MONASH University

MONASH BUSINESS SCHOOL

CENTRE FOR HEALTH ECONOMICS

WORKING PAPERS

Ethnic Inequality and Public Health

Discussion Paper no. 2023-07

Sundar Ponnusamy and Mohammad Abbas Hakeem

Keywords: Inequality, child mortality, health

JEL Classification: I14,I18,J13

Sundar Ponnusamy: Monash University (email: sundar.ponnusamy@monash.edu); Mohammad Abbas Hakeem: University of Adelaide (email: abbas.hakeem@adelaide.edu.au).

Ethnic Inequality and Public Health*

Sundar Ponnusamy† Mohammad Abbas Hakeem‡

Abstract

We examine the association between ethnic inequality and various key health outcomes for a global set of developed and developing countries. Our results show that higher ethnic inequality is associated with a poor state of public health, such as higher child and maternal mortality, increased stillbirths and child stunting, and reduced life expectancy at birth. This set of effects is found to be predominant mainly in developing countries, and Sub-Saharan African countries. Results remain robust to the inclusion of various other measures of inequality, ethnic composition indices, geographic endowments, and other relevant controls. We argue that lower contraceptive usage and poor vaccination rates are potential mechanisms through which ethnic inequality affects health outcomes. Policies targeted at improving public health may need to focus more on these key intermediate channels in ethnic minority regions.

Keywords: Inequality; child mortality; health
JEL Classification: I14, I18, J13.

*We would like to thank Umair Khalil, David Johnston, Armand Sim, Dennis Petrie, Firmin Doko Tchatoka, Yaping Shen, Raul Barreto and seminar participants at the University of Adelaide for their valuable feedback. A part of this research was conducted when the corresponding author Sundar Ponnusamy was a Research Scholar at the University of Adelaide. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. All errors are our own.

†Centre for Health Economics, Monash University, 900 Dandenong Road, VIC 3145, Australia. Email: sundar.ponnusamy@monash.edu.

‡School of Economics and Public Policy, University of Adelaide, Level 4, 10 Pulteney St., Adelaide, SA 5005, Australia. Email: abbas.hakeem@adelaide.edu.au.

1 Introduction

The Millennium Summit of the United Nations in 2000 established eight goals as the Millennium Development Goals (MDGs) to be achieved by 2015, with two of these focusing on reducing child mortality and improving maternal health.[1] Even though there has been an improvement in the overall quality of health in recent years owing to the advancement in medical science accompanied by targeted public policies, developing countries still lag in health indices compared with the developed economies.[2] Studies so far have focused on several determinants of health, such as education, technological development, and economic growth (Cutler et al. 2006; Chen and Li 2009; Baird et al. 2011; Jamison et al. 2016; Brunello et al. 2016). But a potentially significant determinant of health that is yet to be explored in detail is ethnic inequality, which is defined as the inequality in economic development across ethnic groups coexisting in a country (Alesina et al., 2016). In this study, our primary focus is on understanding the association between ethnic inequality (henceforth, EI) and child mortality, while we also explore the relationship between EI and other relevant health outcomes such as maternal mortality, number of stillbirths, life expectancy at birth, and child stunting. We investigate potential mechanisms through which EI can affect public health as well.

In a seminal paper, Alesina et al. (2016) have shown that ethnic inequality is a negative correlate of economic development while also suggesting that the consequences of inequality can be two-faceted. A higher degree of ethnic inequality can be accompanied by biased public policies, which can result in several negative consequences, such as the diversion of resources away from ethnic minorities, the creation of barriers to education or investment for certain segments of society, and higher civil conflicts. In this case, EI can have profound negative effects on health outcomes, such as a higher child or maternal mortality or a lower life expectancy. On the other hand, inequality in development may encourage individuals and thereby help inspire higher-quality innovation or entrepreneurship in a society, which can have positive effects on health outcomes. While both forces can operate simultaneously, which one tends to dominate remains an empirical question.

Despite the established link between EI and outcomes such as economic development or governance (Alesina et al. 2016; Baldwin and Huber 2010; Kyriacou 2013), evidence on the association between EI and health outcomes is still lacking. We contribute to the literature by examining the relationship between ethnic inequality and various public health outcomes in

[1] The MDGs set the target for a reduction in under-five mortality to two-thirds and maternal mortality to three-quarters by 2015 from the 1990 levels. Based on MDG (2015), child mortality reduced by over 50 percent, maternal mortality dropped by around 45 percent, and the increase in contraceptive prevalence among women went from 55 to 64 percent between 1990 and 2015.

[2] Despite the improving global averages, there remain stark gaps in health between countries, especially in child survival. For example, under-five mortality (per 1000 live births) in the year 2000 for developing countries (low and low-middle-income countries) in the sample was 103.55, and for high-income countries, it was only 9.26. While the under-5 mortality rates have gone down for developing countries from 103.55 in 2000 to 62.67 in 2012, the number is still high and nowhere close to the high-income countries, which recorded an average of 7.1 deaths per 1000 live births in 2012. Source: Authors' calculations based on WDI (2020).

this study. We rely on a large database of EI for 173 countries from Alesina et al. (2016), in which nightlight activity (luminosity data) measured from space is used as a proxy for income data along ethnic lines to construct ethnic inequality indicators. By employing a simple cross-country regression analysis, we establish a correlation for the year 2000 initially.[3] Then, we control for various other inequality and ethnic fractionalization measures and a set of relevant controls to isolate the association between EI and health.

We use two different measures of ethnic inequality indices in this study. Our primary EI indicator relies on the location of ethnic groups (Geo-referencing of Ethnic Groups, henceforth, GREG index), whereas we also employ an EI indicator based on the location of linguistic groups (henceforth, Ethnologue index).[4] Our initial focus is on the association between EI and under-5 mortality rates (U5MR). Baseline specifications in which we control for some of the relevant covariates such as spatial inequality (SI),[5] and other necessary inequality indices suggest that there is a significant positive association between ethnic inequality and child mortality – one standard deviation increase in the GREG (Ethnologue) index is associated with an increase of 13.40 (12.87) under-5 deaths per 1000 live births, on average. This equates to 23.74% (22.81%) of the mean under-5 deaths for the overall sample of 173 countries.

Ethnic diversity (also referred to as ethnic heterogeneity) is defined as the presence of multiple ethnic or language groups in a society, which can either impose costs or provide benefits to the society. To ensure that ethnic diversity measures that can be indicative of the degree of ethnic fractionalization in a country do not confound our findings, we control for ethnic-linguistic fragmentation and segregation (Alesina et al. 2003; Desmet et al. 2012; Alesina and Zhuravskaya 2011), cultural fragmentation (Fearon, 2003), ethnic polarization (Montalvo and Reynal-Querol, 2005) and predicted genetic diversity (Ashraf and Galor, 2013) in our specifications.[6] We identify that the positive association between EI and U5MR remains mostly robust to the inclusion of these variables, while the other related measures of ethnic diversity remain insignificant.[7] Therefore, ethnic diversity may affect child mortality through ethnic inequality, and once the latter is controlled for, the other measures of ethnic diversity lose their

[3]In line with Alesina et al. (2016), we focus on the year 2000 and establish the baseline effects, which is also the year MDGs were set. As robustness, we also test the relationship for the years 1992, 2006, and 2012. However, we limit our analysis to these four years owing to the data limitation. As an additional exercise, we pool the data for these four years and perform an analysis subject to specifications with year and region/country fixed effects, to which our findings remain robust.

[4]EI index based on ethnic groups is constructed using the GREG map from Weidmann et al. (2010) that provides the location of around 930 ethnic groups, globally; whereas, EI index based on linguistic groups are constructed using Ethnologue map from Gordon and Raymond (2005) that contains the locations of around 7,600 language groups. A detailed discussion regarding the inequality indices construction is provided in Section 3.2.

[5]Spatial inequality index measures inequality in development across 2.5 x 2.5-degree spatial grids. As the correlation between ethnic and spatial inequality is high with a correlation coefficient of 0.7494 (source: Authors' calculations based on data from Alesina et al. (2016)), it is essential to control for SI to isolate the association between EI and outcomes considered, especially if SI is also correlated with outcomes.

[6]A brief discussion of each of these variables is provided in Section 3.2.

[7]One exception is the estimates of the Ethnologue index being insignificant to the inclusion of linguistic segregation. While the coefficient of the Ethnologue index remains positive, the higher standard errors can be attributed to the smaller sample size as we lose around 50% of the sample.

explanatory power.[8]

It is important to note that while we are unable to employ any random variation in ethnic inequality, we control for various relevant sets of covariates to isolate the relationship between EI and U5MR, in a similar vein to Alesina et al. (2016). First, we include real per capita Gross Domestic Product (GDP), spatial inequality, and other measures of inequality related to the land or population size of a country in our specifications. Second, we account for ethnic diversity measures related to fragmentation, polarization, segregation, and inequality in land and population size. Third, we add geographic features (such as terrain ruggedness, percentage of a country with fertile soil, and tropical climate) and historical features (whether had a British colonial past and historical population density) of a country that can contribute to its contemporaneous inequality. Fourth, we perform a robustness exercise by pooling data for the years 1992, 2000, 2006, and 2012 and analyse the relationship between EI and U5MR subject to year fixed effects and region/country fixed effects. Fifth, we perform various exercises to address any potential migration concerns. The relationship between EI and U5MR remains stable in all these exercises. Controlling for these potential confounders helps reduce the omitted variable bias concerns further, and any potential bias left is likely to be minimal.

Having established the relationship between U5MR and EI, we turn our attention next to other health outcomes such as maternal mortality (MM), life expectancy at birth (LEAB), stillbirths, and child stunting. EI is found to share a detrimental association with these health indicators as well. Next, we explore potential mechanisms through which EI can affect health. Several studies have suggested that vaccination rates (Levine and Rothman, 2006), education (Black et al. 2008; Chen and Li 2009) and income (Pritchett and Summers 1996; Braendle and Colombier 2016) are significant determinants of health. In light of these, we examine the association between EI and potential mechanisms such as the GDP of a country, child vaccination rates, school enrollment, fertility, and contraceptive usage. We also perform an analysis in which we introduce interaction terms between EI and these potential mechanism variables. Findings provide suggestive evidence that lower vaccination rates, reduced contraceptive usage (which can result in unwanted fertility), and higher fertility rates are the channels through which EI can affect health.

Owing to the productivity benefits of heterogeneity in skills and a better ability to cope with conflict elements (Alesina et al. 2003; Alesina and Ferrara 2005; Bluhm and Thomsson 2020), developed countries may find ethnic diversity to be beneficial. Therefore, we focus next on whether the relationship between EI and health outcomes, along with the mechanisms established earlier, are heterogeneous to the income levels of countries. Using the World Bank's income classification for the year 2010, we explore the relationship for the group of countries

[8]In a recent and related work, Desmet et al. (2020) has noted that ethnic-linguistic fractionalization (ELF) is negatively associated with public goods provision, while a stronger local interaction among different groups (local-global ethnolinguistic complementarity, LGC) in society improves the outcome. We also examine the robustness of our results by controlling for the ELF and LGC measures from Desmet et al. (2020). Our results remain robust to this alternative exercise as well.

based on their income levels. Countries are divided into three categories – developing countries (a sample of low- and low-middle-income countries, henceforth, DCs), upper-middle-income countries (UMICs), and high-income countries (HICs). Results suggest that DCs are affected the most, while the strength of association decreases when we move up in the income spectrum. Especially HICs seem mostly unaffected by ethnic inequality in terms of health outcomes. However, this does not imply that higher inequality has no burden on high-income countries, as inequality may have implications beyond health outcomes. We further perform an analysis based on geographic regions and observe that child mortality in Sub-Saharan African (SSA) countries, Latin American and Caribbean countries (LAC), and East Asia and Pacific Countries (EAP) are strongly associated with ethnic inequality.

This study contributes to the literature by providing evidence of the relationship between ethnic inequality and public health across the globe. We also lay out potential mechanisms through which ethnic inequality can affect public health. To the best of our knowledge, this is the first study to focus on the association between ethnic inequality and public health outcomes – especially, we apply ethnic inequality indicators that are constructed based on nightlights measured from space that are less prone to any government manipulations (Martinez, 2022) and are subject to lower measurement errors that are generally found in economic growth measures available from sources such as Penn World Tables (Johnson et al., 2013).[9] The findings of the study can be of significance to policymakers to help better allocate the resources to improve health outcomes.

The rest of the paper is organized as follows: Section 2 provides a brief summary of the literature, Section 3 details the data used, Section 4 explains the estimation method, Section 5 presents the results, and Section 6 concludes.

2 Literature

The studies so far have mainly focused on the association between health and income inequality, i.e., measures of inequality that are not based on ethnic regions (henceforth, inequality measures that are not based on ethnic regions are simply referred to as income inequality or spatial inequality in this study).[10] Income inequality has been found to result in lower economic growth (Alesina and Rodrik 1994; Persson and Tabellini 1994), higher mortality and lower life expectancy (Mayer and Sarin 2005; Elgar 2010), and higher inflation (Albanesi, 2007). The sensitivity of changes in health or poverty to income is found to be a decreasing function of income inequality (Biggs et al. 2010; Fosu 2010). In a recent study, Ponnusamy (2022) has

[9]In a study on the overstatement of economic growth in non-democracies, Martinez (2022) finds that the authoritarian regimes artificially inflate their yearly GDP growth measures by approximately 35%.

[10]Income inequality measure that relies on nightlights measured from space as a proxy for local economic development is generally referred to as spatial inequality. In this study, our ethnic inequality indicators are based on nightlights within ethnic regions, whereas spatial inequality measures rely on nightlights within 2.5x2.5 degree spatial grids.

shown that the effects of droughts on under-five mortality tend to be higher in regions with higher income inequality, especially in low-income group of countries. On the other hand, the studies that focus on the relationship between ethnic inequality and societal or developmental outcomes are limited, with the majority of studies focused on ethnic heterogeneity.

Various studies have demonstrated the beneficial effects of ethnic heterogeneity. Ethnic diversity or heterogeneity is generally referred to as the presence of multiple ethnic or language groups in a society, whereas ethnic inequality is the inequality in economic development across ethnic or linguistic groups in a country.[11] Ethnic diversity is correlated with an abundant talent pool as each ethnic group can be specialized in specific skills, which can result in higher levels of innovation and thereby, the invention of new technologies (Fafchamps 2000; Churchill et al. 2020). Due to the positive association between ethnic heterogeneity and special skills, areas with higher diversity also contain more agricultural markets and heterogeneous products, creating an incentive to exchange goods and help promote economic development near ethnic borders (Michalopoulos, 2012). Ethnic diversity has also been linked to higher economic growth at a smaller geographic level (Montalvo and Reynal-Querol, 2021), and increased productivity along with the availability of more intermediate goods due to skill complementarity (Spolaore et al. 2000; Alesina and Ferrara 2005). However, there is also significant evidence that points to the detrimental effects of ethnic diversity. Ethnic fragmentation is associated with the lack of social and productive capital (Collier and Gunning, 1999), poor economic performance (Easterly and Levine, 1997), a lower institutional quality and schooling attainment (La Porta et al. 1999; Alesina et al. 2003), lower public goods provision (Alesina et al., 1999) and reduced labor productivity (Hjort, 2014). In summary, the evidence of the impact of ethnic diversity is mixed.

As the above studies have focused on ethnic diversity, some of the recent literature has directly examined the relationship between ethnic inequality and relevant outcomes. Especially the development of satellite nightlight data has contributed to the development of ethnic inequality indices that use nightlights measured from space as a proxy for economic development. Using the ethnic inequality index constructed based on satellite data, Alesina et al. (2016) shows that the negative association between inequality and economic growth is a global phenomenon, even though the effects are more pronounced in Africa. Ethnic inequality is also found to be detrimental to public goods provision (Baldwin and Huber, 2010), diminished government quality (Kyriacou, 2013), and an increase in political parties' ethnic transformation (Huber and Suryanarayan, 2020). While most of these studies on ethnic inequality have focused on its association with economic growth or political outcomes, the evidence on the relationship between inequality across ethnic groups and public health outcomes is still scarce, especially using inequality indices constructed based on nightlights, which are considered to

[11] Montalvo and Reynal-Querol (2005) note that the definition of ethnic or racial identities is fluid to some extent, and goes on to define ethnicity as something based on a purely biological or genetic perspective.

be highly reliable sources of economic activity at sub-national levels.[12]

Moreover, the relationship between ethnic inequality and health outcomes can be heterogeneous at different income levels. Especially in developed countries, the detrimental effects of income inequality are alleviated by the demand for skill complementaries and by the better quality of political institutions, which are instrumental in dampening societal conflict elements arising due to the multitude of ethnic groups (Alesina and Ferrara 2005; Bluhm and Thomsson 2020). Therefore, along with examining the association between ethnic inequality and health outcomes for the global set of countries, we also put emphasis on their income levels to identify the heterogeneous effects.

3 Data

3.1 Outcome Variables of Interest

Our objective is to understand the relationship between ethnic inequality and various public health outcomes. The primary outcome variable of interest is U5MR, expressed as the number of under-five deaths per 1,000 live births for a given year. Mortality rates are sourced from WDI (2020) for 173 countries. Following Alesina et al. (2016), we examine the relationship between EI and U5MR for the year 2000 in our baseline analysis. Considering this is the year in which MDGs were set, it can provide us with an understanding of the state of public health in the year of their implementation.

While the primary interest of this study is on the relationship between EI and U5MR, we also examine the association between EI and other health measures that can indicate the health status of the population in a country. First, we use data on maternal mortality, expressed as the number of maternal deaths per 100,000 live births. Second, we focus on the stillbirths rate from Hug et al. (2021), indicated as the number of stillbirths per 1000 total births.[13] Finally, we focus on the life expectancy at birth and the prevalence of child stunting. Child stunting is defined as the percentage of children under the age of five who are two standard deviations below the median in terms of height for their age. To explore potential mechanisms, we focus on child vaccination rates (measles and DPT), primary and secondary schooling enrollment indicative of the level of human capital development in a country, the prevalence of modern contraceptive methods that can help avoid unwanted fertility, and finally, fertility rates of women in a country.

[12] Even the most widely used measure of economic growth and welfare at a country level, i.e., Gross Domestic Product, is not accurately measured, especially in developing countries, economic activity data for ethnic groups or at sub-national levels is generally unavailable (Henderson et al., 2012). Therefore, using income or ethnic inequality indices based on satellite data is preferred and sometimes the only available source, especially considering the uncertainty in many of the income estimates used for international comparisons (Deaton and Heston, 2010), and the inherent measurement error in some of the commonly used economic growth data such as Penn World Tables (Johnson et al., 2013).

[13] Stillbirths is defined as "the birth of a baby following fetal death before labor (antepartum stillbirth) or during labor or birth (intrapartum stillbirth)," see Hug et al. (2021).

Except for the data on stillbirths, the rest are sourced from WDI (2020). We also consider GDP per capita from Penn World Tables (PWT), version 7.0, as a mechanism channel. Table 1 contains summary statistics for the outcome variables considered in this study, along with information on the inequality indices and other health or development indicators used. In the last column, we briefly describe the variables used.

3.2 Inequality Indices

We apply two different ethnic inequality indicators (GINI coefficients),[14] available for 173 countries from Alesina et al. (2016), constructed based on the locations of ethnic and linguistic groups, respectively. The first source is the Geo-referencing of Ethnic Groups (GREG) map, a digitized version of the Soviet Atlas Narodov Mira from Weidmann et al. (2010). GREG contains the locations of 928 ethnic groups globally for the year 1964. Alesina et al. (2016) match the ethnic locations from GREG to the political boundaries of the 2000 Digital Chart of the World to identify the ethnic locations within each country. The second source is the 15th edition of Ethnologue from Gordon and Raymond (2005), which outlines the locations of 7,581 language groups within countries for the mid/late 1990s. Similar to GREG, the Ethnologue boundaries are superimposed on the political boundaries for the year 2000 to identify within-country locations of linguistic groups. Alesina et al. (2016) also provides spatial inequality indices that are measured at different spatial or administrative levels, which we control for in this study to isolate the relationship between EI and health outcomes.[15]

There are a few differences in terms of aggregation between the GREG and Ethnologue maps. As Ethnologue considers the regions where languages overlap, the level of linguistic aggregation in Ethnologue is much more refined than in GREG. On the other hand, while GREG focuses on major immigration groups, Ethnologue focuses on pure indigenous groups. In Canada, for example, Ethnologue provides locations of 77 groups that are primarily indigenous (such as the Babine and the Cayuga). In contrast, GREG provides the locations of 23 groups that are primarily non-indigenous (such as the Swedes and the Norwegians). While GREG and Ethnologue provide different measures of ethnic homelands (at least for some continents), we do not try to distinguish which one is more significant. Instead, in a similar vein to Alesina et al. (2016), we use inequality indices based on both GREG and Ethnologue and examine their association with health outcomes. Panel B of Table 1 provides descriptive statistics

[14]The Gini coefficient is a widely accepted measure of economic inequality which aims to quantify the disparities in income, wealth, and development across a population (Deaton, 1997). It indicates the average difference in income or wealth between all pairs of individuals in the population. Conventionally, disposable income is considered to capture living standards and is used to construct the Gini coefficients. Instead of income or wealth data, we use nightlight data as a proxy for economic development. Hence the Gini coefficients used in this study represent development inequalities across ethnic homelands or the average difference in development between ethnic or language groups in a country. It is a suitable measure for our analysis as it shows the relative inequality among a country's ethnic groups and also due to the common familiarity with its interpretation.

[15]We provide a brief description of the inequality index construction in Section A1 of the online appendix. Refer to Alesina et al. (2016) for a detailed summary.

Table 1: Descriptive Statistics

Variable	Mean	Standard Deviation	Source	Description
Panel A: Outcome Variables of Interest				
Under-5 Mortality Rates	56.439	57.381	WDI (2020)	Under-5 Mortality Rates, expressed as, the number of child deaths per 1000 live births
Maternal Mortality Rates	273.351	391.421	WDI (2020)	Maternal Mortality, expressed as, the number of maternal deaths per 100,000 live births
Number of Stillbirths	14.923	10.119	Hug et al. (2021)	Stillbirths per 1000 total births
Child Stunting	26.635	16.144	WDI (2020)	Prevalence of Stunting, height for age (% of Children under age five)
Life Expectancy at Birth	66.387	10.354	WDI (2020)	Life Expectancy at Birth, total (years)
Panel B: Inequality Measures				
Ethnic Inequality (GREG)	0.424	0.259	Alesina et al. (2016)	Inequality across ethnic groups measured using GREG Mapping
Ethnic Inequality (Ethnologue)	0.446	0.333	Alesina et al. (2016)	Inequality across linguistic groups measured using Ethnologue Mapping
Spatial Inequality	0.421	0.269	Alesina et al. (2016)	Inequality across 2.5 x 2.5 degree pixels
Number of Ethnic Groups (GREG)	10.994	13.579	Alesina et al. (2016)	Total number of ethnic homelands in a country
Number of Ethnic Groups (Ethnologue)	38.620	94.029	Alesina et al. (2016)	Total number of linguistic homelands in a country
Inequality in Population	0.632	0.282	Alesina et al. (2016)	Inequality in population clustering across ethnic groups
Inequality in Land Size	0.615	0.266	Alesina et al. (2016)	Inequality in the size of ethnic homelands
Panel C: Other Health/Economic Indicators				
Measles Vaccination	81.012	18.848	WDI (2020)	% of Children aged 12-23 months who received the measles vaccine 12 months before the survey
DPT Vaccination	81.659	19.477	WDI (2020)	% of Children aged 12-23 months who received the DPT vaccine 12 months before the survey
Contraceptive Use	33.455	20.721	WDI (2020)	% of married women ages 15-49 practicing one modern method of contraception.
Fertility Rates	3.302	1.754	WDI (2020)	Total births per women
Primary Schooling Enrollment	84.241	17.359	WDI (2020)	Ratio of primary school enrollment to the total population of the age group
Secondary Schooling Enrollment	63.304	26.809	WDI (2020)	Ratio of secondary school enrollment to the total population of the age group
Log GDP Per Capita	8.461	1.357	PWT 7.0	Natural log of real per capita GDP (Chain Index)

Note: The table provides the summary statistics for the main outcome variables of interest, inequality measures used, and other relevant health/development indicators for the group of countries used in the analysis for the year 2000. Panel A lists the main outcome variables of interest, panel B contains descriptive statistics for ethnic and spatial inequality indices, and panel C provides information for the other relevant health or development indicators used. Refer to WDI (2020) for a detailed description of the health indicators and schooling enrollment variables.

for the ethnic inequality measured using the GREG and Ethnologue mappings and other inequality measures controlled for in this study. A brief description regarding the construction of inequality indicators is provided in Section A1, online appendix, to preserve space.

Ethnic inequality indices employed in this study measure income inequality across ethnic regions. There are other measures of ethnic heterogeneity, such as ethnic-linguistic fragmentation indices, segregation, and polarization indices, which, if excluded, can confound the association between EI and health outcomes. Therefore, we also control some of these relevant variables to test the sensitivity of our baseline results while also confirming whether other aspects of ethnic diversity matter for public health. First, we employ ethnic-linguistic fractionalization indices from Alesina et al. (2003) and Desmet et al. (2012) that provide the probability that two randomly selected individuals from the population belong to different ethnic or linguistic groups, respectively. The index ranges from zero to one, with a value of zero (one) implying that each individual in the population belongs to the same (a different) ethnic/linguistic group. Second, the cultural fractionalization index adopted from Fearon (2003) captures cultural distances between groups in a country using structural distances between languages as a proxy.

Third, the ethno-linguistic polarization index is sourced from Montalvo and Reynal-Querol (2005). While the ethnic fractionalization index increases when the number of ethnic groups in a country goes up, the polarization index increases when there are "few (equally) large groups with homogeneous characteristics within each group, and differences in a cluster of characteristics among groups," see Østby (2008). The polarization index is at its maximum value when there are two large groups of exactly equal sizes in the population. Next, the ethno-linguistic segregation index from Alesina and Zhuravskaya (2011) measures the extent to which ethno-linguistic groups live spatially separately. If the segregation index is of value one, then each group lives in a separate region in a country, i.e., full segregation, whereas the index takes on a value zero if the fraction of each group in every region is similar to that of the entire country, therefore no segregation. Finally, we employ predicted genetic diversity measures from Ashraf and Galor (2013) that give us the probability that two individuals drawn randomly from a population differ genetically from each other.

4 Estimation Method

To establish the correlation between ethnic inequality and health outcomes, we employ a cross-country level specification estimated by ordinary least squares (OLS) as below:

$$Y_{ij} = \beta_0 + \beta_1 EI_{ij} + \beta_2 SI_{ij} + \delta \mathbf{X}_{ij} + \alpha_j + \varepsilon_{ij} \quad (1)$$

Y_{ij} refers to the under-5 mortality rates in country i from region j for 2000, our primary outcome variable of interest.[16] We choose the year 2000 for our analysis as this was the year

[16]Following Alesina et al. (2016), we classify the countries in our sample into seven different geographical

MDGs were set. However, we perform the analysis individually for 1992, 2006, and 2012 as a robustness exercise. Along with U5MR, we also focus on various other outcomes such as maternal mortality, stillbirths, child stunting, and LEAB, which can be potential indicators of public health status in a country.

Our explanatory variable of interest EI_{ij} indicates the level of ethnic inequality in the country i. The coefficient of interest is β_1, providing the relationship between ethnic inequality in a country and its health status. If the consequences of ethnic inequality are detrimental, then increases in EI can result in worse health outcomes. This can result in β_1 from Equation 1 being positive for all the outcomes except for LEAB, for which the coefficient will be negative, i.e., the higher the EI, the lower the LEAB. Along with ethnic inequality, we also control for SI_{ij}, a measure of the spatial inequality in the country i. In the baseline analysis, spatial inequality measured across equal-sized pixels is used. As robustness, we also control for SI observed across first- and second-level administrative regions.

$\mathbf{X}_{ij}$ is a matrix of other covariates which we control for, to isolate the relationship between EI and health outcomes. In line with the literature, we account for the number of ethnic or linguistic groups in a country, inequality indices based on the population size to account for the potential unequal clustering of the population across ethnic homelands, inequality based on the size of ethnic homelands, along with the land and population size in the baseline specifications. To ensure that other measures of ethnic composition do not confound the estimates of EI, we include ethnic-linguistic fragmentation, indices related to segregation, polarization, cultural fragmentation, and genetic diversity. We also include a set of geographic endowments such as average terrain ruggedness, the percentage of fertile soil in a country, the percentage of a country with a tropical climate, gem-quality diamond extraction index, and the average distance to the nearest ice-free coast, along with a set of historical controls such as whether the country had a British common-law system, population density circa 1500 CE (in log form), and the log of timing since the Neolithic transition.

α_j in Equation 1 refers to the region-fixed effects to account for common factors that can affect all the countries within a geographic region. Finally, ε_{ij} includes time-varying unobservable shocks to the outcome variable. We perform various robustness tests to examine the sensitivity of our findings, including an exercise in which we pool the data for years 1992, 2000, 2006, and 2012 and perform an analysis subject to year-fixed effects and region/country fixed effects. Our results remain robust to all the exercises employed, as explained in Section 5.4.

regions: East Asia and Pacific Countries (EAP), Eastern Europe and Central Asia (ECA), Latin America and Caribbean countries (LAC), Middle-East and North African countries (MENA), Central Asian countries (SA), Sub-Saharan African countries (SSA) and Western European countries (WE).

5 Results

5.1 Baseline Results

5.1.1 Descriptive Evidence

First, we show the association between ethnic inequality and child mortality subject to both unconditional and conditional (on region-fixed effects) specifications graphically. Figure 1 produces the association between the two variables as scatter plots. The top panels use the EI index based on the GREG mapping, and in the bottom panels, the EI index constructed using the Ethnologue mapping is applied. The left (right) panels show unconditional (conditional) linear relationships.[17] From Figure 1, it is evident that there exists a positive, linear relationship between EI and U5MR, irrespective of the type of index or the specification applied. While the strength of the association drops slightly once the regional fixed effects are introduced, a positive correlation between ethnic inequality and child mortality persists.[18]

An interesting observation from Figure 1 is the variation around the trend line. For the same levels of inequality, we observe a large variation in U5MR. For example, around the Gini levels of 0.60 in Figure 1a, several countries perform worse (better) in terms of child health, i.e., lie above (below) the trend line. First, we provide summary statistics that can help explain this relationship. We present mean values of several potential intermediate channels that can affect child health, such as measles and DPT vaccination rates, contraceptive use, fertility rates, schooling enrollment, and the level of economic development in a country for four different samples: the group of countries that are above/below the median in terms of ethnic inequality and above/below the regression line that predicts under-5 mortality.

The summary statistics are provided in Table 2. Comparing columns 1 and 2, which provide statistics for the group of countries with higher EI (above the median), the worse-performing countries in child health (Column 1 of Table 2) are also the ones that perform poorly in the intermediate health/development outcomes. In contrast, the status of indicators in the better-performing group of countries (Column 2 of Table 2) is more favorable. A similar pattern is observed in the last two columns as well. As these statistics provide some suggestive evidence that it is possible to attain a relatively superior health status even with higher levels of inequality, we perform a mechanism analysis in Section 5.3 to identify the health indicators through which

[17]The analysis of the conditional relationship between ethnic inequality and child mortality is performed by netting out the effect of region-fixed effects from ethnic inequality and child mortality, respectively, and then plotting the residuals from the two regressions.

[18]Note that for visual comparison of the unconditional and conditional relationships, the scales of the two plots vary. This is because original variables are used in the unconditional plots, and the residuals (after netting out region-fixed effects from the variables) are used in the conditional plots. To make the scales comparable, we transform the residuals to have the same scale as the raw variables, i.e., ethnic inequality and child mortality. This is done by standardizing the residuals and then re-scaling them to have the same mean and standard deviation as the ethnic inequality and child mortality variables. This can help the visual inspection of the conditional relationship comparable to the unconditional relationship. The re-scaled graphs are provided in Figure A1 in the online appendix to preserve space. A positive association between U5MR and EI persists.

EI affects child health the most.

There are a couple of potential questions. First, how are countries such as US or Brazil that are at similar or even higher levels of inequality but can perform well in development indicators than poorer countries. Using Sri Lanka and Hong Kong as examples, Ray (2014) notes that Hong Kong does better than Sri Lanka simply because of the availability of a higher amount of resources at its disposal. Therefore, developed countries may be able to perform well compared with developing countries that are at similar (or even lower) levels of inequality, owing to the higher levels of economic resources at hand, which can translate into improved health and nutritional standards among the population (Ray, 2014).

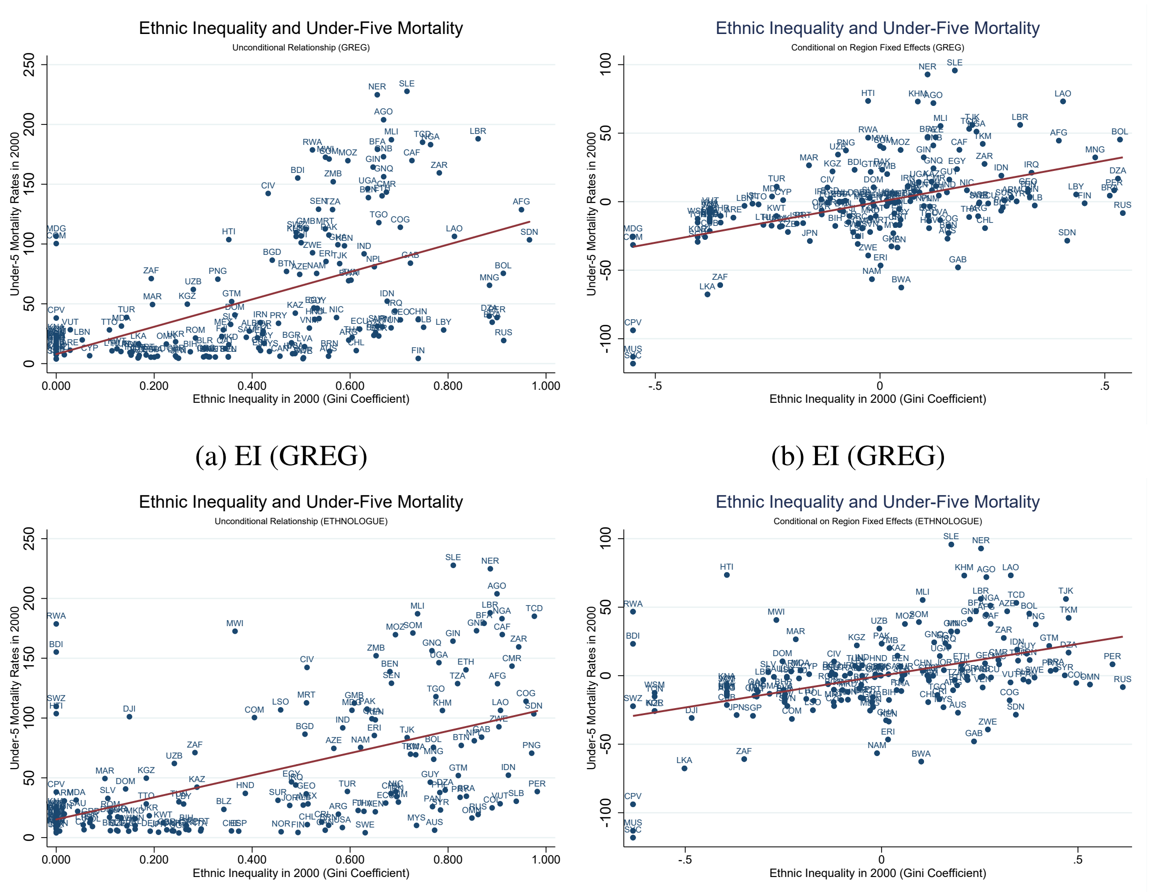
The second possible question is how some countries at lower economic development levels can perform well in indicators such as vaccination rates, schooling enrollment, and other outcomes. An example is that of Bangladesh, a developing country that is similar in economic development compared with low-income countries but has performed strongly in several indicators of development. Compared with the LICs, which have an average under-5 mortality of around 153 deaths per 1000 births, Bangladesh experiences 86.5 deaths per 1000 live births. Likewise, Bangladesh has also performed efficiently in other indicators such as maternal mortality, LEAB, vaccination rates, and schooling enrollments.

There are a couple of explanations, among others, that can be attributed to these differences: first, Bangladesh has relatively low levels of ethnic (or income) inequality compared with LICs – while Bangladesh experiences an EI of 0.441, the EI of LICs are higher with an average of 0.644. This can partially explain how the poorest population in a country is able to access key health resources, which can result in better health status overall. Second, other factors such as social and economic empowerment of women which can help reduce child mortality significantly, the promotion of popular health and science education programs that can result in improved health outcomes in the long run, government policies concerning education and health, and the public demand for such policies may contribute towards better health status on top of the underlying levels of inequality (Ray, 2014). While we estimate the average correlation between ethnic inequality and health outcomes in this study, we acknowledge that there may be other factors that can cause outliers, i.e., a few countries that perform well than others but with similar levels of inequality or economic development.

5.1.2 Regression-based Estimates

Next, we analyze the relationship between U5MR and EI under cross-country OLS regression. The results for specifications with GREG index are provided in Table 3, whereas the findings subject to the Ethnologue inequality index are presented in Table A1, online appendix to preserve space. The dependent variable is U5MR, and we control for real GDP per capita in all the specifications. Column (1) of these tables controls only for the EI indices along with

Figure 1: Relationship between Ethnic Inequality and U5MR



(a) EI (GREG) (b) EI (GREG)

(c) EI (Ethnologue) (d) EI (Ethnologue)

Figure 1: The plot provides the unconditional relationship and also the conditional (on regional fixed effects) relationship between ethnic inequality and child mortality. The top two panels apply the ethnic inequality indicator measured using the GREG mapping, and in the bottom two panels, the inequality indicator based on the Ethnologue mapping is used. The left and right panels are in different scales as the original variables are plotted against each other in the unconditional plots, whereas the residuals after netting out the effect of region-fixed effects from ethnic inequality and child mortality are plotted in the conditional plots.

GDP, and the results suggest a strong, positive association between EI and U5MR, in line with Figure 1. To provide a quantitative interpretation of the estimates, one standard deviation increase in the GREG ethnic inequality index, i.e., an increase of 0.259 units in the GREG index (refer to panel B, Table 1), is associated with an increase of 13.40 child deaths per 1000 births. The magnitude of this change equates to 23.74% of the mean dependent variable.[19]

[19]A potential concern is that the GDP per capita controlled for in our specifications in Table 3 correlates with the nightlights data used to construct the GINI coefficients. Therefore, we perform a robustness exercise using the mean GDP constructed as the average of the first to the fifth lag of the real GDP per capita (i.e., average GDP for the years 1995 to 1999, as the U5MR used in Table 3 is for the year 2000) instead of the current GDP. Based on results from Table A11 in the online appendix, the coefficients of Ethnic Inequality (GREG Index) remain close to the coefficients of EI in Table 3 of the main text. This exercise ensures that using the current GDP or mean of

Table 2: Summary Statistics: Ethnic Inequality and U5MR

Variables	Above Median EI and Above Line U5MR	Above Median EI and Below Line U5MR	Below Median EI and Above Line U5MR	Below Median EI and Below Line U5MR
	(1)	(2)	(3)	(4)
Measles Vaccination	57.77	89.05	85.09	91.20
DPT Vaccination	58.18	88.29	86.15	92.89
Contraceptive Use	16.46	40.62	38.41	51.21
Fertility Rates	5.55	2.74	3.14	2.00
Primary Schooling	62.61	91.78	85.80	93.95
Secondary Schooling	26.04	64.18	64.24	80.97
GDP	6.83	8.68	8.52	9.58

Note: The table provides the mean of the variables for four different samples: the group of countries that are above/below the median in terms of ethnic inequality (GREG Index) and above/below the regression line that predicts under-5 mortality.

To account for any regional disparities in economic development, we next control for Spatial inequality in the model, i.e., GINI coefficients calculated using a spatially overlapped global grid of 2.5 × 2.5-degree pixels. As EI and SI measures are highly correlated, with a correlation coefficient of around 0.75, it is essential to account for the latter to isolate the association between EI and U5MR. In column (2) of Table 3 and Table A1, we account only for the spatial inequality. When SI is controlled for individually, there seems to be a strong relationship between U5MR and SI. But once EI is included in the specification, the significant relationship observed for SI in column (2) ceases to exist (refer to column (3)). This suggests that when regressed in isolation, the coefficient of SI may be capturing the effects of EI.

So far, we controlled for SI that measures inequality across equal-sized spatial grids. Next, we perform a robustness exercise in which we apply spatial inequality indices constructed using nightlights across administrative regions. The qualitative interpretation of the ethnic inequality index remains unaffected, and a detailed discussion is provided in Section A4.3, online appendix, to preserve space.

In columns 4–6, we control for the effect of ethnolinguistic heterogeneity (ethnic fractionalization) on U5MR by accounting for the number of ethnic groups in a country (in log form). In Table 3, we include the number of ethnic groups, and in Table A1, we control for the number of linguistic groups. Results based on column (4) in both tables suggest that the unconditional association between ethnic-linguistic fractionalization and U5MR is positive. However, once we account for the ethnic inequality in columns (5) and (6), the significant effects observed for the fractionalization indices in column (4) disappear, while a strong association between EI and U5MR persists. This suggests that it is income differences along ethnic lines that correlate with child mortality and not ethnic fractionalization.

Further, in Columns 7–8, we include two more inequality measures constructed based on the unequal clustering of populations across different ethnic homelands and asymmetry in the size of ethnic homelands. Controlling for these inequality indices does not affect the estimates

the lagged GDP values has negligible effects on the coefficient of interest. We thank an anonymous referee for providing us with this valuable suggestion.

of the main variable of interest in either of the tables. In Column 9, we introduce the size of land area and population in a country, along with the other variables that were included in columns 1–8. It is essential to account for the size of the countries in terms of their land area and population since bigger countries could have greater diversity and a higher fractionalization (Alesina et al. 2016; de Soysa and Almås 2019; Yong 2019). The coefficient of EI remains unaffected even after controlling for all these potential confounders. Therefore, accounting for various other measures that can affect child health does not have an impact on the association between EI and U5MR.[20]

5.2 Controlling for Other Measures of Ethnic Composition

Several studies have noted the detrimental effects of ethnic diversity, such as growth-reducing public policies (Easterly and Levine, 1997), inferior government performance (La Porta et al., 1999)), lower spending on productive public goods (Alesina et al., 1999), and civil wars or corruption (Alesina and Ferrara, 2005). Moreover, ethnic-linguistic fractionalization is found to be positively associated with ethnic inequality, with a correlation coefficient in the range of 0.35–0.45 (refer to Alesina et al. (2016)). Therefore, other forms of ethnic composition can be correlated with our outcomes of interest. Not accounting for these measures could lead to biased estimates of ethnic inequality, as the detrimental association between EI and U5MR could indeed be confounded by the omitted ethnic fractionalization measures. To ensure it is not the case, we extend our analysis by incorporating other aspects of ethnic composition and present the findings in Table 4.

Columns 1-5 contain results for the EI index constructed based on GREG mappings, and Columns 6-10 provide results for Ethnologue mappings. In columns (1) and (6), we include ethnic and linguistic fractionalization indices from Alesina et al. (2003) and Desmet et al. (2012), respectively. In both cases, EI maintains its statistical significance. In columns 2 and 7, we include cultural distances of ethnic groups from Fearon (2003), whereas, in columns 3 and 8, we control for ethnic polarization index from Montalvo and Reynal-Querol (2005). In all these specifications, estimates of ethnic diversity remain insignificant, while the qualitative interpretation of ethnic inequality indices remains unaffected.[21]

[20] We also control for variables such as the percentage of women in parliament that can act as a proxy for the social attitudes towards women in a society, the level of corruption in a country which can be indicative of the attitudes of government towards development policies, and the amount of Research and Development expenditure in a country (as a % of GDP) which can account for a country's inclination towards investing in scientific advancements. The association between EI and U5MR remains positive and significant to these exercises as well. Data on corruption is from the Transparency International database, whereas the other two variables are sourced from WDI (2020).

[21] Gomes (2020) finds that children of mothers ethnically distant from their neighbors display worse health outcomes. Therefore, controlling for the cultural fractionalization index from Fearon (2003) can help alleviate some of the concerns due to linguistic distance across groups confounding the estimates of EI.

Table 3: Baseline Results - Ethnic Inequality (GREG) and U5MR

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnic Inequality	51.74***		50.76***		49.49***	48.34***	58.77***	54.89***	55.28***
	(9.93)		(12.86)		(12.04)	(15.54)	(16.47)	(16.63)	(18.54)
GDP Per Capita	-0.83***	-1.04***	-0.83***	-0.88***	-0.82***	-0.82***	-0.85***	-0.81***	-0.81***
	(0.25)	(0.28)	(0.25)	(0.29)	(0.26)	(0.26)	(0.25)	(0.26)	(0.26)
Spatial Inequality		36.33***	1.35			1.50	0.09	-0.20	7.84
		(10.30)	(12.50)			(12.63)	(12.42)	(12.51)	(17.58)
Log Number of Ethnicities				9.47***	0.77	0.79		3.86	3.34
				(2.39)	(2.54)	(2.58)		(3.65)	(3.93)
Ethnic inequality in population							-57.03*	-65.75**	-64.81**
							(29.20)	(30.07)	(31.18)
Ethnic inequality in size							54.32*	54.21*	54.62*
							(31.21)	(30.36)	(30.72)
Log Land Area									-2.06
									(2.53)
Log Population (in 2000)									2.00
									(2.31)
Adjusted R-Squared	0.7749	0.7525	0.7735	0.7542	0.7736	0.7722	0.7777	0.7777	0.7761
Observations	173	173	173	173	173	173	173	173	173

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates. The dependent variable is the under-5 mortality rates for the year 2000. Ethnic inequality indicator based on the GREG mappings is used. The coefficient of real GDP per capita is multiplied by 1000. All the regressions include region-fixed effects. Robust standard errors are reported in the parenthesis.

Columns 4 and 9 contain spatial segregation indices from Alesina and Zhuravskaya (2011), who show that the ethnically and linguistically segregated societies also have a lower quality of government. The ethnic segregation index used in column 4 of Table 4 provides us with a measure of how spatially mixed/segregated ethnic communities are in a country. While higher EI implies that certain ethnic groups in a country can be economically disadvantaged compared to the rest, if ethnic groups are geographically mixed, then it can be problematic for the interpretation of the effect of EI. Controlling for the ethnic segregation indicator from Alesina and Zhuravskaya (2011) in column 4 of Table 4 helps alleviate any potential bias due to ethnic groups coexisting in geographically mixed communities. We also perform an additional exercise in which we exclude capital cities while constructing the EI indicator, as ethnic mixing is more likely to occur in the capital cities (Tindale 2019; Yong 2019). Our results remain robust to this exercise as well, and a detailed discussion can be found in Section A4.6 of the online appendix.

In columns 5 and 10, we account for within-country genetic diversity (heterozygosity) from Ashraf and Galor (2013). We also include a squared term of heterozygosity to account for any non-linear relationship. The estimates of EI remain generally robust, qualitatively, in all these specifications.[22] The results from Table 4 suggest that the effect of EI on U5MR is not confounded by other ethnic diversity measures. Therefore the ethnic composition is associated with U5MR across countries mainly via ethnic inequality. As a further robustness exercise, we control for the ELF and LGC indices from Desmet et al. (2020), to which the correlation between EI and U5MR remains significant and positive.

5.3 Other Health Outcomes and Potential Mechanisms

In this section, first, we explore the relationship between EI and other relevant health outcomes such as maternal mortality, the number of stillbirths, LEAB, and child stunting. Refer to columns 1–4 in Table 5 for the results. The table also provides the mean dependent variable for ease of interpretation of estimates. From the results, it is evident that EI shares a significant, detrimental association with maternal mortality, stillbirths, child stunting, and LEAB. The results remain robust to the use of either GREG or Ethnologue mappings as seen in panels A and B of Table 5. To provide a quantitative interpretation of estimates, based on column 1, Panel A – one standard deviation increase in the GREG index is associated with an increase of 92.90 maternal deaths. Compared with the mean dependent variable, this equates to 33.98% of the mean outcome.

[22] An exception is the insignificance of the Ethnologue index in column 9, which can be owed due to the smaller sample size as we lose around 50% of the observations. We perform an additional exercise in which we control for all measures of ethnic diversity (except for linguistic segregation) in the same specification, and the estimates of EI remain positive and significant at 10%.

Table 4: Ethnic Inequality and U5MR - Controlling for Ethnolinguistic Fragmentation, Polarization, and Diversity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	GREG Index					Ethnologue Index				
Ethnic Inequality	45.03***	44.98***	51.07***	36.92*	54.55***	25.59**	25.89**	31.76***	19.22	33.24***
	(12.97)	(15.52)	(13.04)	(21.56)	(14.38)	(11.06)	(11.73)	(10.55)	(22.07)	(12.46)
GDP Per Capita	-0.92***	-1.35***	-0.82***	-0.96**	-0.76***	-0.85***	-1.52***	-0.82***	-2.39***	-0.80**
	(0.23)	(0.32)	(0.26)	(0.45)	(0.29)	(0.28)	(0.33)	(0.30)	(0.52)	(0.34)
Spatial Inequality	-1.12	-4.36	1.71	6.26	-1.29	14.71	5.72	13.54	13.06	9.81
	(12.60)	(13.31)	(12.46)	(17.48)	(12.22)	(10.11)	(10.09)	(10.11)	(15.23)	(10.09)
Ethnic-linguistic Fragmentation	23.55**					11.15				
	(9.52)					(9.45)				
Cultural Fragmentation		13.57					2.89			
		(10.45)					(11.16)			
Ethnolinguistic Polarization			-10.76					-14.37		
			(33.12)					(33.76)		
Ethnic-linguistic Segregation				6.78					-34.76	
				(28.39)					(29.60)	
Genetic Diversity					-1991.42					-2755.71
					(4190.95)					(4640.98)
Genetic Diversity Squared					1473.40					2067.95
					(3022.37)					(3342.85)
Adjusted R-Squared	0.7789	0.7885	0.7711	0.8251	0.7782	0.7657	0.7826	0.7626	0.7929	0.7687
Observations	173	150	172	96	157	173	150	172	92	157

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates. The dependent variable is the under-5 mortality rates for the year 2000. The coefficient of real GDP per capita is multiplied by 1000. All the regressions include region-fixed effects. Robust Standard errors are reported in the parenthesis.

Table 5: Ethnic Inequality and Other Health/Development Outcomes

	Maternal Mortality	Stillbirths	LEAB	Child Stunting	Log GDP	Measles Vaccination	DPT Vaccination	Schooling Primary	Schooling Secondary	Contraceptive Use	Fertility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
						Panel A: GREG Index					
Ethnic Inequality	358.69***	8.62***	-6.65***	16.70**	-1.31***	-8.02	-12.35**	-13.39**	-33.93***	-29.63**	1.26***
	(103.39)	(2.52)	(2.08)	(6.46)	(0.37)	(6.09)	(6.25)	(6.08)	(9.92)	(14.06)	(0.47)
Spatial Inequality	116.14	2.32	-2.09	6.63	-0.50	-9.94	-12.03	4.65	9.90	5.15	0.16
	(162.85)	(2.92)	(2.62)	(6.50)	(0.42)	(7.36)	(7.56)	(5.37)	(11.53)	(10.62)	(0.46)
Mean Dependent Variable	273.351	14.923	66.388	26.636	8.461	81.012	81.659	84.241	63.304	33.455	3.302
Adjusted R-Squared	0.6470	0.7500	0.8305	0.6249	0.6612	0.5116	0.5008	0.6026	0.6851	0.5515	0.7089
Observations	171	173	171	140	173	173	173	114	87	61	171
						Panel B: Ethnologue Index					
Ethnic Inequality	194.54**	4.02*	-3.63**	13.76***	-1.07***	-14.26***	-16.52***	-3.92	-38.09***	-29.56***	1.39***
	(78.20)	(2.23)	(1.79)	(4.43)	(0.28)	(4.78)	(4.84)	(6.94)	(8.45)	(8.23)	(0.37)
Spatial Inequality	219.12	5.16*	-3.99	8.97	-0.69*	-7.16	-10.55	-1.22	8.48	10.51	0.15
	(157.50)	(2.85)	(2.50)	(6.22)	(0.41)	(6.87)	(7.02)	(5.39)	(12.43)	(11.54)	(0.45)
Mean Dependent Variable	273.351	14.923	66.388	26.636	8.461	81.012	81.659	84.241	63.304	33.455	3.302
Adjusted R-Squared	0.6348	0.7369	0.8246	0.6303	0.6640	0.5325	0.5222	0.5878	0.7208	0.5783	0.7233
Observations	171	173	171	140	173	173	173	114	87	61	171

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates for the year 2000. All specifications control for population size and land size. Real GDP per capita is included in all columns, except Column 5. All the regressions include region-fixed effects. Robust Standard errors are reported in the parentheses.

Next, we focus on various possible channels through which EI can have an impact on the public health system. Several factors can determine the status of public health in a country, such as its levels of economic development (Braendle and Colombier 2016; Pritchett and Summers 1996), education (Chen and Li 2009; Black et al. 2008), contraceptive use (Mamdani et al., 1993) and vaccination rates (Levine and Rothman, 2006). Therefore, we examine the relationship between EI and these potential mechanisms and provide results in columns 5–11 of Table 5.[23] Based on the results, it is evident that a higher EI is associated with a poor state of these potential mechanisms such as lower levels of economic development and vaccination rates, reduced contraceptive usage, higher fertility rates, and poor schooling enrollment.

We perform a robustness exercise in which we interact EI with some of these mechanism variables (such as GDP, measles and DPT vaccination rates, total fertility rates, contraceptive usage, and schooling enrollment) to ascertain their role in affecting health outcomes such as U5MR, MM, stillbirths, child stunting and LEAB.[24] First, we create an indicator for each of these mechanism variables that take on one if its status is poor in a country and zero otherwise.[25] Then, we interact these binary variables with our ethnic inequality indicators. The results are provided in Table A2, online appendix, to preserve space. Columns 1–5 present the GREG index, whereas the last five columns apply the Ethnologue index.

Based on the results, the effects of EI are more detrimental in the group of countries that are performing poorly in terms of GDP, contraceptive usage, total fertility, and DPT vaccination.[26] Therefore, it may be possible to enhance the health status in these high inequality countries by improving the child vaccination rates and by increasing contraceptive use among the population. Efficient targeting of resources among ethnically disadvantaged groups might be required. This analysis helps provide further suggestive evidence on the role of mechanism variables through which ethnic inequality can impact key health outcomes.

We further perform a couple of heterogeneity analyses based on the income levels of countries, and their geographical locations. Findings suggest that the group of countries that are lower in terms of economic development, i.e., developing countries, are affected the most due to ethnic inequality. Based on geographical locations, we find that the SSA group of countries is affected the most, followed by Latin American and Caribbean countries and East Asia and Pacific countries. We present the findings and a brief discussion of the results in Sections A2

[23] WDI (2020) defines contraceptive prevalence rates as the percentage of married women between the ages of 15 and 49 who practice any modern contraceptive method, and fertility rates as "the number of children that would be born to a woman if she were to live to the end of her childbearing years and bear children in accordance with age-specific fertility rates of the specified year".

[24] We thank an anonymous referee for providing us with this valuable suggestion.

[25] For example, an indicator for GDP takes on one if the level of GDP in a country is below the median, and zero otherwise. For fertility rates, the indicator takes on one if the total fertility rate is above the median and zero otherwise. Except for vaccination rates, below or above the median is used to create indicator variables, whereas, for the vaccination rates, we use the 25th percentile as the cut-off. Using the 25th percentile for vaccination rates instead of the median is due to the high measles and DPT vaccination levels overall.

[26] While stillbirths seem to be higher in the group of countries performing worse in terms of measles vaccination (column 3 of Table A2), child and maternal mortality, LEAB and stunting outcomes appear to be higher in the group of countries that are performing poorly in terms of DPT vaccination (see columns 6–10 of Table A2).

and A3 of the online appendix to preserve space.

5.4 Robustness Exercises

We subject our findings to various robustness exercises to examine the sensitivity of results. First, we control for various geographic and historical features of a country that can affect the contemporaneous levels of health. Second, we address the potential migration concerns relevant to our study. Third, we apply alternate measures of spatial inequality constructed based on administrative regions. Fourth, we examine the relationship between EI and health for alternate years. We also pool the data for 1992, 2000, 2006, and 2012 and perform an analysis subject to year fixed effects and country/region fixed effects. Fifth, we address a potential measurement error concern in the ethnic or linguistic regions. Sixth, we use inequality indicators constructed by excluding certain regions, such as capital cities or smaller ethnic groups. Our results remain robust to all these exercises employed. A detailed discussion of the findings can be found in Section A4, online appendix.[27]

5.5 Limitations

To isolate the relationship between ethnic inequality and child health, we controlled for various potential confounders. To account for income effects (in levels), we added real GDP per capita. To control for the effects of regional disparities in income, we condition on spatial inequality measured at different spatial/administrative levels. We account for land and population size, which can reflect the effect of country size on inequality/health outcomes. Further, inequalities in land size and population size are included for as regressors. Then we condition on various other measures related to ethnic diversity, such as fragmentation, polarization, segregation, and genetic diversity. As the historical features of a country can affect the current levels of health or development, we condition on whether the country had a colonial past and its historical population density. As the geographical factors of a country might shape its inequality, several features such as terrain ruggedness, percentage of a country with fertile soil or tropical climate, distance to the coast, and the gem quality are controlled for. Finally, we add the percentage of women parliamentarians in a country, the level of corruption, and the amount of Research and Development expenditure in a country which can act as proxies for the social attitudes toward women in society, government attitudes toward development policies, and a country's inclination toward investing in scientific advancements, respectively. We also address any concerns

[27] Results for the specifications in which we control for geographic and historical factors of a country are in Section A4.1 and Table A5 of the online appendix. Migration concerns are addressed in Section A4.2. Results subject to alternate measures of spatial inequality are in Section A4.3 and Table A6 of the online appendix. The findings when alternate years are used are in Section A4.4 and Table A7, the use of pooled data are in Section A4.5 and Table A8, results subject to excluding certain areas in the EI index construction are in Section A4.6 and Table A9, and finally, discussion of findings where we address a measurement error concern in the EI index construction and its associated results are in Section A4.7 and Table A10 of the online appendix.

related to potential migration or measurement errors in ethnic-linguistic locations.

If some unobservables are correlated with both health and inequality, it can induce a bias in the estimates. In our case, inequality is, on average, likely to be correlated with negative unobservables, i.e., those that reduce health (in our case, an increase in child mortality). For example, higher ethnic inequality is positively associated with the level of corruption in a country, which can worsen its health outcomes. The omission of a relevant proxy for corruption can induce an upward bias in the estimates for the effect of EI. However, any bias due to unobservables is likely to be small in this study, as we control for a diverse set of factors related to ethnic inequality and health.

6 Conclusion

We examine the association between ethnic inequality and public health in this study. Specifically, we focus on the correlation between inequality in development across ethnic regions and various key health indicators such as child mortality, maternal mortality, LEAB, stillbirths, and child stunting. Using a large cross-country dataset for 173 developed and developing countries, we provide evidence that ethnic inequality has a detrimental association with the overall health status in a country. We find that one standard deviation increase in the GREG EI index is associated with an increase of 13.40 under-5 deaths per 1000 live births, equivalent to 23.74% of the mean outcome. The association between EI and these health outcomes is found to be stronger in developing countries, while high-income countries remain unaffected. Therefore, ethnic inequality may be detrimental to health, mainly to the group of countries at the lower end of the income distribution, while the developed nations are able to reap the benefits of skill complementarities associated with ethnic diversity. However, this does not imply that higher inequality has no impact on high-income countries, as ethnic inequality may have implications beyond health outcomes.

Along with EI, we control for income inequality measured using 2.5x2.5-degree spatial grids and other measures of ethnic diversity such as ethnolinguistic fragmentation, polarization and segregation indices, and geographic and historical endowments. A stronger association between EI and child mortality persists even after controlling for all these relevant variables, whereas a relationship between U5MR and other measures of ethnic diversity remains largely unseen. This suggests that ethnic diversity may affect health outcomes mainly via ethnic inequality. We also explore potential mechanisms through which EI can affect public health. Lower vaccination rates, schooling enrollment, GDP, and contraceptive use, along with increased fertility rates, are found to be associated with higher ethnic inequality.

This paper makes two major contributions to the literature: First, it is one of the first studies to explore the relationship between ethnic inequality measured using nightlight activity and public health; second, we provide evidence that ethnic inequality is detrimental to health, mainly in countries at the lower levels of economic development. While advancement in technology along with a unified effort of global organizations (such as WHO, UNICEF) and respective countries has enabled reasonable progress in the state of public health (such as reduced child and maternal mortality), developing countries still lag far behind developed nations. In order to improve the state of public health further and achieve the intended sustainable development goals by the year 2030, further insights into the determinants of public health are required. This study contributes to the literature and suggests that the efforts made to improve the state of health in a country may have to focus more on ethnic minority regions. Higher use of essential resources such as vaccinations and modern contraceptive methods, establishing awareness about the benefits of education, and providing better health infrastructure in ethnically disadvantaged regions might be key to achieving the desired targets in public health.

References

Albanesi, S. (2007). Inflation and inequality. *Journal of Monetary Economics*, 54(4), 1088–1114.

Alesina, A., Baqir, R., and Easterly,W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, 114(4), 1243–1284.

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

Alesina, A., and Ferrara, E. L. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3), 762–800.

Alesina, A., Michalopoulos, S., and Papaioannou, E. (2016). Ethnic inequality. *Journal of Political Economy*, 124, 428–488.

Alesina, A., and Rodrik, D. (1994). Distributive Politics and Economic Growth. *The Quarterly Journal of Economics*, 109(2), 465–490.

Alesina, A., and Zhuravskaya, E. (2011). Segregation and the quality of government in a cross section of countries. *The American Economic Review*, 101(5), 1872–1911.

Ashraf, Q., and Galor, O. (2013). The 'out of africa' hypothesis, human genetic diversity, and comparative economic development. *American Economic Review*, 103(1), 1–46.

Baird, S., Friedman, J., and Schady, N. (2011). Aggregate income shocks and infant mortality in the developing world. *The Review of Economics and Statistics*, 93(3), 847–856.

Baldwin, K., and Huber, J. D. (2010). Economic versus cultural differences: Forms of ethnic diversity and public goods provision. *American Political Science Review*, 104(4), 644–662.

Biggs, B., King, L., Basu, S., and Stuckler, D. (2010). Is wealthier always healthier? the impact of national income level, inequality, and poverty on public health in latin america. *Social Science and Medicine*, 71(2), 266–273.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2008). Staying in the classroom and out of the maternity ward? the effect of compulsory schooling laws on teenage births. *The Economic Journal*, 118, 1025–1054.

Bluhm, R., and Thomsson, K. (2020). Holding on? ethnic divisions, political institutions and the duration of economic declines. *Journal of Development Economics*, 144, 102457.

Braendle, T., and Colombier, C. (2016). What drives public health care expenditure growth? evidence from swiss cantons, 1970–2012. *Health Policy*, 120, 1051–1060.

Brunello, G., Fort, M., Schneeweis, N., and Winter-Ebmer, R. (2016). The causal effect of education on health: What is the role of health behaviors? *Health Economics*, 25(3), 314–336.

Chen, Y., and Li, H. (2009). Mother's education and child health: Is there a nurturing effect? *Journal of Health Economics*, 28, 413–426.

Churchill, S. A., Ivanovski, K., and Bhattacharya, M. (2019). The role of ethnic diversity in sustainable environmental growth: new evidence across different income regions. *Applied Economics*, 51(4), 398–408.

Collier, P., and Gunning, J. W. (1999). Explaining african economic performance. *Journal of Economic Literature*, 37(1), 64–111.

Cutler, D., Deaton, A., and Lleras-Muney, A. (2006). The determinants of mortality. *Journal of Economic Perspectives*, 20(3), 97–120.

de Soysa, I., and Almås, S. (2019). Does ethnolinguistic diversity preclude good governance? a comparative study with alternative data, 1990-2015. *Kyklos*, 72(4): 604–636.

Deaton A., (1997), "The Analysis of Household Surveys: A Microeconometric Approach to Development Policy", Baltimore, MD Johns Hopkins University Press.

Deaton, A., and Heston, A. (2010). Understanding ppps and ppp-based national accounts. *American Economic Journal: Macroeconomics*, 2(4), 1–35.

Desmet, K., Gomes, J. F., and Ortuño-Ortín, I. (2020). The geography of linguistic diversity and the provision of public goods. *Journal of Development Economics*, 143, 102384.

Desmet, K., Ortuño-Ortín, I., and Wacziarg, R. (2012). The political economy of linguistic cleavages. *Journal of Development Economics*, 97(2), 322–338.

Easterly, W., and Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *The Quarterly Journal of Economics*, 112(4), 1203–1250.

Elgar, F. J. (2010). Income inequality, trust, and population health in 33 countries. *American Journal of Public Health*, 100(11), 2311–2315.

Fafchamps, M. (2000). Ethnicity and credit in african manufacturing. *Journal of Development Economics*, 61(1), 205–235.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth*, 8, 195–222.

Fosu, A. K. (2010). Inequality, income, and poverty: Comparative global evidence. *Social Science Quarterly*, 91(5), 1432–1446.

Gomes, J.F. (2020). The health costs of ethnic distance: evidence from sub-Saharan Africa. *Journal of Economic Growth*, 25: 195–226.

Gordon, J., Raymond G. (2005). Ethnologue: Languages of the world. 15th Edition.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102, 994–1028.

Hjort, J. (2014). Ethnic Divisions and Production in Firms. *The Quarterly Journal of Economics*, 129(4), 1899–1946.

Huber, J. D., and Suryanarayan, P. (2020). Ethnic inequality and ethnification of political parties: Evidence from india. *World Politics*, 68(1), 149–188.

Hug, L., You, D., Blencowe, H., and et al. (2021). Global, regional, and national estimates and trends in stillbirths from 2000 to 2019: a systematic assessment. *The Lancet*, 398(10302), 772–785.

Jamison, D. T., Murphy, S. M., and Sandbu, M. E. (2016). Why has under-5 mortality decreased at such different rates in different countries? *Journal of Health Economics*, 48, 16–25.

Johnson, S., Larson, W., Papageorgiou, C., and Subramanian, A. (2013). Is newer better? Penn World Table Revisions and their impact on growth estimates. *Journal of Monetary Economics*, 60(2), 255–274.

Kyriacou, A. P. (2013). Ethnic Group Inequalities and Governance: Evidence from Developing Countries. *Kyklos*, 66(1), 78–101.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. (1999). The quality of government. *Journal of Law, Economics, and Organization*, 15(1), 222–79.

Levine, D. I., and Rothman, D. (2006). Does trade affect child health? *Journal of Health Economics*, 25, 538–554.

Mamdani, M., Garner, P., Harpham, T., and Campbell, O. (1993). Fertility and contraceptive use in poor urban areas of developing countries. *Health Policy and Planning*, 8, 1–18.

Martinez, L.R., (2022), How Much Should We Trust the Dictator's GDP Growth Estimates?, *Journal of Political Economy* 130(10): 000-000.

Mayer, S. E., and Sarin, A. (2005). Some mechanisms linking economic inequality and infant mortality. *Social Science and Medicine*, 60, 439–455.

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

Montalvo, J. G., and Reynal-Querol, M. (2005). Ethnic polarization, potential conflict, and civil wars. *American Economic Review*, 95, 796–816.

Montalvo, J. G., and Reynal-Querol, M. (2021). Ethnic Diversity and Growth: Revisiting the Evidence. *The Review of Economics and Statistics*, 103(3), 521–532.

Østby, G. (2008). Polarization, Horizontal Inequalities and Violent Civil Conflict. *Journal of Peace Research*, 45(2), 143-162.

Persson, T., and Tabellini, G. (1994). Is inequality harmful for growth? *American Economic Review*, 84(3), 600–621.

Ponnusamy, S. (2022). Rainfall Shocks, Child Mortality, and Water Infrastructure. *Health Economics*, 31(7): 1317-38

Pritchett, L., and Summers, L. H. (1996). Wealthier is healthier. *The Journal of Human Resources*, 31(4), 841–868.

Ray, D. (2014), Development Economics. Princeton University Press, Princeton.

Spolaore, E., Alesina, A., and Wacziarg, R. (2000). Economic Integration and Political Disintegration. *American Economic Review*, 90(5), 1276–1296.

Tindale, A. (2019). New patterns of ethnic diversity: exploring the residential geographies of mixed-ethnicity individuals in sydney, australia. *Journal of Ethnic and Migration Studies*, 45(15), 2910–2933.

Millenium Development Goals Report, 2015. Accessed on February, 2022. https://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf

World Bank, World Development Indicators (WDI), 2020.

Weidmann, N. B., Roslashd, J. K., and Cederman, L.-E. (2010). Representing ethnic groups in space: A new dataset. *Journal of Peace Research*, 47(4), 491–499.

Yong, E. L. (2019). "Understanding cultural diversity and economic prosperity in europe: a literature review and proposal of a culture–economy framework." *Asian Journal of German and European Studies*, 4.

Online Appendix (Not for Publication)

Ethnic Inequality and Public Health

A1 Inequality Index Construction

This section summarizes the methodology followed by Alesina et al. (2016) in the construction of inequality indicators. We employ two different types of inequality indices available from Alesina et al. (2016) in this study – inequality index constructed using economic activity across ethnic/linguistic homelands, i.e., ethnic inequality (henceforth, EI), and inequality measures developed using economic activity under various spatial/administrative levels, i.e., spatial inequality (SI). Due to the lack of economic activity data at sub-national levels, nightlight intensity measured from space is used as a proxy for economic development by Alesina et al. (2016), in line with Henderson et al. (2012). The inequality indices are then developed using the following equation:

$$G_c = \frac{1}{n}\left[n+1-2\frac{\sum_{i=1}^{n}(n+1-i)y_{ic}}{\sum_{i=1}^{n}y_{ic}}\right] \tag{A1}$$

The independent variable of interest in this study is ethnic inequality (GINI coefficient) measured at the country level. G_c captures the income inequality across ethnic regions for each country c in the sample; n refers to the number of ethnic homelands within a country, and y_{ic} is the nightlight per capita in an ethnic region i, from country c. Average nightlights within each ethnic boundary for a country are extracted first, and then using the gridded population data from CIESEN (2018), per capita nightlight intensity is calculated. The resulting Gini coefficient in Equation A1 accounts for differences in average income between ethnic homelands in a country, as represented by nightlights per capita at the ethnic homeland level. G_c ranges from zero to one, representing the highest level of equality and inequality possible, respectively. Two different measures of ethnic inequality are constructed, using ethnic and linguistic maps from GREG and Ethnologue, respectively, for the years 1992, 2000, 2006, and 2012. While we mainly rely on examining the association between EI and the health outcomes for the year 2000, we also explore the relationship for the other years as robustness.

A potential issue with the ethnic inequality index constructed using the boundaries of ethnic homelands instead of using individual-level income or consumption data is that the relationship between EI and the outcomes might be confounded by the regional disparities in income or public goods provision (Alesina et al., 2016). Therefore, we control for three different measures of spatial inequality indices: mainly, we control for inequality measured using economic activity across 2.5 x 2.5-degree spatial grids, and as robustness, we also employ SI indices based on first and second-level administrative regions. To create the SI index at a spatial grid level, a global grid of 2.5 x 2.5 decimal degrees is overlapped with the 2000 Digital chart of the world and the nightlight per capita is measured for the overlapped regions. Then, the SI index for each country is constructed using Equation A1 for the years 1992, 2000, 2006, and 2012. To compute SI data at the first- and second-level administrative regions, sub-national boundary maps from the Global Administrative Areas Database (GADM) are used. Using nightlights per capita within administrative regions, SI indices for each country are constructed by Alesina et al. (2016).

A2 Heterogeneity Analysis by Income Levels

Existing literature has suggested that the effects of ethnic inequality on societal outcomes is heterogeneous to the income levels of the countries. Due to the demand for skill complementarities and owing to the better political status of the developed countries, the detrimental effects of ethnic inequality may be alleviated (Alesina and Ferrara 2005; Bluhm and Thomsson 2020). Therefore, we focus next on whether the association that has been observed between EI and U5MR varies based on the development status of a country. We divide the sample of countries in our analysis into developing countries (both low- and lower-middle-income countries, DCs), upper-middle-income countries, and high-income countries. We rely on the World Bank's income classification of countries for the year 2010 to perform these groupings.

Table A3 provides the results for the three groups of countries, along with the mean of the outcomes. We only report the results based on the GREG index to preserve space.[1] There are certain notable features that can be observed from Table A3. First, we find that ethnic inequality is detrimental to the key health outcomes (U5MR, MM, LEAB, and child stunting), only in DCs. Second, EI has some negative effects in UMICs in terms of child stunting (column 5, Panel B) and contraceptive usage (column 11, Panel B). Third, health outcomes in HICs do not display any association with EI.

While most of the coefficients in columns 6–12, Panel A suggest a detrimental association between EI and health/development outcomes, the estimates are statistically insignificant, which could be owed to the smaller sample size. Therefore we perform an additional exercise in which we pool data for the years 1992, 2000, 2006, and 2012 and perform the analysis. The results reaffirm our findings so far, i.e., higher EI is significantly associated with a lower GDP and measles/vaccination rates in DCs, whereas no detrimental relationship is observed in HICs. Results not shown.

Therefore, the group of countries that is at the lower end of the income spectrum (i.e., DCs) is found to be affected the most due to ethnic inequality, whereas the association between EI and health outcomes is mostly non-existent in HICs. Developed countries, through their better management of conflicts and differences across ethnic groups, may be able to reap the benefits of skill complementarities associated with the multitude of ethnic groups, and our study provides suggestive evidence of the same. However, our findings do not imply that higher inequality is not detrimental for HICs, as the implications of EI can extend beyond

[1] As robustness, we also examine the association between EI measured using the Ethnologue index and health outcomes subject to the extended specification from Table 3, main text. In the results not shown, findings remain robust to the use of the Ethnologue index as well.

health outcomes.

A3 Heterogeneity Analysis by Geographical Regions

In this section, we conduct a heterogeneity analysis by focusing on different geographic regions, individually, to identify the group of countries affected the most. To preserve space, we focus only on the GREG index. We divide the sample of countries in our analysis into seven regions: East Asia and Pacific Countries (EAP), Eastern Europe and Central Asia (ECA), Latin America and Caribbean countries (LAC), Middle-East and North African countries (MENA), Central Asian countries (SA), Sub-Saharan African countries (SSA) and Western European countries (WE).

In this part, we focus only on U5MR to conserve space. As we divide our sample into seven regions, the sample size can get quite low for some of the specifications. To fix this issue, we pool data for 1992, 2000, 2006, and 2012 to maximize the sample size and increase the statistical power. Based on the results from Table A4, U5MR in the SSA, LAC, and EAP group of countries share a strong, detrimental association with ethnic inequality. In the results not shown, we observe that SSA regions are affected strongly due to EI in terms of MM, LEAB, and stunting as well. Therefore, Sub-Saharan countries that experience the highest levels of ethnic inequality are also the ones affected the most.[2] Instead of using the pooled data, we re-perform this heterogeneity analysis by applying data only for 2000. SSA group of countries are found to be affected the most in this alternate exercise as well.

A4 Robustness Exercises

We perform several robustness tests in this section to test the sensitivity of our findings. First, we control for various geographic and historical endowments in the specifications. Second, we address a potential migration concern. Third, instead of using spatial inequality measures that are constructed based on 2.5 x 2.5 spatial grids, we use inequality measures developed based on first- and second-level administrative boundaries. Fourth, we test the relationship between EI and health/economic outcomes using data from the years 1992, 2006, and 2012. Fifth, we pool the data for the years 1992, 2000, 2006, and 2012 and perform the analysis subject to year-fixed effects and region/country fixed-effects. Sixth, we use EI indices from Alesina et al. (2016) that are constructed by excluding capital cities, and groups that account for less than one percent of the total population, respectively. As a final exercise, we examine a potential measurement error concern in the construction of EI indicators using a two-stage least squares (2SLS) specification.

[2]The mean of GREG EI is in the range of 0.288–0.522, with SSA regions suffering the highest levels of ethnic inequality. Source: Authors' calculations based on data from Alesina et al. (2016).

A4.1 Controlling for Geographic, and Historical Factors

Our estimates of ethnic inequality are quite robust so far to the inclusion of spatial inequality, per capita GDP of a country, simple controls such as population and land size, and other measures of inequality and ethnic diversity. Other potential confounders such as the geographic or historical features of a country are yet to be accounted for which can shape its current levels of economic development or inequality. Therefore, we include some of these variables next, to isolate the association between EI and U5MR.

First, we control for various geographical features of a country such as terrain ruggedness, the percentage of fertile soil in a country, the percentage of the country with a tropical climate, the gem-quality diamond extraction index, and the average distance to the nearest ice-free coast, in line with Nunn and Puga (2012).[3] Next, we control for several historical features of a country such as the population density of a country around 1500 CE, whether the country had a British common-law system, and finally, the duration (in log) since the Neolithic revolution. Data on these variables are sourced from Alesina et al. (2016). We control for the historical population density as it reflects the 'reversal of fortune' in countries due to the European colonization (Acemoglu et al., 2002) and whether a country follows British Common Law as it can act as a proxy for the legal protections of investors (La Porta et al., 1998). In line with Putterman and Weil (2010) and Ashraf and Galor (2013), we control for the timing since the Neolithic revolution as it can act as a proxy for the pre-colonial conditions and help account for the experience of a country's current residents regarding their ancestors' transition to agriculture (Alesina et al., 2016).

Refer to Table A5 for the results. We introduce three sets of controls (simple, geographic, and historical) individually and also simultaneously. Based on columns (4) and (8), even after all the geographic and historical factors are accounted for together, the detrimental association between EI and U5MR persists. This robustness exercise provides more credibility to our baseline estimates and helps ensure that any leftover bias is likely to be minimal.

A4.2 Migration Concern

In the construction of the GREG index, Alesina et al. (2016) applies ethnic boundaries for the year 1964. A possible concern here is that there could have been demographic shifts, for example, the migration of people, which could have altered the distribution. This can induce some bias in our estimates as the outcomes we use are from 1992 onwards. However, this concern is likely minimal as people usually migrate for short-term periods (Banerjee and Duflo, 2007). A potential reason for temporary migration is that people emphasize a higher value on remaining close to their social network, while earning more money may not be the top priority (Munshi and Rosenzweig, 2005). Even in the case of migration, the workers that emigrate to urban areas remit a fraction of their income back to their ethnic homelands (Stark

[3] Refer to Alesina et al. (2016) for a further description and sources of these geographic factors.

and Yitzhaki 1982; Stark et al. 1986; Du et al. 2005) and these remittances would be reflected in the nightlight activity (Alesina et al., 2016), which further reduces any biases due to migration. Nevertheless, we address this potential migration concern in several ways.

First of all, we have already shown in Table A1 that the relationship between the Ethnologue based EI indicator and U5MR for the year 2000 is detrimental. In this case, the year 2000 for which U5MR is observed is close to the period in which the locations of language groups were available (i.e. mid/late 1990s). Therefore, the migration concern is not a major one to start with. Second, we use an EI indicator that is constructed after the exclusion of capital regions, where ethnic mixing is likely to occur the most. EI indicator that includes capital regions in its construction may be affected due to ethnic mixing, as migrant workers are more likely to reside in capital cities. EI and U5MR have a strong positive association with the use of this alternative EI indicator and a detailed discussion of the findings is provided in section A4.6, below.

Economic transformation in developing countries is responsible for a large rural-to-urban within-country migration, which is mainly responsible for the increase in urban population (Freeman, 2006). This migration toward urban centers can also induce demographic shifts in the population of ethnic groups as they might migrate toward the urban areas in anticipation of better economic opportunities. We already discussed that migrant workers remit their income back to their homelands, which will be reflected in the nightlights. Nevertheless, as a third exercise, we control for the level of the urban population (as a % of the total) in our specification (data from WDI (2020)). Fourth, we omit the top ten percentile of countries that have experienced the highest percentage of growth in terms of urban population between 1964 and 2000, and that are also more likely to have experienced the most demographic shifts in the population. The association between ethnic inequality and under-5 mortality remains unchanged to these exercises as well. Results not provided.

A4.3 Alternate Measures of Spatial Inequality

The spatial inequality measure used in the analysis so far relied on a spatial grid of 2.5×2.5 decimal degrees to capture regional differences within a country. An alternative way of measuring SI is by evaluating the differences across administrative boundaries. This would reveal any effects of within-country variation in economic activity or provision of public goods across administrative units on public health. Following Equation A1, GINI coefficients for each country are constructed using first and second-level administrative boundaries from GADM by Alesina et al. (2016), who notes that administrative boundaries were formed based on the ethnic settlements in some countries. Therefore, we are also able to examine whether the relationship between EI and U5MR still holds to the controlling of inequality based on politically defined units.

Results for specifications involving various levels of administration are provided in Table A6. Panel A contains spatial inequality measures based on first-level administrative boundaries (henceforth, SI1) and in Panel B, the GINI index subject to second-level administrative units (henceforth, SI2) is used. In columns (1) and (5), we include only EI and SI (based on administrative levels), along with the real per capita GDP. In the remaining columns, we account for simple controls and geographic controls. We also add ethnic-linguistic fragmentation from Alesina et al. (2003) and Desmet et al. (2012), respectively. Results for the GREG index from Table A6 provide strong evidence of the detrimental association between EI and U5MR. The coefficients of Ethnologue EI remain positive in all columns and become significant once the full set of controls is accounted for (refer to columns 7 and 8 of Panel B).

In contrast with Table 3 from the main text, SI based on administrative units shares a strong correlation with U5MR, even after controlling for EI. Ali et al. (2018) discuss that colonists used a 'divide and rule strategy' in which administrative boundaries were drawn based on ethnic divisions. Pierskalla (2019) argues that administrative boundaries have been shaping and reshaping around ethnic group territories, especially in the developing world. Following such arguments, SI based on administrative units explaining variations in U5MR would mean that SI1 and SI2 are capturing some effect of ethnic inequality that is manifesting as differences in nightlights across administrative units. And this is a potential reason for the estimates of the Ethnologue index being insignificant in some of the specifications.

A4.4 Examining Relationship for Alternate Years

So far, we have used data for the year 2000 to explore the relationship between ethnic inequality and outcomes. Now, we follow Alesina et al. (2016) and examine the association individually for the years 1992, 2006, and 2012. Especially the analysis for 2012 will help us understand whether the policies implemented under Millennium Development Goals (MDGs) helped reduce the association between ethnic inequality and public health. If a stronger association is found for 2012 as well, then it provides suggestive evidence that the level of public good provision for ethnic minority regions may not be on par with the rest of the ethnic groups.

Results for this analysis are provided in Table A7. Panel A contains results for the year 1992, whereas panel B provides results for the year 2012. We do not provide results for 2006 for brevity, however, findings remain the same. Only the results for the GREG index are shown to preserve space. Results for several variables are missing in Panel A due to the lack of data availability. From the table, it is evident that a strong correlation exists even to the use of different years, especially for the health outcomes. Therefore, to achieve Sustainable Development Goals by the year 2030, it might be essential to focus more on ethnic minority areas.

A4.5 Using Pooled Data

As a final robustness exercise, we pool data for the years 1992, 2000, 2006, and 2012 and perform the analysis subject to specifications with year-fixed effects and region/country-fixed effects. Refer to Table A8 for the results. In Panel A, we control for region-fixed effects, whereas in Panel B, we control for country-fixed effects. Irrespective of the type of fixed effects employed, the significant association between ethnic inequality and outcomes such as U5MR, MM, secondary schooling, and contraceptive use persists. In terms of LEAB, GDP, and child stunting, spatial inequality shares a significant relationship, as evident from columns 3–6 of Panel B. To summarize, there is a detrimental association between EI and several outcomes to the use of pooled data as well.[4]

We conduct an additional exercise by manually extracting nightlight data for the years 1992–2013 and construct an ethnic inequality index using the GREG mappings. We also constructed a spatial inequality index using the first-level administrative boundary shapefile from GADM. Then we perform an analysis in which we control for year-fixed effects and country-fixed effects, along with the spatial inequality. We find that the relationship between EI and under-5 mortality remains positive with a coefficient of 17.58 and statistically significant at 1% to the use of the panel data.

A4.6 Excluding Certain Areas

In Table A9, we re-examine the relationship between EI and U5MR by using ethnic inequality indices that exclude certain regions in its construction. In line with Alesina et al. (2016), we include simple controls such as the size of land area and population in the specifications, along with the real per capita GDP of a country and spatial inequality. Columns 1–3 of Table A9 provide results for the GREG index, whereas the last three columns contain results for the Ethnologue index. In columns 1 and 4, we include the EI index constructed based on all ethnic and linguistic groups. The association between EI and U5MR is stronger in both magnitude and significance, compared with the estimates of SI.

In columns 2 and 5, ethnic regions that include capital cities are dropped when calculating ethnic GINI. Capital cities are regions where most ethnic mixing would occur in terms of different nationalities, cultures, and ethnicities (Tindale 2019; Yong 2019). Hence, attributing the observed nightlights to ethnic groups in capital regions can become tricky. Further, capital regions are relatively more developed in terms of infrastructure as compared to non-capital regions within the same country. The migrant workers are more likely to reside in the capital regions as well. These could lead to differences in nightlights across the capital and non-capital regions that do not arise due to ethnic differences. Therefore, attributing the nightlights to the most prevalent ethnic group in these regions would capture differences in rural and urban areas more than the differences across ethnic groups. After dropping such ethnic groups, the coefficient for EI is smaller in magnitude as captured by GREG but not very different when Ethnologue mapping is used. The coefficient of EI retains its significance even when capital

[4] We also apply a Box-Cox transformation following Box and Cox (1964) that can account for non-normality in residuals. We perform this analysis for the year 2000, and also for the pooled data with years 1992, 2000, 2006, and 2012. The coefficient of ethnic inequality remains positive and significant at 1% to this extra exercise as well. We thank an anonymous referee for raising this valuable suggestion.

cities are excluded from the analysis.

In columns 3 and 6, ethnic groups that form less than one percent of the population in a country are dropped when calculating the ethnic Gini to alleviate concerns of skewed results due to such ethnicities. Now GREG EI remains statistically significant, while the coefficient of Ethnologue EI is close to column (5), but now insignificant at conventional levels with a p-value of 0.102. Based on the table, SI can significantly explain variation in U5MR only when small groups are taken out of the analysis (columns 3 and 6), whereas the estimates of EI remain robust in almost all the specifications.

A4.7 Addressing Measurement Error Concern

Our findings in the main text remained similar to the use of two distinct inequality indices – the GREG index based on ethnic locations and the Ethnologue index based on linguistic groups' locations. A potential concern here is that there are measurement errors in the mappings of ethnic or linguistic group boundaries. For example, the location of ethnolinguistic homelands available in these maps may not be accurate, which can induce a bias in our estimates. To address this potential issue, we follow a similar approach to Alesina et al. (2016) and perform an analysis in which we instrument the GREG EI with the Ethnologue EI, and vice versa, subject to a 2SLS specification.

Here, we take advantage of the fact that the maps used to construct the GREG and Ethnologue indices are different from each other. While the GREG index is based on the ethnic location maps that were published by the Soviet ethnographers in the 1960s, the Ethnologue index is based on the language group locations developed by SIL International, a Christian linguistic service organization. As far as the measurement errors in these two maps are not correlated with each other, the use of these EI indices as instruments for each other in a 2SLS specification can help us address this potential measurement error concern (Alesina et al., 2016).

Another advantage of this analysis is that it can help us address the within-country migration concern to some extent. While the Ethnologue maps refer to the linguistic groups in the mid/late 1990s, GREG maps locate ethnic groups in the 1960s. And our GREG EI indicator is constructed based on nightlights for the year 2000 subject to the ethnic locations from the 1960s. Hence, there may be some internal migration of people between 1960–2000 which can cause some errors in the inequality measurement. While several studies have already suggested that the migration workers remit a fraction of their income to their families in their ethnic homelands (Stark and Yitzhaki 1982; Stark et al. 1986; Du et al. 2005), and that migration is usually temporary (Banerjee and Duflo, 2007), this concern is less likely to start with. Nevertheless, using the Ethnologue indicator based on the most recent mappings as an instrument for the GREG indicator can help allay some of the concerns regarding the use of ethnic boundaries from the 1960s.

The results for 2SLS specifications are presented in Table A10. In columns 1–5, the results for the specifications in which the Ethnologue index is used as an instrument for the GREG index are provided, and vice-versa, in the last five columns. The table also presents the first-stage F-statistic which suggests that the first-stage fit is quite strong in all the specifications. The relationship between EI and U5MR observed in Tables 2 and 3 in the main text remains strong to the application of the 2SLS method as well, qualitatively. This exercise help allays any measurement error concern in the construction of EI indices, while also addressing the within-country migration concern to some extent as the results for the GREG index in the first five columns remain strong to the use of the Ethnologue indicator as an instrument.

References

Acemoglu, D., Johnson, S., and Robinson, J.A., (2002), Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics*, 117(4): 1231–1294.

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

Alesina, A., and Ferrara, E. L. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3), 762–800.

Alesina, A., Michalopoulos, S., and Papaioannou, E. (2016). Ethnic inequality. *Journal of political Economy*, 124, 428–488.

Ali, M., Fjeldstad, O.H., Jiang, B., and Shifa, A.B. (2018). "Colonial legacy, state-building and the salience of ethnicity in sub-saharan africa." *The Economic Journal*, 129(619): 1048-1081.

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

Banerjee, A. V., and Duflo, E., (2007), The Economic Lives of the Poor. *Journal of Economic Perspectives* 21(1): 141-168.

Bluhm, R., and Thomsson, K. (2020). Holding on? ethnic divisions, political institutions and the duration of economic declines. *Journal of Development Economics*, 144, 102457.

Box, G.E.P., and Cox, D.R., (1964), An analysis of transformations. *Journal of the Royal Statistical Society, Series B* 26: 211–252.

CIESEN (2018). Gridded population of the world, version 4 (gpwv4): Population density. Revision 11.

Desmet, K., Ortuño-Ortín, I., and Wacziarg, R. (2012). The political economy of linguistic cleavages. *Journal of Development Economics*, 97(2), 322–338.

Du, Y., Park, A., and Wang, S., (2005), Migration and rural poverty in China, *Journal of Comparative Economics*, 33(4): 688-709.

Freeman, R.B., (2006), People Flows in Globalization. *Journal of Economic Perspectives*, 20(2): 145-170.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102, 994–1028.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R., (1998), Law and Finance. *Journal of Political Economy*, 106(6): 1113-1155.

Munshi, K. and Rosenzweig, M., (2005), "Why is Social Mobility in India so Low? Social Insurance, Inequality, and Growth." BREAD [Bureau for Research and Economic Analysis of Development] Working Paper 097.

Nunn, N., and Puga, D. (2012), Ruggedness: The Blessing of Bad Geography in Africa. *The Review of Economics and Statistics* 94(1), 20–36.

Pierskalla, J.H. (2019). "The Proliferation of Decentralized Governing Units," (p. 115–143). Cambridge University Press.

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

Stark, O., and Yitzhaki, S., (1982), Migration, growth, distribution and welfare. *Economic Letters* 10: 243-49.

Stark, O., Taylor, J.E., and Yitzhaki, S., (1986), Remittances and Inequality. *Economic Journal* 28: 309–22.

Tindale, A. (2019). New patterns of ethnic diversity: exploring the residential geographies of mixed-ethnicity individuals in sydney, australia. *Journal of Ethnic and Migration Studies*, 45(15), 2910–2933.

World Bank, World Development Indicators (WDI), 2020.

Yong, E. L. (2019). "Understanding cultural diversity and economic prosperity in europe: a literature review and proposal of a culture–economy framework." *Asian Journal of German and European Studies*, 4.

Figure A1: Relationship between Ethnic Inequality and U5MR

(a) EI (GREG)

(b) EI (Ethnologue)

Figure A1: The plot provides the conditional (on regional fixed effects) relationship between ethnic inequality and child mortality. X-axis and Y-axis contain the rescaled residuals after netting out the effect of region-fixed effects from ethnic inequality and child mortality. Rescaling is done by standardising the residuals and then rescaling them to have the same mean and standard deviation as the ethnic inequality and child mortality variables.

Table A1: Ethnic Inequality (Ethnologue) and U5MR

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnic Inequality	38.66***		31.26***		39.57***	30.57**	41.32***	37.70***	44.76***
	(9.91)		(10.57)		(13.02)	(13.83)	(12.01)	(14.04)	(15.18)
GDP Per Capita	-0.82***	-1.04***	-0.83***	-0.94***	-0.82***	-0.83***	-0.81***	-0.81***	-0.78***
	(0.30)	(0.28)	(0.29)	(0.31)	(0.30)	(0.29)	(0.27)	(0.27)	(0.26)
Spatial Inequality		36.33***	13.21			13.31	14.41	15.19	20.80
		(10.30)	(9.98)			(10.11)	(10.17)	(10.31)	(15.30)
Log Number of Ethnicities				6.14***	-0.23	0.16		1.27	-0.26
				(1.82)	(2.22)	(2.22)		(2.38)	(3.02)
Ethnic inequality in population							-42.81	-44.35	-44.33
							(28.62)	(29.43)	(29.56)
Ethnic inequality in size							33.09	32.04	29.91
							(29.37)	(29.04)	(29.30)
Log Land Area									-2.15
									(2.39)
Log Population (in 2000)									3.13
									(2.41)
Adjusted R-Squared	0.7646	0.7525	0.7651	0.7506	0.7632	0.7636	0.7654	0.7643	0.7638
Observations	173	173	173	173	173	173	173	173	173

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates. The dependent variable is the under-5 mortality rates for the year 2000. Ethnic inequality indicator based on the Ethnologue mappings is used. The coefficient of real GDP per capita is multiplied by 1000. All the regressions include region-fixed effects. Robust standard errors are reported in the parenthesis.

Table A2: Ethnic Inequality, EI-Mechanism Variable Interactions, and Health Outcomes

Dependent Variable	U5MR	MM	Stillbirths	LEAB	Stunting	U5MR	MM	Stillbirths	LEAB	Stunting
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	GREG Index					Ethnologue Index				
Ethnic Inequality	-14.902	-91.294	-3.263	2.769	-9.134	-16.384	-149.627**	-5.887***	3.597*	-2.618
	(11.024)	(82.163)	(2.252)	(2.048)	(7.941)	(9.910)	(67.149)	(2.157)	(1.994)	(5.231)
Spatial Inequality	-3.000	50.180	-0.646	-0.974	2.554	9.749	123.956	3.094	-3.456	6.328
	(12.635)	(137.825)	(2.786)	(2.467)	(6.230)	(12.578)	(132.128)	(2.741)	(2.499)	(6.151)
Ethnic Fragmentation	3.588	123.593*	0.037	-1.111	0.242	6.145	109.789	1.289	-2.174	-0.551
	(7.938)	(64.048)	(2.243)	(1.752)	(4.806)	(8.185)	(67.029)	(2.121)	(1.801)	(4.053)
EI-GDP-Interaction	43.628***	240.505**	11.242***	-5.342***	13.475***	31.767***	213.164**	7.429***	-2.448	6.800
	(10.277)	(94.794)	(1.906)	(1.703)	(3.885)	(10.287)	(85.968)	(1.838)	(1.501)	(4.895)
EI-Measles-Interaction	26.333	64.201	6.225**	1.825	-4.225	16.412	-56.420	5.046*	2.239	-8.950**
	(17.523)	(113.323)	(2.900)	(2.402)	(6.180)	(17.958)	(101.829)	(2.739)	(2.256)	(4.406)
EI-DPT-Interaction	21.176	139.571	0.637	-2.411	1.217	26.315*	225.080**	1.655	-3.942**	10.357***
	(14.140)	(94.769)	(2.611)	(1.842)	(6.338)	(15.739)	(94.698)	(2.397)	(1.600)	(3.772)
EI-Fertility-Interaction	21.857**	125.701	1.672	-6.382***	15.116***	4.814	38.995	-0.341	-4.211**	7.224
	(10.929)	(83.848)	(2.095)	(1.808)	(4.350)	(10.513)	(80.003)	(1.898)	(1.645)	(4.839)
EI-Contraception-Interaction	31.253**	369.203**	4.429**	-4.421**	7.871**	20.480*	284.786**	3.324**	-1.620	4.030
	(13.590)	(145.482)	(1.832)	(1.873)	(3.544)	(10.802)	(135.412)	(1.642)	(1.513)	(3.126)
EI-PrimarySch-Interaction	5.391	-18.014	2.681	-1.873	3.510	3.330	-56.778	0.630	-0.884	1.423
	(9.898)	(97.759)	(1.738)	(1.486)	(3.108)	(9.537)	(99.186)	(1.573)	(1.383)	(3.019)
EI-SecondarySch-Interaction	-16.528	-109.343	-0.730	0.764	3.863	-7.474	-4.845	1.923	-0.358	5.765*
	(10.351)	(86.272)	(1.811)	(1.587)	(3.670)	(7.931)	(70.278)	(1.390)	(1.182)	(3.028)
Adjusted R-Squared	0.860	0.737	0.809	0.853	0.664	0.833	0.705	0.799	0.844	0.665
Observations	173	171	173	171	140	173	171	173	171	140

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates for the year 2000. All specifications control for population size, land size, and Real GDP per capita. All the regressions include region-fixed effects. Robust Standard errors are reported in the parenthesis.

Table A3: Heterogeneity Analysis by the Income Classification of Countries

	U5MR	Maternal Mortality	Stillbirths	LEAB	Child Stunting	GDP	Measles Vaccination	DPT Vaccination	Schooling Primary	Schooling Secondary	Contraceptive Use	Fertility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Developing Countries												
Ethnic Inequality	51.51**	496.13**	7.55*	-7.46**	3.64	-0.51	-4.48	-7.54	-9.57	13.23	-11.89	0.80
	(20.16)	(188.47)	(4.24)	(3.62)	(7.82)	(0.38)	(11.17)	(11.08)	(11.39)	(25.08)	(10.34)	(0.69)
Spatial Inequality	-12.17	197.45	-0.31	2.77	2.00	-1.18**	-14.69	-19.07	4.38	-56.12	4.81	-0.45
	(35.22)	(441.08)	(5.17)	(5.47)	(7.91)	(0.48)	(15.42)	(15.73)	(11.02)	(32.79)	(16.60)	(0.79)
Mean Dependent Variable	103.549	562.861	23.546	58.083	38.025	7.210	68.986	69.361	71.950	35.155	28.664	4.715
Adjusted R-Squared	0.6907	0.5349	0.5876	0.7253	0.5017	0.5514	0.4372	0.4198	0.4767	0.5984	0.6427	0.5962
Observations	72	72	72	72	72	72	72	72	46	28	36	72
Panel B: Upper-Middle-Income Countries												
Ethnic Inequality	30.95	63.56	3.36	2.59	15.58**	-0.47	-7.91	-14.70	-6.18	-22.11	-46.46**	0.68
	(19.94)	(50.96)	(3.68)	(2.30)	(5.78)	(0.33)	(8.89)	(11.34)	(8.07)	(16.28)	(17.91)	(0.74)
Spatial Inequality	-4.87	-32.48	0.52	-2.98	-6.18	-0.19	-5.87	-4.87	-7.38	6.04	7.95	-0.61
	(15.80)	(36.74)	(4.08)	(3.15)	(7.63)	(0.35)	(10.07)	(11.41)	(7.96)	(21.86)	(25.14)	(0.51)
Mean Dependent Variable	34.844	97.319	12.295	69.087	18.349	8.675	87.729	87.521	91.146	65.291	35.721	2.631
Adjusted R-Squared	0.4929	0.7601	0.2143	0.7709	0.1808	0.0151	0.2869	0.2002	0.1586	0.0798	0.6244	0.4792
Observations	48	47	48	47	43	48	48	48	31	26	19	47
Panel C: High Income Countries												
Ethnic Inequality	-0.57	-6.37	-2.60	0.12	1.76	-0.01	-6.47	-8.85	-5.51	5.51	-	-0.42
	(3.93)	(12.99)	(2.05)	(2.72)	(17.04)	(0.40)	(5.81)	(5.71)	(7.23)	(7.86)	-	(0.59)
Spatial Inequality	-0.64	6.70	0.51	-0.94	3.50	0.38	11.72	9.66	13.87**	5.75	-	0.54
	(4.58)	(11.98)	(1.98)	(2.65)	(13.20)	(0.43)	(9.43)	(7.35)	(5.99)	(6.66)	-	(0.71)
Mean Dependent Variable	9.257	16.400	5.025	76.030	6.557	10.038	91.863	93.922	95.000	86.691	-	1.857
Adjusted R-Squared	0.5466	0.6331	0.5965	0.5156	-0.0421	0.6084	0.1115	0.0377	0.1713	0.2424	-	0.5253
Observations	51	50	51	50	23	51	51	51	35	32	-	50

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates. All specifications control for population size and land size. Real GDP per capita is included in all columns, except Column 6. All the regressions include region-fixed effects. Robust Standard errors are reported in the parentheses.

Table A4: Heterogeneity Analysis by the Geographical Location of Countries

	EAP	ECA	LAC	MENA	SA	SSA	WE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnic Inequality	40.70**	10.69	38.64***	13.65	22.69	84.65***	1.24
	(15.90)	(10.33)	(13.52)	(8.39)	(51.62)	(30.69)	(1.56)
Spatial Inequality	-10.70	42.80***	-11.41	-1.23	-3.44	18.11	0.54
	(10.90)	(12.03)	(10.86)	(9.62)	(32.20)	(19.60)	(1.36)
Mean Dependent Variable	31.772	25.260	30.467	21.278	73.615	119.948	5.360
Adjusted R-Squared	0.4244	0.4449	0.3315	0.4436	0.4886	0.3768	0.1913
Observations	74	97	99	65	26	151	57

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates using pooled data for the years 1992, 2000, 2006, and 2012. The dependent variable is the under-5 mortality rates. The ethnic inequality index based on the GREG mapping is used in this table. All the regressions include population size, land size, and real GDP per capita as controls. Robust Standard errors are reported in the parenthesis. EAP refers to East Asia and Pacific Countries, ECA stands for Eastern Europe and Central Asia, LAC is Latin America and Caribbean countries, MENA refers to Middle-East and North African countries, SA denotes Central Asian countries, SSA stands for Sub-Saharan African countries, and WE refer to Western European countries.

Table A5: Ethnic Inequality and U5MR - Accounting for Other Relevant Covariates

	GREG Index				Ethnologue Index			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic Inequality	42.66***	34.26**	37.05**	34.52**	22.58*	18.07	17.63	23.99*
	(15.22)	(13.49)	(16.02)	(15.06)	(11.78)	(11.76)	(12.54)	(12.31)
Spatial Inequality	22.45	-2.06	8.15	7.53	35.67**	6.40	19.22	14.94
	(17.05)	(13.01)	(14.87)	(16.48)	(14.88)	(12.27)	(13.04)	(15.80)
Simple Controls	Yes	No	No	Yes	Yes	No	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Historical Controls	No	No	Yes	Yes	No	No	Yes	Yes
Adjusted R-Squared	0.8034	0.8214	0.8039	0.8229	0.7958	0.8166	0.7976	0.8204
Observations	151	151	151	151	151	151	151	151

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates. The dependent variable is the under-5 mortality rates for the year 2000. Simple controls refer to the log of land area and population; geographic controls account for terrain ruggedness, the percentage of fertile soil in a country, the percentage of country with a tropical climate, gem-quality diamond extraction index, and the average distance to the nearest ice-free coast; historical controls refer to whether the country had a British common-law system, population density circa 1500 CE (in log form), and the log of timing since the Neolithic transition. All the regressions include real GDP per capita, along with region fixed effects. Robust Standard errors are reported in the parenthesis.

Table A6: Robustness Exercise: Alternative Measures of Spatial Inequality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Spatial Inequality based on First-level Administrative Units							
	Greg Index				Ethnologue Index			
Ethnic Inequality	33.92***	28.21***	32.63***	29.61***	22.23**	16.07*	13.61	17.21
	(9.21)	(8.75)	(10.78)	(10.62)	(9.49)	(9.37)	(10.79)	(10.56)
Spatial Inequality	52.47***	50.42***	63.33***	55.95***	56.87***	58.42***	71.17***	58.93***
	(14.32)	(14.49)	(14.65)	(14.75)	(15.16)	(15.02)	(15.31)	(15.66)
Ethnic-linguistic Fragmentation		20.03**	22.17**	14.79*		13.17	14.71*	9.07
		(9.47)	(9.26)	(8.56)		(8.56)	(8.46)	(8.16)
Adjusted R-Squared	0.7959	0.7995	0.8064	0.8174	0.7894	0.7911	0.7974	0.8132
Observations	173	173	173	173	173	173	173	173
Simple Controls	No	No	Yes	Yes	No	No	Yes	Yes
Geographic Controls	No	No	No	Yes	No	No	No	Yes
	Panel B: Spatial Inequality based on Second-Level Administrative Units							
	Greg Index				Ethnologue Index			
Ethnic Inequality	26.90**	23.86**	28.74***	19.85*	12.74	9.87	20.49*	20.52*
	(11.00)	(10.66)	(9.72)	(10.41)	(10.46)	(11.59)	(11.07)	(10.94)
Spatial Inequality	33.38***	34.10***	60.96***	52.79***	38.37***	39.23***	63.67***	51.93***
	(11.24)	(11.00)	(13.30)	(14.43)	(11.65)	(11.56)	(13.95)	(14.06)
Ethnic-linguistic Fragmentation		12.39	20.78*	16.13		6.04	6.65	6.20
		(12.47)	(11.14)	(10.55)		(10.28)	(9.48)	(9.34)
Adjusted R-Squared	0.8170	0.8169	0.8305	0.8438	0.8129	0.8119	0.8249	0.8450
Observations	135	135	135	135	135	135	135	135
Simple Controls	No	No	Yes	Yes	No	No	Yes	Yes
Geographic Controls	No	No	No	Yes	No	No	No	Yes

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates for the year 2000. The dependent variable is the under-5 mortality rates. Ethnic-linguistic fragmentation refers to the probability that two individuals randomly chosen in a country are not from the same group. Log land area and population for the year 2000 constitute 'Simple Controls', whereas the percentage of a country with fertile soil, terrain ruggedness, the average distance to the nearest ice-free coast, and an index of gem-quality diamond extraction are used as 'Geographic controls'. All the regressions include region-fixed effects, along with the real GDP per capita. Robust Standard errors are reported in the parenthesis.

Table A7: Relationship Between Ethnic Inequality (GREG) and Health/Development Outcomes for Alternate Years

	U5MR	Maternal Mortality	Stillbirths	LEAB	Child Stunting	GDP	Measles Vaccination	DPT Vaccination	Schooling Primary	Schooling Secondary	Contraceptive Use	Fertility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Association for the year 1992												
Ethnic Inequality	53.07***	-	-	-6.89***	-	-1.75***	-16.39**	-12.52*	0.17	-41.34**	3.37	1.16**
	(15.84)	-	-	(2.67)	-	(0.42)	(7.27)	(6.34)	(8.92)	(16.06)	(6.47)	(0.45)
Spatial Inequality	41.99	-	-	-6.46	-	-1.35***	0.65	-11.37	-14.23	-25.42	-13.97	1.31**
	(25.80)	-	-	(4.12)	-	(0.50)	(9.72)	(9.53)	(12.75)	(31.46)	(12.88)	(0.64)
Mean Dependent Variable	70.632	-	-	64.762	-	8.085	75.235	77.201	81.206	62.941	33.525	3.880
Adjusted R-Squared	0.6910	-	-	0.7437	-	0.6942	0.2470	0.3730	0.5595	0.7575	0.8844	0.7112
Observations	154	-	-	154	-	154	149	149	56	27	27	154
Panel B: Association for the year 2012												
Ethnic Inequality	24.05***	211.17***	5.50**	-4.21*	6.33	-0.78**	-2.73	-0.15	-6.88	-13.79	3.53	0.95**
	(8.60)	(66.90)	(2.20)	(2.14)	(5.44)	(0.36)	(5.23)	(4.81)	(4.95)	(8.31)	(12.95)	(0.39)
Spatial Inequality	14.87*	58.85	4.73**	-3.03	10.35**	-1.01***	-1.49	-2.45	3.98	11.78*	-42.13**	0.60
	(8.51)	(61.76)	(2.17)	(2.39)	(5.02)	(0.33)	(4.84)	(4.76)	(4.42)	(6.52)	(15.17)	(0.41)
Mean Dependent Variable	34.085	175.425	11.714	70.733	20.558	8.594	89.053	90.018	90.889	74.627	40.697	2.874
Adjusted R-Squared	0.7396	0.6476	0.7556	0.7935	0.6256	0.8180	0.2923	0.2463	0.3223	0.7821	0.7362	0.7110
Observations	169	167	169	167	136	169	169	169	129	87	37	167

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates. Specifications control for population and real GDP Per Capita for the respective year, along with the size of land area. All the regressions include region-fixed effects. Robust Standard errors are reported in the parenthesis.

Table A8: Relationship Between Ethnic Inequality (GREG) and Health/Development Outcomes - Using Pooled Data

	U5MR	Maternal Mortality	LEAB	Child Stunting	GDP	Measles Vaccination	DPT Vaccination	Schooling Primary	Schooling Secondary	Contraceptive Use	Fertility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Panel A: Region-Fixed Effects										
Ethnic Inequality	43.277***	268.881***	-5.428***	11.761***	-0.807***	-9.014***	-10.419***	-9.383***	-26.793***	-10.960*	0.958***
	(6.289)	(45.791)	(1.159)	(3.369)	(0.155)	(2.927)	(2.789)	(3.018)	(5.103)	(6.519)	(0.225)
Spatial Inequality	7.397	-14.048	-2.502**	3.458	-0.091	-4.080	-7.609***	0.686	5.027	6.547	0.457**
	(5.933)	(39.275)	(1.149)	(2.745)	(0.148)	(2.681)	(2.659)	(2.616)	(4.206)	(6.316)	(0.191)
Mean Dependent Variable	50.449	217.598	67.655	23.601	8.365	83.272	84.432	87.037	68.968	37.036	3.245
Observations	662	502	656	409	662	657	657	411	283	176	656
	Panel B: Country-Fixed Effects										
Ethnic Inequality	27.133**	178.967**	-0.020	1.037	-0.242	-7.502	-8.311	-9.024	-36.757***	-31.920***	0.264
	(13.437)	(86.005)	(1.659)	(2.511)	(0.186)	(7.024)	(6.555)	(9.262)	(8.437)	(10.400)	(0.327)
Spatial Inequality	16.629	157.339	-3.777*	6.374**	-0.521***	-2.375	-1.358	-18.512	4.653	0.700	0.438
	(21.976)	(164.186)	(2.089)	(3.222)	(0.179)	(8.696)	(7.996)	(13.284)	(13.884)	(10.589)	(0.324)
Mean Dependent Variable	50.449	217.598	67.655	23.601	8.365	83.272	84.432	87.037	68.968	37.036	3.245
Observations	662	502	656	409	662	657	657	411	283	176	656

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates. Pooled data for the years 1992, 2000, 2006, and 2012 is used. All the regressions include year-fixed effects, along with real GDP per capita. In Panel A, we control for region-fixed effects, whereas in Panel B, we use country-fixed effects. Robust Standard errors are reported in the parenthesis.

Table A9: Robustness Exercise: Alternative Measures of Ethnic Inequality

	GREG Index			Ethnologue Index		
	All Ethnic Areas	Excluding Capitals	Excluding Small Groups	All Ethnic Areas	Excluding Capitals	Excluding Small Groups
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic Inequality	50.87***	28.99**	44.83***	31.92***	33.07***	24.73
	(13.01)	(11.38)	(15.92)	(10.62)	(9.85)	(15.03)
Spatial Inequality	10.38	24.10*	54.21**	22.89	18.35	64.38***
	(16.20)	(13.36)	(24.55)	(15.04)	(14.05)	(23.48)
Mean Dependent Variable	56.439	60.068	56.439	56.439	60.494	56.439
Adjusted R-Squared	0.7722	0.7985	0.7853	0.7644	0.8087	0.7768
Observations	173	155	173	173	147	173

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates for the year 2000. The dependent variable is the under-5 mortality rates. Specifications control for population size, land size, and real GDP per capita. All the regressions include region-fixed effects. Robust Standard errors are reported in the parenthesis.

Table A10: Ethnic Inequality and U5MR - 2SLS Estimates - Accounting for Measurement Error in the Ethnic Group Mappings

	GREG Index					Ethnologue Index				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ethnic Inequality	67.57***	94.83***	81.65***	99.06***	55.96**	64.68***	95.75***	82.03***	97.24***	65.33***
	(15.37)	(24.10)	(26.06)	(34.96)	(28.03)	(12.10)	(21.96)	(21.56)	(29.53)	(23.32)
Spatial Inequality				-28.78					-22.71	
				(22.81)					(23.09)	
Administrative Unit Inequality					46.87***					38.37**
					(16.89)					(18.19)
F-Statistic	100.881	32.338	29.253	16.967	22.750	143.095	34.821	32.103	17.345	26.140
LM Statistic	43.724	22.347	21.541	14.689	20.074	43.724	22.347	21.541	14.689	20.074
Simple Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Observations	173	173	173	173	173	173	173	173	173	173

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country two-stage least squares (2SLS) estimates. The dependent variable is the under-5 mortality rates for the year 2000. Simple controls refer to the log of land area and population; geographic controls account for terrain ruggedness, the percentage of fertile soil in a country, the percentage of country with a tropical climate, gem-quality diamond extraction index, and the average distance to the nearest ice-free coast. All the regressions include region fixed effects and real GDP per capita. Robust Standard errors are reported in the parenthesis.

Table A11: Robustness Exercise - Ethnic Inequality (GREG) and U5MR - With Mean GDP Values

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnic Inequality	52.03***		51.96***		48.22***	47.96***	59.08***	54.65***	56.16***
	(9.94)		(13.02)		(12.02)	(15.62)	(16.50)	(16.66)	(18.60)
GDP Per Capita	-0.95***	-1.18***	-0.95***	-1.02***	-0.93***	-0.93***	-0.96***	-0.91***	-0.90***
	(0.25)	(0.28)	(0.25)	(0.28)	(0.26)	(0.26)	(0.25)	(0.26)	(0.26)
Spatial Inequality		35.97***	0.10			0.34	-1.08	-1.48	5.52
		(10.30)	(12.64)			(12.78)	(12.57)	(12.65)	(17.63)
Log Number of Ethnicities				9.83***	1.31	1.31		4.51	3.69
				(2.42)	(2.60)	(2.63)		(3.70)	(3.95)
Ethnic inequality in population							-55.10*	-65.50**	-65.11**
							(29.10)	(29.97)	(31.08)
Ethnic inequality in size							53.52*	53.71*	53.87*
							(31.15)	(30.16)	(30.54)
Log Land Area									-1.90
									(2.52)
Log Population (in 2000)									2.16
									(2.31)
Adjusted R-Squared	0.7760	0.7527	0.7746	0.7565	0.7748	0.7734	0.7784	0.7788	0.7773
Observations	171	171	171	171	171	171	171	171	171

Note: *, **, and *** denote significance at 10%, 5%, and 1% level, respectively. The table reports cross-country OLS estimates. The dependent variable is the under-5 mortality rates for the year 2000. Ethnic inequality indicator based on the GREG mappings is used. Real GDP per capita refers to the five-year average GDP constructed using GDP values for the years 1995–1999. The coefficient of real GDP per capita is multiplied by 1000. All the regressions include region-fixed effects. Robust standard errors are reported in the parenthesis.