sum of other diversity measures					0.022	0.004	0.018	0.012		0.017		
First-stage adjusted R^2											0.759	0.772
First-stage partial R^2 of migratory distance											0.450	0.441
First-stage F statistic											75.018	60.063
Adjusted R^2	0.028	0.170	0.163	0.140	0.151	0.136	0.143	0.242	0.195	0.261		
Effect of 10th–90th %ile move in diversity	0.016***	0.024***	0.023**	0.025*	0.023*	0.026*	0.023*	0.025*	0.047**	0.054**	0.048**	0.054**
	(0.006)	(0.008)	(0.010)	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)	(0.020)	(0.022)	(0.021)	(0.022)

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of civil conflict onsets, as shown in Table 3. Specifically, it establishes robustness to *additionally* accounting for the potentially confounding influence of other deep-rooted determinants of comparative economic development, including (i) the time elapsed since the onset of the Neolithic Revolution (e.g., Ashraf and Galor, 2013a); (ii) an index of experience with institutionalized statehood since antiquity (e.g., Bockstette, Chanda and Putterman, 2002); (iii) the time elapsed since initial human settlement in prehistory (e.g., Ahlerup and Olsson, 2012); and (iv) the great-circle distance to the closest regional technological frontier in the year 1500 (e.g., Ashraf and Galor, 2013a). The specifications examined in this table are otherwise identical to those reported in Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the 2SLS regressions. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.4: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – Robustness to Accounting for Ethnic and Spatial Inequality

Cross-country sample:	Global								Old World		Global	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS	(11) 2SLS	(12) 2SLS
Panel A	Log number of new PRIO25 civil conflict onsets per year, 1960–2008											
Population diversity (ancestry adjusted)	0.178**	0.426***	0.336**	0.394**	0.381**	0.415**	0.403**	0.411**	0.670**	0.907**	0.670**	0.881***
	(0.078)	(0.129)	(0.167)	(0.184)	(0.182)	(0.194)	(0.198)	(0.197)	(0.291)	(0.353)	(0.265)	(0.292)
Ethnic fractionalization					0.009		0.005	0.008		0.012		−0.002
					(0.012)		(0.012)	(0.013)		(0.014)		(0.013)
Ethnolinguistic polarization						0.010	0.008	0.008		0.006		0.019
						(0.013)	(0.014)	(0.013)		(0.015)		(0.014)
Ethnic inequality in luminosity		0.024	0.022	0.030*	0.030*	0.030*	0.030	0.026	0.022	0.030	0.026*	0.033*
		(0.015)	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)	(0.020)	(0.016)	(0.017)
Spatial inequality in luminosity		−0.003	0.002	−0.006	−0.006	−0.005	−0.005	−0.009	0.007	−0.010	0.007	−0.007
		(0.017)	(0.016)	(0.017)	(0.018)	(0.018)	(0.018)	(0.017)	(0.020)	(0.020)	(0.016)	(0.016)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	140	140	140	140	140	140	140	140	116	116	140	140
Partial R^2 of population diversity		0.120	0.049	0.064	0.060	0.069	0.062	0.068	0.098	0.154		
Partial R^2 sum of other diversity measures					0.004	0.006	0.004	0.006		0.008		
First-stage adjusted R^2											0.749	0.757
First-stage partial R^2 of migratory distance											0.448	0.418
First-stage F statistic											166.545	84.765
Adjusted R^2	0.018	0.193	0.206	0.207	0.203	0.204	0.198	0.235	0.263	0.313		
Effect of 10th–90th %ile move in diversity	0.012**	0.028***	0.022**	0.026**	0.025**	0.028**	0.027**	0.027**	0.029**	0.039**	0.044**	0.058***
	(0.005)	(0.009)	(0.011)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.012)	(0.015)	(0.018)	(0.019)
Panel B	Log number of new WCM09 ethnic civil conflict onsets per year, 1960–2005											
Population diversity (ancestry adjusted)	0.229***	0.419***	0.434***	0.480**	0.453**	0.496**	0.455**	0.453**	0.876***	1.063***	0.853***	0.976***
	(0.081)	(0.127)	(0.163)	(0.191)	(0.186)	(0.198)	(0.201)	(0.200)	(0.282)	(0.347)	(0.266)	(0.295)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography and inequality measures		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for ethnic diversity					×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	138	138	138	138	138	138	138	138	114	114	138	138
Partial R^2 of population diversity		0.103	0.071	0.083	0.075	0.086	0.070	0.074	0.141	0.183		
Partial R^2 sum of other diversity measures					0.013	0.003	0.010	0.014		0.023		
First-stage adjusted R^2											0.747	0.758
First-stage partial R^2 of migratory distance											0.445	0.419
First-stage F statistic											160.765	79.866
Adjusted R^2	0.029	0.173	0.220	0.215	0.219	0.211	0.212	0.256	0.263	0.322		
Effect of 10th–90th %ile move in diversity	0.015***	0.028***	0.029***	0.032**	0.031**	0.034**	0.031**	0.031**	0.037***	0.045***	0.058***	0.066***
	(0.005)	(0.009)	(0.011)	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)	(0.012)	(0.015)	(0.018)	(0.020)

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of civil conflict onsets, as shown in Table 3. Specifically, it establishes robustness to *additionally* accounting for the potentially confounding influence of measures of intrastate economic inequality (e.g., Alesina, Michalopoulos and Papaioannou, 2016), as captured by the subnational spatial distribution of per-capita adjusted nighttime luminosity in the year 2000 across either (i) the georeferenced homelands of ethnic groups; or (ii) 2.5×2.5-degree geospatial grid cells. The specifications examined in this table are otherwise identical to those reported in Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the 2SLS regressions. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.5: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – The Analysis under Linguistic Fractionalization

Cross-country sample:	Global								Old World		Global	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS	(11) 2SLS	(12) 2SLS
Panel A	Log number of new PRIO25 civil conflict onsets per year, 1960–2008											
Population diversity (ancestry adjusted)	0.186**	0.448***	0.352**	0.397**	0.407**	0.420**	0.417**	0.449**	0.622**	0.889***	0.593**	0.838***
	(0.080)	(0.130)	(0.168)	(0.190)	(0.189)	(0.197)	(0.196)	(0.197)	(0.259)	(0.327)	(0.236)	(0.270)
Linguistic fractionalization					0.017*		0.015	0.016		0.023**		0.015
					(0.009)		(0.010)	(0.010)		(0.012)		(0.010)
Ethnolinguistic polarization						0.013	0.006	0.006		0.002		0.011
						(0.013)	(0.014)	(0.014)		(0.015)		(0.013)
Absolute latitude		−0.424***	−0.410	−0.332	−0.205	−0.352	−0.232	0.216	−0.294	0.399	−0.481**	0.065
		(0.119)	(0.253)	(0.262)	(0.271)	(0.256)	(0.255)	(0.265)	(0.297)	(0.340)	(0.244)	(0.260)
Land area		0.729	1.720	1.643	2.046	1.643	1.994	1.565	4.211	4.540*	1.487	1.160
		(2.129)	(2.374)	(2.480)	(2.510)	(2.529)	(2.543)	(2.833)	(2.809)	(2.720)	(2.304)	(2.751)
Ruggedness		0.039	0.025	0.026	0.030	0.028	0.031	0.056	0.041	0.085	0.032	0.062
		(0.039)	(0.045)	(0.045)	(0.045)	(0.046)	(0.046)	(0.045)	(0.052)	(0.053)	(0.044)	(0.040)
Mean elevation		−0.017*	−0.015	−0.017	−0.017	−0.018	−0.017	−0.020*	−0.020	−0.025**	−0.017*	−0.024**
		(0.009)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.013)	(0.012)	(0.010)	(0.010)
Range of elevation		0.010**	0.009**	0.009**	0.009**	0.008*	0.008*	0.005	0.009*	0.003	0.010**	0.005
		(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)	(0.004)	(0.004)
Mean land suitability		0.015	0.018	0.015	0.016	0.018	0.017	0.004	0.018	0.003	0.019	0.003
		(0.013)	(0.013)	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)	(0.019)	(0.013)	(0.015)
Range of land suitability		0.012	0.016	0.013	0.010	0.016	0.012	0.008	0.019	0.001	0.018	0.010
		(0.009)	(0.011)	(0.013)	(0.012)	(0.015)	(0.015)	(0.017)	(0.013)	(0.017)	(0.011)	(0.015)
Distance to nearest waterway		0.008	0.004	0.007	0.003	0.007	0.003	−0.000	0.000	−0.005	0.004	−0.002
		(0.009)	(0.010)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.013)	(0.009)	(0.011)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	139	139	139	139	139	139	139	139	118	118	139	139
Partial R^2 of population diversity		0.127	0.050	0.060	0.064	0.066	0.066	0.080	0.091	0.159		
Partial R^2 sum of other diversity measures					0.022	0.010	0.016	0.018		0.034		
First-stage adjusted R^2											0.788	0.781
First-stage partial R^2 of migratory distance											0.517	0.475
First-stage F statistic											220.899	112.736
Adjusted R^2	0.019	0.204	0.208	0.203	0.213	0.204	0.208	0.253	0.260	0.333		
Effect of 10th–90th %ile move in diversity	0.012**	0.029***	0.023**	0.026**	0.027**	0.027**	0.027**	0.029**	0.030**	0.042***	0.039**	0.055***
	(0.005)	(0.008)	(0.011)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.012)	(0.016)	(0.015)	(0.018)
Panel B	Log number of new WCM09 ethnic civil conflict onsets per year, 1960–2005											
Population diversity (ancestry adjusted)	0.205**	0.440***	0.408**	0.416**	0.434**	0.439**	0.431**	0.458**	0.757***	0.953***	0.699***	0.863***
	(0.085)	(0.129)	(0.169)	(0.198)	(0.191)	(0.204)	(0.201)	(0.197)	(0.278)	(0.343)	(0.258)	(0.283)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for ethnolinguistic diversity					×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	137	137	137	137	137	137	137	137	116	116	137	137
Partial R^2 of population diversity		0.098	0.054	0.052	0.060	0.057	0.058	0.071	0.104	0.146		
Partial R^2 sum of other diversity measures					0.063	0.008	0.057	0.057		0.080		
First-stage adjusted R^2											0.786	0.779
First-stage partial R^2 of migratory distance											0.515	0.474
First-stage F statistic											215.900	109.242
Adjusted R^2	0.020	0.129	0.166	0.133	0.181	0.132	0.174	0.246	0.196	0.283		
Effect of 10th–90th %ile move in diversity	0.013**	0.029***	0.027**	0.027**	0.028**	0.028**	0.028**	0.030**	0.036***	0.045***	0.045***	0.056***
	(0.006)	(0.008)	(0.011)	(0.013)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.016)	(0.017)	(0.018)

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of civil conflict onsets, as shown in Table 3. Specifically, it establishes robustness to accounting for the potentially confounding influence of linguistic rather than ethnic fractionalization (e.g., Alesina et al., 2003), as a baseline control for subnational intergroup cultural fragmentation. The specifications examined in this table are otherwise identical to those reported in Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the other baseline covariates considered by the current analysis as well as the identification strategy employed by the 2SLS regressions. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.6: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – The Analysis under Initial Values of Time-Varying Covariates

Cross-country sample:	Global								Old World		Global	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS	(11) 2SLS	(12) 2SLS
Panel A	Log number of new PRIO25 civil conflict onsets per year, 1960–2008											
Population diversity (ancestry adjusted)	0.182**	0.422***	0.322**	0.358**	0.339**	0.377**	0.361**	0.365**	0.639**	0.830**	0.599***	0.789***
	(0.077)	(0.123)	(0.150)	(0.169)	(0.163)	(0.177)	(0.178)	(0.181)	(0.257)	(0.326)	(0.231)	(0.280)
Ethnic fractionalization					0.012		0.007	0.015		0.030*		0.007
					(0.012)		(0.012)	(0.013)		(0.016)		(0.013)
Ethnolinguistic polarization						0.012	0.009	0.009		0.006		0.017
						(0.013)	(0.014)	(0.014)		(0.016)		(0.013)
Executive constraints in initial year				0.002	0.001	0.002	0.001	0.001		−0.001		0.002
				(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		(0.003)		(0.002)
Democracy score in initial year				−0.002	−0.001	−0.001	−0.001	−0.001		0.001		−0.002
				(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		(0.002)		(0.002)
Autocracy score in initial year				−0.001	−0.001	−0.001	−0.001	−0.001		−0.001		−0.001
				(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		(0.002)		(0.002)
Log oil production per capita in initial year								0.001		0.002		0.001
								(0.002)		(0.002)		(0.002)
Log population in initial year								0.003		0.006**		0.003
								(0.002)		(0.003)		(0.002)
Log GDP per capita in initial year								−0.004		−0.006		−0.006
								(0.004)		(0.004)		(0.004)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Legal origin dummies				×	×	×	×	×		×		×
Colonial history dummies				×	×	×	×	×		×		×
Observations	143	143	143	143	143	143	143	143	119	119	143	143
Partial R^2 of population diversity		0.117	0.046	0.053	0.047	0.058	0.050	0.049	0.094	0.124		
Partial R^2 sum of other diversity measures					0.007	0.009	0.006	0.012		0.029		
First-stage adjusted R^2											0.755	0.774
First-stage partial R^2 of migratory distance											0.478	0.444
First-stage F statistic											211.910	95.474
Adjusted R^2	0.019	0.190	0.197	0.183	0.182	0.183	0.178	0.171	0.256	0.253		
Effect of 10th–90th %ile move in diversity	0.012**	0.027***	0.021**	0.023**	0.022**	0.025**	0.023**	0.024**	0.032**	0.041**	0.039***	0.051***
	(0.005)	(0.008)	(0.010)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.013)	(0.016)	(0.015)	(0.018)
Panel B	Log number of new WCM09 ethnic civil conflict onsets per year, 1960–2005											
Population diversity (ancestry adjusted)	0.217***	0.418***	0.375**	0.423**	0.390**	0.439**	0.397**	0.461**	0.780***	1.000***	0.707***	0.930***
	(0.082)	(0.121)	(0.152)	(0.179)	(0.171)	(0.185)	(0.185)	(0.192)	(0.277)	(0.358)	(0.254)	(0.314)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for ethnic diversity					×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	141	141	141	141	141	141	141	141	117	117	141	141
Partial R^2 of population diversity		0.092	0.050	0.060	0.051	0.063	0.050	0.063	0.108	0.138		
Partial R^2 sum of other diversity measures					0.013	0.004	0.009	0.016		0.039		
First-stage adjusted R^2											0.753	0.773
First-stage partial R^2 of migratory distance											0.475	0.432
First-stage F statistic											206.014	87.576
Adjusted R^2	0.024	0.127	0.165	0.158	0.162	0.155	0.155	0.160	0.194	0.199		
Effect of 10th–90th %ile move in diversity	0.014***	0.027***	0.024**	0.027**	0.025**	0.029**	0.026**	0.030**	0.038***	0.049***	0.046***	0.060***
	(0.005)	(0.008)	(0.010)	(0.012)	(0.011)	(0.012)	(0.012)	(0.012)	(0.014)	(0.018)	(0.016)	(0.020)

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of civil conflict onsets, as shown in Table 3. Specifically, it establishes robustness to considering the initial or year-1960 values of the time-dependent baseline controls for institutions (i.e., the degree of executive constraints and indicators for democracy and autocracy), oil production per capita, total population, and GDP per capita, rather than their respective temporal averages over the relevant post-1960 time period. The methodology exploited by the current analysis aims to reduce any ex ante bias in the baseline estimates of the influence of population diversity, arising from the fact that the temporal averages of the aforementioned time-varying controls may well vary more endogenously across countries with the contemporaneous measure of civil conflict onsets. In order to maintain a cross-country sample that is consistent with that in the baseline analysis, observations of the time-dependent covariates from the earliest available year after 1960 are used for the subset of countries with missing 1960 data. The specifications examined in this table are otherwise identical to those reported in Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the other baseline covariates considered by the current analysis as well as the identification strategy employed by the 2SLS regressions. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.7: Interpersonal Population Diversity and the *Count* of Civil Conflict Onset across Countries

Cross-country sample:	Global								Old World		Global	
	(1) NegBin	(2) NegBin	(3) NegBin	(4) NegBin	(5) NegBin	(6) NegBin	(7) NegBin	(8) NegBin	(9) NegBin	(10) NegBin	(11) Poisson	(12) Poisson
Panel A	Total count of new PRIO25 civil conflict onsets, 1960–2008											
Population diversity (ancestry adjusted)	9.089*	20.850***	17.044***	16.890***	16.804***	18.321***	18.295***	17.849***	24.826***	27.124***	17.044***	17.851***
	(4.842)	(4.535)	(6.242)	(6.016)	(6.010)	(5.995)	(6.013)	(5.499)	(7.299)	(5.849)	(6.243)	(5.499)
Ethnic fractionalization					0.248		0.038	0.248		0.559		0.248
					(0.518)		(0.540)	(0.537)		(0.581)		(0.537)
Ethnolinguistic polarization						0.647	0.638	0.852*		1.194**		0.852*
						(0.487)	(0.502)	(0.471)		(0.474)		(0.471)
Absolute latitude		−34.899***	−33.663***	−24.860***	−23.122**	−29.109***	−28.797**	−4.009	−31.388***	1.258	−33.665***	−4.006
		(6.575)	(8.831)	(9.154)	(10.221)	(10.014)	(11.351)	(11.643)	(9.533)	(12.285)	(8.832)	(11.643)
Land area		77.536	117.418**	101.731**	107.864**	104.634**	105.528**	43.343	159.022***	89.997	117.406**	43.320
		(48.510)	(49.640)	(51.751)	(51.181)	(52.308)	(52.604)	(54.553)	(50.194)	(54.979)	(49.642)	(54.554)
Ruggedness		2.723*	2.579*	2.677*	2.742*	2.844*	2.852*	3.815**	2.869*	4.681***	2.579*	3.815**
		(1.419)	(1.560)	(1.553)	(1.569)	(1.544)	(1.555)	(1.580)	(1.695)	(1.653)	(1.561)	(1.580)
Mean elevation		−0.963***	−0.952***	−0.988***	−0.983***	−1.020***	−1.019***	−1.061***	−1.065***	−1.155***	−0.952***	−1.061***
		(0.302)	(0.337)	(0.309)	(0.308)	(0.317)	(0.316)	(0.300)	(0.385)	(0.356)	(0.337)	(0.300)
Range of elevation		0.445***	0.406***	0.362***	0.355***	0.357***	0.356***	0.098	0.380***	−0.011	0.406***	0.098
		(0.100)	(0.085)	(0.083)	(0.086)	(0.086)	(0.089)	(0.116)	(0.083)	(0.136)	(0.085)	(0.116)
Mean land suitability		0.405	0.572	0.568	0.661	0.711	0.722	0.083	0.503	0.066	0.572	0.083
		(0.453)	(0.468)	(0.477)	(0.507)	(0.479)	(0.510)	(0.548)	(0.476)	(0.588)	(0.468)	(0.548)
Range of land suitability		0.739*	0.785*	0.868*	0.807*	0.939**	0.928**	0.467	1.047**	0.474	0.785*	0.467
		(0.430)	(0.456)	(0.449)	(0.443)	(0.443)	(0.443)	(0.465)	(0.486)	(0.446)	(0.456)	(0.465)
Distance to nearest waterway		0.296	0.141	0.339	0.322	0.372	0.369	0.325	−0.013	0.156	0.141	0.325
		(0.249)	(0.253)	(0.316)	(0.320)	(0.323)	(0.328)	(0.301)	(0.284)	(0.334)	(0.253)	(0.301)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	143	143	143	143	143	143	143	143	119	119	143	143
Pseudo R^2	0.012	0.171	0.197	0.216	0.217	0.220	0.220	0.257	0.217	0.292	0.240	0.298
Effect of 10th–90th %ile move in diversity	0.473**	1.123***	0.884**	0.877**	0.872**	0.959**	0.958**	0.932***	1.169**	1.303***	0.884**	0.932***
	(0.241)	(0.332)	(0.369)	(0.359)	(0.358)	(0.375)	(0.376)	(0.338)	(0.472)	(0.420)	(0.369)	(0.338)
Panel B	Total count of new WCM09 ethnic civil conflict onsets, 1960–2005											
Population diversity (ancestry adjusted)	11.835**	21.687***	22.008**	18.459**	17.256**	19.191**	17.756**	22.372***	33.459***	36.069***	24.968***	22.373***
	(5.825)	(5.860)	(9.224)	(8.501)	(8.560)	(8.091)	(8.233)	(7.879)	(8.131)	(7.434)	(8.427)	(7.878)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for ethnic diversity					×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	141	141	141	141	141	141	141	141	117	117	141	141
Pseudo R^2	0.011	0.121	0.191	0.218	0.226	0.221	0.226	0.280	0.173	0.274	0.302	0.389
Effect of 10th–90th %ile move in diversity	0.477**	0.893***	0.901**	0.761**	0.717*	0.799**	0.740**	0.906**	1.542**	1.776**	1.030**	0.906**
	(0.212)	(0.287)	(0.430)	(0.384)	(0.381)	(0.375)	(0.371)	(0.380)	(0.619)	(0.732)	(0.437)	(0.380)

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on civil conflict onsets, as shown in Table 3. Specifically, it establishes robustness to considering the *total count* rather than the annual frequency of civil conflict onsets over the relevant post-1960 time period as the outcome variable. In line with the standard for analyzing over-dispersed count data, the regressions are estimated using the negative binomial as opposed to a least squares estimator. Given the absence of a negative binomial estimator that permits instrumentation, however, the current analysis is unable to implement the strategy of exploiting prehistoric migratory distance from East Africa to the indigenous (precolonial) population of a country as an excluded instrument for the country's contemporary population diversity. Thus, in lieu of implementing the instrument-based identification strategy in the global sample of countries, Columns 11 and 12 examine robustness to employing the Poisson rather than the negative binomial estimator for estimating the key specifications from Columns 7 and 8, respectively. The specifications examined in this table are otherwise identical to the OLS specifications reported in Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the total number of new conflict onsets over the relevant post-1960 time period. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.8: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – Robustness to Accounting for Spatial Dependence using the Conley Correction

Cross-country sample:	Global								Old World		Global	
	(1) Conley OLS	(2) Conley OLS	(3) Conley OLS	(4) Conley OLS	(5) Conley OLS	(6) Conley OLS	(7) Conley OLS	(8) Conley OLS	(9) Conley OLS	(10) Conley OLS	(11) Conley GMM	(12) Conley GMM
Panel A	Log number of new PRIO25 civil conflict onsets per year, 1960–2008											
Population diversity (ancestry adjusted)	0.182***	0.422***	0.322**	0.366**	0.350**	0.390**	0.377**	0.398**	0.639***	0.855***	0.599***	0.805***
	(0.029)	(0.086)	(0.150)	(0.166)	(0.161)	(0.157)	(0.160)	(0.162)	(0.137)	(0.163)	(0.138)	(0.160)
Ethnic fractionalization					0.011**		0.006	0.007*		0.012***		−0.002
					(0.005)		(0.004)	(0.004)		(0.004)		(0.008)
Ethnolinguistic polarization						0.013*	0.010	0.010		0.007		0.019**
						(0.007)	(0.007)	(0.006)		(0.007)		(0.008)
Absolute latitude		−0.404***	−0.440***	−0.331***	−0.225**	−0.356***	−0.292***	0.149*	−0.333*	0.255***	−0.529**	−0.116
		(0.055)	(0.161)	(0.082)	(0.115)	(0.074)	(0.099)	(0.077)	(0.193)	(0.081)	(0.208)	(0.177)
Land area		0.765	1.825*	1.709	1.972*	1.719	1.862*	1.586	4.177***	4.114***	1.626	1.311
		(1.047)	(0.978)	(1.104)	(1.132)	(1.109)	(1.101)	(1.112)	(1.013)	(0.801)	(1.187)	(1.338)
Ruggedness		0.038***	0.028**	0.030**	0.036***	0.032**	0.035**	0.056***	0.041**	0.080***	0.034**	0.054***
		(0.014)	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)	(0.013)	(0.017)	(0.020)	(0.017)	(0.020)
Mean elevation		−0.016**	−0.015***	−0.017***	−0.018***	−0.018***	−0.018***	−0.020***	−0.019***	−0.025***	−0.016***	−0.023***
		(0.006)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
Range of elevation		0.009***	0.009***	0.009***	0.008***	0.008***	0.008***	0.004**	0.009***	0.003*	0.010***	0.005***
		(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Mean land suitability		0.013***	0.018***	0.016***	0.019***	0.019***	0.020***	0.006	0.018***	0.006	0.018***	0.003
		(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Range of land suitability		0.013***	0.014***	0.012***	0.011***	0.014***	0.013***	0.010**	0.019***	0.007	0.017***	0.015**
		(0.002)	(0.005)	(0.004)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)	(0.006)
Distance to nearest waterway		0.008**	0.005	0.007	0.006	0.007	0.007	0.002	0.000	−0.001	0.004	0.001
		(0.004)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	143	143	143	143	143	143	143	143	119	119	119	119
Adjusted R^2	0.325	0.443	0.447	0.444	0.443	0.445	0.441	0.473	0.496	0.531		
Panel B	Log number of new WCM09 ethnic civil conflict onsets per year, 1960–2005											
Population diversity (ancestry adjusted)	0.217***	0.418***	0.375**	0.385**	0.352**	0.408***	0.366**	0.391**	0.780***	0.904***	0.707***	0.795***
	(0.051)	(0.087)	(0.172)	(0.170)	(0.160)	(0.158)	(0.158)	(0.161)	(0.103)	(0.109)	(0.125)	(0.122)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for ethnic diversity					×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	141	141	141	141	141	141	141	141	117	117	117	117
Adjusted R^2	0.239	0.319	0.349	0.328	0.334	0.327	0.329	0.381	0.398	0.427		

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of civil conflict onsets, as shown in Table 3. Specifically, it establishes robustness of the standard-error estimates to accounting for spatial dependence across observations, following the methodology of Conley (1999). To perform this robustness check, the spatial distribution of observations is specified on the Euclidean plane using the full set of pairwise geodesic distances between country centroids, and the spatial autoregressive process across residuals is modeled as varying inversely with distance from each observation up to a maximum threshold of 25,000 kilometers, thus admitting the possibility of spatial dependence at a global scale. The GMM specifications in this table correspond to the 2SLS specifications from Table 3, exploiting prehistoric migratory distance from East Africa to the indigenous (precolonial) population of a country as an excluded instrument for the country's contemporary population diversity. The specifications examined in this table are otherwise identical to those reported in Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis. Standard errors, corrected for spatial autocorrelation, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.9: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – Robustness to Accounting for Spatial Dependence using SARAR Estimation

Cross-country sample:	Global								Old World		Global	
	(1) SARAR OLS	(2) SARAR OLS	(3) SARAR OLS	(4) SARAR OLS	(5) SARAR OLS	(6) SARAR OLS	(7) SARAR OLS	(8) SARAR OLS	(9) SARAR OLS	(10) SARAR OLS	(11) SARAR 2SLS	(12) SARAR 2SLS
Panel A	Log number of new PRIO25 civil conflict onsets per year, 1960–2008											
Population diversity (ancestry adjusted)	0.323**	0.435***	0.379**	0.383**	0.368**	0.406**	0.394**	0.405**	0.815**	0.898***	0.669**	0.575***
	(0.136)	(0.125)	(0.174)	(0.168)	(0.166)	(0.175)	(0.180)	(0.172)	(0.328)	(0.321)	(0.272)	(0.208)
Ethnic fractionalization					0.012		0.007	0.007		0.014		0.004
					(0.011)		(0.011)	(0.011)		(0.012)		(0.012)
Ethnolinguistic polarization						0.013	0.010	0.010		0.007		0.014
						(0.012)	(0.013)	(0.012)		(0.013)		(0.012)
Absolute latitude		−0.401***	−0.449*	−0.332	−0.219	−0.354	−0.284	0.148	−0.347	0.272	−0.532**	0.048
		(0.125)	(0.247)	(0.246)	(0.302)	(0.238)	(0.285)	(0.258)	(0.296)	(0.319)	(0.243)	(0.287)
Land area		0.648	1.501	1.604	1.873	1.629	1.769	1.571	3.949	4.061*	1.134	1.428
		(2.102)	(2.222)	(2.171)	(2.184)	(2.199)	(2.224)	(2.395)	(2.656)	(2.468)	(2.334)	(2.502)
Ruggedness		0.038	0.024	0.028	0.033	0.030	0.032	0.055	0.042	0.078*	0.024	0.052
		(0.037)	(0.038)	(0.040)	(0.042)	(0.040)	(0.042)	(0.042)	(0.045)	(0.047)	(0.036)	(0.042)
Mean elevation		−0.016*	−0.014	−0.017*	−0.017*	−0.018*	−0.018*	−0.020**	−0.019*	−0.025**	−0.014	−0.021**
		(0.008)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.009)	(0.011)	(0.011)	(0.009)	(0.009)
Range of elevation		0.009**	0.009**	0.009**	0.008**	0.008**	0.008*	0.004	0.009	0.003	0.009**	0.005
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)
Mean land suitability		0.014	0.018	0.016	0.019	0.019	0.020	0.006	0.018	0.007	0.017	0.006
		(0.012)	(0.013)	(0.014)	(0.013)	(0.014)	(0.013)	(0.015)	(0.014)	(0.018)	(0.012)	(0.014)
Range of land suitability		0.013	0.014	0.012	0.010	0.014	0.013	0.010	0.019	0.006	0.017	0.012
		(0.008)	(0.011)	(0.011)	(0.011)	(0.013)	(0.013)	(0.013)	(0.012)	(0.014)	(0.011)	(0.013)
Distance to nearest waterway		0.008	0.005	0.007	0.006	0.007	0.007	0.002	0.001	−0.001	0.005	0.002
		(0.009)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)
Spatial lag AR(1) of conflict (λ)	−2.851*	−0.269	−1.006	−0.371	−0.415	−0.369	−0.436	−0.196	−1.273	−0.502	−1.470**	−0.458
	(1.476)	(0.641)	(0.725)	(0.676)	(0.683)	(0.657)	(0.671)	(0.561)	(0.838)	(0.652)	(0.715)	(0.582)
Spatial lag AR(1) of error (ρ)	4.084***	0.349	1.136	0.356	0.462	0.265	0.393	0.066	1.221**	0.383	1.444**	0.211
	(1.008)	(0.742)	(0.774)	(0.815)	(0.840)	(0.788)	(0.798)	(0.564)	(0.616)	(0.470)	(0.603)	(0.559)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	143	143	143	143	143	143	143	143	119	119	143	143
Effect of 10th–90th %ile move in diversity	0.021**	0.028***	0.025**	0.025**	0.024**	0.026**	0.026**	0.026**	0.040**	0.044***	0.044**	0.037***
	(0.009)	(0.008)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.016)	(0.016)	(0.018)	(0.014)
Panel B	Log number of new WCM09 ethnic civil conflict onsets per year, 1960–2005											
Population diversity (ancestry adjusted)	0.319***	0.421***	0.440**	0.468**	0.434**	0.487**	0.440**	0.426**	1.057***	1.032***	0.767***	0.672***
	(0.116)	(0.124)	(0.179)	(0.197)	(0.188)	(0.200)	(0.200)	(0.191)	(0.361)	(0.370)	(0.289)	(0.237)
Spatial lag AR(1) of conflict (λ)	−2.051	−0.049	−1.045	−1.489**	−1.516*	−1.435**	−1.404*	−0.757	−1.967**	−1.120	−1.433*	−1.188*
	(1.373)	(0.725)	(0.806)	(0.756)	(0.811)	(0.725)	(0.787)	(0.621)	(0.772)	(0.710)	(0.744)	(0.620)
Spatial lag AR(1) of error (ρ)	0.368	0.033	0.956	1.347	1.576*	1.318	1.454*	0.778	1.302**	0.815*	1.197*	0.999*
	(1.030)	(0.737)	(0.777)	(0.819)	(0.891)	(0.811)	(0.867)	(0.541)	(0.508)	(0.489)	(0.625)	(0.516)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for ethnic diversity					×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	141	141	141	141	141	141	141	141	117	117	141	141
Effect of 10th–90th %ile move in diversity	0.021***	0.027***	0.029**	0.030**	0.028**	0.032**	0.029**	0.028**	0.052***	0.051***	0.050***	0.044***
	(0.008)	(0.008)	(0.012)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)	(0.018)	(0.018)	(0.019)	(0.015)

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of civil conflict onsets, as shown in Table 3. Specifically, it establishes robustness to accounting for spatial dependence across observations by estimating spatial-autoregressive models with spatial-autoregressive disturbances (SARAR(1,1)) using the generalized spatial two-stage least squares (GS2SLS) estimator (e.g., Drukker, Prucha and Raciborski, 2013). To perform this robustness check, involving the estimation of the AR(1) coefficients, λ and ρ, respectively associated with the spatial lags in the outcome variable and the error term, the estimator exploits an inverse-distance spatial weighting matrix for the regression sample, based on the great-circle distances between the geodesic centroids of country pairs. The specifications examined in this table are otherwise identical to those reported in Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the 2SLS regressions in the last two columns. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.10: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – Robustness to Accounting for Population Diversity as a Generated Regressor

Cross-country sample:	Global								Old World		Global	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS	(11) 2SLS	(12) 2SLS
Panel A	Log number of new PRIO25 civil conflict onsets per year, 1960–2008											
Population diversity (ancestry adjusted)	0.182**	0.422***	0.322**	0.366**	0.350**	0.390**	0.377**	0.398**	0.639***	0.855***	0.599***	0.805***
	(0.075)	(0.118)	(0.146)	(0.167)	(0.166)	(0.174)	(0.180)	(0.183)	(0.241)	(0.312)	(0.229)	(0.306)
Ethnic fractionalization					0.011		0.006	0.007		0.012		−0.002
					(0.013)		(0.014)	(0.014)		(0.016)		(0.017)
Ethnolinguistic polarization						0.013	0.010	0.010		0.007		0.019
						(0.013)	(0.014)	(0.014)		(0.016)		(0.015)
Absolute latitude		−0.404***	−0.440	−0.331	−0.225	−0.356	−0.292	0.149	−0.333	0.255	−0.529*	−0.116
		(0.122)	(0.272)	(0.279)	(0.346)	(0.274)	(0.342)	(0.325)	(0.304)	(0.398)	(0.280)	(0.381)
Land area		0.765	1.825	1.709	1.972	1.719	1.862	1.586	4.177	4.114	1.626	1.311
		(2.451)	(2.887)	(3.053)	(3.073)	(3.132)	(3.187)	(3.472)	(6.044)	(5.654)	(2.948)	(3.731)
Ruggedness		0.038	0.028	0.030	0.036	0.032	0.035	0.056	0.041	0.080	0.034	0.054
		(0.040)	(0.047)	(0.048)	(0.052)	(0.049)	(0.053)	(0.052)	(0.054)	(0.057)	(0.047)	(0.053)
Mean elevation		−0.016*	−0.015	−0.017	−0.018	−0.018	−0.018	−0.020*	−0.019	−0.025*	−0.016	−0.023*
		(0.009)	(0.010)	(0.011)	(0.011)	(0.012)	(0.012)	(0.011)	(0.013)	(0.014)	(0.010)	(0.012)
Range of elevation		0.009**	0.009**	0.009**	0.008*	0.008*	0.008*	0.004	0.009*	0.003	0.010**	0.005
		(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)
Mean land suitability		0.013	0.018	0.016	0.019	0.019	0.020	0.006	0.018	0.006	0.018	0.003
		(0.012)	(0.013)	(0.015)	(0.014)	(0.015)	(0.014)	(0.016)	(0.015)	(0.021)	(0.013)	(0.017)
Range of land suitability		0.013	0.014	0.012	0.011	0.014	0.013	0.010	0.019	0.007	0.017	0.015
		(0.009)	(0.012)	(0.013)	(0.013)	(0.015)	(0.015)	(0.016)	(0.013)	(0.017)	(0.013)	(0.016)
Distance to nearest waterway		0.008	0.005	0.007	0.006	0.007	0.007	0.002	0.000	−0.001	0.004	0.001
		(0.009)	(0.010)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.010)	(0.012)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	143	143	143	143	143	143	143	143	119	119	143	143
Adjusted R^2	0.019	0.190	0.197	0.192	0.190	0.193	0.188	0.235	0.256	0.308		
Effect of 10th–90th %ile move in diversity	0.012**	0.027***	0.021**	0.024**	0.023**	0.025**	0.025*	0.026**	0.032**	0.042**	0.039**	0.052**
	(0.005)	(0.008)	(0.010)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.014)	(0.018)	(0.016)	(0.022)
Panel B	Log number of new WCM09 ethnic civil conflict onsets per year, 1960–2005											
Population diversity (ancestry adjusted)	0.217***	0.418***	0.379**	0.390**	0.357**	0.417**	0.377**	0.402**	0.780***	0.904**	0.880***	0.895**
	(0.083)	(0.120)	(0.156)	(0.184)	(0.180)	(0.186)	(0.190)	(0.191)	(0.273)	(0.357)	(0.319)	(0.357)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for geography		×	×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for ethnic diversity					×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	141	141	141	141	141	141	141	141	117	117	141	141
Adjusted R^2	0.024	0.127	0.161	0.138	0.146	0.140	0.141	0.210	0.194	0.233		
Effect of 10th–90th %ile move in diversity	0.014**	0.027***	0.025**	0.025**	0.023*	0.027**	0.024*	0.026**	0.038**	0.045**	0.057***	0.058**
	(0.006)	(0.008)	(0.010)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.016)	(0.020)	(0.021)	(0.024)

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of civil conflict onsets, as shown in Table 3. Specifically, it establishes robustness of the standard-error estimates to accounting for the fact that the country-level measure of contemporary population diversity is a generated regressor in the empirical specifications, because it is projected from an implicit zeroth-stage relationship between prehistoric migratory distance from East Africa and expected heterozygosity in the HGDP-CEPH sample of 53 ethnic groups. To perform this robustness check, the current analysis adopts the two-step bootstrapping technique implemented by Ashraf and Galor (2013a) for computing the standard-error estimates, so the reader is referred to that work for additional details on the technique. The specifications examined in this table are otherwise identical to those reported in Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the 2SLS regressions in the last two columns. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Bootstrap standard errors, accounting for the use of a generated regressor, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.11: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – The Analysis in the Ethnic Civil Conflict Sample

Cross-country sample:	Global								Old World		Global	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS	(8) OLS	(9) OLS	(10) OLS	(11) 2SLS	(12) 2SLS
	Log number of new WCM09 civil conflict onsets per year, 1960–2005											
Population diversity (ancestry adjusted)	0.219**	0.529***	0.598***	0.615***	0.593**	0.642***	0.623**	0.665***	1.017***	1.225***	0.908***	1.100***
	(0.110)	(0.151)	(0.191)	(0.231)	(0.227)	(0.241)	(0.247)	(0.240)	(0.345)	(0.417)	(0.316)	(0.352)
Ethnic fractionalization					0.014		0.008	0.007		0.016		−0.004
					(0.015)		(0.017)	(0.018)		(0.022)		(0.019)
Ethnolinguistic polarization						0.014	0.011	0.008		−0.002		0.018
						(0.018)	(0.020)	(0.018)		(0.020)		(0.017)
Absolute latitude		−0.741***	−0.922**	−0.688*	−0.554	−0.718*	−0.632	0.077	−0.889*	0.024	−1.024***	−0.219
		(0.179)	(0.373)	(0.391)	(0.453)	(0.376)	(0.421)	(0.386)	(0.473)	(0.515)	(0.331)	(0.375)
Land area		−2.269	−1.114	−0.465	−0.141	−0.438	−0.252	−0.214	1.732	2.600	−1.326	−0.490
		(2.405)	(2.784)	(2.858)	(2.913)	(2.895)	(2.970)	(3.161)	(3.293)	(3.371)	(2.748)	(3.058)
Ruggedness		0.068	0.031	0.037	0.044	0.040	0.043	0.070	0.055	0.108	0.037	0.068
		(0.055)	(0.060)	(0.059)	(0.063)	(0.060)	(0.063)	(0.063)	(0.071)	(0.075)	(0.058)	(0.056)
Mean elevation		−0.026**	−0.019	−0.023	−0.023	−0.024	−0.024	−0.028*	−0.026	−0.037**	−0.021	−0.031**
		(0.012)	(0.013)	(0.016)	(0.016)	(0.016)	(0.016)	(0.014)	(0.017)	(0.018)	(0.013)	(0.014)
Range of elevation		0.013***	0.010**	0.010*	0.009	0.009*	0.009	0.005	0.010	0.003	0.011**	0.006
		(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.005)	(0.006)
Mean land suitability		0.018	0.021	0.011	0.015	0.014	0.015	−0.007	0.018	−0.014	0.021	−0.011
		(0.017)	(0.019)	(0.021)	(0.020)	(0.021)	(0.021)	(0.023)	(0.023)	(0.028)	(0.018)	(0.021)
Range of land suitability		0.022*	0.030*	0.030*	0.029*	0.033*	0.032*	0.032	0.038**	0.027	0.033**	0.037*
		(0.012)	(0.016)	(0.016)	(0.016)	(0.018)	(0.018)	(0.021)	(0.019)	(0.024)	(0.017)	(0.020)
Distance to nearest waterway		0.022	0.017	0.012	0.011	0.012	0.011	0.004	0.012	−0.002	0.016	0.002
		(0.014)	(0.016)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)	(0.019)	(0.015)	(0.016)
Continent dummies			×	×	×	×	×	×	×	×	×	×
Controls for institutions				×	×	×	×	×		×		×
Controls for oil, population, and income								×		×		×
Observations	141	141	141	141	141	141	141	141	117	117	141	141
Partial R^2 of population diversity		0.095	0.079	0.079	0.073	0.084	0.075	0.092	0.119	0.145		
Partial R^2 sum of other diversity measures					0.005	0.006	0.004	0.003		0.005		
First-stage adjusted R^2											0.753	0.760
First-stage partial R^2 of migratory distance											0.475	0.438
First-stage F statistic											206.014	97.246
Adjusted R^2	0.012	0.210	0.237	0.220	0.217	0.218	0.213	0.277	0.285	0.326		
Effect of 10th–90th %ile move in diversity	0.014**	0.035***	0.039***	0.041***	0.039***	0.042***	0.041**	0.044***	0.051***	0.061***	0.060***	0.072***
	(0.007)	(0.010)	(0.013)	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)	(0.021)	(0.021)	(0.023)

Notes: This table conducts a robustness check on the results from the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of *overall* civil conflict onsets, as shown in Panel A of Table 3. Specifically, it establishes robustness to exploiting the same regression sample and conflict data source (i.e., WCM09) as those employed for the analysis of *ethnic* civil conflict in Panel B of Table 3. The current analysis thus provides an appropriate benchmark for assessing the influence of contemporary population diversity on overall versus ethnic civil conflict. The specifications examined in this table are otherwise identical to those reported in Panel A of Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the 2SLS regressions. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.12: Interpersonal Population Diversity and the Frequency of Civil Conflict Onset across Countries – Robustness to the Elimination of Regions from the Global Sample

Omitted region:	None		SSA		MENA		EAP		LAC	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS	(9) OLS	(10) 2SLS
Panel A	Log number of new PRIO25 civil conflict onsets per year, 1960–2008									
Population diversity (ancestry adjusted)	0.411***	0.613***	0.505***	1.210***	0.437***	0.637***	0.340**	0.616***	0.578**	0.736***
	(0.142)	(0.201)	(0.179)	(0.379)	(0.154)	(0.204)	(0.151)	(0.223)	(0.226)	(0.248)
Controls for geography	×	×	×	×	×	×	×	×	×	×
Controls for institutions	×	×	×	×	×	×	×	×	×	×
Controls for ethnic diversity	×	×	×	×	×	×	×	×	×	×
Controls for oil, population, and income	×	×	×	×	×	×	×	×	×	×
Observations	143	143	102	102	127	127	128	128	123	123
Partial R^2 of population diversity	0.100		0.107		0.111		0.060		0.089	
Partial R^2 sum of other diversity measures	0.015		0.007		0.016		0.029		0.006	
First-stage adjusted R^2		0.688		0.635		0.686		0.697		0.804
First-stage partial R^2 of migratory distance		0.510		0.282		0.519		0.508		0.703
First-stage F statistic		71.295		25.274		72.047		56.085		69.174
Adjusted R^2	0.233		0.169		0.236		0.239		0.255	
Effect of 10th–90th %ile move in diversity	0.027***	0.040***	0.031***	0.074***	0.030***	0.044***	0.021**	0.037***	0.027**	0.034***
	(0.009)	(0.013)	(0.011)	(0.023)	(0.011)	(0.014)	(0.009)	(0.014)	(0.011)	(0.012)
Panel B	Log number of new WCM09 ethnic civil conflict onsets per year, 1960–2005									
Population diversity (ancestry adjusted)	0.350**	0.575***	0.376*	0.875**	0.375**	0.634***	0.354**	0.586***	0.451*	0.516**
	(0.141)	(0.185)	(0.195)	(0.348)	(0.151)	(0.183)	(0.144)	(0.211)	(0.244)	(0.238)
Controls for geography	×	×	×	×	×	×	×	×	×	×
Controls for institutions	×	×	×	×	×	×	×	×	×	×
Controls for ethnic diversity	×	×	×	×	×	×	×	×	×	×
Controls for oil, population, and income	×	×	×	×	×	×	×	×	×	×
Observations	141	141	101	101	126	126	126	126	121	121
Partial R^2 of population diversity	0.056		0.053		0.066		0.049		0.042	
Partial R^2 sum of other diversity measures	0.007		0.028		0.005		0.007		0.002	
First-stage adjusted R^2		0.690		0.639		0.688		0.700		0.807
First-stage partial R^2 of migratory distance		0.495		0.288		0.496		0.491		0.704
First-stage F statistic		68.639		25.768		67.661		53.825		69.922
Adjusted R^2	0.163		0.185		0.168		0.135		0.153	
Effect of 10th–90th %ile move in diversity	0.023**	0.037***	0.023*	0.053**	0.026**	0.043***	0.021**	0.036***	0.021*	0.024**
	(0.009)	(0.012)	(0.012)	(0.021)	(0.010)	(0.013)	(0.009)	(0.013)	(0.011)	(0.011)

Notes: This table conducts a robustness check on the results associated with the fully specified empirical models in the baseline cross-country analysis of the reduced-form impact of contemporary population diversity on the annual frequency of civil conflict onsets, as shown in Columns 8 and 12 of Table 3. Specifically, it establishes robustness to the one-at-a-time elimination of world regions from the global sample, including Sub-Saharan Africa (SSA), Middle East and North Africa (MENA), East Asia and Pacific (EAP), and Latin America and the Caribbean (LAC). Due to the lower degrees of freedom afforded by the regression samples with eliminated regions, the current analysis omits continent dummies from the empirical models in order to preserve as much of the cross-country variation in conflict as possible. The regressions in Columns 1–2 should therefore be viewed as the relevant baselines for assessing the robustness results presented in the remaining columns. The set of covariates, however, is otherwise identical to those reported in Columns 8 and 12 of Table 3. The reader is therefore referred to Table 3 and the corresponding table notes for additional details on the set of covariates considered by the current analysis as well as the identification strategy employed by the 2SLS regressions. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the number of new conflict onsets per year. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

A.2 Robustness Checks for the Analysis of Civil Conflict Incidence or Onset in Repeated Cross-Country Data

In this appendix section, we present several robustness checks for our repeated cross-country analysis of the influence of interpersonal population diversity —as proxied by an ancestry-adjusted measure of population diversity— on civil conflict incidence and onset in the post-1960 time horizon.

Robustness to Ecological/Climatic Covariates A nascent interdisciplinary literature (e.g., Burke et al., 2009; Hsiang, Burke and Miguel, 2013; Burke, Hsiang and Miguel, 2015) has emphasized the role of climatic factors, like temperature and precipitation, as important correlates of the risk of civil conflict. Fenske (2014) shows that ecological diversity facilitated state centralization in pre-colonial African states. To keep our main specifications from becoming too unwieldy, we chose to exclude the aforementioned climatic and ecological variables from our baseline set of covariates, especially because this set already included a sizable vector of geographical factors that are known to be correlated with the former. In Table A.13, however, we establish that population diversity remains a significant predictor of overall and ethnic civil conflict incidence (Panel A) as well as onset (Panel B), when we augment our baseline set of covariates in Table 4 with controls for (i) time-invariant fractionalization and polarization measures of the ecological diversity of land (e.g., Fenske, 2014); and (ii) climatic experience in the recent past (e.g., Burke, Hsiang and Miguel, 2015), as captured by either (a) the temporal mean and volatility of annual temperature and annual precipitation over the previous 5-year interval for the quinquennial incidence regressions; or (b) the lagged values of annual temperature and annual precipitation as well as their temporal volatility over the previous 5 years for the annual onset regressions. The specifications examined in this table are otherwise identical to those reported in Table 4 For further details, see the corresponding table notes.

Robustness to Accounting for Deep-Rooted Determinants of Economic Development In Table A.14, we establish robustness of our baseline repeated cross-country analyses of civil conflict incidence and onset to *additionally* accounting for the potentially confounding influence of other deep-rooted determinants of comparative economic development. We augment the analyses in Table 4 with controls for (i) the time elapsed since the onset of the Neolithic Revolution (e.g., Ashraf and Galor, 2013a); (ii) an index of experience with institutionalized statehood since antiquity (e.g., Bockstette, Chanda and Putterman, 2002); (iii) the time elapsed since initial human settlement in prehistory (e.g., Ahlerup and Olsson, 2012); and (iv) the great-circle distance to the closest regional technological frontier in the year 1500 (e.g., Ashraf and Galor, 2013a). Regardless of the sample and specification, contemporary population diversity remains a significant predictor of quinquennial likelihood of a conflict incidence (Panel A) or the annual likelihood of a conflict onset (Panel B).

Robustness to Accounting for Ethnic and Spatial Inequality In Table A.15 we conduct robustness checks to *additionally* account for intrastate economic inequality (e.g., Alesina, Michalopoulos and Papaioannou, 2016), as captured by the subnational spatial distribution of per-capita adjusted nighttime luminosity in the year 2000 across either (i) the georeferenced homelands of ethnic groups (ethnic inequality); or (ii) 2.5×2.5-degree geospatial grid cells (spatial inequality). The specifications examined in this table are otherwise identical to those reported in Table 4. Findings indicate that ethnic —but not spatial— inequality in luminosity appears to be a significant predictor of ethnic —but not overall— civil conflict incidence and onset. Nonetheless, the positive and significant influence of ancestry-adjusted population diversity on both overall and ethnic civil conflict incidence as well as onset remains qualitatively intact.

Robustness to Accounting for Alternative Correlates of Conflict Incidence Table A.16 conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on the quinquennial incidence of civil conflict in repeated cross-country data, as shown in Panel A of Table 4. Specifically, we additionally control for the potentially confounding influence of *alternative* distributional indices of intergroup diversity (e.g., Fearon, 2003; Alesina et al., 2003; Esteban, Mayoral and Ray, 2012) and *additional* geographical correlates of conflict (e.g., Fearon and Laitin, 2003; Cervellati, Sunde and Valmori, 2017). The specifications examined in this table are identical to the fully specified baseline models reported in Panel A of Table 4, with the exception that in Columns 1–3 and 6–8 of the current analysis, each of the reported control variables is employed in lieu of the baseline control for ethnic fractionalization (Alesina et al., 2003), whereas in Columns 4 and 9, the set of reported control variables replaces the baseline controls for both ethnic fractionalization and ethnolinguistic polarization (Desmet, Ortuño-Ortín and Wacziarg, 2012), in the interest of mitigating multicollinearity. Further, in Columns 5 and 10 of the current analysis, the set of reported geographical controls augment the fully specified baseline models from Panel A of Table 4. Our baseline findings regarding the role of population diversity remain qualitatively unaltered across all specifications. Among the additional controls, ethnolinguistic polarization (Esteban, Mayoral and Ray, 2012) and contiguous state dummy enter the global sample IV Probit regressions for overall conflict incidence with a positive and significant coefficient.

Robustness to Employing Alternative Measures of Conflict Incidence Population diversity predicts a higher quinquennial incidence of intrastate conflict in repeated cross-country data (as shown in Panel A of Table 4). Table A.17 probes whether this result carries through when we consider as alternative outcomes the temporal incidence of (i) a high-intensity PRIO1000 civil war in any given 5-year interval during the 1960–2008 time period (Columns 1–4); (ii) a high-intensity WM0609 ethnic civil war in any given 5-year interval during the 1960–2001 time period (Columns 5–8); and (iii) a low-intensity conflict involving nonstate actors in any given 5-year interval during the 1990–2008 time period (Columns 9–12). The findings indicate that, regardless of the identification strategy, population diversity exerts an adverse and significant influence on the aforementioned measures of intrastate conflict.

Analysis of Conflict Incidence in *Annually* Repeated Cross-Country Data In Table A.18 we conduct a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on the temporal incidence of civil conflict in repeated cross-country data, as shown in Panel A of Table 4. Specifically, we consider *annual* rather than quinquennial repetitions of the cross-section, thereby demonstrating the significant positive reduced-form impact of population diversity on the annual likelihood of a conflict incidence during the relevant post-1960 time period. The specifications examined in this table are essentially identical to those reported in Panel A of Table 4, with the exception that the time-dependent baseline controls are all appropriately adjusted to assume their respective lagged annual values rather than the values corresponding to the previous 5-year interval. Our baseline incidence findings are qualitatively not affected by this modification.

Robustness to Employing the Logit and Rare-Events Logit Estimators Table A.19 checks the robustness of our quinquennial conflict incidence and annual onset regressions (reported in Table 4) to employing the ordinary logit and rare-events logit (King and Zeng, 2001) estimators, rather than the probit estimator, for estimating the relevant empirical models of conflict incidence and onset. This table conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on either the quinquennial incidence or the annual onset of civil conflict in repeated cross-sectional data for the Old World sample of countries, as shown in Columns 1–2 and 5–6 of Panels A and B of Table 4. Given the absence of readily available ordinary logit and rare-events logit estimators that permit instrumentation, the current analysis is unable to implement the global-sample identification strategy of exploiting prehistoric migratory distance from East Africa to the indigenous (precolonial) population of a country as an excluded instrument for the country's contemporary population diversity. The findings attest to the robustness of the reduced-form influence of population diversity on our baseline outcome measures of overall and ethnic conflict incidence as well as onset. If anything, rare events logit estimates appear somewhat larger than their counterparts under the regular logit estimator.

Robustness to Accounting for Spatiotemporal Dependence using Two-Way Clustering of Standard Errors Table A.20 conducts a robustness check on the results from the baseline probit and logit analyses of the reduced-form impact of contemporary population diversity on either the quinquennial incidence or the annual onset of civil conflict in repeated cross-sectional data for the Old World sample of countries, as shown in Columns 1–2 and 5–6 of Table 4 and in odd-numbered columns of Table A.19. Specifically, we probe the robustness of the standard-error estimates to accounting for spatiotemporal dependence across country-time observations by implementing multi-dimensional clustering of standard errors, following the methodology of Cameron, Gelbach and Miller (2011). To implement this robustness check, the standard errors across country-time observations are clustered in two dimensions: (i) the country level, which allows for temporal dependence within a country over time (i.e., across 5-year intervals or years); and (ii) the time level, which allows for spatial dependence across countries within a given time period (i.e., within a 5-year interval or year). Ultimately, our baseline estimates for population diversity remain statistically significant (either at the level of 1 percent or 5 percent) when two-way clustering is allowed for.

Robustness to Accounting for Alternative Correlates of Conflict Onset In Table A.21 we conduct a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on the annual onset of civil conflict in repeated cross-country data, as shown in Panel B of Table 4. Specifically, we account for the potentially confounding influence of an *additional* distributional index of intergroup diversity (ethnic dominance) (e.g., Collier and Hoeffler, 2004) and *additional* lagged values of time-varying institutional correlates of conflict (political instability and a new state dummy) (e.g., Fearon and Laitin, 2003). In light of sample restrictions imposed by the availability of data on these additional control variables, the specification presented in each odd-numbered column of the table is intended to provide a relevant baseline for the robustness check in the subsequent even-numbered column (i.e., by holding fixed the regression sample). Although the additional time-varying institutional covariates enter some of the specifications with a significant and positive sign, the coefficients on population diversity remain positive and precisely estimated.

Robustness to Accounting for Commodity Export Price Shocks In Table A.22 we conduct a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on the annual onset of a PRIO25 civil conflict in repeated cross-country data, as shown in Columns 1–4 of Panel B of Table 4. In the table, we *additionally* account for the potentially confounding "income effect" of commodity export price shocks (e.g., Bazzi and Blattman, 2014), as captured by the contemporaneous, lagged, and twice lagged values of either an annual price shock that has been aggregated across commodity export types (Columns 1–2 and 5–6) or annual price shocks disaggregated by type of commodity export, including export price shocks associated with annual crops, perennial crops, and extractive crops (Columns 3–4 and 7–8). These export price shock variables are all obtained from the data set of Bazzi and Blattman (2014), so the reader is referred to that work for additional details on these variables. Regardless of the identification strategy, population diversity remains a significant predictor of conflict onset.

Table A.13: Interpersonal Population Diversity and the Incidence or Onset of Civil Conflict in Repeated Cross-Country Data – Robustness to Accounting for Ecological/Climatic Covariates

Cross-country sample:	Old World		Global		Old World		Global	
	(1) Probit	(2) Probit	(3) IV Probit	(4) IV Probit	(5) Probit	(6) Probit	(7) IV Probit	(8) IV Probit
Panel A	Quinquennial PRIO25 civil conflict incidence, 1960–2008				Quinquennial WCM09 ethnic civil conflict incidence, 1960–2005			
Population diversity (ancestry adjusted)	16.129***	13.288***	16.486***	14.788***	26.834***	28.032***	24.652***	26.063***
	(4.965)	(5.039)	(4.610)	(4.976)	(6.896)	(7.416)	(6.314)	(6.724)
Ecological fractionalization	−0.782	−0.892	−0.717	−0.852*	−0.010	−0.153	−0.628	−0.530
	(0.513)	(0.558)	(0.475)	(0.496)	(0.712)	(0.833)	(0.758)	(0.800)
Ecological polarization	1.342***	1.289**	1.280***	1.354***	1.414**	1.395**	1.516***	1.514**
	(0.500)	(0.543)	(0.471)	(0.509)	(0.609)	(0.697)	(0.588)	(0.650)
Average annual temperature, lagged	0.092***	0.041	0.071***	0.037	0.097***	0.084**	0.122***	0.118***
	(0.024)	(0.030)	(0.022)	(0.025)	(0.033)	(0.042)	(0.034)	(0.044)
Average annual precipitation, lagged	0.189	0.060	0.183	0.117	0.434*	0.493*	0.627**	0.703**
	(0.183)	(0.188)	(0.181)	(0.182)	(0.250)	(0.272)	(0.288)	(0.281)
Volatility of annual temperature, lagged	−0.426	−0.382	−0.117	−0.100	−0.010	0.050	0.113	0.113
	(0.449)	(0.435)	(0.391)	(0.378)	(0.595)	(0.557)	(0.505)	(0.494)
Volatility of annual precipitation, lagged	−0.354	−0.519	0.565	0.457	−1.208	−1.404	−0.818	−0.766
	(0.970)	(1.025)	(0.688)	(0.736)	(1.105)	(1.196)	(1.254)	(1.392)
Controls for time dependence and geography	×	×	×	×	×	×	×	×
Controls for all other baseline covariates		×		×		×		×
Observations	944	944	1,154	1,154	927	927	1,039	1,039
Countries	119	119	141	141	117	117	129	129
Pseudo R^2	0.441	0.464			0.530	0.556		
Marginal effect of diversity	2.938***	2.327***	3.287***	2.841***	3.379***	3.352***	3.452***	3.512***
	(0.882)	(0.867)	(0.939)	(0.984)	(0.891)	(0.915)	(1.012)	(1.015)
Panel B	Annual PRIO25 civil conflict onset, 1960–2008				Annual WCM09 ethnic civil conflict onset, 1960–2005			
Population diversity (ancestry adjusted)	10.980***	11.454***	10.815***	11.352***	19.631***	21.546***	16.856***	18.505***
	(3.672)	(3.861)	(3.508)	(3.995)	(4.288)	(4.862)	(4.299)	(4.907)
Ecological fractionalization	−0.481	−0.607	−0.446	−0.499	−0.203	0.247	−0.521	−0.051
	(0.369)	(0.428)	(0.387)	(0.428)	(0.571)	(0.676)	(0.626)	(0.714)
Ecological polarization	1.078***	1.079**	1.020***	1.057**	1.177**	1.130*	1.202**	1.247**
	(0.385)	(0.435)	(0.396)	(0.423)	(0.537)	(0.579)	(0.553)	(0.581)
Annual temperature, lagged	0.025	0.013	0.016	0.004	0.088***	0.118***	0.107***	0.139***
	(0.017)	(0.020)	(0.017)	(0.020)	(0.026)	(0.033)	(0.026)	(0.036)
Annual precipitation, lagged	0.078	0.042	0.029	−0.016	0.242	0.247	0.361*	0.383**
	(0.101)	(0.119)	(0.116)	(0.119)	(0.163)	(0.168)	(0.194)	(0.193)
Volatility of annual temperature, last 5 years	0.227	0.154	0.223	0.174	0.217	0.108	0.254	0.163
	(0.336)	(0.343)	(0.304)	(0.309)	(0.378)	(0.403)	(0.353)	(0.381)
Volatility of annual precipitation, last 5 years	−0.378	−0.316	−0.257	−0.305	1.007	1.460*	0.301	0.632
	(0.722)	(0.772)	(0.759)	(0.759)	(0.850)	(0.883)	(0.907)	(0.939)
Controls for time dependence and geography	×	×	×	×	×	×	×	×
Controls for all other baseline covariates		×		×		×		×
Observations	3,849	3,828	4,896	4,874	3,607	3,585	4,038	4,016
Countries	119	119	141	141	117	117	129	129
Pseudo R^2	0.128	0.147			0.169	0.201		
Marginal effect of diversity	0.545***	0.561***	0.530***	0.544**	0.899***	0.962***	0.886***	1.044***
	(0.188)	(0.199)	(0.203)	(0.225)	(0.227)	(0.247)	(0.283)	(0.332)

Notes: This table conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on either the quinquennial incidence or the annual onset of civil conflict in repeated cross-country data, as shown in Panels A and B of Table 4. Specifically, it establishes robustness to *additionally* accounting for the potentially confounding influence of (i) time-invariant fractionalization and polarization measures of the ecological diversity of land (e.g., Fenske, 2014); and (ii) climatic experience in the recent past (e.g., Burke, Hsiang and Miguel, 2015), as captured by either (a) the temporal mean and volatility of annual temperature and annual precipitation over the previous 5-year interval for the quinquennial incidence regressions; or (b) the lagged values of annual temperature and annual precipitation as well as their temporal volatility over the previous 5 years for the annual onset regressions. The specifications examined in this table are otherwise identical to those reported in Table 4. The reader is therefore referred to Table 4 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis, the identification strategy employed by the IV probit regressions, and the estimation and interpretation of the marginal effect of population diversity on the incidence or onset of conflict. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.14: Interpersonal Population Diversity and the Incidence or Onset of Civil Conflict in Repeated Cross-Country Data – Robustness to Accounting for Deep-Rooted Determinants of Economic Development

Cross-country sample:	Old World		Global		Old World		Global	
	(1) Probit	(2) Probit	(3) IV Probit	(4) IV Probit	(5) Probit	(6) Probit	(7) IV Probit	(8) IV Probit
Panel A	Quinquennial PRIO25 civil conflict incidence, 1960–2008				Quinquennial WCM09 ethnic civil conflict incidence, 1960–2005			
Population diversity (ancestry adjusted)	14.501***	14.580***	17.247***	17.285***	23.152***	25.322***	18.346***	19.675***
	(5.418)	(5.294)	(5.644)	(5.961)	(6.484)	(7.125)	(7.017)	(7.434)
Log years since Neolithic Revolution	−0.038	−0.135	−0.331	−0.358	−0.131	−0.283	−0.233	−0.355
	(0.315)	(0.282)	(0.304)	(0.292)	(0.300)	(0.318)	(0.367)	(0.399)
Log index of state antiquity	0.169	0.178	0.247**	0.257**	0.101	0.076	0.174	0.177
	(0.107)	(0.109)	(0.115)	(0.117)	(0.124)	(0.159)	(0.122)	(0.146)
Log duration of human settlement	−0.043	0.024	0.023	0.049	−0.093	0.033	0.006	0.085
	(0.117)	(0.097)	(0.104)	(0.097)	(0.148)	(0.138)	(0.120)	(0.112)
Log distance from regional frontier in 1500	−0.040	−0.058	−0.065	−0.067	−0.060	−0.093	−0.100	−0.120*
	(0.061)	(0.049)	(0.045)	(0.045)	(0.073)	(0.067)	(0.061)	(0.065)
Controls for time dependence and geography	×	×	×	×	×	×	×	×
Controls for all other baseline covariates		×		×		×		×
Observations	873	873	1,083	1,083	873	873	985	985
Countries	109	109	131	131	109	109	121	121
Pseudo R^2	0.430	0.460			0.517	0.548		
Marginal effect of diversity	2.767***	2.646***	3.551***	3.429***	3.080***	3.160***	2.650**	2.715***
	(0.998)	(0.937)	(1.163)	(1.204)	(0.874)	(0.915)	(1.031)	(1.046)
Panel B	Annual PRIO25 civil conflict onset, 1960–2008				Annual WCM09 ethnic civil conflict onset, 1960–2005			
Population diversity (ancestry adjusted)	9.618***	11.359***	14.998***	16.797***	13.281***	14.997***	11.265**	12.234*
	(3.440)	(3.369)	(4.300)	(4.579)	(4.887)	(5.405)	(5.395)	(6.293)
Log years since Neolithic Revolution	0.100	0.061	−0.380	−0.381	−0.081	−0.082	−0.234	−0.229
	(0.202)	(0.204)	(0.263)	(0.263)	(0.261)	(0.289)	(0.327)	(0.409)
Log index of state antiquity	0.075	0.096*	0.167**	0.187**	0.011	0.028	0.087	0.109
	(0.049)	(0.054)	(0.070)	(0.075)	(0.099)	(0.107)	(0.105)	(0.109)
Log duration of human settlement	0.054	0.080*	0.095*	0.107*	0.140	0.152	0.207*	0.216**
	(0.042)	(0.047)	(0.053)	(0.054)	(0.117)	(0.105)	(0.114)	(0.091)
Log distance from regional frontier in 1500	0.037	0.019	−0.038	−0.047	0.042	−0.001	−0.037	−0.082
	(0.042)	(0.046)	(0.032)	(0.036)	(0.071)	(0.078)	(0.065)	(0.073)
Controls for time dependence and geography	×	×	×	×	×	×	×	×
Controls for all other baseline covariates		×		×		×		×
Observations	3,468	3,449	4,482	4,462	3,390	3,369	3,821	3,800
Countries	109	109	131	131	109	109	121	121
Pseudo R^2	0.126	0.145			0.149	0.183		
Marginal effect of diversity	0.512***	0.597***	0.816***	0.906***	0.646***	0.710***	0.656*	0.788*
	(0.179)	(0.175)	(0.291)	(0.314)	(0.250)	(0.260)	(0.352)	(0.452)

Notes: This table conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on either the quinquennial incidence or the annual onset of civil conflict in repeated cross-country data, as shown in Panels A and B of Table 4. Specifically, it establishes robustness to *additionally* accounting for the potentially confounding influence of other deep-rooted determinants of comparative economic development, including (i) the time elapsed since the onset of the Neolithic Revolution (e.g., Ashraf and Galor, 2013a); (ii) an index of experience with institutionalized statehood since antiquity (e.g., Bockstette, Chanda and Putterman, 2002); (iii) the time elapsed since initial human settlement in prehistory (e.g., Ahlerup and Olsson, 2012); and (iv) the great-circle distance to the closest regional technological frontier in the year 1500 (e.g., Ashraf and Galor, 2013a). The specifications examined in this table are otherwise identical to those reported in Table 4. The reader is therefore referred to Table 4 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the IV probit regressions. The estimated marginal effect of a 1 percentage point increase in population diversity is the average marginal effect across the entire cross-section of observed diversity values, and it reflects the increase in either the quinquennial likelihood of a conflict incidence (Panel A) or the annual likelihood of a conflict onset (Panel B), both expressed in percentage points. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.15: Interpersonal Population Diversity and the Incidence or Onset of Civil Conflict in Repeated Cross-Country Data – Robustness to Accounting for Ethnic and Spatial Inequality

Cross-country sample:	Old World		Global		Old World		Global	
	(1) Probit	(2) Probit	(3) IV Probit	(4) IV Probit	(5) Probit	(6) Probit	(7) IV Probit	(8) IV Probit
Panel A	Quinquennial PRIO25 civil conflict incidence, 1960–2008				Quinquennial WCM09 ethnic civil conflict incidence, 1960–2005			
Population diversity (ancestry adjusted)	12.345***	12.058**	14.675***	14.585***	20.178***	22.469***	15.274***	16.416***
	(4.399)	(4.725)	(4.346)	(4.786)	(5.776)	(6.318)	(5.270)	(5.752)
Ethnic inequality in luminosity	0.112	0.179	0.176	0.198	0.264	1.102***	0.527	1.169***
	(0.341)	(0.359)	(0.342)	(0.350)	(0.426)	(0.426)	(0.407)	(0.435)
Spatial inequality in luminosity	0.215	0.227	0.504	0.626	0.024	−0.040	−0.122	−0.030
	(0.417)	(0.403)	(0.409)	(0.412)	(0.506)	(0.537)	(0.479)	(0.550)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions		×		×		×		×
Controls for oil, population, and income		×		×		×		×
Controls for ethnic diversity		×		×		×		×
Observations	916	916	1,126	1,126	899	899	1,011	1,011
Countries	116	116	138	138	114	114	126	126
Pseudo R^2	0.414	0.452			0.505	0.551		
Marginal effect of diversity	2.338***	2.147***	3.000***	2.849***	2.604***	2.668***	2.182***	2.241***
	(0.797)	(0.827)	(0.891)	(0.964)	(0.742)	(0.762)	(0.745)	(0.775)
Panel B	Annual PRIO25 civil conflict onset, 1960–2008				Annual WCM09 ethnic civil conflict onset, 1960–2005			
Population diversity (ancestry adjusted)	8.694***	10.032***	9.927***	11.427***	14.756***	15.265***	10.697***	10.816***
	(3.177)	(3.530)	(3.090)	(3.644)	(3.535)	(3.691)	(3.558)	(4.097)
Ethnic inequality in luminosity	0.028	0.205	0.145	0.213	1.224***	1.579***	1.272***	1.591***
	(0.240)	(0.266)	(0.246)	(0.258)	(0.371)	(0.388)	(0.377)	(0.419)
Spatial inequality in luminosity	−0.031	−0.228	−0.027	−0.077	−0.315	−0.550	−0.437	−0.575
	(0.272)	(0.322)	(0.292)	(0.328)	(0.345)	(0.458)	(0.352)	(0.477)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Peace duration cubic splines	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions		×		×		×		×
Controls for oil, population, and income		×		×		×		×
Controls for ethnic diversity		×		×		×		×
Observations	3,641	3,620	4,773	4,751	3,360	3,339	3,776	3,755
Countries	116	116	138	138	114	114	126	126
Pseudo R^2	0.126	0.149			0.158	0.197		
Marginal effect of diversity	0.437***	0.495***	0.478***	0.543***	0.672***	0.671***	0.658***	0.750**
	(0.165)	(0.183)	(0.174)	(0.205)	(0.176)	(0.173)	(0.223)	(0.295)

Notes: This table conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on either the quinquennial incidence or the annual onset of civil conflict in repeated cross-country data, as shown in Panels A and B of Table 4. Specifically, it establishes robustness to *additionally* accounting for the potentially confounding influence of measures of intrastate economic inequality (e.g., Alesina, Michalopoulos and Papaioannou, 2016), as captured by the subnational spatial distribution of per-capita adjusted nighttime luminosity in the year 2000 across either (i) the georeferenced homelands of ethnic groups; or (ii) 2.5×2.5-degree geospatial grid cells. The specifications examined in this table are otherwise identical to those reported in Table 4. The reader is therefore referred to Table 4 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the IV probit regressions. The estimated marginal effect of a 1 percentage point increase in diversity is the average marginal effect across the entire cross-section of observed diversity values, and it reflects the increase in either the quinquennial likelihood of a conflict incidence (Panel A) or the annual likelihood of a conflict onset (Panel B), both expressed in percentage points. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.16: Interpersonal Population Diversity and the Incidence of Civil Conflict in Repeated Cross-Country Data – Robustness to Accounting for Alternative Correlates of Conflict Incidence

Cross-country sample:	Old World					Global				
	(1) Probit	(2) Probit	(3) Probit	(4) Probit	(5) Probit	(6) IV Probit	(7) IV Probit	(8) IV Probit	(9) IV Probit	(10) IV Probit
Panel A	Quinquennial PRIO25 civil conflict incidence, 1960–2008									
Population diversity (ancestry adjusted)	12.406***	12.674***	13.284***	10.813**	12.161**	12.501***	13.033***	13.008***	11.867**	14.753***
	(4.658)	(4.670)	(4.818)	(5.446)	(4.925)	(4.642)	(4.861)	(4.717)	(5.533)	(5.644)
Ethnic fractionalization (Fearon, 2003)	0.006					0.018				
	(0.306)					(0.292)				
Linguistic fractionalization (Alesina et al., 2003)		0.259					0.198			
		(0.307)					(0.278)			
Religious fractionalization (Alesina et al., 2003)			−0.551**					−0.646**		
			(0.276)					(0.262)		
Ethnolinguistic fractionalization (Esteban et al., 2012)				0.287					0.259	
				(0.367)					(0.345)	
Ethnolinguistic polarization (Esteban et al., 2012)				2.358					3.504**	
				(1.605)					(1.485)	
Gini index of ethnolinguistic diversity (Esteban et al., 2012)				−1.479					−1.440	
				(1.448)					(1.340)	
Log percentage mountainous terrain					0.124*					0.089
					(0.072)					(0.074)
Noncontiguous state dummy					0.367*					0.559***
					(0.217)					(0.199)
Disease richness					−0.007					−0.003
					(0.011)					(0.009)
Controls for all baseline covariates	×	×	×	×	×	×	×	×	×	×
Observations	938	935	944	883	928	1,148	1,117	1,154	1,083	1,138
Countries	118	118	119	107	117	140	137	141	128	139
Pseudo R^2	0.448	0.456	0.458	0.460	0.460					
Marginal effect of diversity	2.270***	2.265***	2.376***	1.941**	2.157**	2.478***	2.509***	2.531***	2.317**	2.859**
	(0.839)	(0.817)	(0.846)	(0.974)	(0.860)	(0.932)	(0.943)	(0.927)	(1.094)	(1.120)
Panel B	Quinquennial WCM09 ethnic civil conflict incidence, 1960–2005									
Population diversity (ancestry adjusted)	20.307***	23.605***	20.602***	18.694***	24.620***	14.105**	19.464***	14.843***	13.137**	18.033***
	(6.022)	(6.183)	(5.959)	(6.535)	(6.384)	(5.836)	(5.939)	(5.718)	(6.689)	(6.430)
Ethnic fractionalization (Fearon, 2003)	0.669*					0.536				
	(0.354)					(0.346)				
Linguistic fractionalization (Alesina et al., 2003)		0.892**					0.707*			
		(0.412)					(0.374)			
Religious fractionalization (Alesina et al., 2003)			−0.007					−0.417		
			(0.374)					(0.399)		
Ethnolinguistic fractionalization (Esteban et al., 2012)				0.632					0.353	
				(0.461)					(0.415)	
Ethnolinguistic polarization (Esteban et al., 2012)				−0.087					5.397**	
				(1.864)					(2.486)	
Gini index of ethnolinguistic diversity (Esteban et al., 2012)				−4.652					−3.212	
				(3.269)					(2.864)	
Log percentage mountainous terrain					−0.000					0.021
					(0.095)					(0.087)
Noncontiguous state dummy					0.336					0.157
					(0.233)					(0.252)
Disease richness					−0.003					0.022
					(0.015)					(0.017)
Controls for all baseline covariates	×	×	×	×	×	×	×	×	×	×
Observations	936	918	927	882	927	1,048	1,002	1,039	984	1,039
Countries	118	116	117	107	117	130	125	129	118	129
Pseudo R^2	0.548	0.557	0.550	0.551	0.550					
Marginal effect of diversity	2.450***	2.775***	2.476***	2.295***	2.961***	1.858**	2.443***	1.968***	1.727*	2.423***
	(0.732)	(0.732)	(0.727)	(0.803)	(0.789)	(0.758)	(0.750)	(0.755)	(0.910)	(0.873)

Notes: This table conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on the quinquennial incidence of civil conflict in repeated cross-country data, as shown in Panel A of Table 4. Specifically, it establishes robustness to accounting for the potentially confounding influence of *alternative* distributional indices of intergroup diversity (e.g., Fearon, 2003; Alesina et al., 2003; Esteban, Mayoral and Ray, 2012) and *additional* geographical correlates of conflict (e.g., Fearon and Laitin, 2003; Cervellati, Sunde and Valmori, 2017). The specifications examined in this table are identical to the fully specified baseline models reported in Panel A of Table 4, with the exception that in Columns 1–3 and 6–8 of the current analysis, each of the reported control variables is employed in lieu of the baseline control for ethnic fractionalization (Alesina et al., 2003), whereas in Columns 4 and 9, the set of reported control variables replaces the baseline controls for both ethnic fractionalization and ethnolinguistic polarization (Desmet, Ortuño-Ortín and Wacziarg, 2012), in the interest of mitigating multicollinearity. Further, in Columns 5 and 10 of the current analysis, the set of reported geographical controls augment the fully specified baseline models from Panel A of Table 4. The reader is therefore referred to Table 4 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the IV probit regressions. The estimated marginal effect of a 1 percentage point increase in population diversity is the average marginal effect across the entire cross-section of observed diversity values, and it reflects the increase in the quinquennial likelihood of a conflict incidence, expressed in percentage points. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.17: Interpersonal Population Diversity and the Incidence of Civil Conflict in Repeated Cross-Country Data – Robustness to Employing Alternative Measures of Conflict Incidence

Cross-country sample:	Old World		Global		Old World		Global		Old World		Global	
	(1) Probit	(2) Probit	(3) IV Probit	(4) IV Probit	(5) Probit	(6) Probit	(7) IV Probit	(8) IV Probit	(9) Probit	(10) Probit	(11) IV Probit	(12) IV Probit
	Quinquennial PRIO1000 civil war incidence, 1960–2008				Quinquennial WM0609 ethnic civil war incidence, 1960–2001				Quinquennial UCDP nonstate conflict incidence, 1990–2008			
Population diversity (ancestry adjusted)	14.799***	14.024***	16.021***	16.725***	27.920***	33.714***	22.611***	25.920***	26.809***	35.278***	24.973***	30.326***
	(4.839)	(5.141)	(4.587)	(5.509)	(7.716)	(8.440)	(6.564)	(7.576)	(6.592)	(7.867)	(5.391)	(6.180)
Lagged conflict incidence	×	×	×	×	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×	×	×	×	×
Controls for institutions		×		×		×		×		×		×
Controls for oil, population, and income		×		×		×		×		×		×
Controls for ethnic diversity		×		×		×		×		×		×
Observations	944	928	1,154	1,138	764	764	860	860	447	447	535	535
Countries	119	119	141	141	114	114	126	126	118	118	140	140
Pseudo R^2	0.387	0.408			0.456	0.472			0.432	0.488		
Marginal effect of diversity	1.730***	1.615***	1.906***	1.952**	1.830***	2.147***	1.664***	1.869***	4.046***	4.796***	4.233***	4.893***
	(0.602)	(0.627)	(0.650)	(0.759)	(0.588)	(0.671)	(0.587)	(0.652)	(0.959)	(1.071)	(0.958)	(1.159)

Notes: This table conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on the quinquennial incidence of intrastate conflict in repeated cross-country data, as shown in Panel A of Table 4. Specifically, it establishes robustness to considering the temporal incidence of alternative forms of intrastate conflict as the outcome variable, including the incidence of (i) a high-intensity PRIO1000 civil war in any given 5-year interval during the 1960–2008 time period (Columns 1–4); (ii) a high-intensity WM0609 ethnic civil war in any given 5-year interval during the 1960–2001 time period (Columns 5–8); and (iii) a low-intensity conflict involving nonstate actors in any given 5-year interval during the 1990–2008 time period (Columns 9–12). The specifications examined in this table are otherwise identical to those reported in Panel A of Table 4. The reader is therefore referred to Table 4 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the IV probit regressions. The estimated marginal effect of a 1 percentage point increase in population diversity is the average marginal effect across the entire cross-section of observed diversity values, and it reflects the increase in the quinquennial likelihood of a conflict incidence, expressed in percentage points. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.18: Interpersonal Population Diversity and the Incidence of Civil Conflict in *Annually* Repeated Cross-Country Data

Cross-country sample:	Old World		Global		Old World		Global	
	(1) Probit	(2) Probit	(3) IV Probit	(4) IV Probit	(5) Probit	(6) Probit	(7) IV Probit	(8) IV Probit
	Annual PRIO25 civil conflict incidence, 1960–2008				Annual WCM09 ethnic civil conflict incidence, 1960–2005			
Population diversity (ancestry adjusted)	9.439***	10.033***	10.631***	12.154***	12.916***	12.265***	9.379**	8.196*
	(3.627)	(3.609)	(3.521)	(3.974)	(3.958)	(4.137)	(3.662)	(4.361)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions		×		×		×		×
Controls for oil, population, and income		×		×		×		×
Controls for ethnic diversity		×		×		×		×
Observations	4,756	4,756	5,797	5,797	4,331	4,316	4,855	4,840
Countries	119	119	141	141	117	117	129	129
Pseudo R^2	0.617	0.627			0.785	0.792		
Marginal effect of diversity	0.885**	0.921***	1.045***	1.177***	0.592***	0.549***	0.541**	0.484*
	(0.346)	(0.343)	(0.381)	(0.434)	(0.184)	(0.188)	(0.219)	(0.262)

Notes: This table conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on the temporal incidence of civil conflict in repeated cross-country data, as shown in Panel A of Table 4. Specifically, it establishes robustness to considering *annual* rather than quinquennial repetitions of the cross-section, thereby demonstrating the significant positive reduced-form impact of population diversity on the annual likelihood of a conflict incidence during the relevant post-1960 time period. The specifications examined in this table are essentially identical to those reported in Panel A of Table 4, with the exception that the time-dependent baseline controls for institutions (i.e., executive constraints, indicators for the type of political regime, and indicators for colonial experience by identity of the colonizing power), oil production per capita, total population, GDP per capita, and conflict persistence are all appropriately adjusted to assume their respective lagged annual values rather than the values corresponding to the previous 5-year interval. The reader is therefore referred to Table 4 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the IV probit regressions. The estimated marginal effect of a 1 percentage point increase in population diversity is the average marginal effect across the entire cross-section of observed diversity values, and it reflects the increase in the annual likelihood of a conflict incidence, expressed in percentage points. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.19: Interpersonal Population Diversity and the Incidence or Onset of Civil Conflict in Repeated Cross-Country Data – Robustness to Employing the Logit and Rare-Events Logit Estimators

	(1) Logit	(2) RELogit	(3) Logit	(4) RELogit	(5) Logit	(6) RELogit	(7) Logit	(8) RELogit
Panel A	Quinquennial PRIO25 civil conflict incidence, 1960–2008				Quinquennial WCM09 ethnic civil conflict incidence, 1960–2005			
Population diversity (ancestry adjusted)	22.264***	21.485***	21.464***	20.122**	38.213***	36.234***	40.717***	36.533***
	(8.137)	(7.952)	(8.332)	(8.034)	(11.150)	(10.891)	(11.820)	(11.390)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions			×	×			×	×
Controls for oil, population, and income			×	×			×	×
Controls for ethnic diversity			×	×			×	×
Observations	944	944	944	944	927	927	927	927
Countries	119	119	119	119	117	117	117	117
Pseudo R^2	0.423		0.457		0.514		0.546	
Marginal effect of diversity	3.319***	3.586**	2.909**	3.149**	2.668***	3.297***	2.343***	3.057***
	(1.150)	(1.397)	(1.149)	(1.340)	(0.699)	(1.047)	(0.669)	(1.075)
Panel B	Annual PRIO25 civil conflict onset, 1960–2008				Annual WCM09 ethnic civil conflict onset, 1960–2005			
Population diversity (ancestry adjusted)	21.501***	20.458***	23.444***	21.423***	29.656***	28.034***	33.499***	30.012***
	(7.296)	(7.205)	(8.090)	(8.000)	(8.505)	(8.258)	(9.138)	(9.040)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Peace duration cubic splines	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions			×	×			×	×
Controls for oil, population, and income			×	×			×	×
Controls for ethnic diversity			×	×			×	×
Observations	3,849	4,756	3,828	4,756	3,607	4,331	3,585	4,331
Countries	119	119	119	119	117	117	117	117
Pseudo R^2	0.120		0.142		0.146		0.182	
Marginal effect of diversity	0.289***	0.255**	0.283***	0.259**	0.315***	0.375**	0.279***	0.369***
	(0.095)	(0.116)	(0.097)	(0.123)	(0.094)	(0.146)	(0.078)	(0.141)

Notes: This table conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on either the quinquennial incidence or the annual onset of civil conflict in repeated cross-sectional data for the Old World sample of countries, as shown in Columns 1–2 and 5–6 of Panels A and B of Table 4. Specifically, it establishes robustness to employing the ordinary logit and rare-events logit (King and Zeng, 2001) estimators, rather than the probit estimator, for estimating the relevant empirical models of conflict incidence and onset. The specifications examined in this table are otherwise identical to those reported in Columns 1–2 and 5–6 of Table 4. The reader is therefore referred to Table 4 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis. Given the absence of readily available ordinary logit and rare-events logit estimators that permit instrumentation, the current analysis is unable to implement the global-sample identification strategy of exploiting prehistoric migratory distance from East Africa to the indigenous (precolonial) population of a country as an excluded instrument for the country's contemporary population diversity. The estimated marginal effect of a 1 percentage point increase in population diversity is the marginal effect at the mean value of diversity in the cross-section, and it reflects the increase in either the quinquennial likelihood of a conflict incidence (Panel A) or the annual likelihood of a conflict onset (Panel B), both expressed in percentage points. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.20: Interpersonal Population Diversity and the Incidence or Onset of Civil Conflict in Repeated Cross-Country Data – Robustness to Accounting for Spatiotemporal Dependence using Two-Way Clustering of Standard Errors

	(1) Probit	(2) Logit	(3) Probit	(4) Logit	(5) Probit	(6) Logit	(7) Probit	(8) Logit
Panel A	Quinquennial PRIO25 civil conflict incidence, 1960–2008				Quinquennial WCM09 ethnic civil conflict incidence, 1960–2005			
Population diversity (ancestry adjusted)	11.883***	22.264***	12.043***	21.464***	20.171***	38.213***	21.488***	40.717***
	(3.748)	(6.716)	(3.965)	(6.122)	(6.427)	(11.766)	(6.217)	(11.809)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions			×	×			×	×
Controls for oil, population, and income			×	×			×	×
Controls for ethnic diversity			×	×			×	×
Observations	944	944	944	944	927	927	927	927
Countries	119	119	119	119	117	117	117	117
Pseudo R^2	0.423	0.423	0.457	0.457	0.516	0.514	0.549	0.546
Panel B	Annual PRIO25 civil conflict onset, 1960–2008				Annual WCM09 ethnic civil conflict onset, 1960–2005			
Population diversity (ancestry adjusted)	9.088**	21.501**	10.302**	23.444**	13.251***	29.656***	14.710***	33.499***
	(3.812)	(8.536)	(4.230)	(9.904)	(3.784)	(8.489)	(4.165)	(9.349)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Peace duration cubic splines	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions			×	×			×	×
Controls for oil, population, and income			×	×			×	×
Controls for ethnic diversity			×	×			×	×
Observations	3,849	3,849	3,828	3,828	3,607	3,607	3,585	3,585
Countries	119	119	119	119	117	117	117	117
Pseudo R^2	0.120	0.120	0.141	0.142	0.146	0.146	0.180	0.182

Notes: This table conducts a robustness check on the results from the baseline probit and logit analyses of the reduced-form impact of contemporary population diversity on either the quinquennial incidence or the annual onset of civil conflict in repeated cross-sectional data for the Old World sample of countries, as shown in Columns 1–2 and 5–6 of Table 4 and in odd-numbered columns of Table A.19. Specifically, it establishes robustness of the standard-error estimates to accounting for spatiotemporal dependence across country-time observations by implementing multi-dimensional clustering of standard errors, following the methodology of Cameron, Gelbach and Miller (2011). To implement this robustness check, the standard errors across country-time observations are clustered in two dimensions: (i) the country level, which allows for temporal dependence within a country over time (i.e., across 5-year intervals or years); and (ii) the time level, which allows for spatial dependence across countries within a given time period (i.e., within a 5-year interval or year). The specifications examined in this table are otherwise identical to those reported in Columns 1–2 and 5–6 of Table 4. The reader is therefore referred to Table 4 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis. Given the absence of readily available probit and logit estimators that not only allow for multi-dimensional clustering of standard errors but also permit instrumentation, the current analysis is unable to implement the global-sample identification strategy of exploiting prehistoric migratory distance from East Africa to the indigenous (precolonial) population of a country as an excluded instrument for the country's contemporary population diversity. Heteroskedasticity robust standard errors, clustered multi-dimensionally at the country and time levels, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.21: Interpersonal Population Diversity and the Onset of Civil Conflict in Repeated Cross-Country Data – Robustness to Accounting for Alternative Correlates of Conflict Onset

Cross-country sample:	Old World		Global		Old World		Global	
	(1) Probit	(2) Probit	(3) IV Probit	(4) IV Probit	(5) Probit	(6) Probit	(7) IV Probit	(8) IV Probit
	Annual PRIO25 civil conflict onset, 1960–1999				Annual WCM09 ethnic civil conflict onset, 1960–1999			
Population diversity (ancestry adjusted)	8.284**	7.822**	10.520**	9.836**	13.203***	15.207***	10.321**	11.447**
	(3.873)	(3.925)	(4.147)	(4.217)	(4.192)	(5.000)	(4.375)	(4.880)
Ethnic dominance		0.091		0.108		−0.205		0.013
		(0.134)		(0.120)		(0.144)		(0.158)
Political instability, lagged		0.203		0.139		0.254*		0.122
		(0.127)		(0.121)		(0.146)		(0.138)
New state dummy, lagged		0.196		−0.136		1.182**		0.962**
		(0.538)		(0.524)		(0.510)		(0.477)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Peace duration cubic splines	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions	×	×	×	×	×	×	×	×
Controls for oil, population, and income	×	×	×	×	×	×	×	×
Controls for ethnic diversity	×	×	×	×	×	×	×	×
Observations	2,561	2,561	3,368	3,368	2,676	2,676	3,007	3,007
Countries	95	95	116	116	95	95	106	106
Pseudo R^2	0.141	0.144			0.185	0.197		
Marginal effect of diversity	0.491**	0.462*	0.597**	0.551**	0.643***	0.730***	0.681**	0.743**
	(0.239)	(0.240)	(0.274)	(0.271)	(0.221)	(0.262)	(0.310)	(0.333)

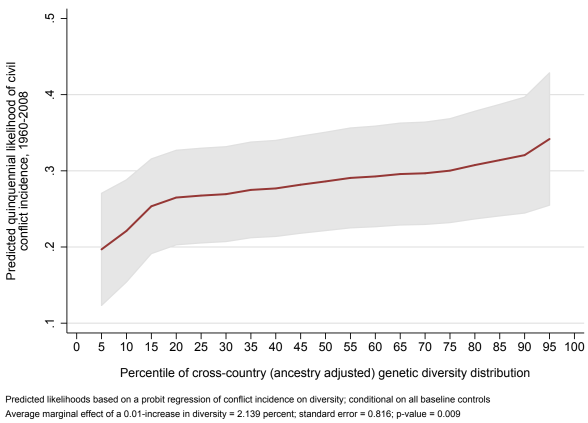
Notes: This table conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on the annual onset of civil conflict in repeated cross-country data, as shown in Panel B of Table 4. Specifically, it establishes robustness to accounting for the potentially confounding influence of an *additional* distributional index of intergroup diversity (e.g., Collier and Hoeffler, 2004) and *additional* time-varying institutional correlates of conflict (e.g., Fearon and Laitin, 2003). In light of sample restrictions imposed by the availability of data on these additional control variables, the specification presented in each odd-numbered column of the table is intended to provide a relevant baseline for the robustness check in the subsequent even-numbered column (i.e., by holding fixed the regression sample). The specifications examined in this table are otherwise identical to the fully specified baseline models reported in Panel B of Table 4. The reader is therefore referred to Table 4 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the IV probit regressions. The estimated marginal effect of a 1 percentage point increase in population diversity is the average marginal effect across the entire cross-section of observed diversity values, and it reflects the increase in the annual likelihood of a conflict onset, expressed in percentage points. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.22: Interpersonal Population Diversity and the Onset of Civil Conflict in Repeated Cross-Country Data – Robustness to Accounting for Commodity Export Price Shocks

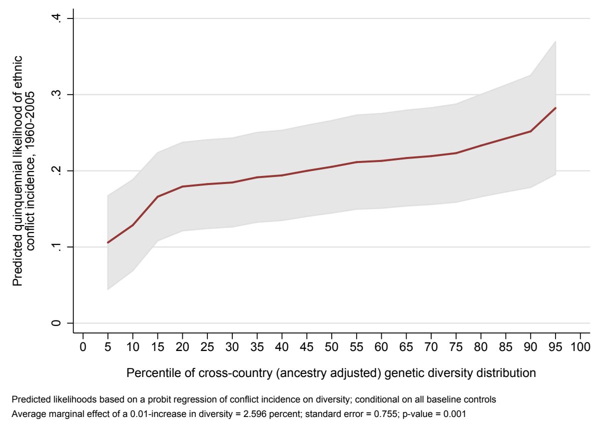
Cross-country sample:	Old World				Global			
	(1) Probit	(2) Probit	(3) Probit	(4) Probit	(5) IV Probit	(6) IV Probit	(7) IV Probit	(8) IV Probit
	Annual PRIO25 civil conflict onset, 1960–2007							
Population diversity (ancestry adjusted)	6.173*	7.121**	6.153*	6.895**	7.063**	8.453**	7.079**	8.355**
	(3.337)	(3.358)	(3.271)	(3.385)	(2.874)	(3.638)	(2.852)	(3.692)
Aggregate price shock	−0.152***	−0.161***			−0.191***	−0.202***		
	(0.057)	(0.062)			(0.053)	(0.057)		
Aggregate price shock, lagged	0.032	0.039			0.047	0.053		
	(0.062)	(0.070)			(0.055)	(0.061)		
Aggregate price shock, twice lagged	−0.161***	−0.181***			−0.086	−0.089		
	(0.062)	(0.067)			(0.065)	(0.070)		
Annual crop price shock			−0.185**	−0.204**			−0.224***	−0.243***
			(0.081)	(0.087)			(0.070)	(0.075)
Annual crop price shock, lagged			−0.038	−0.029			−0.011	−0.003
			(0.086)	(0.096)			(0.081)	(0.089)
Annual crop price shock, twice lagged			−0.155*	−0.180*			−0.059	−0.066
			(0.091)	(0.097)			(0.094)	(0.102)
Perennial crop price shock			−0.150**	−0.143*			−0.162***	−0.160***
			(0.071)	(0.073)			(0.058)	(0.060)
Perennial crop price shock, lagged			0.129**	0.133**			0.127***	0.124**
			(0.052)	(0.056)			(0.046)	(0.049)
Perennial crop price shock, twice lagged			−0.131***	−0.146***			−0.063	−0.061
			(0.050)	(0.054)			(0.049)	(0.052)
Extractive crop price shock			−0.236***	−0.249***			−0.273***	−0.290***
			(0.089)	(0.096)			(0.082)	(0.087)
Extractive crop price shock, lagged			0.068	0.080			0.082	0.092
			(0.091)	(0.099)			(0.084)	(0.090)
Extractive crop price shock, twice lagged			−0.311***	−0.329***			−0.198*	−0.198*
			(0.107)	(0.113)			(0.109)	(0.115)
Lagged conflict incidence	×	×	×	×	×	×	×	×
Peace duration cubic splines	×	×	×	×	×	×	×	×
Continent dummies	×	×	×	×	×	×	×	×
Time dummies	×	×	×	×	×	×	×	×
Controls for geography	×	×	×	×	×	×	×	×
Controls for institutions		×		×		×		×
Controls for ethnic diversity		×		×		×		×
Observations	2,574	2,553	2,574	2,553	3,527	3,505	3,527	3,505
Countries	81	81	81	81	103	103	103	103
Pseudo R^2	0.115	0.147	0.128	0.159				
Marginal effect of diversity	0.381*	0.428**	0.375*	0.410*	0.387**	0.446**	0.383**	0.437**
	(0.214)	(0.211)	(0.209)	(0.210)	(0.176)	(0.212)	(0.174)	(0.213)

Notes: This table conducts a robustness check on the results from the baseline analysis of the reduced-form impact of contemporary population diversity on the annual onset of a PRIO25 civil conflict in repeated cross-country data, as shown in Columns 1–4 of Panel B of Table 4. Specifically, it establishes robustness to *additionally* accounting for the potentially confounding "income effect" of commodity export price shocks (e.g., Bazzi and Blattman, 2014), as captured by the contemporaneous, lagged, and twice lagged values of either an annual price shock that has been aggregated across commodity export types (Columns 1–2 and 5–6) or annual price shocks disaggregated by type of commodity export, including export price shocks associated with annual crops, perennial crops, and extractive crops (Columns 3–4 and 7–8). These export price shock variables are all obtained from the data set of Bazzi and Blattman (2014), so the reader is referred to that work for additional details on these variables. The specifications examined in this table are otherwise identical to those reported in Columns 1–4 of Panel B of Table 4, with the exception that the fully specified models in the current analysis omit the controls for oil production per capita, total population, and GDP per capita, in the interest of minimizing endogeneity with the export price shock variables and maximizing degrees of freedom. The reader is therefore referred to Table 4 and the corresponding table notes for additional details on the baseline set of covariates considered by the current analysis as well as the identification strategy employed by the IV probit regressions. The estimated marginal effect of a 1 percentage point increase in population diversity is the average marginal effect across the entire cross-section of observed diversity values, and it reflects the increase in the annual likelihood of a conflict onset, expressed in percentage points. Heteroskedasticity robust standard errors, clustered at the country level, are reported in parentheses. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

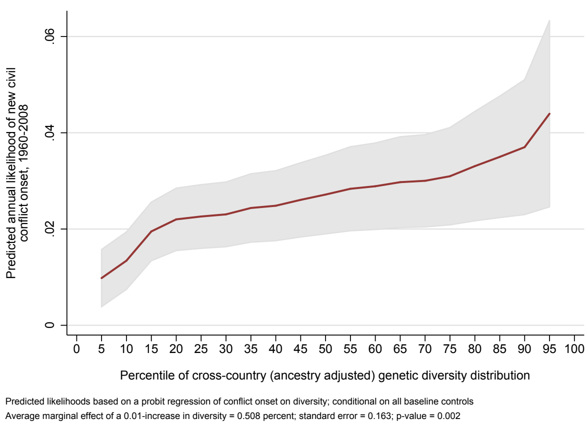
A.3 Supplementary Figures



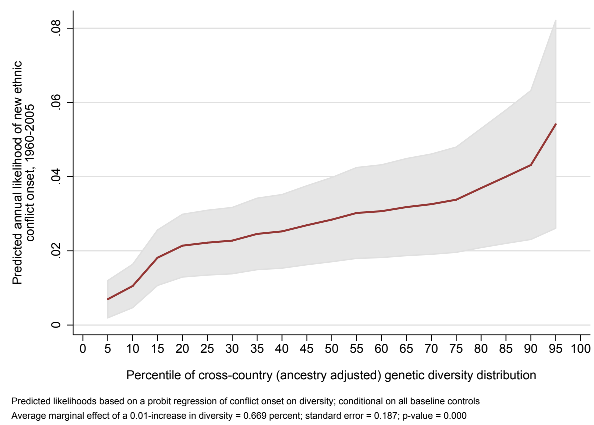
(a) Civil conflict incidence



(b) Ethnic civil conflict incidence



(c) Civil conflict onset



(d) Ethnic civil conflict onset

Figure A.1: Interpersonal Population Diversity and the Incidence or Onset of Civil Conflict in the Old World

Notes: This figure depicts the influence of contemporary population diversity at the country level on the *predicted* likelihood of observing (i) the incidence of a PRIO25 civil conflict in any given 5-year interval during the 1960–2008 time period (Panel (a)); (ii) the incidence of a WCM09 ethnic civil conflict in any given 5-year interval during the 1960–2005 time period (Panel (b)); (iii) the onset of a new PRIO25 civil conflict in any given year during the 1960–2008 time period (Panel (c)); and (iv) the onset of a new WCM09 ethnic civil conflict in any given year during the 1960–2005 time period (Panel (d)), conditional on other well-known diversity measures, the proximate geographical, institutional, and development-related correlates of conflict, controls for temporal dependence in conflict outcomes, and continent and time dummies, as considered by the specifications in Columns 2 and 6 of Panels A and B of Table 4 for the Old-World sample of countries. In each panel, the predicted likelihood of the conflict outcome is illustrated as a function of the percentile of the cross-country population diversity distribution, and the shaded area reflects the 95-percent confidence-interval region of the depicted relationship.

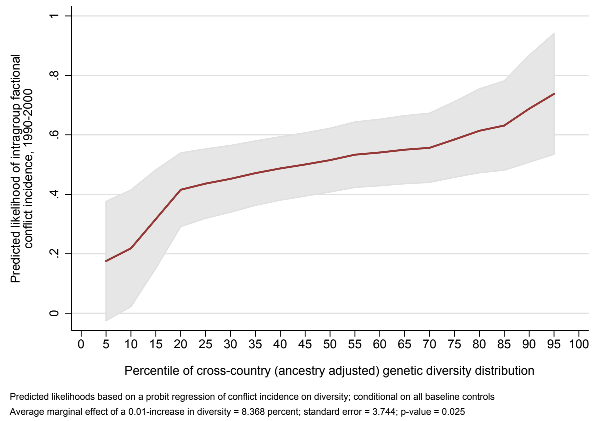
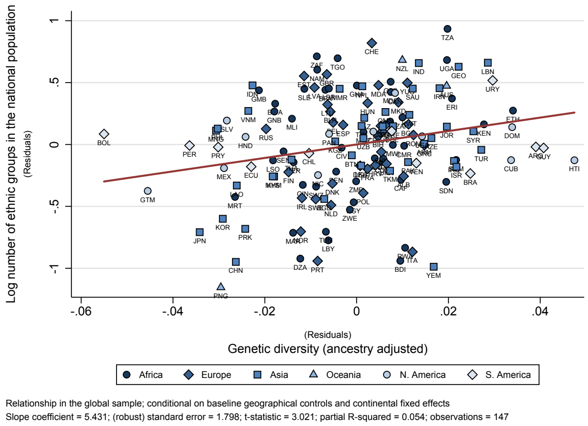
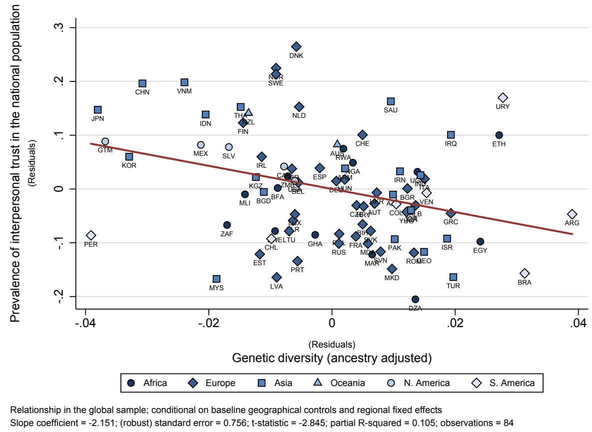


Figure A.2: Interpersonal Population Diversity and the Incidence of Intragroup Factional Conflict in the Old World

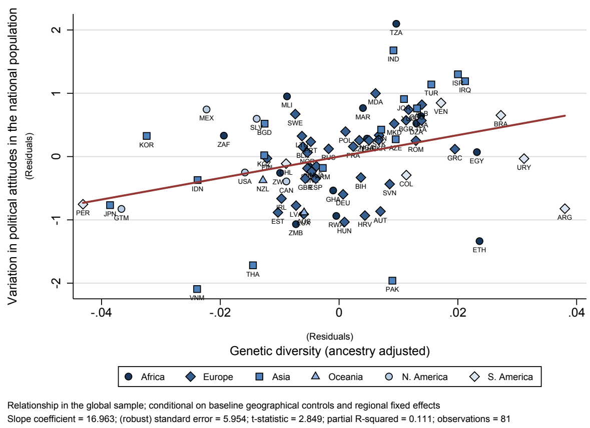
Notes: This figure depicts the influence of contemporary population diversity at the country level on the *predicted* likelihood of observing one or more factional conflicts *within* the "minorities at risk" (MAR) groups of a country's population during the 1990–1999 time period, conditional on other well-known diversity measures, the proximate geographical, institutional, and development-related correlates of conflict, and continent dummies, as considered by the specification in Column 2 of Table 6 for the Old-World sample of countries. The predicted likelihood of observing one or more intragroup factional conflicts is illustrated as a function of the percentile of the cross-country diversity distribution, and the shaded area reflects the 95-percent confidence-interval region of the depicted relationship.



(a) Number of ethnic groups



(b) Prevalence of interpersonal trust



(c) Variation in political attitudes

FIGURE A.3: Interpersonal Population Diversity and Proximate Determinants of the Frequency of Civil Conflict Onset across Countries

Notes: This figure depicts the global cross-country relationship between contemporary population diversity and each of three potentially conflict-augmenting proximate channels, including (i) the degree of cultural fragmentation, as reflected by the number of ethnic groups in the national population (Panel (a)); (ii) the prevalence of generalized interpersonal trust at the country level (Panel (b)); and (iii) the extent of heterogeneity in preferences for redistribution and public-goods provision, as reflected by the intracountry dispersion in individual political attitudes on a politically "left"–"right" categorical scale (Panel (c)), conditional on the baseline geographical correlates of conflict, as considered by the analysis in Table 10. Each of Panels (a), (b), and (c) presents an added-variable plot with a partial regression line, corresponding to the estimated coefficient associated with population diversity in Columns 1, 4, and 7, respectively, of Table 10.

A.4 Descriptive Statistics

Table A.23: Descriptive Statistics of the Regression Sample for Explaining the Frequency of Overall Civil Conflict across Countries

(a) Summary statistics

						Percentile	
		Mean	SD	Min	Max	10^{th}	90^{th}
(1)	Log PRIO25 civil conflict onsets per year, 1960–2008	0.020	0.030	0.000	0.186	0.000	0.054
(2)	Genetic diversity (ancestry adjusted)	0.728	0.027	0.628	0.774	0.688	0.752
(3)	Ethnic fractionalization	0.461	0.259	0.002	0.930	0.107	0.792
(4)	Ethnolinguistic polarization	0.450	0.241	0.000	0.957	0.097	0.734
(5)	Absolute latitude	0.028	0.017	0.001	0.064	0.007	0.051
(6)	Land area	0.001	0.002	0.000	0.016	0.000	0.002
(7)	Ruggedness	0.122	0.117	0.004	0.585	0.018	0.278
(8)	Mean elevation	0.585	0.537	0.001	2.837	0.104	1.232
(9)	Range of elevation	1.685	1.370	0.040	6.176	0.322	3.442
(10)	Mean land suitability	0.388	0.248	0.003	0.951	0.046	0.718
(11)	Range of land suitability	0.717	0.266	0.000	0.999	0.317	0.994
(12)	Distance to nearest waterway	0.359	0.463	0.014	2.386	0.038	1.010
(13)	Ever a U.K. colony dummy	0.245	0.431	0.000	1.000	0.000	1.000
(14)	Ever a French colony dummy	0.196	0.398	0.000	1.000	0.000	1.000
(15)	Ever a non-U.K./non-French colony dummy	0.308	0.463	0.000	1.000	0.000	1.000
(16)	British legal origin dummy	0.238	0.427	0.000	1.000	0.000	1.000
(17)	French legal origin dummy	0.462	0.500	0.000	1.000	0.000	1.000
(18)	Executive constraints, 1960–2008 average	3.939	1.884	1.000	7.000	1.537	7.000
(19)	Fraction of years under democracy, 1960–2008	0.377	0.384	0.000	1.000	0.000	1.000
(20)	Fraction of years under autocracy, 1960–2008	0.393	0.337	0.000	1.000	0.000	0.918
(21)	Log oil production per capita, 1960–2008 average	2.507	2.758	0.000	9.748	0.000	6.501
(22)	Log population, 1960–2008 average	9.163	1.405	5.863	13.841	7.513	10.951
(23)	Log GDP per capita, 1960–2008 average	8.051	1.057	6.210	9.943	6.664	9.596

(b) Pairwise correlations

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
(1)	Log PRIO25 civil conflict onsets per year, 1960–2008	1.000																					
(2)	Genetic diversity (ancestry adjusted)	0.162	1.000																				
(3)	Ethnic fractionalization	0.203	0.198	1.000																			
(4)	Ethnolinguistic polarization	0.048	0.096	0.243	1.000																		
(5)	Absolute latitude	-0.138	0.083	-0.549	0.202	1.000																	
(6)	Land area	0.184	-0.056	-0.003	0.020	0.110	1.000																
(7)	Ruggedness	0.061	-0.201	-0.264	0.021	0.190	-0.115	1.000															
(8)	Mean elevation	0.138	-0.009	0.118	0.133	-0.050	0.054	0.645	1.000														
(9)	Range of elevation	0.293	-0.324	0.102	0.100	-0.028	0.360	0.408	0.604	1.000													
(10)	Mean land suitability	-0.023	-0.234	-0.348	-0.340	0.063	-0.251	0.281	-0.122	-0.142	1.000												
(11)	Range of land suitability	0.195	-0.198	0.001	-0.158	0.161	0.276	0.153	0.196	0.429	0.112	1.000											
(12)	Distance to nearest waterway	0.209	0.191	0.264	0.140	-0.000	0.443	-0.023	0.405	0.335	-0.403	0.129	1.000										
(13)	Ever a U.K. colony dummy	0.050	0.240	0.246	0.182	-0.223	0.099	-0.201	-0.019	-0.018	-0.236	-0.150	0.019	1.000									
(14)	Ever a French colony dummy	0.060	0.144	0.247	-0.055	-0.243	0.086	-0.180	-0.150	-0.097	-0.194	-0.031	0.118	-0.076	1.000								
(15)	Ever a non-U.K./non-French colony dummy	0.123	-0.345	0.173	-0.304	-0.511	0.053	-0.057	0.011	0.210	0.136	-0.026	-0.152	-0.062	-0.062	1.000							
(16)	British legal origin dummy	0.062	0.164	0.186	0.093	-0.195	0.103	-0.125	0.062	0.015	-0.144	0.003	0.003	0.714	-0.151	-0.052	1.000						
(17)	French legal origin dummy	-0.012	-0.124	0.178	-0.131	-0.413	-0.080	-0.148	-0.082	0.022	-0.039	-0.164	-0.084	-0.299	0.321	0.416	-0.517	1.000					
(18)	Executive constraints, 1960–2008 average	-0.101	-0.207	-0.349	-0.009	0.503	0.050	0.170	-0.092	-0.025	0.244	0.277	-0.309	-0.035	-0.281	-0.074	0.123	-0.289	1.000				
(19)	Fraction of years under democracy, 1960–2008	-0.157	-0.212	-0.371	-0.036	0.468	0.042	0.075	-0.192	-0.066	0.248	0.220	-0.328	-0.070	-0.272	-0.045	0.093	-0.203	0.919	1.000			
(20)	Fraction of years under autocracy, 1960–2008	0.048	0.238	0.223	0.076	-0.276	0.015	-0.204	-0.010	-0.021	-0.309	-0.287	0.215	0.079	0.239	-0.048	-0.120	0.151	-0.849	-0.804	1.000		
(21)	Log oil production per capita, 1960–2008 average	-0.004	0.005	0.006	0.182	0.103	0.282	-0.232	-0.175	0.111	-0.386	-0.050	0.077	0.076	0.004	-0.026	-0.085	0.084	-0.132	-0.094	0.238	1.000	
(22)	Log population, 1960–2008 average	0.311	-0.186	-0.120	-0.111	0.059	0.466	0.036	0.081	0.522	0.115	0.586	0.092	0.038	-0.024	0.045	0.080	-0.073	0.105	0.063	-0.063	0.063	1.000
(23)	Log GDP per capita, 1960–2008 average	-0.289	-0.152	-0.485	0.092	0.661	0.117	0.043	-0.271	-0.092	0.004	0.078	-0.298	-0.131	-0.302	-0.171	-0.124	-0.150	0.597	0.599	-0.394	0.437	0.006

Appendix B Appendix for the Ethnicity-Level Analyses

B.1 Construction of the Georeferenced Dataset

The novel geo-referenced data set of population diversity across ethnic groups is based on several sources. It links the measurements of observed genetic diversity of the 232 ethnic group (as provided Pemberton, DeGiorgio and Rosenberg (2013)), as well as the measurement of predicted diversity for the entire set of ethnic groups in the *Ethnographic Atlas* (as constructed in the current paper) to: (i) the geographical area of the historical homelands of these ethnic groups, (ii) the ethnographic characteristics of these ethnic groups, (as reported by the *Ethnographic Atlas* and the *Standard Cross-Cultural Sample*), and (iii) the geographical characteristics of the homelands of these ethnic groups.

The link between population diversity of each ethnic group and the geographical area of the historical homeland of these ethnic groups exploits several sources. Polygons for observations in the *Ethnographic Atlas* is based on Fenske (2013), who linked observations in the *Ethnographic Atlas* to the: (i) polygons found in Murdock (1959), (ii) the Handbook of North American Indians (Heizer, 1978), (iii) Global Mapping International's (GMI) World Language Mapping System, (iv) the Geo-Referencing Ethnic Groups (GREG) map of Weidmann, Rød and Cederman (2010), and (v) data for modern administrative boundaries. We used the link between observations in the *Ethnographic Atlas* and James Fenske's collection of polygons that was implied by the reported centroid coordinates in the data by Fenske (2013).

The matching process of observed population diversity for the 232 ethnic groups in Pemberton, DeGiorgio and Rosenberg (2013) was based on four phases. First, 65 observations from the Pemberton data was merged with name-based matches with the *Ethnographic Atlas* and via that to James Fenske's polygons.[1] Second, the geo-coded points of the ethnic groups reported in Pemberton, DeGiorgio and Rosenberg (2013) was overlaid with the map of James Fenske's polygons and proximate pairs of polygons and points were classified as as either separate, similar, or disparate groups, yielding 84 matches between polygons and points. Third, an additional 97 merges were achieved using a similar method with polygons from the GMI data set and their associated Ethnologue information. Fourth, for some remaining ethnic groups, a plausible polygon could be constructed based on secondary information about the ethnic group.[2]

[1] This matching process required the use of the various names given to each group in different sources.

[2] For instance, Tuscans were merged to the modern region of Toscana, Orcadians were merged to the South Orkney Islands, the Zenú were merged to the Zenú reserve, and the Sengwer were merged to the Embobut Forest area.

B.2 Robustness Checks

Table B.1: Population Diversity and the Spatiotemporal Prevalence of Conflict across Ethnic Homelands – Robustness to Accounting for Alternative Distances

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
	Log spatio-temporal prevalence of UCDP GED conflicts, 1989–2008					
Observed diversity	21.858**	22.919**	22.412**			
	[10.412]	[10.633]	[10.526]			
Predicted diversity				74.511***	70.830***	73.109***
				[6.279]	[6.134]	[6.237]
Distance to technological frontier in year 1 (in 1000 kms)	0.078			0.007		
	[0.159]			[0.057]		
Distance to technological frontier in year 1000 (in 1000 kms)		-0.104			-0.191***	
		[0.157]			[0.065]	
Distance to technological frontier in year 1500 (in 1000 kms)			-0.086			-0.069
			[0.142]			[0.061]
Latitude	-0.111***	-0.107**	-0.111***	-0.012	-0.009	-0.013
	[0.040]	[0.041]	[0.040]	[0.017]	[0.017]	[0.017]
Ruggedness	0.105	0.078	0.077	0.163**	0.174**	0.162**
	[0.182]	[0.180]	[0.181]	[0.078]	[0.077]	[0.077]
Elevation	-0.012	0.027	0.042	0.065	-0.008	0.044
	[0.463]	[0.465]	[0.477]	[0.190]	[0.187]	[0.190]
S.D. of elevation	-1.101	-1.073	-1.102	-0.971**	-0.754	-0.886*
	[0.930]	[0.910]	[0.918]	[0.488]	[0.489]	[0.489]
Mean land suitability (native crops)	-0.146*	-0.139	-0.146*	0.016	0.005	0.012
	[0.085]	[0.087]	[0.085]	[0.023]	[0.023]	[0.023]
S.D. of land suitability (native crops)	0.391*	0.406*	0.414*	0.062	0.046	0.059
	[0.236]	[0.238]	[0.236]	[0.094]	[0.091]	[0.093]
Change in land suitability (Columbian Exchange)	0.000	0.013	0.014	-0.241***	-0.231***	-0.235***
	[0.097]	[0.091]	[0.091]	[0.038]	[0.038]	[0.038]
Share desert	-2.872**	-2.910**	-2.880**	-1.670***	-1.685***	-1.667***
	[1.160]	[1.140]	[1.135]	[0.405]	[0.394]	[0.401]
Nearest waterway	-0.048	-0.050	-0.050	0.040	0.042	0.042
	[0.078]	[0.079]	[0.079]	[0.043]	[0.043]	[0.044]
Average temperature	0.160**	0.153**	0.153**	0.079***	0.065**	0.072***
	[0.063]	[0.063]	[0.063]	[0.026]	[0.025]	[0.025]
Volatility of temperature	7.245***	7.190***	7.250***	1.581**	1.039	1.419**
	[1.856]	[1.804]	[1.811]	[0.651]	[0.651]	[0.651]
Average precipitation	0.018***	0.017***	0.017***	0.003	0.002	0.002
	[0.006]	[0.006]	[0.006]	[0.002]	[0.002]	[0.002]
Volatility of precipitation	-0.133***	-0.125***	-0.126***	0.012	0.016	0.012
	[0.041]	[0.042]	[0.040]	[0.012]	[0.013]	[0.012]
Regional dummies	×	×	×	×	×	×
Additional climatic covariates	×	×	×	×	×	×
Sample	Observed	Observed	Observed	Extended	Extended	Extended
Observations	230	230	230	1251	1251	1251
Effect of increasing genetic diversity	0.323**	0.339**	0.331**	1.217***	1.157***	1.194***
from the 10th to the 90th percentile	[0.154]	[0.157]	[0.156]	[0.103]	[0.100]	[0.102]
Adjusted R^2	0.480	0.480	0.480	0.444	0.449	0.444

Notes: This table exploits cross-ethnicity variations to establish a significant positive impact of observed and predicted population on the log spatio-temporal prevalence of UCDP/PRIO conflicts during the 1989–2008 period, conditional on migratory distances from historical technological frontiers as well as the baseline geographical characteristics. The set of continent and regional dummies includes indicators for Europe, Asia, North America, South America, Oceania, North Africa, and Sub-Saharan Africa. Additional climatic covariates refer to the average diurnal temperature range, average cloud cover, and average temperature range in the homeland. Robust standard errors are reported in square brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table B.2: Population Diversity and the Spatiotemporal Prevalence of Conflict across Ethnic Homelands – Robustness to Accounting for Measures of Ecological Diversity

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) 2SLS	(8) 2SLS
	Log spatio-temporal prevalence of UCDP GED conflicts, 1989–2008							
Observed diversity	31.263***	24.997**	18.327*	20.909*			90.477***	88.013***
	[10.194]	[9.939]	[10.508]	[10.897]			[33.148]	[27.495]
Predicted diversity					82.223***	82.756***		
					[5.899]	[5.982]		
Ecological diversity	-1.710	-0.622	0.490	0.227	-0.154	0.217	-0.463	-0.193
	[1.228]	[1.355]	[1.208]	[1.202]	[0.597]	[0.601]	[1.478]	[1.258]
Ecological polarization	1.290	1.591	0.374	0.885	0.002	-0.154	-0.074	0.325
	[0.957]	[0.985]	[0.892]	[0.938]	[0.467]	[0.455]	[1.072]	[0.987]
Latitude		-0.092***	-0.107***	-0.100**	0.000	0.025	-0.112***	-0.088**
		[0.018]	[0.040]	[0.046]	[0.017]	[0.018]	[0.040]	[0.043]
Ruggedness		0.216	0.065	0.121	0.125	0.121	0.407*	0.372*
		[0.145]	[0.180]	[0.197]	[0.083]	[0.085]	[0.228]	[0.203]
Elevation		-0.829***	0.025	0.133	0.201	0.208	-0.276	-0.170
		[0.305]	[0.459]	[0.468]	[0.215]	[0.219]	[0.495]	[0.502]
S.D. of elevation		-1.503*	-0.775	-0.361	-1.239**	-1.234**	-1.414	-0.727
		[0.849]	[0.964]	[1.034]	[0.597]	[0.589]	[1.141]	[1.138]
Mean land suitability (native crops)		-0.149*	-0.145*	-0.123	0.012	0.031	-0.230**	-0.182**
		[0.088]	[0.085]	[0.087]	[0.023]	[0.024]	[0.095]	[0.086]
S.D. of land suitability (native crops)		0.636***	0.442*	0.355	0.081	0.121	0.398	0.290
		[0.196]	[0.239]	[0.224]	[0.097]	[0.100]	[0.260]	[0.247]
Change in land suitability (Columbian Exchange)		0.272***	0.019	0.097	-0.241***	-0.237***	0.050	0.091
		[0.081]	[0.093]	[0.102]	[0.038]	[0.038]	[0.094]	[0.096]
Share desert		2.773***	-2.523**	-2.450*	-1.710***	-1.733***	-3.146**	-3.030**
		[1.030]	[1.223]	[1.264]	[0.427]	[0.446]	[1.357]	[1.332]
Nearest waterway		-0.064	-0.043	-0.097	0.040	0.019	-0.051	-0.090
		[0.080]	[0.081]	[0.085]	[0.044]	[0.043]	[0.073]	[0.078]
Average temperature			0.160**	0.164**	0.094***	0.134***	0.171***	0.201***
			[0.063]	[0.071]	[0.025]	[0.025]	[0.065]	[0.070]
Volatility of temperature			7.608***	7.918***	1.684**	1.316**	8.626***	8.631***
			[1.838]	[1.949]	[0.691]	[0.664]	[1.804]	[1.864]
Average precipitation			0.017***	0.015**	0.006**	0.005*	0.020**	0.018**
			[0.006]	[0.007]	[0.003]	[0.003]	[0.008]	[0.009]
Volatility of precipitation			-0.118***	-0.102**	0.004	0.011	-0.131**	-0.120**
			[0.040]	[0.044]	[0.015]	[0.015]	[0.051]	[0.052]
Time since initial settlement				0.617		0.637***		1.100*
				[0.480]		[0.193]		[0.563]
Malaria endemicity				1.464		-0.879		0.144
				[1.340]		[0.556]		[1.422]
Log luminosity				-0.181*		-0.194***		-0.358***
				[0.107]		[0.041]		[0.132]
Regional dummies	×	×	×	×	×	×	×	×
Additional climatic covariates			×	×	×	×	×	×
Decile of population density dummies				×		×		×
Sample	Observed	Observed	Observed	Observed	Extended	Extended	Observed	Observed
Observations	228	228	228	228	1215	1215	228	228
Effect of increasing genetic diversity	0.470***	0.376**	0.276*	0.314*	1.385***	1.394***	1.361***	1.324***
from the 10th to the 90th percentile	[0.153]	[0.149]	[0.158]	[0.164]	[0.099]	[0.101]	[0.499]	[0.414]
First-stage F statistic							12.406	20.516
Adjusted R^2	0.165	0.327	0.488	0.511	0.447	0.464		

Notes: This table exploits cross-ethnicity variations to establish a significant positive impact of contemporary population diversity on the log spatio-temporal prevalence of UCDP/PRIO conflicts during the 1989–2008 period, conditional on ecological diversity and ecological polarization as well as the baseline control variables. The set of continent and regional dummies includes indicators for Europe, Asia, North America, South America, Oceania, North Africa, and Sub-Saharan Africa. Additional climatic covariates refer to the average diurnal temperature range, average cloud cover, and average temperature range in the homeland. The 2SLS regressions exploit prehistoric migratory distance from East Africa to each ethnic homeland as an excluded instrument for the observed population diversity of this ethnic group. The estimated effect associated with increasing population diversity from the tenth to the ninetieth percentile of its cross-country distribution is expressed in terms of the change in the average yearly share of the area of each ethnic homeland that was within the boundaries of internal armed conflict over the period 1989–2008. Robust standard errors are reported in square brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent.

Table B.3: Population Diversity and the Spatiotemporal Prevalence of Conflict across Ethnic Homelands – Robustness to Accounting for Spatial Dependence using SARAR Estimation

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) 2SLS	(8) 2SLS
	Log spatio-temporal prevalence of UCDP GED conflicts, 1989–2008							
Genetic diversity (observed)	36.221***	33.464***	29.650***	35.235***			32.034***	35.035***
	[9.565]	[8.633]	[8.228]	[8.452]			[10.877]	[9.769]
Genetic diversity (predicted)					54.045***	59.075***		
					[4.847]	[5.362]		
Latitude		-0.125***	-0.152***	-0.125***	-0.006	-0.027*	-0.174***	-0.149***
		[0.019]	[0.039]	[0.039]	[0.014]	[0.014]	[0.041]	[0.042]
Ruggedness		0.108	0.139	0.122	-0.021	0.017	0.045	0.020
		[0.160]	[0.165]	[0.167]	[0.060]	[0.062]	[0.164]	[0.165]
Elevation		-0.223	-0.136	-0.117	0.421***	0.364**	-0.034	-0.075
		[0.309]	[0.436]	[0.430]	[0.152]	[0.155]	[0.426]	[0.424]
S.D. of elevation		-1.176	-1.597*	-1.015	-0.150	-0.316	-1.587*	-0.805
		[0.772]	[0.905]	[0.904]	[0.339]	[0.353]	[0.858]	[0.874]
Mean land suitability (native crops)		-0.143*	-0.148**	-0.143*	-0.034	-0.026	-0.125*	-0.124*
		[0.077]	[0.075]	[0.074]	[0.024]	[0.024]	[0.075]	[0.074]
S.D. of land suitability (native crops)		0.270	0.094	0.022	0.039	0.117	0.090	0.031
		[0.201]	[0.199]	[0.197]	[0.069]	[0.073]	[0.187]	[0.190]
Change in land suitability (Columbian Exchange)		0.231***	0.030	0.056	-0.008	-0.012	0.046	0.070
		[0.082]	[0.090]	[0.088]	[0.035]	[0.034]	[0.091]	[0.089]
Share desert		1.332	-2.512**	-2.831**	-0.833**	-0.120	-1.980*	-2.574**
		[0.956]	[1.182]	[1.186]	[0.364]	[0.377]	[1.135]	[1.154]
Nearest waterway		-0.052	-0.109	-0.176**	0.017	0.004	-0.133	-0.198**
		[0.089]	[0.086]	[0.086]	[0.036]	[0.037]	[0.083]	[0.083]
Average temperature			0.086	0.122**	0.123***	0.083***	0.039	0.080
			[0.060]	[0.061]	[0.024]	[0.024]	[0.061]	[0.063]
Volatility of temperature			5.663***	5.775***	-0.597	-0.275	5.786***	5.782***
			[1.533]	[1.532]	[0.642]	[0.611]	[1.536]	[1.516]
Average precipitation			0.013**	0.012**	-0.001	-0.003	0.012**	0.014**
			[0.006]	[0.006]	[0.002]	[0.002]	[0.006]	[0.006]
Volatility of precipitation			-0.070*	-0.066*	-0.004	0.015	-0.039	-0.066*
			[0.037]	[0.039]	[0.011]	[0.011]	[0.036]	[0.040]
Time since initial settlement				0.047		0.024		-0.254
				[0.487]		[0.160]		[0.464]
Malaria endemicity				1.376		4.332***		1.018
				[1.234]		[0.559]		[1.407]
Log Luminosity				-0.261***		-0.044		-0.241***
				[0.094]		[0.032]		[0.092]
Regional dummies	×	×	×	×	×	×	×	×
Additional climatic covariates			×	×	×	×	×	×
Decile of population density dummies				×		×		×
Direct impact of genetic diversity	37.020***	35.051***	30.293***	35.952***	54.217***	58.580***	32.433***	36.039***
	[9.795]	[9.137]	[8.425]	[8.662]	[4.858]	[5.338]	[11.027]	[10.144]
Sample	Observed	Observed	Observed	Observed	Extended	Extended	Observed	Observed
Observations	202	202	202	202	1108	1108	202	202

Notes: This table exploits cross-ethnicity variations to establish a significant positive reduced-form impact of contemporary population diversity on the log spatio-temporal prevalence of UCDP/PRIO conflicts during the 1989–2008 period, conditional on the baseline control variables (i.e., proximate geographical and development-related correlates of conflict) and accounting for spatial dependence using a spatial autoregressive (SAR) model, with a spectral-normalized inverse-distance weighting matrix, estimated with maximum-likelihood estimation, with a spatial lag of the dependent variable and a spatially lagged error. The model treat errors as heteroskedastic. Variables relating to observations associated with the same homeland polygon are averaged and a single observation is kept for each polygon. The set of continent and regional dummies includes indicators for Europe, Asia, North America, South America, Oceania, North Africa, and Sub-Saharan Africa. Additional climatic covariates refer to the average diurnal temperature range, average cloud cover, and average temperature range in the homeland. Robust standard errors are reported in square brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table B.4: Population Diversity and the Number of Conflicts or Conflict-Related Deaths across Ethnic Homelands

	(1) Poisson	(2) Poisson	(3) Poisson	(4) Poisson	(5) Poisson	(6) Poisson	(7) Poisson	(8) Poisson	(9) Poisson	(10) Poisson
	Number of conflicts					Number of deaths				
Observed diversity	123.941***	96.373***	104.235***	120.420***		218.539**	350.771***	224.638***	191.334**	
	[34.156]	[22.381]	[39.102]	[35.813]		[101.997]	[113.355]	[56.568]	[78.383]	
Predicted diversity					66.825***					66.825***
					[20.631]					[20.631]
Latitude		-0.004	0.088	0.158***	-0.163***		-0.252***	-0.143*	-0.039	-0.163***
		[0.023]	[0.057]	[0.053]	[0.062]		[0.062]	[0.081]	[0.086]	[0.062]
Ruggedness		0.103	-0.098	-0.038	0.922***		2.169***	1.446***	1.732***	0.922***
		[0.166]	[0.514]	[0.354]	[0.229]		[0.507]	[0.395]	[0.479]	[0.229]
Elevation		-0.873**	0.583	1.543**	0.869*		-2.700***	-2.776***	-0.687	0.869*
		[0.416]	[1.055]	[0.639]	[0.485]		[1.024]	[0.784]	[1.063]	[0.485]
S.D. of elevation		1.935*	2.970**	3.143**	-5.933***		-0.519	2.738	0.827	-5.933***
		[1.127]	[1.205]	[1.515]	[1.671]		[2.156]	[2.164]	[2.269]	[1.671]
Mean land suitability (native crops)		-0.007	-0.046	-0.174	-0.069		-0.126	-0.378**	-0.303	-0.069
		[0.124]	[0.151]	[0.169]	[0.097]		[0.165]	[0.180]	[0.245]	[0.097]
S.D. of land suitability (native crops)		0.770***	0.862***	0.864***	1.020***		1.531***	1.271***	1.057***	1.020***
		[0.218]	[0.174]	[0.202]	[0.209]		[0.492]	[0.361]	[0.261]	[0.209]
Change in land suitability (Columbian Exchange)		0.020	-0.080	-0.092	0.123		0.395***	0.323**	0.097	0.123
		[0.105]	[0.116]	[0.087]	[0.112]		[0.150]	[0.162]	[0.140]	[0.112]
Share desert		-0.244	-2.809**	-4.171***	-3.547***		2.113	-2.106	-3.683*	-3.547***
		[1.288]	[1.247]	[1.398]	[1.361]		[2.459]	[3.287]	[1.999]	[1.361]
Nearest waterway		-0.270**	-0.257**	-0.202	0.511***		0.656***	-0.080	0.172	0.511***
		[0.112]	[0.111]	[0.132]	[0.087]		[0.242]	[0.206]	[0.224]	[0.087]
Average temperature			0.390***	0.566***	0.296***			0.154	0.552***	0.296***
			[0.121]	[0.103]	[0.097]			[0.168]	[0.177]	[0.097]
Volatility of temperature			3.803**	3.006*	3.263*			1.143	5.634**	3.263*
			[1.884]	[1.628]	[1.757]			[2.703]	[2.425]	[1.757]
Average precipitation			0.018	0.039***	0.031***			0.054***	0.071***	0.031***
			[0.011]	[0.012]	[0.008]			[0.013]	[0.017]	[0.008]
Volatility of precipitation			-0.179***	-0.304***	-0.353***			-0.499***	-0.455***	-0.353***
			[0.064]	[0.095]	[0.067]			[0.128]	[0.131]	[0.067]
Time since initial settlement				2.132**	3.213***				4.503***	3.213***
				[0.889]	[0.649]				[0.847]	[0.649]
Malaria endemicity				-3.381*	-4.370***				-3.437	-4.370***
				[2.049]	[1.206]				[2.337]	[1.206]
Log Luminosity				0.045	0.712***				0.161	0.712***
				[0.112]	[0.131]				[0.136]	[0.131]
Regional dummies	×	×	×	×	×	×	×	×	×	×
Additional climatic covariates			×	×	×			×	×	×
Decile of population density dummies				×	×				×	×
Sample	Observed	Observed	Observed	Observed	Extended	Observed	Observed	Observed	Observed	Extended
Observations	230	230	230	230	1251	230	230	230	230	1251

Notes: This table exploits cross-ethnicity variations to establish a significant positive reduced-form impact of contemporary population diversity on the log spatio-temporal prevalence of UCDP/PRIO conflicts and conflict-related deaths during the 1989–2008 period, conditional on the baseline control variables (i.e., proximate geographical and development-related correlates of conflict). The set of continent and regional dummies includes indicators for Europe, Asia, North America, South America, Oceania, North Africa, and Sub-Saharan Africa. Additional climatic covariates refer to the average diurnal temperature range, average cloud cover, and average temperature range in the homeland. Robust standard errors are reported in square brackets. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

B.3 Supplementary Figures

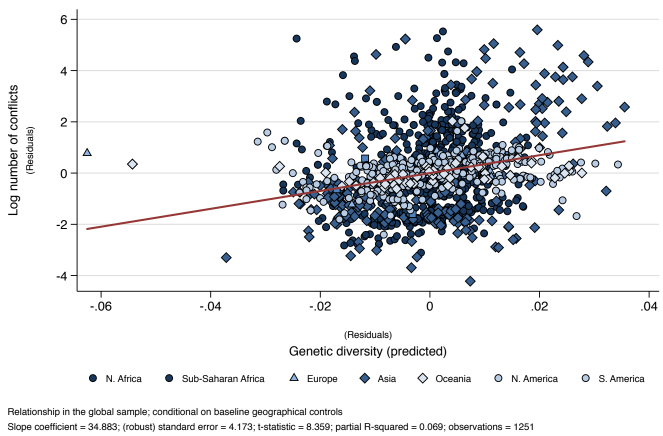


FIGURE B.1: Predicted Population Diversity and the Number of Conflicts across Ethnic Homelands

Notes: This figure depicts the global cross-ethnic homeland relationship between contemporary predicted population diversity and the log number of UCDP/GED conflicts, conditional on the baseline geographical correlates of conflict, as considered by the specifications in Column 3 of Table 9.

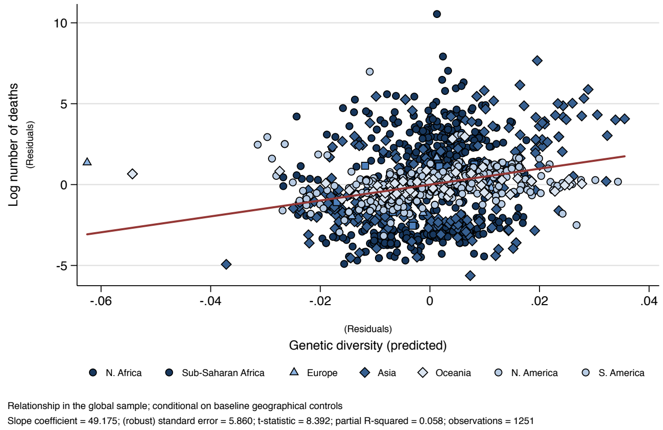


Figure B.2: Predicted Population Diversity and the Number of Conflict-Related Deaths across Ethnic Homelands

Notes: This figure depicts the global cross-ethnic homeland relationship between predicted population diversity and the log prevelance of UCDP/GED conflict-related deaths, conditional on the baseline geographical correlates of conflict, as considered in Column 7 of Table 9

B.4 Descriptive Statistics

Table B.5: Summary Statistics

	Mean	S.D.	Min	Max	P10	P90
Panel A: Sample of observed diversity ($N = 230$)						
Observed diversity	0.72	0.05	0.56	0.77	0.65	0.76
Spatio-temporal prevalence of UCDP/PRIO conflicts, 1989–2008	0.13	0.25	0.00	1.28	0.00	0.55
Latitude	14.44	14.77	0.04	68.92	1.77	37.72
Ruggedness	1.35	1.46	0.01	9.96	0.16	2.91
Elevation	0.74	0.75	-0.84	4.06	0.04	1.67
S.D. of elevation	0.32	0.27	0.01	1.44	0.05	0.68
Mean land suitability (native crops)	6.03	2.79	0.00	11.80	2.78	9.85
S.D. of land suitability (native crops)	1.07	0.90	0.00	4.40	0.08	2.14
Change in land suitability (Columbian Exchange)	2.33	2.03	-0.00	7.20	0.00	5.13
Share desert	0.07	0.21	0.00	1.00	0.00	0.24
Nearest waterway	1.03	1.92	0.00	15.21	0.04	2.67
Average temperature	21.40	7.57	-13.71	28.89	9.73	27.23
Volatility of temperature	0.41	0.18	0.14	1.16	0.19	0.67
Average precipitation	129.36	106.84	9.99	480.08	31.30	297.66
Volatility of precipitation	19.52	13.94	3.02	70.29	7.55	42.13
Diurnal temperature range	11.25	2.40	5.63	18.19	8.21	14.30
Cloud cover	60.63	15.37	20.48	86.64	37.66	77.49
Temperature range	5.58	6.27	0.00	35.03	0.63	14.17
Time since initial settlement	0.86	0.31	-0.11	1.00	0.25	1.00
Malaria endemicity	0.17	0.19	0.00	0.67	0.00	0.50
Log luminosity	-2.11	2.17	-4.61	3.13	-4.61	1.18
Panel B: Sample of predicted diversity ($N = 1251$)						
Predicted diversity	0.71	0.05	0.59	0.77	0.65	0.76
Spatio-temporal prevalence of UCDP/PRIO conflicts, 1989–2008	0.15	0.29	0.00	1.85	0.00	0.60
Latitude	19.97	16.45	0.02	78.07	3.29	44.46
Ruggedness	1.23	1.33	0.00	9.96	0.11	2.99
Elevation	0.65	0.72	-2.34	4.53	0.03	1.64
S.D. of elevation	0.28	0.23	0.00	1.48	0.05	0.58
Mean land suitability (native crops)	6.20	3.60	0.00	18.96	0.95	10.34
S.D. of land suitability (native crops)	0.91	0.85	0.00	6.64	0.07	2.00
Change in land suitability (Columbian Exchange)	2.05	2.18	-0.00	12.37	0.00	5.32
Share desert	0.07	0.22	0.00	1.28	0.00	0.22
Nearest waterway	0.85	1.50	0.00	15.21	0.04	1.98
Average temperature	19.87	8.40	-18.91	29.54	6.69	27.26
Volatility of temperature	0.47	0.19	0.04	1.16	0.25	0.73
Average precipitation	110.07	75.69	0.88	592.87	26.84	218.86
Volatility of precipitation	16.20	9.69	0.00	81.66	6.14	29.12
Diurnal temperature range	11.68	2.83	4.57	19.38	8.14	15.22
Cloud cover	59.08	13.19	16.01	89.62	40.79	74.70
Temperature range	4.50	4.46	0.00	32.22	0.47	9.93
Time since initial settlement	0.88	0.30	-0.20	1.00	0.40	1.00
Malaria endemicity	0.16	0.21	0.00	0.73	0.00	0.52
Log luminosity	-1.75	2.24	-4.61	3.69	-4.61	1.43