Navarra Center for International Development

Working Paper nº 04/2017

The Health Costs of Ethnic Distance: Evidence from Sub-Saharan Africa

Joseph Gomes

University of Navarra

Navarra Center for International Development
WP-04/2017

The Health Costs of Ethnic Distance: Evidence from Sub-Saharan Africa

Joseph Flavian Gomes*
University of Navarra

November 2017

Abstract

We show that ethnic distances lead to worse child health outcomes by impeding access to health-related information. We combine individual level micro data from DHS surveys for fourteen sub-Saharan African countries, with a high-resolution dataset on the spatial distribution of ethnic groups at the 1×1 sq. km level constructed using an Iterative Proportional Fitting algorithm. We show that children whose mothers are linguistically more distant to their neighbours face higher mortality rates and are shorter in size. Linguistically distant mothers are also less likely to know about the oral rehydration product for treating children with diarrhoea.

Keywords: ethnic distance, linguistic distance, linguistic diversity, ethnic inequality, child mortality, African development, health inequalities.
JEL Codes: I14, O10, O15, Z10, Z13

1 Introduction

Ethnic distances matter for economic development. A recent macro literature has established that these distances can act as a barrier to trade, and the diffusion of innovation and technology, hence negatively affecting economic development (Guiso et al., 2009; Spolaore and Wacziarg, 2016, 2009). These macro studies shed little light on what these distances and barriers mean or how they operate at the micro level. We provide individual level micro evidence on how ethnic distances can be a barrier to health-related knowledge transmission resulting in higher child mortality rates for individuals who are more ethnically distant to their neighbours.

We combine high quality individual level micro data from the Demographic and Health Surveys (DHS) for fourteen sub-Saharan African countries with a novel dataset on the spatial

*University of Navarra, ICS, Navarra Center for International Development, Pamplona, Spain, E-mail: jgomes@unav.es. The author is grateful to Klaus Desmet, Ignacio Ortuño-Ortin, Ulrich Wagner, Jesus Carro, Sonia Bhalotra, Irma Clots-Figueras, Damian Clarke, Juan Jose Dolado, Joan Maria Esteban, Jim Fearon, James Fenske, Jed Friedman, Paola Giuliano, Oded Galor, Saumitra Jha, Eliana La Ferrara, Edward Leamer, Matilde Machado, Stelios Michalopoulos, Ricardo Mora, Owen Ozier, Dan Posner, Diego Puga, Richard Scheffler, Alessandro Tarozzi, Nico Voigtlaender, Romain Wacziarg, Yuya Kudo and all seminar/conference participants at the World Bank ABCDE 2016 conference, NOVAFRICA conference, ISI Delhi ACEGD 2015 Conference, NCID Research Workshop 2014, Warwick Summer Workshop in Economic Growth 2014, UC3M, ISER, Nottingham, Essex Govt Dept, UPNA, and Kent for their comments and suggestions; to Jim Fearon for generously sharing the data on ethnicity and languages matching; to UCLA Anderson and UPF Barcelona for their hospitality.

distribution of ethnic groups at the level of approximately 1 x 1 square *km* constructed using an Iterative Proportional Fitting (IPF) algorithm from Desmet et al. (2016). Exploiting the individual mother's location, ethnicity (which we map to languages from the Ethnologue database) and the spatial distribution of language groups, we are able to construct individual level ethnic distances of the mothers from people living around them. This allows us to study how the mother's ethnic distance from her neighbours affects her children's health outcomes.

Our primary focus is on the concept of individual level ethnic distance, which measures how ethnically different an individual is from others living in the same neighbourhood or region. Instead of taking a stand on what the appropriate neighbourhood or region for calculating these distances should be, we calculate these distances drawing circles of different radii around the mothers. Following a burgeoning literature (Fearon, 2003; Desmet et al., 2012, 2009; Esteban et al., 2012a,b; Laitin and Ramachandran, 2016), the ethnic distance between any two ethnic groups is measured by how different the languages that the two groups speak are. The actual distance metric is based on the number of shared branches between any two languages according to the Ethnologue language trees.

Our individual level measure of ethnic distance stands in contrast to the more aggregate measures of ethnic fractionalization (Alesina et al., 2003) or genetic diversity (Ashraf and Galor, 2013). For instance, ethnic fractionalization gives us the probability that two randomly selected individuals from a given region belong to two different ethnic groups. All individuals, regardless of their ethnicity, face the same level of ethnic fractionalization in a region. However, ethnic distance is individual-ethnicity specific and tells us how ethnically different any particular individual is to others living in the same region. We are able to identify the effects of individual level ethnic distances while controlling for more aggregated measures of diversity.

Our primary finding is that children of mothers who are ethnically distant from their neighbours have a higher probability of dying as children. This result holds regardless of whether we focus on child mortality (dying before reaching age 5), infant mortality (dying before reaching age one) or neonatal mortality (dying in the first month after birth). Our results are robust to a rich set of controls, including several birth specific variables like child gender and birth order; mother specific variables like education and wealth; ethnicity and religion fixed effects; and time varying region and ethnicity fixed effects; apart from measures of aggregate diversity like fractionalization or polarization.

Our results are much stronger and more robust for individuals who have never moved from

their village of residence. Digging deeper into this heterogeneity, we are able to get some indications of the channels via which ethnic distance affects child mortality. Restricting our sample to individuals who have never migrated from their village of residence, we see that linguistically distant individuals are less likely to have heard of the oral rehydration product (ORS) for treating children with diarrhoea.[2] This shows that information does not flow smoothly across ethnic lines and individuals who are ethnically distant to their neighbours lose out.[3]

We hence argue that individual ethnic distances act as barriers to accessing health-related knowledge and information, which in turn leads to worse health outcomes. The fact that our results are stronger for individuals who have never migrated could also be reflecting this. Individuals who have moved from other places are more likely to have acquired health information in their previous place of residence, which we cannot observe. However, individuals who have never moved have acquired health-related knowledge in their current place of residence. Hence, being linguistically distant affects them more than they affect individuals who have moved.[4]

We also find strong and robust negative effects of mother's linguistic distance on child height, measured either by the child's height-for-age Z-score i.e. HAZ or by stunting, but no statistically significant effects on child weight. This is pertinent to our linguistic distance being a barrier to information interpretation since information is crucial for child height which is in turn an important marker of child health. For instance, Thomas et al. (1991) show that almost all the impact of maternal education on child height can be explained by access to information. We also find evidence of linguistically distant individuals who have never moved from their place of residence to be less likely to have received tetanus injections and iron tables during pregnancy.

Since we control for ethnicity and religion specific fixed effects our results are not driven by heterogeneity in unobservable characteristics across ethnic or religious groups. The use of time varying region fixed effects allows us to discard explanations based on locational, geographic and environmental advantages. The inclusion of time varying ethnicity fixed effects lets us abstract from explanations based on political ethnic favouritism.

The ethnic favouritism literature has highlighted the importance of the ethnicity of the countries' leaders on different outcomes like mortality rates in different ethnic groups (Kudamatsu,

[2]We measure ethnic distance using linguistic distance. Hence, the terms ethnic distance and linguistic distance will be used interchangeably throughout the paper.

[3]This might also be indicative of lack of other types of information including feeding-practices which are crucial for child health (Malhotra, 2012; UNICEF, 2012), but which we cannot measure in the current set-up.

[4]Also, individuals with the possibility to move, might choose to move to more favourable locations. However, we find that migrants generally have higher linguistic distance to their neighbours, and also face higher child mortality rates.

2009; Franck and Rainer, 2012), schooling outcomes for children (Kramon and Posner, 2016) and road building (Burgess et al., 2015). Again, De Luca et al. (2016) and Dickens (2016) show how region specific transfers from the centre can benefit certain ethnic groups at the cost of others. We are able to control for such ethnic favouritism, and focus on how ethnic distance can act as a barrier over and above its possible effects through ethnic favouritism.

Another alternative explanation could be that linguistically distant individuals face more discrimination in access to healthcare and other public goods in general, rather than (or in addition to) having lower access to information. We do not find any significant effects of linguistic distance on the access to different public goods including education, water and electricity. Hence, it is unlikely that discrimination is a channel.

Finally, another novel finding is that linguistic distances driven by splits that occurred thousands of years ago explain the child health outcomes better than more recent splits. We find this by varying the values of a parameter that determines how fast the distance between any two languages declines as the number of shared branches increases. This is reminiscent of the findings of Spolaore and Wacziarg (2009, 2016) who argue genetic/ancestral distances act as barriers to the diffusion of development.

As far as the marginal effects are concerned, considering a circle of 50 km radius around the mother, a one SD increase in linguistic distance leads to approximately 8 additional child deaths per 1000 live births which is around 2% SD deaths in the sample. If we restrict our sample to non-migrants then the corresponding number goes up to around 15 additional child deaths or 3.5% SD deaths.

While we include a rich set of controls in our specifications, to allay remaining concerns about endogeneity we turn to recently developed methods by Altonji et al. (2005). Using their heuristics and incorporating insights from Oster (2013), we show that our results are unlikely to be driven by selection on unobservables. If anything, selection on unobservable variables drives our main coefficient of interest away from zero.

Let us consider an illustrative example from the data, to help clarify our results. Consider two Tamasheq speaking mothers, A and B, residing in the Timbuktu region of Mali in two nearby villages. Mother A lives in a village where the Tamasheq speakers share their linguistic homeland with Arabic speakers. While Mother B lives in a predominantly Songhay Koyra Chini speaking village. Tamasheq and Arabic are both Afro-Asiatic languages, while Songhay Koyra Chini is a Nilo-Saharan language. Languages belonging to the same language family share more

branches in common and have split from each other more recently than languages in different language families. This implies that Mother A is linguistically less distant to her neighbours compared to Mother B.

Both these mothers live in the same region, in nearby locations, and hence face similar geographic and environmental conditions. In the data however, we observe that a Tamasheq speaker living in a predominantly Afro-Asiatic language speaking village is more likely to have heard of ORS and faces lower child mortality rates compared to a Tamasheq speaker living in a predominantly Nilo-Saharan language speaking village, not far from each other. Following our analysis, we argue that these mothers have differential access to health-related information due to differences in their ethnic distance from their neighbours. This is reflected by the differences in knowledge about ORS, which in turn has profound implications for their children's healths.

We contribute to four different strands of the literature. First, we contribute to the literature that demonstrates the role of ethnic and cultural distances for economic outcomes (Spolaore and Wacziarg, 2009, 2016; Guiso et al., 2009; Desmet et al., 2017). In particular, Spolaore and Wacziarg (2009) argue that more closely related individuals and hence societies can more easily learn from each other and adopt each other's innovations. However, they are largely agnostic about specific mechanisms, and how exactly these barriers operate. We take this to the micro-level and show how ethnic distances might act as a barrier to health-related information leading to higher child mortality rates among linguistically distant groups.

Second, we contribute to a small but burgeoning literature that has highlighted the role of ethnic distances in explaining economic development via human capital accumulation (Laitin and Ramachandran, 2016; Shastry, 2012), trade flows (Isphording and Otten, 2013), literacy and labour market outcomes of immigrants (Isphording, 2013), and market integration (Fenske et al., 2017).[5] We are the first to relate individual level health outcomes to the ethnic distance of the individual from her neighbours exploiting high quality data on the spatial distribution of ethnic groups, particularly focussing on the information channel.[6]

Third, we contribute to a huge literature investigating the effects of ethnic diversity on different political economy outcomes.[7] While most of this literature is at the cross country level,

[5] See also Desmet et al. (2012), Desmet et al. (2009), Gomes (2013), Esteban et al. (2012a,b) for aggregate cross-country measures incorporating ethnic distances.

[6] See also Fisman et al. (2012), who show how cultural proximity mitigates problems of asymmetric information in lending. Similarly, Pongou (2009) points out that information circulates more easily within ethnic groups than across and highlights the implications for HIV/AIDS in Africa.

[7] Ashraf and Galor (2013), Miguel and Gugerty (2005), Habyarimana et al. (2007), Alesina et al. (2003), Desmet et al. (2012), La Porta et al. (1999)

there has been a recent surge in the number of studies looking at different political economy outcomes at the local level. These include: Alesina et al. (1999) (U.S. cities); Dahlberg et al. (2012) (Swedish municipalities); Munshi and Rosenzweig (2015) (wards in India); Algan et al. (2016) (apartment blocks in France) and Montalvo and Reynal-Querol (2016) (1 x 1 degree pixels in Africa); among others. Focussing on 5 x 5 sq. *km* cells Desmet et al. (2016) show how local interaction affects public goods outcomes at the national level. In contrast to this literature, we focus on individual level ethnic distances, controlling for ethnic diversity at the local level in addition to a rich set of controls.

Finally, we contribute to the literature on ethnic inequality. Ethnic inequality defined as the inequality in well-being across ethnic groups that coexist, is bad for economic growth (Alesina et al., 2012), provision of public goods (Baldwin and Huber, 2010), and can lead to civil conflicts (Mitra and Ray, 2010; Gomes, 2015). We show how ethnic distances might lead to ethnic inequality in health outcomes in Africa. While the existence of ethnic inequality in child mortality rates has been pointed out previously (Gyimah, 2002; Brockerhoff and Hewett, 2000), we are the first to underscore the importance of ethnic distance.

The rest of the paper is organized as follows. In Section 2 we discuss the data sources and how the different variables are constructed. In Section 3 we present our empirical analysis and results. In Section 4 we provide some evidence on the channels through which ethnic distance affects child health. In Section 5 we conclude.

2 Data

In this paper we aim to estimate the effects of the ethnic distance of the mother from people living around her on her children's health outcomes. For this purpose, we require the mother's GPS location, her ethnicity, and the spatial distribution of ethnic groups around her. In this section we explain how the different variables used in our analysis were constructed and discuss their data sources.[8]

2.1 Spatial Distribution of Ethnic Groups

In order to construct the ethnic distance of the mother from people living around her, we need the distribution of ethnic groups across space. Until recently there was no comprehensive

[8]Our sample is comprised of fourteen countries (see Figure C.1). See Appendix Section A.1 for more details.

database on the spatial distribution of ethnic groups available at a geographically disaggregated level. Desmet et al. (2016) fill this gap by constructing a comprehensive database on the spatial distribution of ethnic groups for the entire world (223 countries) at a resolution of approximately 5 x 5 sq. *km*. Moreover, they generate a base dataset and make available codes using which such data can be generated at different levels of geographic disaggregation.

Desmet et al. (2016) construct their database using two different sources of data. For the spatial distribution of population they use the LandScan database. At a resolution of 30 arc seconds by 30 arc seconds (approximately 1 x 1 sq. *km* at the equator), LandScan is the finest resolution population distribution database available for the entire world.[9] For information on ethnic groups they use the 17^{th} edition of the Ethnologue database (Lewis et al., 2014), from the World Language Mapping system (WLMS),[10] which maps 6905 distinct linguistic groups for the whole world and is the most comprehensive database on linguistic groups available. The linguistic groups are represented in the form of polygons across space where each polygon represents the homeland of a particular linguistic group. Areas where multiple languages are spoken are represented via overlapping polygons. The total population pertaining to a particular linguistic group within a particular political boundary is also provided.

Then using an Iterative Proportional Fitting (IPF) algorithm, they combine the information from the above two sources to come up with a distribution of languages for each 5 x 5 sq. *km* cell, in every country in the world. The IPF algorithm, which is widely used in statistics, ensures that while allocating languages to cells, the total population of each country, the population of each of the cells and the population speaking each of the languages in every country exactly add up to consistent totals.[11] We exploit their methods and base data to construct our database on the spatial distribution of linguistic groups at the 1 x 1 sq. *km* level for the fourteen countries in our sample.

In Figure C.2 we plot the polygons representing the linguistic regions in the fourteen countries in our sample based on the Ethnologue database. The polygons of different colours represent the different language groups. There are many regions in these countries where multiple languages are spoken, which are represented by overlapping polygons. Unfortunately, these overlapping polygons are not distinguishable in the map. In order to illustrate this possibility let us consider an example from Mali.

[9] For more details see http://web.ornl.gov/sci/landscan/.

[10] WLMS Version 17, World Geo Datasets

[11] Please refer to Desmet et al. (2016) for more details on the method. For more details on IPF please refer to Bishop et al. (1975), Deming and Stephan (1940), and Fienberg (1970).

Figure C.3 gives the linguistic map of Mali and the area highlighted with a blue border in the south-eastern corner of the country is the linguistic homeland of the Mamara Senoufo speakers. In Figure C.4, we zoom into this area. Notice that while Mamara Senoufo is spoken in this entire area bordered in blue, there are different possible overlaps with other languages. First, in the light blue shaded polygon in the south-eastern corner of the map, there are no other languages spoken apart from Mamara Senoufo. In the polygon with a darker shade of blue, just north of this area, both Mamara Senoufo and Northern Bobo Madare are spoken. In the green shaded polygon in the centre of the map, Mamara Senoufo and Maasina Fulfulde are spoken. Finally, in the pink shaded polygon in the west, Mamara Senoufo is spoken with two other languages viz. Maasina Fulfulde and Bamanankan. Our IPF algorithm takes into account all these possibilities.

Let us now explore the process of generating the final spatial distribution of language groups in some more detail expanding on this example of Mali. While Figure C.3 gave the map of linguistic regions in Mali, Figure C.5 plots Mali's population distribution at the 1 x 1 sq. *km* level from LandScan. In Figure C.6 we overlay the language polygons on the population distribution for Mali. Based on the data generated from this combined map, we allocate languages to the 1 x 1 sq. *km* pixel level following the IPF algorithm of Desmet et al. (2016). The other input to the algorithm is the information on total populations pertaining to each of the language groups in each country from Ethnologue. We repeat this exercise for each of the fourteen countries in our sample to come up with the spatial distribution of linguistic groups for these countries.

Other possible sources of sub-national data on ethnic diversity include Alesina and Zhuravskaya (2011) (district level for 92 countries), and Gershman and Rivera (2016) (around 400 first level administrative regions in 36 countries of sub-Saharan Africa). However, these data are at the administrative region level and for the purposes of our paper we require data at the disaggregated cell level.[12]

Matuszeski and Schneider (2006) have also generated pixel level data on the spatial distribution of languages. However, they do not take into account languages which are considered widespread by Ethnologue, nor languages for which Ethnologue only provides a point as the location rather than a polygon. Also, their data construction does not ensure consistency of language population values, which the IPF algorithm ensures.

Another alternative to the Ethnologue data would be the Weidmann et al. (2010) GREG

[12] Desmet et al. (2016) demonstrate correlations of 0.80 at the regional level and 0.96 at the country level with the Gershman and Rivera (2016) database based on census and survey data.

(Geo-Referencing of Ethnic Groups) database based on the Atlas Narodov Mira. However, these data are a lot less detailed containing information on only 929 language groups compared to Ethnologue's nearly 7000 groups.

2.2 Linguistic Distance

We measure ethnic distances using the linguistic distance between the languages that the different ethnic groups speak. For the fourteen countries in our sample, DHS data provides us with the ethnicity of the sampled mother. We first match the ethnicity of the mother to the unique language that the ethnicity speaks.[13] We follow a wide stream of papers in the recent literature including Fearon (2003), Desmet et al. (2009), Desmet et al. (2012), and most recently Desmet et al. (2016) and Gershman and Rivera (2016), among others, which use linguistic tree diagrams to measure distances between languages. The distance between two languages j and k using this approach is defined as:

$$\tau_{jk} = 1 - \left(\frac{l}{m}\right)^{\delta} \tag{1}$$

where l is the number of shared branches between languages j and k, m is the maximum number of branches between any two languages, and δ is the decay factor, which is a parameter that determines how fast the distance declines as the number of shared branches increases. Data on language trees come from the Ethnologue database.[14]

The decay factor δ measures, "how much more distant should we consider two languages from different families to be relative to languages that belong to the same family" (Desmet et al., 2009). There is no consensus in the literature on what the value of δ should be. While in their empirical exploration Desmet et al. (2009) find that values of δ between 0.04 and 0.10 perform well and choose a δ of 0.05, Fearon (2003) uses a δ of 0.5. Since, there is no theoretical basis for choosing one value of delta or the other, we let the data tell us which values of δ perform better than others and find that lower values of δ perform better than higher values and fix δ at 0.0025. As we will later show, choosing a δ of 0.05 like Desmet et al. (2009) leads to qualitatively similar

[13] We provide the exact procedure of the ethnicity to language mapping in the appendix Section A.2. And in appendix Section A.1 we provide the list of countries and DHS surveys used in the paper.

[14] There are other ways of measuring linguistic distances. For example, Dyen et al. (1992) use the proportion of cognates in any two languages. Again, Isphording (2013) uses not only cognates but also the number of sounds that need to be changed between two words that have the same meaning (say, *Tu* in Spanish and *You* in English) in two different languages. However, distance calculated using language tree diagrams are more useful since the data is a lot more comprehensive and exists for all countries. See Desmet et al. (2009) for a discussion.

results. However, a high value of δ, like the one chosen by Fearon (2003) leads to insignificant results. We discuss the implications of this finding in more detail in the empirical section.

In order to understand what the different values of δ imply in practice, let us consider the two Indo-European languages Greek and Italian.[15] Following the language tree from Ethnologue, these two languages share one common branch. Taking a δ of 0.5 like Fearon (2003), the distance between them is 0.74. Again, if we consider Chinese and Italian which belong to completely different families and thus share no branches in common, the distance between them is one. On the other hand, if we take a δ of 0.05 following Desmet et al. (2009), the distance between Greek and Italian becomes 0.13, whereas that between Chinese and Italian continues to be one. Finally, if we choose a δ of 0.0025, the distance between Greek and Italian is 0.007 while that between Chinese and Italian is still one. [16]

For our final analysis we need to calculate the average linguistic distance of each mother in our sample to all other individuals living around her in circles of different radii. The linguistic distance LD_j, for mother j (who speaks language j) to all other individuals in the circle is given by,

$$LD_j = \frac{1}{n}\sum_{k=1}^{n} \tau_{jk} \tag{2}$$

where there are n individuals living in the circle and k represents the language groups of each of those n individuals. The function τ_{jk} is defined by the formula 1. The geographic distance between individuals are calculated using the formula for the great circle distance. The geographic distance between any two points in space ℓ and k, denoted by $|\ell, k|$, is computed as the great circle distance:

$$|\ell, k| = r_E \arccos(\sin(lat_\ell)\sin(lat_k) + \cos(lat_\ell)\cos(lat_k)\cos(long_\ell - long_k) \tag{3}$$

2.3 Linguistic Diversity

Our primary measure of ethnic diversity is the commonly used measure of ethno-linguistic fractionalization (ELF). ELF has been found to have a negative effect on a host of socio economic outcomes (Alesina et al., 2003) and has often been blamed for Africa's poor economic performance (Easterly and Levine, 1997). However, a recent literature has emphasized that for certain

[15] Example from Desmet et al. (2009).

[16] In Figure C.9 we provide simulations of how distances between any two languages change as they share different number of branches ranging from 0 to 15, for different values of δ.

outcomes like civil conflicts, ethnic polarization (ELP) rather than fractionalization is more relevant (Esteban and Ray, 1994, 1999; Montalvo and Reynal-Querol, 2005). Hence, in some of our specifications we explicitly control for polarization rather than fractionalization.

The fractionalization measure ELF_j^k gives the probability that two randomly selected individuals from a given region j speak two different languages. The polarization measure ELP_j^k on the other hand measures how far the distribution of the linguistic groups in region j is from the bipolar distribution (i.e. the (1/2, 0, 0, ... , 0, 1/2) distribution) which represents the highest level of polarization (Montalvo and Reynal-Querol, 2005). The fractionalization index is maximized when each individual in the region belongs to a different linguistic group, while the polarization index is maximized when there are only two groups in the region and they are of equal size.[17]

The other issue while calculating these ELF/ELP indices is whether two closely related languages should be considered as two different groups or as the same group. For instance, let us consider the two languages of Gikuyu and Kiembu which are both spoken in Kenya. They are both Bantu languages belonging to the broader Niger-Congo language family. Their language family structure according to Ethnologue is formed by the following branches: "Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Narrow Bantu, Central, E, Kikuyu-Kamba." Now consider the language Dholuo, which is another language spoken in Kenya, but is a Nilotic language belonging to the broader Nilo-Saharan language family. More specifically, its language family structure is as follows: "Nilo-Saharan, Eastern Sudanic, Nilotic, Western, Luo, Southern, Luo-Acholi, Luo."

According to Ethnologue, Gikuyu has 73% lexical similarity with Kiembu, which is not surprising given that they share many branches in common and belong to the same broader language family.[18] On the other hand, Dholuo does not share any branches in common with either of these two languages. The question is should we consider Gikuyu, Kiembu and Dholuo as three different languages while constructing our ELF measures or should we club Gikuyu and Kiembu as the same language given their similarity.

We follow the recent literature (Desmet et al., 2012, 2016; Gershman and Rivera, 2016) and use the Ethnologue language trees to calculate these measures at different levels of aggregation in order to take into account the distances between the language groups. There are 15 possible

[17] The reader is directed to Montalvo and Reynal-Querol (2005) for a detailed discussion and comparison of the two measures.

[18] See https://www.ethnologue.com/language/kik

levels with Level 15 (Level 1) representing the most disaggregated (aggregated) level. While at Level 15, Gikuyu and Kiembu are treated as two different languages, at higher levels these two languages are treated as the same linguistic group since they are both Niger-Congo languages belonging to the Kikyu-Kiemba sub branch.

Formally, the two measures of ELF and ELP, in region j and at linguistic aggregation level k, are defined as follows:

$$\text{Fractionalization: } ELF_j^k = 1 - \Sigma[S_{i(j)}^k]^2. \tag{4}$$

$$\text{Polarization: } ELP_j^k = 4\Sigma[S_{i(j)}^k]^2[1 - S_{i(j)}^k]. \tag{5}$$

where $S_{i(j)}^k$ is the proportion of the population speaking language i at linguistic aggregation level k in the geographic region j. The disaggregated nature of our data allows us to calculate diversity at different levels of geographic aggregation. Instead of taking a stand on what region or geographic aggregation should be more appropriate for calculating these indices, we calculate these measures of diversity, and our measures of linguistic distance, at the circle level, drawing circles of different radii around the mothers.

2.4 Child Mortality

Our individual level child survival data are based on the Demographic and Health surveys (DHS). These individual level data are available for many developing countries from across the world. Funded by the the U.S. Agency for International Development (USAID), the DHS has been conducting surveys in several developing countries since the 1980s. By interviewing a nationally representative sample of women of child bearing age (15 to 49), the DHS collect data on all the children they have ever given birth to in the past including the children who did not survive till the time of the interview. The standardized components of the DHS questionnaires can be used to compile cross country micro data sets.

Child mortality is the death of a child before reaching the age of five. If the child dies before reaching the age of one then it is termed as infant mortality while if the child does not survive for a month after its birth, then it is termed as neo-natal mortality. In Figure C.7 we plot the 28,993 DHS clusters which show the geographic locations of the 208,898 individual mothers whose children's survival outcomes we use in this study. In Figure C.8 we show the locations

of the individual mothers in the case of Mali along with the 25 km circles around the mothers' locations and the language groups in the background.

2.5 Other DHS Data

2.5.1 Other Health Outcomes

We use a host of other child and mother level variables which are constructed from the DHS data. These variables include, the height-for-age Z-score (HAZ), the weight-for-age Z-score (WAZ), whether the child is stunted (defined as the child being less than 2 standard deviations of HAZ), immunizations received (polio, DPT, measles, and tetanus), and whether the mother received iron tablets during pregnancy. The other individual level variables used as controls in the empirical section are also constructed using the DHS data. These include different child, mother and household level variables which the literature has found to be important for understanding child mortality.[19]

2.5.2 Migration

The DHS provides us with a variable that gives us the "number of years the respondent has lived in the village, town, or city where she was interviewed." Exploiting this question we are able to determine which individuals have always lived in the DHS cluster where they were interviewed and which individuals have moved there from elsewhere. The migration status variable is unfortunately available only for 13 of the 14 countries in our sample and also for some of the surveys rather than the full set of surveys we use in the complete analysis. In total it is available for 25 of the 30 surveys used in the study.[20] Of the 208,898 mothers in the sample, the migrant status variable is available for 167,130 of the mothers, of which another 2,822 mothers are identified as temporary visitors rather than residents and are dropped from the sample. Hence, finally we have information on the migrant status of 164,209 mothers.

2.5.3 Access to Information and Public Goods

The DHS surveys ask each respondent whether they have either heard or used the oral rehydration product (ORS) for treating children with diarrhoea. Using the responses to this question,

[19] The full list of these variables is provided in Section 3.1.

[20] Missing for one of three surveys for Burkina Faso, and Ethiopia; one of the two surveys for Guinea, and Senegal; and for Uganda.

we create a 0-1 binary variable called ORS which takes the value 1 if the individual has either heard or used the Oral Rehydration Solution used for treating children with diarrhoea, and 0 otherwise. This question will serve as a test for access to health-related knowledge or information.

We also exploit information on whether the respondent's household has access to electricity, whether the household has access to water i.e. they take less than 30 minutes to a source of water, the individual's educational attainment and whether the individual is literate or not. Among these, electricity access, water access, and literacy are binary variables, taking the value 1, if the individual has access to these variables in the first two cases or is literate in the last case, and 0 otherwise. Educational attainment is a categorical variable taking the values 0 (no education), 1 (incomplete primary education), 2 (completed primary education), 3 (incomplete secondary education), 4 (completed secondary education), and 5 (higher education). These variables allow us to measure access to public goods in general.

3 Empirical Analysis

3.1 Econometric Specification

Our primary relationship of interest is that between the individual child level mortality outcome and the linguistic distance (LD) of the mother from her neighbours, while controlling for the overall linguistic diversity of the neighbourhood, and a host of other controls. We provide our baseline specification in equation 6.

$$y_{iet} = \alpha_w + \alpha_{rel} + \alpha_{et} + \alpha_{rt} + \beta_1 LD_{ie} + \beta_2 ELF_i + \beta_3 X_{it} + \beta_4 X_i + \epsilon_{iet} \tag{6}$$

where y_{iet} is the mortality outcome of child 'i' born to mother belonging to ethnicity 'e' in birth year 't'. It is a binary variable which takes the value 1 if the child dies before reaching the age of five and 0 otherwise. In some of the analyses we will replace child mortality by other child health variables as the dependent variable y_{iet}. These include infant mortality, neonatal mortality, height-for-age Z-score (HAZ), stunting, and the weight-for-age Z-score (WAZ).

The LD_{ie} variable is our primary variable of interest and it gives the linguistic distance of the mother of child 'i' belonging to ethnicity 'e' from people living within circles of different radii around her, constructed as explained in section 2.3. The ELF_i variable gives the ethno-

linguistic fractionalization in the circles of different radii around the mother. We will replace ELF_i by ethno-linguistic polarization, ELP_i in some of the specifications later as a robustness check. For calculating the linguistic distance, ELF and ELP variables, we have used circles of different radii, viz. 25, 50, 75, 100, 125, 150, 175, 200, 250 km around the mother.

The variables X_{it} and X_i come from the literature on child mortality and have been found to be important for child mortality.[21] X_{it} includes birth specific variables viz. female child dummy, mother's age at birth, mother's age at birth squared, multiple birth indicator, birth order, birth order squared, short birth spacing prior to the birth, and short birth spacing after the birth. X_i includes mother specific variables viz. the location of the mother in the form of an urban dummy, dummies for her educational attainment and her families' wealth index.[22] We also control for the distance of the mother's location from the capital and the logged population in the circle.

We include time varying region fixed effects α_{rt} which allows for the non-parametric evolution of year effects for each of the 109 DHS regions in the data. These time varying region fixed effects ensure that our results are not driven by the geographic and environmental advantages of some regions; region specific shocks like conflict and natural calamities; or region specific transfers from the centre which benefit certain ethnic groups at the cost of others (De Luca et al., 2016; Dickens, 2016). α_{rel} represents the religion specific fixed effects, which controls for differences in religious beliefs and practices across different individuals.

We include time varying ethnicity fixed effects, α_{et}, to control for unobserved heterogeneity across ethnic groups. This allows us to identify the effect of ethnic distance on child mortality that is not driven by ethnicity specific characteristics like ethnic dominance of certain groups or cultural differences leading to differences in health practices between different groups. Moreover, since these ethnic group specific fixed effects are time varying, having a co-ethnic as the head of the country does not affect our results (Kramon and Posner, 2016). Finally, α_w controls for the survey wave specific fixed effects.

We use a linear probability model for our regressions. β_1 is our coefficient of interest since it gives the effect of linguistic distance of the mother on the probability of death of the child before reaching the age of five. Due to different possible endogeneity concerns, giving a causal interpretation to β_1 is not straightforward. Also, we cannot use mother specific fixed effects since the ethnic distance variable does not vary across time for the same mother. However, we

[21] See Kudamatsu (2009), Baird et al. (2011), Franck and Rainer (2012) for example.

[22] The wealth index is a categorical variable taking the values 1 (lowest wealth level) to 5 (highest wealth level).

are able to control for a host of maternal and birth characteristics which alleviate endogeneity concerns to a great extent. Moreover, we later use the insights from Altonji et al. (2005) and show that our results are not driven by selection on unobservables. This increases our faith in the causal interpretation of β_1. The standard errors are clustered at the region level for the 109 regions in the sample.

3.2 Summary Statistics

In Tables B.1-B.5 we provide the descriptive statistics of the variables used in the study. First, in Table B.1 we provide the summary statistics for the variables used in the child level regressions. And in Table B.2 we provide the summary statistics of the variables used in the mother level regressions in Section 4. We have fourteen countries and a total of thirty surveys with information on the births and deaths of over 860,000 children.[23] For the child mortality variable we can consider only the children who would have already reached the age of five by the day of the sampling, since we do not know if the others are going to survive till the age of five or not.

Thus, as can be seen from Table B.1, in the child mortality sample we have information on the births of 654,672 children out of which about 23% do not survive until their fifth birthday. About 12% of the 816,268 children, for whom we have infant survival data, do not survive until their first birthday. And finally, from 862,358 children, 5.5% die within the first month of their birth. The sample is made of 49% female children. And an overwhelming majority of the sample is rural with only 22.5% of the births taking place in urban areas. Births and deaths in the sample span from 1955-2011.[24]

In Table B.3 we provide the summary statistics of the linguistic distance variables for different values of the decay factor δ, and for circles of different radii around the mother. Correspondingly in Tables B.4 and B.5 we provide the descriptive statistics for the ELF and ELP variables at different levels of linguistic aggregation and for circles of different radii around the mother. The linguistic distance and fractionalization variables all lie between 0 and 1.

In Table B.6 we provide the correlations between our diversity measures ELF/ELP (at 4 different levels of aggregation), and the LD variables (for the three alternative values of δ) at the individual mother level, for the 206,076 mothers in our sample. Finally in Table B.7 we provide the correlation between ELF and ELP at different levels of aggregation.

[23] See appendix Section A.1 for a list of countries and DHS surveys used.

[24] We have dropped the information of mothers who were identified as temporary visitors in the sample.

3.3 Mother's Ethnic Distance and Child mortality

In Table 1 we present our first set of results. In this table, we regress child mortality on the linguistic distance of the mother from people living around her, while controlling for overall ELF. In this baseline specification we calculate the linguistic distance and the ELF variables by considering a radius of 50 *km* around the mother, and a decay factor δ of 0.0025 for the LD variable.

In Column 1 we have a parsimonious specification, controlling only for the survey wave, and time varying region fixed effects. In column 2, we add ethnic group fixed effects. In column 3, we add individual level controls including gender of the child, age of the mother at birth of the child and its square term, multiple birth dummy, urban residence dummy, birth order and its square term, short birth spacing prior and post the birth of the child, educational attainment, dummies for the wealth index and religion fixed effects. In column 4, we add the logged population of the circle and logged distance to the capital. These variables respectively control for population density and geographic isolation. And finally in column 5, we add time varying ethnic group fixed effects. Column 5 is our most complete and hence most preferred specification. From here on, unless otherwise specified, we will use this specification to present the other results of the paper.

From Table 1 we notice that LD significantly increases the probability of child death, and this effect is robust to a host of controls. We also see that ELF if anything has a negative effect on child mortality. This implies that on the one hand, children of mothers who are linguistically distant to others living around them, have a higher mortality rate. On the other hand, children of mothers living in more linguistically fractionalized localities face lower rates of mortality. However, while LD has a significant effect on child mortality, ELF does not.

In Table 1, we chose to calculate linguistic distance of the mother from all individuals living in a 50 *km* circle around her. Using the complete specification of column 5 from Table 1, in Table 2 we present results for alternative radii ranging from 25 *km* to 250 *km*. Our results remain quite similar, though the effect size goes up marginally for higher radii. Again, as discussed earlier, for calculating our linguistic distance variable LD, we also need to choose the decay factor δ. For the results presented in Tables 1 and 2 we have calculated LD using a decay factor δ of 0.0025. In the following section we will discuss this choice of δ and its implications in more detail.

In terms of marginal effects, a one SD increase in LD increases child mortality by 1.6-2.6 %

SD child deaths, depending on the radius of the circle. We have considered circles ranging from 25 *km* to 250 *km* radii as presented in Table 2. This implies that a one SD increase in LD leads to 6.6 to 10.5 additional child deaths per 1000 live births. If we consider a radius of 50 *km* like we have done in Table 1, then a one SD increase in LD leads to around 8.2 additional deaths per 1000 live births, which is about 2% SD deaths in the sample.[25]

In our analysis we have included all the births in the entire maternal history of the mother. One possible concern with using retrospective data is recall bias. This stems from the fact that women might be less likely to accurately remember more distant births and deaths. To minimize recall bias we replicate our baseline results using births and deaths occurring in the ten years preceding the date of the survey (following Baird et al. (2011) and Kudamatsu et al. (2012)). The results remain qualitatively similar.[26]

3.4 Varying the Decay Factor δ

For Tables 1 and 2, we chose to calculate linguistic distance using a decay factor δ of 0.0025. In Appendix Tables B.8 and B.9 we replicate the results from Tables 1 and 2 for alternative values of δ. In particular we present our results for three alternative values of δ viz. $\delta = 0.0025$ (Panel 1), $\delta = 0.05$ *à la* Desmet et al. (2009) (Panel 2) and $\delta = 0.50$ *à la* Fearon (2003) (Panel 3). From the results in the three panels of Tables B.8 and B.9, we notice that our results are a lot more robust for a $\delta = 0.0025$ compared to a $\delta = 0.05$, which in turn leads to more robust results compared to a $\delta = 0.50$. Our choice of a lower $\delta = 0.0025$ for presenting our main results is based on this.

From the above it is clear that LD calculated using lower values of δ explain child mortality better than higher values. What does this mean? As explained in section 2.3 the decay factor δ is a parameter that determines how fast the distance between any two languages declines as the number of shared branches increases. Under lower values of δ, as soon as two languages share a single branch their distance falls more rapidly than under higher values of δ. However, post that as the number of shared branches go up the drop in distance is not as drastic and the drop is comparable to higher values of δ even though the actual magnitudes of the distances are different.[27] This implies that the results are driven more by the divisions in the broad language

[25] In the leftmost panel of Table B.10 we provide the marginal effects for each of the 9 circles of alternative radii.

[26] Results not provided and are available upon request from the author.

[27] In Figure C.9 we provide simulations of how distances between any two languages change as they share different number of branches ranging from 0 to 15, for different values of δ.

families. In other words, the splits that occurred thousands of years ago are more important compared to more recent splits.

Rather than mere differences across dialects of the same languages or differences in closely related languages, our results show that deeper cleavages matter more. We interpret this as a sign of linguistic distance acting as a barrier to information and networks. This is in line with Spolaore and Wacziarg (2009, 2016) who point out how genetic/ ancestral distances act as barriers to development. This is also in line with Desmet et al. (2012) who show that for civil conflict higher values of aggregation matter more than lower values of aggregation of ELF. Also, Fenske et al. (2017) find that the highest level of distinction drives the results in their data for explaining market integration in colonial India.

For the analyses that follow we will fix the decay factor δ to 0.0025. Also, for the sake of brevity, the LD and ELF variables in the following sections are always calculated using a radius of 50 km around the mother, unless otherwise specified.[28] And finally, we will present results based on the most comprehensive specification of column 5 from Table 1, unless otherwise specified.

3.5 Other Health Outcomes

In this section we move beyond child mortality and try to identify the effects of LD on other child health variables, as well as outcomes like immunizations which are crucial for child health.

3.5.1 Infant and Neonatal Mortality

So far our focus has been on child mortality which is the event that the child dies before reaching the age of five. Other related variables could be infant mortality and neonatal mortality. The former is defined as the child dying before reaching the age of one, while the latter is the event of the child dying before the first month of their birth. In columns 1 and 2 of Table 3 we provide results for infant and neonatal mortality respectively. In general our results are quite similar. In particular we find that LD significantly increases both infant and neonatal mortality. ELF continues to have a negative effect on the mortality outcomes, and is significant at the 10% level for the neonatal mortality variable.[29]

[28] The results are qualititatively similar for circles of alternative radii. These results are available upon request from the author.

[29] In results not provided we find that the ELF variable is not robust to changing the circle radius. These results are available upon request from the author.

3.5.2 Height and Weight

In columns 3 to 5 of Table 3 we respectively provide regressions studying the impact of LD on the child's height-for-age Z-score (HAZ), probability of child being stunted, and finally the child's weight-for-age Z-score (WAZ). We see that linguistic distance has a strong and significant effect on child height measured either by HAZ or the stunting status of the child. LD also reduces child weight as measured by WAZ, but this effect is not statistically significant. ELF continues to have a benign effect on the different variables and significantly improves WAZ.

3.5.3 Immunization and Some Other Variables

In the different columns of Appendix Table B.11 we respectively present results for whether the mother received tetanus injections during pregnancy, if the child received measles immunization, polio vaccination, DPT immunization and finally if the mother received iron tablets during her pregnancy. We notice that among these variables, LD significantly (at the 10% level) reduces the probability of the mother taking iron tables during pregnancy. LD does have a significant effect on any of the other variables.

3.6 Are the results driven by Ethnic Diversity?

Ethnic diversity usually measured by ethno-linguistic fractionalization or ELF has often been found to have a negative effect on different socio-economic outcomes.[30] In contrast to the LD variable, our circle level ELF variable seems to have a more benign effect on the different health outcomes. However, the effect is almost never statistically significant. The recent literature has underscored the importance of the level at which the linguistic groups enter the ELF calculations as crucial (Desmet et al., 2012). In order to incorporate this insight, we follow the recent literature (Desmet et al., 2012, 2016; Gershman and Rivera, 2016), and calculate ELF at different levels of aggregation based on Ethnologue language trees, with 15 possible levels.[31]

Without taking any a priori decision on which is the right level of aggregation, we consider a range of levels of aggregation including levels 15, 10, 5 and 2. This ensures that we have a high level of aggregation given by Level 2, a medium level of aggregation given by Level 5, and a lower level of aggregation given by Level 10. We also include results for the most disaggregated

[30] See Alesina et al. (2003), Alesina et al. (1999), Easterly and Levine (1997) for example. See Alesina and La Ferrara (2000) for a review of the literature.

[31] See Section 2.3 for more details.

level of ELF given by Level 15, which is also the basic ELF used in the previous tables.

In columns 1 to 4 of Panel 1 of Table 4, we show that our results do not change if we control for ELF at different levels of aggregation. In column 5, we show that ELF in general is not significant even if we do not control for LD. However, LD continues to be significant regardless of the level of aggregation at which ELF is calculated and its absolute magnitude hardly changes. In column 6, we include a quadratic term for ELF following Ashraf and Galor (2013), who argue that diversity has a hump shaped effect on economic development. We notice that LD continues to have a significant effect on child mortality, whereas ELF does not have any significant effects. Finally, in column 7, we show that our results are qualitatively unchanged even if we do not control for ELF.

While ELF has been traditionally used to measure ethnic diversity, some papers have highlighted the relevance of ethnic polarization (ELP) rather than fractionalization, particularly in the context of intergroup conflict (Montalvo and Reynal-Querol, 2005). Like in the case of ELF, we calculate ELP at different levels of aggregation (Desmet et al., 2012). First, from Appendix Table B.7 we notice that in our context the polarization and fractionalization measures are highly correlated. This is particularly true at higher levels of aggregation and in circles of lower radii. In a circle of radius 25 *km* the correlation between ELF and ELP at aggregation level 15 is 0.78, and this goes up substantially to 0.98 at the aggregation level 2.

In Panel 2 of Table 4 we rerun our estimations as described for Panel 1, but using ELP instead of ELF. In columns 1 to 4, we control ELP at different levels of aggregation. In column 5, we include ELP by leaving LD out. In column 6, we include a quadratic term for ELP and finally, in column 7, we include a specification controlling for both ELF and ELP together following Montalvo and Reynal-Querol (2005). We see that LD continues to have a significant and robust effect on child mortality.

3.7 Are the results explained by Ethnic Favouritism?

In recent work Kramon and Posner (2016) show that having a co-ethnic as president during one's school-age years leads to better schooling outcomes for children. Again, Franck and Rainer (2012) provide evidence of similar ethnic favouritism for educational and child mortality outcomes of ethnic groups in 18 Sub-Sharan African countries. Hence, one could argue that our results might be explained by women who have co-ethnics as the countries' leaders.[32] The

[32]In a study of Guinea, Kudamatsu (2009) did not find any such ethnic favouritism effects on infant mortality.

inclusion of time varying ethnicity fixed effects, rules out our results being driven by such ethnic favouritism.

The recent literature has also talked about region specific transfers from the centre which benefit certain ethnic groups at the cost of others (De Luca et al., 2016; Dickens, 2016). Again, Burgess et al. (2015) show that during less democratic periods in Kenya, there is ethnic favouritism in road building in regions that share the ethnicity of the president. We have included region specific year effects in all our specifications. This allows for the non-parametric evolution of year effects differently for each region in addition to time varying ethnic groups fixed effects. This gives us confidence that our results are not driven by region or ethnic group specific transfers from the centre.

3.8 Heterogeneous Effects

Up to this point we have assumed that the linguistic distance variable has a homogeneous effect on all children. There are several reasons why this might not be the case. For example, wealthier or more educated mothers might be better able to insulate their children from the negative effects of linguistic distance. Again, linguistic distance might have different implications for male and female children. In the different columns of Appendix Table B.12 we try to identify the heterogeneity in the effects of linguistic distance by the following variables respectively: child gender, place of residence (urban or rural), mother's years of education, ELF, ELP, population, distance from the capital and wealth. We do not find any evidence in favour of heterogeneity in the effects of LD by any of these variables.

3.9 Migration

A possible concern in estimating the effects of ethnic distance on child mortality is spatial sorting. If individuals realize that being linguistically distant is bad for them then they might try to sort themselves into neighbourhoods where they are less distant to others. Given various barriers to movement (eg. transportation costs), perfect sorting is not observed in reality. In fact, in spite of population movements, ethnic populations tend to reside in their respective historical homelands (Michalopoulos and Papaioannou, 2014). Even in the face of large scale population displacements caused by civil wars, individuals tend to try and return to their historical ethnic homeland (Glennerster et al., 2013).[33]

[33] Almost 55% of the Afrobarometer Survey respondents currently live in their ethnic group's ancestral homeland (Nunn and Wantchekon, 2009). Again, Gershman and Rivera (2016) show how sub-national ethnic diversity is

However, if individuals actually are able to move to places where they are less distant to others then if anything we are underestimating the effects of ethnic distance on child mortality and the effects of linguistic distance would be even stronger. This is exactly what we find in our data.

In Table 5 we investigate heterogeneity of our results by migrant status. In column 1 of Table 5, we first use a 0-1 binary variable indicating migrant status as the dependent variable. We notice that being a migrant reduces the effect of LD on child mortality. In other words, the effects of being linguistically distant are worse for children of mothers who have never moved from their village of residence. In column 2, we directly check for heterogeneity by the continuous variable which tells us how many years the mother has lived in the village of residence. We again notice that effects of LD on child mortality is much stronger for individuals who have lived a longer number of years in their village of residence.

Finally, in columns 3 and 4, we respectively restrict the sample to individuals who have moved and individuals who have never moved from their village of residence. From these two columns we notice that our results are driven by non-migrants rather than migrants. The coefficient on LD is much bigger and highly statistically significant for the non-migrant sample. For a circle of 50 km radius around the mother, a one SD increase in LD leads to around 14 additional child deaths per 1000 live births, which is around 3.3% SD deaths in the non-migrant sample. This stands in sharp contrast to around 4 additional child deaths per 1000 live births (around 1% SD deaths) in the migrant sample, for a similar one SD increase in LD. The corresponding figures in the full sample are 8.2 deaths per 1000 live births which is 2% of the SD deaths.[34]

Next, in the two panels of Tables B.14 and B.15 we respectively provide the differences in the effects of LD for other variables by splitting the sample by migrants and non-migrants. In particular, in the different columns of Table B.14, we study the impact of LD on infant mortality, neonatal mortality, the height-for-age Z-score (HAZ), whether the child is stunted, and the weight-for-age Z-score (WAZ). In Table B.15, we look at the effects of LD on whether the mother received tetanus injections during pregnancy, if the child received measles immunization, polio vaccination, DPT immunization and finally if the mother received iron tablets during her pregnancy. We notice that LD consistently worsens the several different health outcomes in

stable across several decades in sub-Saharan Africa. More importantly, they find that changes in diversity at the sub-national level are not correlated with changes in economic conditions (Gershman and Rivera, 2016).

[34]For circles of alternative radii ranging from 25 km to 250 km, a one SD increase in LD leads to around 11.5 to 19.2 (3.3 to 4.2) additional child deaths per 1000 live births in the non-migrant sample (migrant sample). Please refer to Table B.10 for more details.

the non-migrant sample rather than the migrant sample. The results indicate that linguistic distance has a more harmful effect on the health outcomes of mothers who have never moved from their village of residence, and our results are driven by this non-migrant sample.

It is possible that migrants choose to relocate to places where they are less linguistically distant to others. Hence, the migrant sample might in general have on average a lower linguistic distance to their neighbours than the non-migrant sample. In Appendix Table B.13, we investigate the correlates of migrant status and find the opposite. We see that individuals who are migrants have a higher linguistic distance to their neighbours than individuals who are not.[35] Clearly, if anything migration is biasing our results away from zero and without migration our results would have been much stronger.

3.10 Selection on Unobservables

Our identification strategy relies on controlling for a rich set of observable control variables and fixed effects. In order to understand how selection on unobservable variables might be driving our results we turn to the methodology developed by Altonji et al. (2005) and Bellows and Miguel (2009) who present new estimation strategies that can be used when strong prior information regarding the exogeneity of the variable of interest is unavailable. Following their heuristics, we check for coefficient stability while moving from a specification with a parsimonious set of controls to the full set of controls.

If anything, we find that our coefficients become substantially larger controlling for more observables, which implies that selection on unobservables actually pushes our estimates away from zero. Following Oster (2013), we have also verified that the R^2 becomes substantially larger moving from the restricted to unrestricted regressions. See for instance, the movement in the coefficient for LD and R^2 while moving from columns 1 to 5 in Table 1.[36] Hence, if we could have actually controlled for the unobserved variables that might be biasing our results, our estimated beta coefficients would become much larger and our results would be further strengthened.

[35] We also find that individuals who are migrants tend to be more concentrated in urban areas and are wealthier.
[36] The full set of results from this section are not provided and are available upon request from the author.

4 Some Evidence on Channels: Linguistic Distance as a Barrier to Information

One possible explanation of why linguistic distance worsens health outcomes could be that it acts as a barrier to health-related information. For instance, linguistically distant mothers might not receive the information on best practices about how to rear their children, due to perhaps lack of communication with groups who are very different to them. The other possibility is that linguistically distant mothers have worse access to public goods in general, arising from for instance discrimination, which harms their children's health. In this section we provide some evidence in favour of the former.

In order to understand whether linguistic distance acts as a barrier to information we exploit the DHS question about whether the respondent has heard of the oral rehydration product (ORS) for treating children with diarrhoea. Diarrhoea is big child killer. "Of the 10.6 million yearly deaths in children younger than age 5 years: 1.9 million (18%) are caused by diarrhoea. Of the 6.6 million deaths among children aged 28 days to five years: 1.7 million (26%) are caused by diarrhoea" (`http://rehydrate.org/facts/child-deaths.htm`). Oral Rehydration therapy is the cornerstone treatment for treating diarrhoea (Victora et al., 2000).

For measuring access to public goods in general, we exploit information on access to four different public goods: access to electricity, access to water, the individual's educational attainment and whether the individual is literate or not. For instance, lower literacy or educational attainment among linguistically distant mothers might indicate lower access to schools.

In the previous section we have established that our results were stronger for non-migrants. Also, given the possibility that migrants might have acquired the knowledge on ORS elsewhere rather than where they currently reside, we split the sample by migrant status. In the different columns of Table 6, we check whether linguistic distance impedes educational attainment, literacy, access to water, electricity, and finally knowledge about ORS. In the two panels we split the sample by migrants and non-migrants. The LD variable does not have a significant effect on any of these variables except for the ORS variable in the non-migrant sample. In other words, while LD does not impede general access to public goods, it poses a barrier to information about ORS, in particular for the individuals who have never moved from their place of residence.

Thus, the results from this section do indicate that individuals who are linguistically distant to others living around them have lower access to information which leads to higher rates of

mortality for their children. On the other hand, linguistically distant individuals do not necessarily face lower access to public goods in general. Moreover, we see that our results are driven by individuals who have never moved from their place of residence rather than individuals who have migrated from their place of birth.

The fact that our results are driven by non-migrants rather than migrants also supports our interpretation of access to information being the channel. If ethnic distance is a barrier to knowledge and information about how to take care of one's children, it is important to understand where individuals might acquire such information. If individuals moved from some other place to their current place of residence, then it is likely that they already acquired such information elsewhere. Hence, linguistic distance in the place of their current residence does little to affect their children's health outcomes unless LD affects discrimination in general rather than imposing information barriers.

5 Conclusion

Child mortality rates are still unnecessarily high, particularly in sub-Saharan Africa. Nineteen thousand children die worldwide every day before reaching the age of five. The highest rates of child mortality are still concentrated in sub-Saharan Africa, where 1 in 9 children die before reaching the age of five, which is not only more than 16 times the average for developed regions (1 in 152), but also a lot higher than in South Asia (1 in 16), which has the second highest rates of child mortality (UNICEF, 2012). Not surprisingly, reducing child mortality was part of the Millennium Developmental Goals and is currently part of the Sustainable Development Goals.

In this paper we put together a high quality individual level micro database from the Demographic and Health Surveys and combine it with a novel dataset on the spatial distribution of ethnic groups at the level of approximately 1 x 1 sq. *km* for fourteen sub-Saharan African countries. We map individual level ethnicities to languages and calculate how ethnically distant an individual is to her neighbours. Then we go on to show that children of mothers who are ethnically distant from their neighbours face a higher probability of dying before reaching the age of five, and those who survive are shorter in size. We also show that ethnically distant mothers are less likely to know about oral rehydration therapy which is beneficial to their children.[37]

[37] Singleton and Krause (2009) point out how Spanish speaking patients face barriers to accessing health care even in the US. In Mexico indigenous people don't go to the hospital in fear that their language and customs will not be understood and due to lack of trust between groups http://www.nytimes.com/video/2013/08/13/world/americas/100000002373842/a-chiapas-medicine-man.html.

One clear policy implication from our paper is that in order to reduce child mortality rates in Africa, we need to target ethnic minorities who might be losing out solely because they speak a distant language compared to their neighbours. This could help moving closer to the sustainable development goal of reducing child mortality.

References

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

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

Alesina, A. and E. La Ferrara (2000). Participation in heterogeneous communities. *The Quarterly Journal of Economics, MIT Press 115(3) August*, 847–904.

Alesina, A. F., S. Michalopoulos, and E. Papaioannou (2012). Ethnic inequality. *NBER Working Paper No. 18512, November*.

Alesina, A. F. and E. Zhuravskaya (2011). Segregation and the quality of government in a cross section of countries. *American Economic Review, American Economic Association vol. 101(5), August*.

Algan, Y., C. Hémet, and D. D. Laitin (2016). The social effects of ethnic diversity at the local level: A natural experiment with exogenous residential allocation. *Journal of Political Economy 124*(3), 696–733.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy Vol. 113, No. 1, February*, 151–184.

Ashraf, Q. and O. Galor (2013). The 'out of africa' hypothesis, human genetic diversity, and comparative economic development. *American Economic Review Vol. 103 No. 1, February*, 1–46.

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

Baldwin, K. and J. D. Huber (2010). Ecnomic versus cultural differences: Forms of ethnic diversity and public good provision. *American Political Science Review 104, No. 4, November.*

Bellows, J. and E. Miguel (2009). War and local collective action in sierra leone. *Journal of Public Economics 93*, 1144—-1157.

Bishop, Y., S. Fienberg, and P. Holland (1975). Discrete multivariate analysis: Theory and practicemit press. *Cambridge, Massachusetts*.

Brockerhoff, M. and P. Hewett (2000). Inequality of child mortality among ethnic groups in sub-saharan africa. *Bulletin of the World Health Organization 78*(1), 30–41.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, et al. (2015). The value of democracy: evidence from road building in kenya. *The American Economic Review 105*(6), 1817–1851.

Dahlberg, M., K. Edmark, and H. Lundqvist (2012). Ethnic diversity and preferences for redistribution. *Journal of Political Economy 120*(1), 41–76.

De Luca, G., R. Hodler, P. A. Raschky, and M. Valsecchi (2016). Ethnic favoritism: An axiom of politics?

Deming, W. E. and F. F. Stephan (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics 11*(4), 427–444.

Desmet, K., J. Gomes, and I. Ortuño (2016). The geography of linguistic diversity and the provision of public goods. Technical report, CEPR Discussion Paper.

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

Desmet, K., I. Ortuno-Ortin, and I. Weber (2009). Linguistic diversity and redistribution. *Journal of European Economic Association 7(6)*, 1291–1318.

Desmet, K., I. Ortuño-Ortín, and R. Wacziarg (2017). Culture, ethnicity, and diversity. *American Economic Review 107*(9).

Dickens, A. (2016). Ethnolinguistic favoritism in african politics. Technical report, York University, mimeo,(February).

Dyen, I., J. B. Kruskal, and P. Black (1992). An indo-european classification, a lexicostatistical experiment. 1. *Transactions of the American Philosophical Society 82*, 1–132.

Easterly, W. and R. Levine (1997). Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics 112, no.4 November*, 1203–1250.

Esteban, J., L. Mayoral, and D. Ray (2012a). Ethnicity and conflict: An empirical study. *American Economic Review 102, No.4*, 1310–1342.

Esteban, J., L. Mayoral, and D. Ray (2012b). Ethnicity and conflict: Theory and facts. *Science 336, 858*.

Esteban, J. and D. Ray (1999). Conflict and distribution. *Journal of Economic Theory 87(2)*, 379–415.

Esteban, J.-M. and D. Ray (1994). On the measurement of polarization. *Econometrica: Journal of the Econometric Society*, 819–851.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth 8*(2), 195–222.

Fenske, J., N. Kala, et al. (2017). Linguistic distance and market integration in india. Technical report, Competitive Advantage in the Global Economy (CAGE).

Fienberg, S. E. (1970). An iterative procedure for estimation in contingency tables. *The Annals of Mathematical Statistics*, 907–917.

Fisman, R., D. Paravisini, and V. Vig (2012). Cultural proximity and loan outcomes. Technical report, National Bureau of Economic Research.

Franck, R. and I. Rainer (2012). Does the leader's ethnicity matter? ethnic favoritism, education and health in sub-saharan africa. *American Political Science Review 106*(2, May).

Gershman, B. and D. Rivera (2016). Subnational diversity in sub-saharan africa: Insights from a new dataset.

Glennerster, R., E. Miguel, and A. D. Rothenberg (2013). Collective action in diverse sierra leone communities. *The Economic Journal 123*(568), 285–316.

Gomes, J. F. (2013). Religious diversity, intolerance and civil conflicts. *UC3M Working Paper 13-11, Economic Series, May*.

Gomes, J. F. (2015). The political economy of the maoist conflict in india: an empirical analysis. *World Development 68*, 96–123.

Guiso, L., P. Sapienza, and L. Zingales (2009). Cultural biases in economic exchange? *The Quarterly Journal of Economics 124*(3), 1095–1131.

Gyimah, S. O. (2002). Ethnicity and infant mortality in sub-saharan africa: The case of ghana. *PSC Discussion Papers Series 16*(10), 1.

Habyarimana, J., M. Humphreys, D. N. Posner, and J. M. Weinstein (2007). Why does ethnic diversity undermine public goods provision? *American Political Science Review 101*(04), 709–725.

Isphording, I. E. (2013). Disadvantages of linguistic origin: Evidence from immigrant literacy scores. Technical report, IZA Discussion Paper.

Isphording, I. E. and S. Otten (2013). The costs of babylon—linguistic distance in applied economics. *Review of International Economics 21*(2), 354–369.

Kramon, E. and D. N. Posner (2016). Ethnic favoritism in education in kenya. *Quarterly Journal of Political Science 11*(1).

Kudamatsu, M. (2009). *Ethnic Favoritism: Micro Evidence from Guinea*. unpublished.

Kudamatsu, M., T. Persson, and D. Strmberg (2012). Weather and infant mortality in africa. *Prepared for the conference on Climate and the Economy in Stockholm, September 5-8*.

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

Laitin, D. D. and R. Ramachandran (2016). Language policy and human development. *American Political Science Review 110*(3), 457–480.

Lewis, M. P., G. F. Simons, C. D. Fennig, et al. (2014). *Ethnologue: Languages of the world*, Volume 17. SIL international Dallas, TX.

Malhotra, N. (2012). Inadequate feeding of infant and young children in india: lack of nutritional information or food affordability? *Public Health Nutrition 1*(1), 1–9.

Matuszeski, J. and F. Schneider (2006). Patterns of ethnic group segregation and civil conflict. *unpublished, Harvard University*.

Michalopoulos, S. and E. Papaioannou (2014). National institutions and subnational development in africa. *The Quarterly Journal of Economics 129*(1), 151–213.

Miguel, E. and M. K. Gugerty (2005). Ethnic diversity, social sanctions, and public goods in kenya. *Journal of Public Economics 89*(11), 2325–2368.

Mitra, A. and D. Ray (2010). Implications of an economic thory of conflict: Hindu-muslim violence in india. *unpublished*.

Montalvo, J. and M. Reynal-Querol (2005). Ethnic polarization, potential conflict and civil war. *The American Economic Review 95(3) June*, 796–816.

Montalvo, J. G. and M. Reynal-Querol (2016). Ethnic diversity and growth: Revisiting the evidence. *Manuscript, Universat Pompeu Frabra-ICREA*.

Munshi, K. and M. Rosenzweig (2015). Insiders and outsiders: local ethnic politics and public goods provision. Technical report, National Bureau of Economic Research.

Nunn, N. and L. Wantchekon (2009). The slave trade and the origins of mistrust in africa. *American Economic Review*.

Oster, E. (2013). Unobservable selection and coefficient stability: Theory and validation. Technical report, National Bureau of Economic Research.

Pongou, R. (2009). Anonymity and infidelity: Ethnic identity, strategic cross-ethnic sexual network formation, and hiv/aids in africa. *Unpublished paper, Department of Economics, Brown University*.

Shastry, G. K. (2012). Human capital response to globalization education and information technology in india. *Journal of Human Resources 47*(2), 287–330.

Singleton, K. and E. Krause (2009). Understanding cultural and linguistic barriers to health literacy. *OJIN: The online journal of issues in nursing 14*(3).

Spolaore, E. and R. Wacziarg (2009). The diffusion of development. *The Quarterly Journal of Economics 124*(2), 469–529.

Spolaore, E. and R. Wacziarg (2016). Ancestry and development: New evidence. *Discussion Papers Series, Department of Economics, Tufts University 820*.

Thomas, D., J. Strauss, and M.-H. Henriques (1991). How does mother's education affect child height? *Journal of human resources*, 183–211.

UNICEF (2012). Levels trends in child mortality: Report 2012. *UN Inter-agency Group for Child Mortality Estimation, United Nations Children's Fund, UNICEF*.

Victora, C. G., J. Bryce, O. Fontaine, and R. Monasch (2000). Reducing deaths from diarrhoea through oral rehydration therapy. *Bulletin of the World Health Organization 78*(10), 1246–1255.

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

Tables

Table 1: Mother's Linguistic Distance and Child mortality: 50 *km* Radius

	(1)	(2)	(3)	(4)	(5)
$\delta = 0.0025$					
Linguistic Distance 50 KM	0.0271***	0.0402**	0.0430***	0.0437***	0.0435***
	(0.00908)	(0.0158)	(0.0135)	(0.0135)	(0.0133)
ELF 50 KM	-0.00303	-0.00396	-0.00540	-0.00653	-0.00739
	(0.0101)	(0.00969)	(0.00893)	(0.00829)	(0.00823)
Observations	654506	654502	654237	654237	653666
R^2	0.090	0.092	0.146	0.146	0.154
Survey Wave FE	Y	Y	Y	Y	Y
Region x Year FE	Y	Y	Y	Y	Y
Ethnicity FE	N	Y	Y	Y	N
Religion FE	N	N	Y	Y	Y
Individual Controls	N	N	Y	Y	Y
Geographic isolation	N	N	N	Y	Y
Ethnicity x Year FE	N	N	N	N	Y

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The dependent variable is the individual child level mortality outcome. The numbers after linguistic distance and ELF indicate the radius of the circle around the mother in which these variables have been calculated. The individual controls include female child dummy, mother's age at birth, mother's age at birth squared, multiple birth indicator, birth order, birth order squared, short birth spacing prior to the birth, short birth spacing after the birth, the location of the mother in the form of an urban dummy, dummies for her educational attainment and her families' wealth index. Geographical isolation controls include the distance of the mother's location from the capital and the logged population in the circle.

Table 2: Mother's Linguistic Distance and Child mortality: Alternative radii

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	25 km	75 km	100 km	125 km	150 km	175 km	200 km	250 km
$\delta = 0.0025$								
Linguistic Distance	0.0347***	0.0474***	0.0487**	0.0528**	0.0537**	0.0540**	0.0543**	0.0522**
	(0.00964)	(0.0177)	(0.0192)	(0.0205)	(0.0226)	(0.0229)	(0.0227)	(0.0236)
ELF	-0.00372	-0.00893	-0.0112	-0.00540	0.00393	0.0104	0.0134	-0.0135
	(0.00620)	(0.00996)	(0.0122)	(0.0131)	(0.0138)	(0.0156)	(0.0184)	(0.0258)
Observations	653666	653666	653666	653666	653666	653666	653666	653666
R^2	0.154	0.154	0.154	0.154	0.154	0.154	0.154	0.154

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The dependent variable is the individual child level mortality outcome. The numbers in the column headings indicate the radius of the circle around the mother in which the linguistic distance and ELF have been calculated. All columns include controls for survey wave FE, region x year FE, ethnicity x year FE, religion FE, individual controls and geographic isolation controls described in the notes of Table 1.

Table 3: Mother's Linguistic Distance and Other Child Health Variables

	(1) infant	(2) neonatal	(3) HAZ	(4) stunted	(5) WAZ
Linguistic Distance	0.0204***	0.00725***	-0.0875**	0.0313**	-0.0497
	(0.00601)	(0.00235)	(0.0381)	(0.0132)	(0.0317)
ELF	-0.00291	-0.00357*	0.0233	-0.00605	0.101**
	(0.00445)	(0.00194)	(0.0504)	(0.0169)	(0.0409)
Observations	815267	861386	141475	141475	141475
R^2	0.097	0.069	0.205	0.153	0.161

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The column headings indicate the individual child level dependent variable for each specification. These are: infant mortality, neonatal mortality, height-for-age Z-score (HAZ), stunting, and the weight-for-age Z-score (WAZ). A circle of radius 50 km has been considered for calculating the linguistic distance and ELF variables. All columns include controls for survey wave FE, region x year FE, ethnicity x year FE, religion FE, individual controls and geographic isolation controls described in the notes of Table 1.

Table 4: Mother's Linguistic Distance and child mortality: Robustness for aggregate diversity

	(1) ELFL15	(2) ELFL10	(3) ELFL5	(4) ELFL2	(5) ONLYELF	(6) ELFSQ	(7) NOELF
Linguistic Distance	0.0435***	0.0426***	0.0430***	0.0435***		0.0428***	0.0414***
	(0.0133)	(0.0133)	(0.0125)	(0.0119)		(0.0131)	(0.0143)
ELF	-0.00739	-0.00403	-0.00584	-0.00733	-0.00272	0.00339	
	(0.00823)	(0.00757)	(0.0115)	(0.0135)	(0.00955)	(0.0213)	
ELF squared						-0.00956	
						(0.0254)	
Observations	653666	653666	653666	653666	653666	653666	653666
R^2	0.154	0.154	0.154	0.154	0.154	0.154	0.154
	ELPL15	ELPL10	ELPL5	ELPL2	ONLYELP	ELPSQ	BOTH
Linguistic Distance	0.0410***	0.0417***	0.0425***	0.0442***		0.0411***	0.0444***
	(0.0140)	(0.0140)	(0.0136)	(0.0131)		(0.0140)	(0.0131)
ELP	0.00177	-0.00170	-0.00369	-0.00697	0.00396	-0.00679	0.0146*
	(0.00699)	(0.00630)	(0.00688)	(0.00572)	(0.00745)	(0.0201)	(0.00854)
ELP squared						0.00929	
						(0.0192)	
ELF							-0.0201**
							(0.0101)
Observations	653666	653666	653666	653666	653666	653666	653666
R^2	0.154	0.154	0.154	0.154	0.154	0.154	0.154

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The dependent variable is the individual child level mortality outcome. In Panel 1 (Panel 2): column 1 controls for ELF (ELP) at aggregation Level 15; column 2 for ELF (ELP) at aggregation Level 10; column 3 for ELF (ELP) at aggregation Level 5; column 4 for ELF (ELP) at aggregation Level 2; column 5 for ELF (ELP) at aggregation Level 15, without LD; column 6 for ELF (ELP) at aggregation Level 15,and its square term. In column 7 of Panel 1, we do not control for ELF or ELP. In column 7 of Panel 2, we include both ELF and ELP. A circle of radius 50 km has been considered for calculating the linguistic distance and ELF variables. All columns include controls for survey wave FE, region x year FE, ethnicity x year FE, religion FE, individual controls and geographic isolation controls described in the notes of Table 1.

Table 5: Mother's Linguistic Distance and Child mortality: Heterogeneity by Migration Status

	(1) HetMigrant	(2) HetYearsLived	(3) Migrants	(4) NMigrants
Linguistic Distance	0.0577***	0.0341***	0.0198**	0.0758***
	(0.0148)	(0.0114)	(0.00935)	(0.0137)
Het. Variable	0.00610***	-0.000348***		
	(0.00166)	(0.0000718)		
Linguistic Distance × Het. Variable	-0.0185**	0.000555**		
	(0.00819)	(0.000270)		
Observations	521217	521217	278952	241309
R^2	0.163	0.163	0.177	0.167

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The dependent variable is the individual child level mortality outcome. In column 1 the Het. Variable refers to the 0-1 migrant status of the mother; in column 2 it refers to the continuous variable indicating how many years the mother has been living in the village where she was interviewed. In column 3 (column 4), we restrict the sample to only children of mothers who are migrants (non-migrants). A circle of radius 50 km has been considered for calculating the linguistic distance and ELF variables. All columns include controls for survey wave FE, region x year FE, ethnicity x year FE, religion FE, individual controls and geographic isolation controls described in the notes of Table 1.

Table 6: Public Goods, Information and Linguistic Distance: Migrants vs. Non-Migrants

	(1) education	(2) literacy	(4) water	(5) electricity	(6) ORS
Migrants					
Linguistic Distance	0.0277	0.0101	-0.0352	0.0312	0.00154
	(0.0795)	(0.0362)	(0.0271)	(0.0190)	(0.0178)
ELF	-0.0198	-0.0278**	0.0565*	-0.0171	-0.0190
	(0.0311)	(0.0138)	(0.0294)	(0.0127)	(0.0330)
Observations	90456	71929	74748	89536	88542
R^2	0.458	0.423	0.112	0.548	0.206
Non-Migrants					
Linguistic Distance	0.000477	0.0121	-0.0352	0.00731	-0.0841**
	(0.0667)	(0.0265)	(0.0271)	(0.0186)	(0.0374)
ELF	0.0704	0.0561*	0.0565*	-0.00774	-0.00560
	(0.0545)	(0.0294)	(0.0294)	(0.00942)	(0.0308)
Observations	73681	60445	74748	72936	72648
R^2	0.474	0.391	0.112	0.546	0.242

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The column headings indicate the individual mother level dependent variable for each specification. These are: educational attainment, literacy, access to water, access to electricity and knowledge about ORS. A circle of radius 50 km has been considered for calculating the linguistic distance and ELF variables. All columns include controls for survey wave FE, region FE, ethnicity FE, religion FE, year of birth FE, and geographic isolation controls described in the notes of Table 1.

A Data Appendix

A.1 DHS Countries and Surveys Used

In this study, we use 30 DHS surveys from 14 sub-Saharan African countries. These are listed in Table A.1. These countries and surveys were chosen based on the availability of GPS coordinates and ethnicities of the mothers, and other covariates. In particular, countries and surveys for which a one to one matching for a large number of the ethnicities was not possible had to be left out of the sample.

For example, Cameroon has multiple DHS surveys viz. 1991, 1998, 2004 and 2011. For the 1991 survey the only two ethnicities provided are Cameroonian and others. Hence, for our purposes, this survey is unusable. The 1998 survey did not collect GPS data. The 2011 survey has a more disaggregated division of ethnic groups, but they are still very broad to allow a one to one matching with languages. For instance, some of the ethnicities comprise of three of four groups clubbed together e.g. "arab-choa/peulh/haoussa/kanuri", "ctier/ngoe/oroko" or "beti/bassa/mbam." This makes a one to one ethnicity to language mapping impossible. The 2004 survey had more disaggregated data on ethnic groups, but the a huge proportion of the respondents were still left unmapped. Hence, Cameroon had to be discarded altogether. Several other countries and surveys also had to be discarded for similar reasons. A full list is available from the authors on request.

Table A.1: Study Sample

Country	DHS surveys used
Benin	1996, 2001
Burkina Faso	1993, 1998-99, 2003, 2010
Ethiopia	2000, 2005, 2011
Ghana	1993, 2998, 2003, 2008
Guinea	1999, 2005
Kenya	2003, 2008
Malawi	2000, 2004, 2010
Mali	1995-96, 2001, 2006
Namibia	2000
Niger	1998
Senegal	2005, 2010-11
Sierra Leone	2008
Uganda	2011
Zambia	2007

A.2 Ethnicity and Languages Matching

In this section we list the different steps taken to match ethnic groups from the DHS to languages from Ethnologue. A full list of ethnicity and languages matching is available upon request.

- If the name of an ethnicity from the DHS is identical to a language name from Ethnologue, then we already have the language we need and no further mapping is required. E.g. Kalenjin is both an ethnicity and a language spoken in Kenya.
- If the name of an ethnicity from the DHS is an alternative name for a language group from Ethnologue, then we just rename the former to the latter. Then we already have the language we need and no further mapping is required. E.g. Kissi is the name of an ethnic group in Kenya, which is actually also an alternative name for the Ekgussi language (http://www.ethnologue.com/language/guz). Similarly Peulh is an alternative name for the Borgu Fulfulde language spoken in Benin (https://www.ethnologue.com/language/fue).
- Again there are ethnic groups from the DHS which are also names of languages from Ethnologue, but the spellings differ across the two sources. In this case the language assignation is trivial. E.g. Afar and Amharic language groups from Ethniopia are spelt as Affar and Amhara respectively in the DHS.
- In some instances, the DHS provides macro language groups in the ethnicity field. In these cases, we assign one of the actual languages that form part of the macro language group to the entire group. Since distances are based on the number of shared branches, assigning a different language from the same group does not change the actual distance. E.g. For the Luhya group in Kenya we assign the Lubukusu language.
- For some groups we follow Jim Fearon's classifications, originally from Fearon (2003). E.g. the San group in Namibia is assigned the Hai||om language. Again the Diola group in Senegal is assigned the Jola-Fonyi language.
- In a very few cases there was some ambiguity, since the same ethnicity name from the DHS could have been referring to multiple closely related languages from Ethnologue. For instance, Limba in Sierra Leone could refer to East Limba or West-Central Limba. We randomly assign it East Limba. But since both East Limba and West-Cental Limba are closely related and share the exact same number of branches with any other language, this should not make a difference in the actual linguistic distance calculations.

B Appendix Tables

B.1 Descriptive Statistics

Table B.1: Child Level Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Child Death	0.228	0.42	0	1	654672
Infant Death	0.12	0.325	0	1	816268
Neonatal Death	0.055	0.229	0	1	862358
Height-for-Age Score	-1.571	1.746	-6	5.99	141673
Weight-for-Age Score	-1.182	1.304	-5.98	4.97	141673
Stunting	0.409	0.492	0	1	141673
Tetanus Vaccine	0.703	0.457	0	1	154814
Measles	0.844	0.810	0	3	196789
Polio Vaccine	0.278	0.448	0	1	182048
DPT Vaccine	0.466	0.499	0	1	161085
Iron Tablets	0.697	0.459	0	1	115590
Migrant	0.539	0.498	0	1	686231
Years lived in cluster	23.078	14.461	0	50	686231
Urban Residence	0.224	0.417	0	1	862358
Female Child	0.49	0.5	0	1	862358
Age At Birth	25	6.425	8	50	862358
Age At Birth Squared	666.292	348.533	64	2500	862358
Multiple Birth	0.032	0.177	0	1	862358
Birth Order Number	3.448	2.317	1	18	862358
Birth Order Number Squared	17.256	22.557	1	324	862358
Short Birth Spacing Prior	0.209	0.407	0	1	862358
Short Birth Spacing Post	0.209	0.407	0	1	862358
Highest educational level	0.424	0.664	0	3	862352
Educational Attainment	0.582	1.02	0	5	862352
Years of Education	1.997	3.405	0	26	862028
Log(Distance to the Capital)	5.13	1.217	-2.614	7.221	862358
Wealth Index	2.867	1.399	1	5	862358
Child's Birth Year	1992.274	9.366	1955	2011	862358
Mother's Birth Year	1968.416	9.572	1943	1996	862358
Log(Population) in 25 km	12.09	1.307	3.849	15.238	862358
Log(Population) in 50 km	13.272	1.164	6.137	15.665	862358
Log(Population) in 75 km	13.958	1.093	6.971	16.011	862358
Log(Population) in 100 km	14.428	1.045	7.467	16.346	862358
Log(Population) in 125 km	14.771	1.007	8.176	16.57	862358
Log(Population) in 150 km	15.036	0.974	8.582	16.87	862358
Log(Population) in 175 km	15.258	0.943	8.893	17.046	862358
Log(Population) in 200 km	15.437	0.915	9.332	17.188	862358
Log(Population) in 250 km	15.728	0.871	10.161	17.433	862358

Table B.2: Mother Level Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
ORS Knowledge	0.761	0.426	0	1	205794
Educational Attainment	0.782	1.189	0	5	208896
Water Access	0.656	0.475	0	1	182482
Electricity Access	0.186	0.389	0	1	204920

Table B.3: Summary Statistics for LD variables

Variable	Mean	Std. Dev.	Min.	Max.
$\delta = 0.0025$				
Linguistic distance in 25 km	0.073	0.191	0	1
Linguistic distance in 50 km	0.077	0.189	0	1
Linguistic distance in 75 km	0.081	0.19	0	1
Linguistic distance in 100 km	0.083	0.191	0	1
Linguistic distance in 125 km	0.087	0.194	0	1
Linguistic distance in 150 km	0.089	0.196	0	1
Linguistic distance in 175 km	0.092	0.197	0	1
Linguistic distance in 200 km	0.095	0.199	0	1
Linguistic distance in 250 km	0.1	0.202	0	1
$\delta = 0.05$ *à la* Desmet et al. (2012)				
Linguistic distance in 25 km	0.094	0.178	0	1
Linguistic distance in 50 km	0.1	0.176	0	1
Linguistic distance in 75 km	0.105	0.177	0	1
Linguistic distance in 100 km	0.109	0.179	0	1
Linguistic distance in 125 km	0.113	0.182	0	1
Linguistic distance in 150 km	0.117	0.184	0	1
Linguistic distance in 175 km	0.12	0.186	0	1
Linguistic distance in 200 km	0.124	0.188	0	1
Linguistic distance in 250 km	0.131	0.192	0	1
$\delta = 0.50$ *à la* Fearon (2003)				
Linguistic distance in 25 km	0.277	0.247	0	1
Linguistic distance in 50 km	0.294	0.235	0	1
Linguistic distance in 75 km	0.31	0.228	0	1
Linguistic distance in 100 km	0.323	0.224	0	1
Linguistic distance in 125 km	0.337	0.221	0	1
Linguistic distance in 150 km	0.349	0.219	0	1
Linguistic distance in 175 km	0.359	0.216	0	1
Linguistic distance in 200 km	0.369	0.213	0	1
Linguistic distance in 250 km	0.388	0.208	0.002	1
N		862358		

Table B.4: Summary statistics for ELF variables

Variable	Mean	Std. Dev.	Min.	Max.
ELF at Level 2 in 25 km	0.179	0.187	0	0.786
ELF at Level 2 in 50 km	0.207	0.193	0	0.792
ELF at Level 2 in 75 km	0.224	0.198	0	0.793
ELF at Level 2 in 100 km	0.238	0.199	0	0.802
ELF at Level 2 in 125 km	0.251	0.2	0	0.776
ELF at Level 2 in 150 km	0.262	0.2	0	0.779
ELF at Level 2 in 175 km	0.272	0.201	0	0.772
ELF at Level 2 in 200 km	0.281	0.201	0	0.773
ELF at Level 2 in 250 km	0.295	0.202	0	0.758
ELF at Level 5 in 25 km	0.245	0.236	0	0.871
ELF at Level 5 in 50 km	0.285	0.242	0	0.868
ELF at Level 5 in 75 km	0.313	0.247	0	0.87
ELF at Level 5 in 100 km	0.333	0.25	0	0.873
ELF at Level 5 in 125 km	0.353	0.251	0	0.866
ELF at Level 5 in 150 km	0.37	0.251	0	0.864
ELF at Level 5 in 175 km	0.384	0.25	0	0.854
ELF at Level 5 in 200 km	0.398	0.249	0	0.84
ELF at Level 5 in 250 km	0.42	0.248	0.001	0.838
ELF at Level 10 in 25 km	0.384	0.26	0	0.917
ELF at Level 10 in 50 km	0.457	0.25	0	0.906
ELF at Level 10 in 75 km	0.507	0.238	0	0.9
ELF at Level 10 in 100 km	0.542	0.225	0	0.898
ELF at Level 10 in 125 km	0.572	0.213	0	0.907
ELF at Level 10 in 150 km	0.596	0.199	0	0.913
ELF at Level 10 in 175 km	0.618	0.185	0	0.918
ELF at Level 10 in 200 km	0.637	0.17	0	0.921
ELF at Level 10 in 250 km	0.667	0.144	0.008	0.921
ELF at Level 15 in 25 km	0.408	0.27	0	0.918
ELF at Level 15 in 50 km	0.487	0.262	0	0.921
ELF at Level 15 in 75 km	0.54	0.25	0	0.937
ELF at Level 15 in 100 km	0.576	0.236	0	0.932
ELF at Level 15 in 125 km	0.607	0.222	0	0.938
ELF at Level 15 in 150 km	0.631	0.207	0	0.941
ELF at Level 15 in 175 km	0.653	0.192	0	0.944
ELF at Level 15 in 200 km	0.672	0.178	0	0.944
ELF at Level 15 in 250 km	0.702	0.152	0.008	0.938
N		862358		

Table B.5: Summary statistics for ELP variables

Variable	Mean	Std. Dev.	Min.	Max.
ELP at Level 2 in 25 km	0.329	0.33	0	1
ELP at Level 2 in 50 km	0.377	0.339	0	1
ELP at Level 2 in 75 km	0.407	0.341	0	1
ELP at Level 2 in 100 km	0.428	0.339	0	1
ELP at Level 2 in 125 km	0.45	0.337	0	1
ELP at Level 2 in 150 km	0.469	0.335	0	1
ELP at Level 2 in 175 km	0.485	0.333	0	1
ELP at Level 2 in 200 km	0.5	0.331	0	1
ELP at Level 2 in 250 km	0.521	0.328	0	1
ELP at Level 5 in 25 km	0.372	0.327	0	1
ELP at Level 5 in 50 km	0.426	0.326	0	1
ELP at Level 5 in 75 km	0.456	0.321	0	1
ELP at Level 5 in 100 km	0.476	0.312	0	1
ELP at Level 5 in 125 km	0.495	0.305	0	0.999
ELP at Level 5 in 150 km	0.512	0.299	0	0.998
ELP at Level 5 in 175 km	0.526	0.295	0	0.994
ELP at Level 5 in 200 km	0.538	0.293	0	0.995
ELP at Level 5 in 250 km	0.555	0.289	0.002	0.994
ELP at Level 10 in 25 km	0.504	0.292	0	1
ELP at Level 10 in 50 km	0.567	0.253	0	1
ELP at Level 10 in 75 km	0.599	0.219	0	1
ELP at Level 10 in 100 km	0.616	0.189	0	1
ELP at Level 10 in 125 km	0.626	0.161	0	0.99
ELP at Level 10 in 150 km	0.633	0.142	0	0.986
ELP at Level 10 in 175 km	0.64	0.133	0	0.978
ELP at Level 10 in 200 km	0.646	0.129	0	0.975
ELP at Level 10 in 250 km	0.651	0.128	0.016	0.974
ELP at Level 15 in 25 km	0.504	0.286	0	1
ELP at Level 15 in 50 km	0.555	0.246	0	1
ELP at Level 15 in 75 km	0.577	0.217	0	1
ELP at Level 15 in 100 km	0.587	0.193	0	0.994
ELP at Level 15 in 125 km	0.592	0.173	0	0.982
ELP at Level 15 in 150 km	0.595	0.163	0	0.979
ELP at Level 15 in 175 km	0.597	0.161	0	0.978
ELP at Level 15 in 200 km	0.599	0.162	0	0.975
ELP at Level 15 in 250 km	0.596	0.164	0.016	0.974
N		862358		

Table B.6: Correlations of Linguistic Distance and diversity (206,076 observations (mothers))

	Correlation of LD with ELF					Correlation of LD with ELP				
	25 *km*	50 *km*	75 *km*	100 *km*	125 *km*	25 *km*	50 *km*	75 *km*	100 *km*	125 *km*
Aggregation										
Level 2	0.34	0.37	0.38	0.39	0.40	0.31	0.32	0.33	0.33	0.34
Level 5	0.26	0.28	0.30	0.31	0.32	0.27	0.28	0.28	0.28	0.28
Level 10	0.15	0.15	0.16	0.17	0.18	0.14	0.11	0.07	0.03	-0.01
Level 15	0.13	0.13	0.13	0.14	0.16	0.13	0.09	0.05	0.01	-0.03

	25 *km*	50 *km*	75 *km*	100 *km*	125 *km*	25 *km*	50 *km*	75 *km*	100 *km*	125 *km*
Level 2	0.39	0.42	0.44	0.44	0.44	0.35	0.37	0.37	0.37	0.36
Level 5	0.33	0.36	0.38	0.38	0.38	0.32	0.32	0.32	0.31	0.30
Level 10	0.23	0.23	0.24	0.24	0.24	0.19	0.15	0.09	0.03	-0.04
Level 15	0.22	0.22	0.22	0.22	0.22	0.17	0.12	0.05	-0.01	-0.08

$\delta = 0.5$	25 *km*	50 *km*	75 *km*	100 *km*	125 *km*	25 *km*	50 *km*	75 *km*	100 *km*	125 *km*
Level 2	0.58	0.61	0.63	0.64	0.65	0.57	0.59	0.61	0.63	0.63
Level 5	0.58	0.60	0.63	0.64	0.65	0.57	0.58	0.59	0.60	0.60
Level 10	0.47	0.47	0.47	0.47	0.47	0.42	0.38	0.32	0.25	0.17
Level 15	0.47	0.47	0.47	0.47	0.48	0.40	0.33	0.24	0.13	0.02

Table B.7: Correlations of ELF and ELP (28,839 DHS clusters)

	25 km	50 km	75 km	100 km	125 km
Level 1	0.99	1.00	0.99	0.99	0.99
Level 2	0.98	0.98	0.97	0.97	0.96
Level 3	0.98	0.97	0.97	0.96	0.96
Level 4	0.95	0.95	0.94	0.94	0.93
Level 5	0.94	0.94	0.93	0.93	0.92
Level 6	0.94	0.93	0.92	0.91	0.91
Level 7	0.92	0.90	0.89	0.88	0.87
Level 8	0.91	0.89	0.88	0.86	0.85
Level 9	0.87	0.80	0.73	0.65	0.56
Level 10	0.84	0.76	0.65	0.52	0.37
Level 11	0.83	0.74	0.62	0.49	0.32
Level 12	0.83	0.75	0.62	0.49	0.32
Level 13	0.83	0.75	0.62	0.49	0.32
Level 14	0.83	0.75	0.62	0.49	0.32
Level 15	0.78	0.63	0.44	0.25	0.03

B.2 Other Tables

Table B.8: Mother's Linguistic Distance and Child mortality: 50 *km* Radius

	(1)	(2)	(3)	(4)	(5)
$\delta = 0.0025$					
Linguistic Distance 50 KM	0.0271***	0.0402**	0.0430***	0.0437***	0.0435***
	(0.00908)	(0.0158)	(0.0135)	(0.0135)	(0.0133)
ELF 50 KM	-0.00303	-0.00396	-0.00540	-0.00653	-0.00739
	(0.0101)	(0.00969)	(0.00893)	(0.00829)	(0.00823)
Observations	654506	654502	654237	654237	653666
R^2	0.090	0.092	0.146	0.146	0.154
$\delta = 0.05$ *à la* Desmet et al. (2012)					
Linguistic Distance 50 KM	0.0201*	0.0349*	0.0388**	0.0400**	0.0398**
	(0.0102)	(0.0183)	(0.0158)	(0.0159)	(0.0157)
ELF 50 KM	-0.00327	-0.00471	-0.00640	-0.00762	-0.00847
	(0.00989)	(0.00934)	(0.00877)	(0.00805)	(0.00800)
Observations	654506	654502	654237	654237	653666
R^2	0.090	0.092	0.146	0.146	0.154
$\delta = 0.50$ *à la* Fearon (2003)					
Linguistic Distance 50 KM	-0.0130	-0.00712	0.00876	0.00989	0.01000
	(0.00958)	(0.0140)	(0.0123)	(0.0124)	(0.0123)
ELF 50 KM	0.00521	0.00315	-0.00438	-0.00585	-0.00678
	(0.00995)	(0.00974)	(0.00882)	(0.00810)	(0.00815)
Observations	654506	654502	654237	654237	653666
R^2	0.090	0.092	0.146	0.146	0.154
Survey Wave FE	Y	Y	Y	Y	Y
Region x Year FE	Y	Y	Y	Y	Y
Ethnicity FE	N	Y	Y	Y	N
Religion FE	N	N	Y	Y	Y
Individual Controls	N	N	Y	Y	Y
Geographic isolation	N	N	N	Y	Y
Ethnicity x Year FE	N	N	N	N	Y

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The dependent variable is the individual child level mortality outcome. The numbers after linguistic distance and ELF indicate the radius of the circle around the mother in which these variables have been calculated. The three panels use three different decay factors δ for calculating LD as indicated in the panel headings. The individual controls include female child dummy, mother's age at birth, mother's age at birth squared, multiple birth indicator, birth order, birth order squared, short birth spacing prior to the birth, short birth spacing after the birth, the location of the mother in the form of an urban dummy, dummies for her educational attainment and her families' wealth index. Geographical isolation controls include the distance of the mother's location from the capital and the logged population in the circle.

Table B.9: Mother's Linguistic Distance and Child mortality: Alternative radii

	(1) 25 km	(2) 75 km	(3) 100 km	(4) 125 km	(5) 150 km	(6) 175 km	(7) 200 km	(8) 250 km
$\delta = 0.0025$								
Linguistic Distance	0.0347***	0.0474***	0.0487**	0.0528**	0.0537**	0.0540**	0.0543**	0.0522**
	(0.00964)	(0.0177)	(0.0192)	(0.0205)	(0.0226)	(0.0229)	(0.0227)	(0.0236)
ELF	-0.00372	-0.00893	-0.0112	-0.00540	0.00393	0.0104	0.0134	-0.0135
	(0.00620)	(0.00996)	(0.0122)	(0.0131)	(0.0138)	(0.0156)	(0.0184)	(0.0258)
Observations	653666	653666	653666	653666	653666	653666	653666	653666
R^2	0.154	0.154	0.154	0.154	0.154	0.154	0.154	0.154
$\delta = 0.05$ *à la* Desmet et al. (2012)								
Linguistic Distance	0.0299**	0.0439**	0.0445**	0.0476*	0.0465*	0.0463*	0.0449	0.0466*
	(0.0118)	(0.0204)	(0.0223)	(0.0240)	(0.0268)	(0.0276)	(0.0279)	(0.0276)
ELF	-0.00407	-0.0104	-0.0125	-0.00695	0.00257	0.00943	0.0132	-0.0136
	(0.00596)	(0.00956)	(0.0117)	(0.0126)	(0.0132)	(0.0149)	(0.0177)	(0.0251)
Observations	653666	653666	653666	653666	653666	653666	653666	653666
R^2	0.154	0.154	0.154	0.154	0.154	0.154	0.154	0.154
$\delta = 0.50$ *à la* Fearon (2003)								
Linguistic Distance	0.00553	0.0119	0.0129	0.0128	0.0107	0.0105	0.0112	0.0240
	(0.00987)	(0.0143)	(0.0156)	(0.0179)	(0.0206)	(0.0223)	(0.0243)	(0.0262)
ELF	-0.00227	-0.00868	-0.0110	-0.00445	0.00629	0.0140	0.0182	-0.0119
	(0.00571)	(0.00970)	(0.0121)	(0.0134)	(0.0142)	(0.0158)	(0.0189)	(0.0255)
Observations	653666	653666	653666	653666	653666	653666	653666	653666
R^2	0.154	0.154	0.154	0.154	0.154	0.154	0.154	0.154

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The dependent variable is the individual child level mortality outcome. The numbers in the column headings indicate the radius of the circle around the mother in which the linguistic distance and ELF have been calculated. The three panels use three different decay factors δ for calculating LD as indicated in the panel headings. All columns include controls for survey wave FE, region x year FE, ethnicity x year FE, religion FE, individual controls and geographic isolation controls described in the notes of Table 1.

Table B.10: Marginal Effects

	Full Sample		Migrants		Non-Migrants	
Circle Radius	Child Deaths	% of SD	Child Deaths	% of SD	Child Deaths	% of SD
25 km	6.62	1.6%	3.31	0.79%	11.49	2.67%
50 km	8.22	2.0%	4.19	1.00%	14.09	3.28%
75 km	9.00	2.1%	4.15	0.99%	16.37	3.80%
100 km	9.30	2.2%	4.26	1.01%	16.86	3.92%
125 km	10.21	2.4%	4.08	0.97%	18.62	4.33%
150 km	10.48	2.5%	3.82	0.91%	18.96	4.41%
175 km	10.61	2.5%	3.54	0.84%	19.20	4.46%
200 km	10.75	2.6%	3.97	0.95%	18.67	4.34%
250 km	10.51	2.5%	4.23	1.01%	17.35	4.03%

Notes: This table provides the marginal effects for the most comprehensive specification indicated in Table 1 for circles of alternative radii around the mother. The leftmost panel includes the full sample of mothers, the middle panel restricts the sample to only migrant mothers, while the last panel restricts the sample to only non-migrant mothers.

Table B.11: Mother's Linguistic Distance and Other Variables:

	(2) tetanus	(3) measles	(4) polio	(5) dpt	(6) iron
Linguistic Distance	0.0103	0.0233	0.0113	0.0232	-0.0449*
	(0.0214)	(0.0279)	(0.0152)	(0.0173)	(0.0248)
ELF	0.00982	0.0341	-0.0239	-0.0266	0.000799
	(0.0212)	(0.0242)	(0.0209)	(0.0234)	(0.0198)
Observations	154650	196627	181890	160914	115498
R^2	0.264	0.329	0.249	0.366	0.360

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The column headings indicate the individual level dependent variable for each specification. These are: tetanus vaccination, measles immunization, polio vaccination, DPT vaccination, and if the mother received iron tablets during pregnancy. A circle of radius 50 km has been considered for calculating the linguistic distance and ELF variables. All columns include controls for survey wave FE, region FE, ethnicity FE, religion FE, year of birth, and geographic isolation controls described in the notes of Table 1.

Table B.12: Mother's Linguistic Distance and Child mortality: Heterogeneity

	(1) female	(2) urban	(3) education	(4) ELF	(5) ELP	(6) population	(7) lndist2cap	(8) wealth
Linguistic Distance	0.0439***	0.0470***	0.0454***	0.0365***	0.0364***	0.0456*	0.0245	0.0598**
	(0.0150)	(0.0173)	(0.0156)	(0.0114)	(0.0136)	(0.0263)	(0.0182)	(0.0292)
Het. Variable	-0.0180***	-0.0148***	-0.00398***	-0.00827	0.00140	0.00581**	0.00405*	-0.00911***
	(0.00136)	(0.00302)	(0.000405)	(0.00797)	(0.00700)	(0.00260)	(0.00219)	(0.000998)
Interaction Term	-0.000938	-0.00826	-0.000624	0.0139	0.00802	-0.000171	0.00356	-0.00477
	(0.00572)	(0.0141)	(0.00166)	(0.0312)	(0.0197)	(0.00247)	(0.00247)	(0.00583)
Observations	653666	653666	653413	653666	653666	653666	653666	653666
R^2	0.154	0.154	0.154	0.154	0.154	0.154	0.154	0.154

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The dependent variable is the individual child level mortality outcome. A circle of radius 50 km has been considered for calculating the linguistic distance and ELF variables. The column headings indicate the Het. Variable included in the specification. The individual controls include female child dummy, mother's age at birth, mother's age at birth squared, multiple birth indicator, birth order, birth order squared, short birth spacing prior to the birth, short birth spacing after the birth, the location of the mother in the form of an urban dummy, dummies for her educational attainment and her families' wealth index. Geographical isolation controls include the distance of the mother's location from the capital and the logged population in the circle.

Table B.13: Correlates of Migrant status

	(1) 25 km	(2) 50 km	(3) 75 km	(4) 100 km	(5) 125 km
Linguistic Distance	0.0543***	0.0646***	0.0758***	0.0827***	0.0874***
	(0.0196)	(0.0220)	(0.0250)	(0.0265)	(0.0275)
ELF	0.00873	0.0234	0.00887	-0.00671	-0.0103
	(0.0217)	(0.0208)	(0.0294)	(0.0405)	(0.0526)
urban	0.0616***	0.0575***	0.0569***	0.0570***	0.0570***
	(0.0162)	(0.0160)	(0.0161)	(0.0161)	(0.0161)
population	-0.0104	-0.00225	0.00381	0.00287	-0.00140
	(0.00631)	(0.00725)	(0.00871)	(0.0101)	(0.0111)
lndist2cap	-0.0106	-0.00717	-0.00459	-0.00469	-0.00594
	(0.0113)	(0.0115)	(0.0114)	(0.0111)	(0.0108)
wealth_index=2	0.0116	0.0110	0.0108	0.0108	0.0109
	(0.0105)	(0.0104)	(0.0105)	(0.0104)	(0.0104)
wealth_index=3	0.0280**	0.0274**	0.0274**	0.0275**	0.0276**
	(0.0126)	(0.0124)	(0.0125)	(0.0124)	(0.0124)
wealth_index=4	0.0693***	0.0681***	0.0679***	0.0681***	0.0682***
	(0.0177)	(0.0174)	(0.0174)	(0.0173)	(0.0174)
wealth_index=5	0.159***	0.158***	0.158***	0.158***	0.158***
	(0.0359)	(0.0353)	(0.0353)	(0.0353)	(0.0354)
incomplete primary	0.00526	0.00480	0.00484	0.00501	0.00506
	(0.0115)	(0.0114)	(0.0114)	(0.0113)	(0.0114)
complete primary	0.00289	0.00229	0.00223	0.00232	0.00233
	(0.0192)	(0.0189)	(0.0189)	(0.0189)	(0.0190)
incomplete secondary	-0.00331	-0.00374	-0.00387	-0.00388	-0.00389
	(0.0253)	(0.0251)	(0.0251)	(0.0251)	(0.0251)
complete secondary	0.0346	0.0341	0.0342	0.0344	0.0344
	(0.0386)	(0.0383)	(0.0383)	(0.0383)	(0.0383)
higher	0.0194	0.0186	0.0184	0.0186	0.0187
	(0.0337)	(0.0335)	(0.0335)	(0.0335)	(0.0334)
Observations	164141	164141	164141	164141	164141
R^2	0.122	0.122	0.122	0.122	0.122
Observations	164141	164141	164141	164141	164141
R^2	0.122	0.122	0.122	0.122	0.122

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The dependent variable is the migrant status of the mother. The column headings indicate the radii of the circles around the mother used to construct the linguistic distance, ELF and population variables. The lndist2cap variable gives the Log(Distance to the capital). "No education" is the excluded category for the educational attainment variable. All columns include controls for survey wave FE, region FE, ethnicity FE, religion FE, year of birth, and geographic isolation controls described in the notes of Table 1.

Table B.14: Mother's Linguistic Distance and Other Variables 1: Migrants vs. Non-Migrants

	(1) infant	(2) neonatal	(3) HAZ	(4) stunted	(5) WAZ
Migrants					
Linguistic Distance	0.00574	0.00297	-0.0468	0.0159	-0.0437
	(0.00502)	(0.00312)	(0.0761)	(0.0252)	(0.0465)
ELF	-0.00561	-0.00510*	0.00662	-0.00842	0.0644
	(0.00512)	(0.00290)	(0.0548)	(0.0189)	(0.0422)
Observations	348759	368880	66230	66230	66230
R^2	0.107	0.079	0.214	0.162	0.177
Non-Migrants					
Linguistic Distance	0.0396***	0.0134***	-0.129**	0.0426***	-0.0465
	(0.00869)	(0.00329)	(0.0619)	(0.0134)	(0.0440)
ELF	-0.00315	-0.00351	0.106*	-0.0218	0.163***
	(0.00530)	(0.00282)	(0.0582)	(0.0220)	(0.0580)
Observations	299028	315688	52477	52477	52477
R^2	0.111	0.082	0.218	0.169	0.169

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The column headings indicate the individual child level dependent variable for each specification. These are: infant mortality, neonatal mortality, height-for-age Z-score (HAZ), stunting, and the weight-for-age Z-score (WAZ). Panel 1 (Panel 2) restricts the sample to only migrants (non-migrants). A circle of radius 50 km has been considered for calculating the linguistic distance and ELF variables. All columns include controls for survey wave FE, region x year FE, ethnicity x year FE, religion FE, individual controls and geographic isolation controls described in the notes of Table 1.

Table B.15: Mother's Linguistic Distance and Other Variables 2: Migrants vs. Non-Migrants

	(1) tetanus	(2) measles	(3) polio	(4) dpt	(5) iron
Migrants					
Linguistic Distance	0.0620**	0.00764	0.0252	0.0387*	0.0271
	(0.0261)	(0.0292)	(0.0199)	(0.0225)	(0.0256)
ELF	0.00164	0.0209	-0.0336	-0.0513*	0.00384
	(0.0260)	(0.0281)	(0.0255)	(0.0279)	(0.0255)
Observations	67000	83952	77421	69444	48693
R^2	0.249	0.336	0.217	0.359	0.310
Non-Migrants					
Linguistic Distance	-0.0639*	-0.0166	-0.00954	0.00510	-0.108***
	(0.0323)	(0.0288)	(0.0251)	(0.0227)	(0.0343)
ELF	0.0196	0.0224	-0.0170	-0.0178	-0.00390
	(0.0272)	(0.0370)	(0.0215)	(0.0274)	(0.0249)
Observations	54496	69350	64774	58565	39198
R^2	0.300	0.340	0.239	0.384	0.351

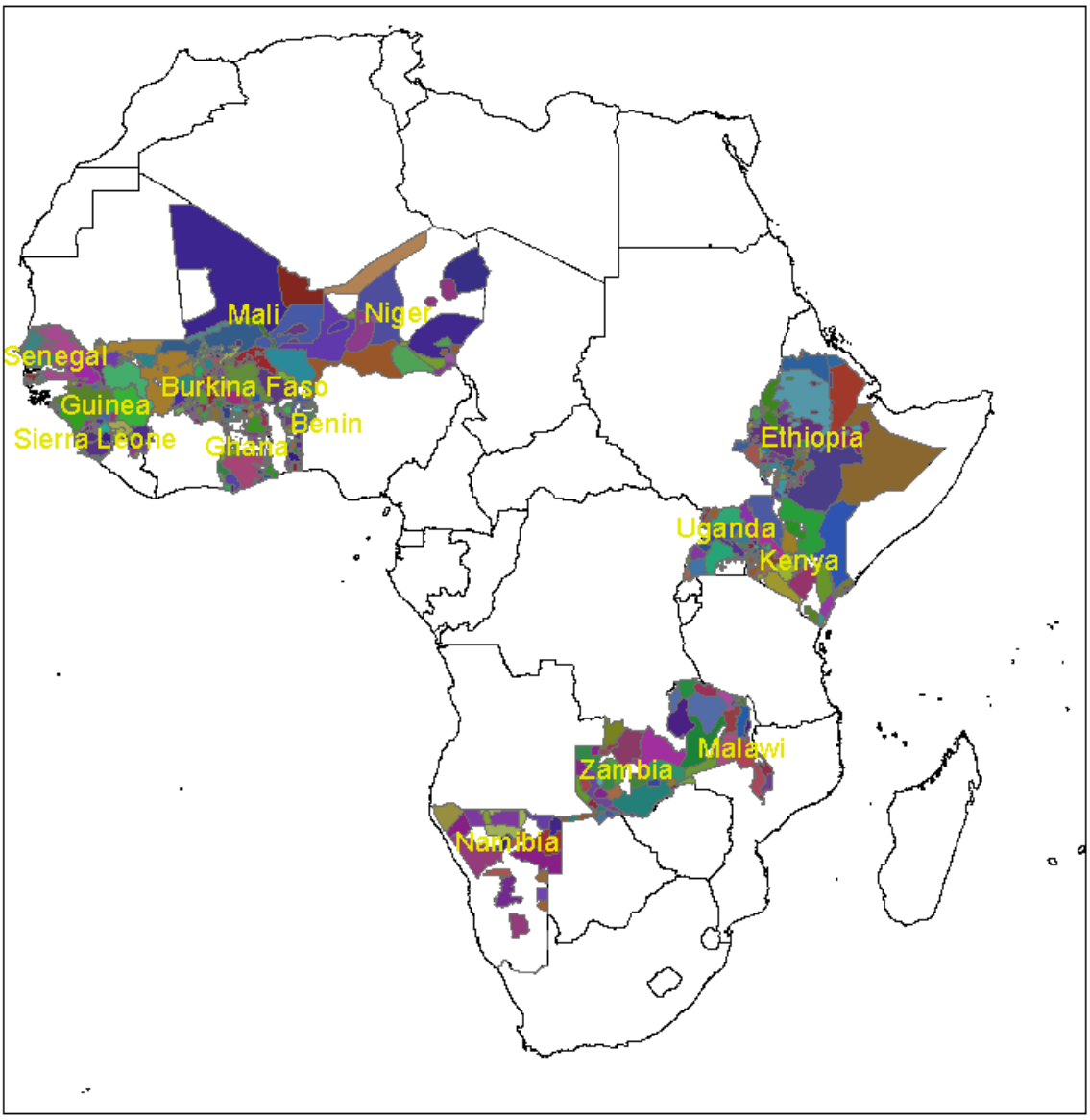
Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the region level. The column headings indicate the individual level dependent variable for each specification. These are: tetanus vaccination, measles immunization, polio vaccination, DPT vaccination, and if the mother received iron tablets during pregnancy. Panel 1 (Panel 2) restricts the sample to only migrants (non-migrants). A circle of radius 50 km has been considered for calculating the linguistic distance and ELF variables. All columns include controls for survey wave FE, region x year FE, ethnicity x year FE, religion FE, individual controls and geographic isolation controls described in the notes of Table 1.

C Appendix Figures

Figure C.1: Countries used

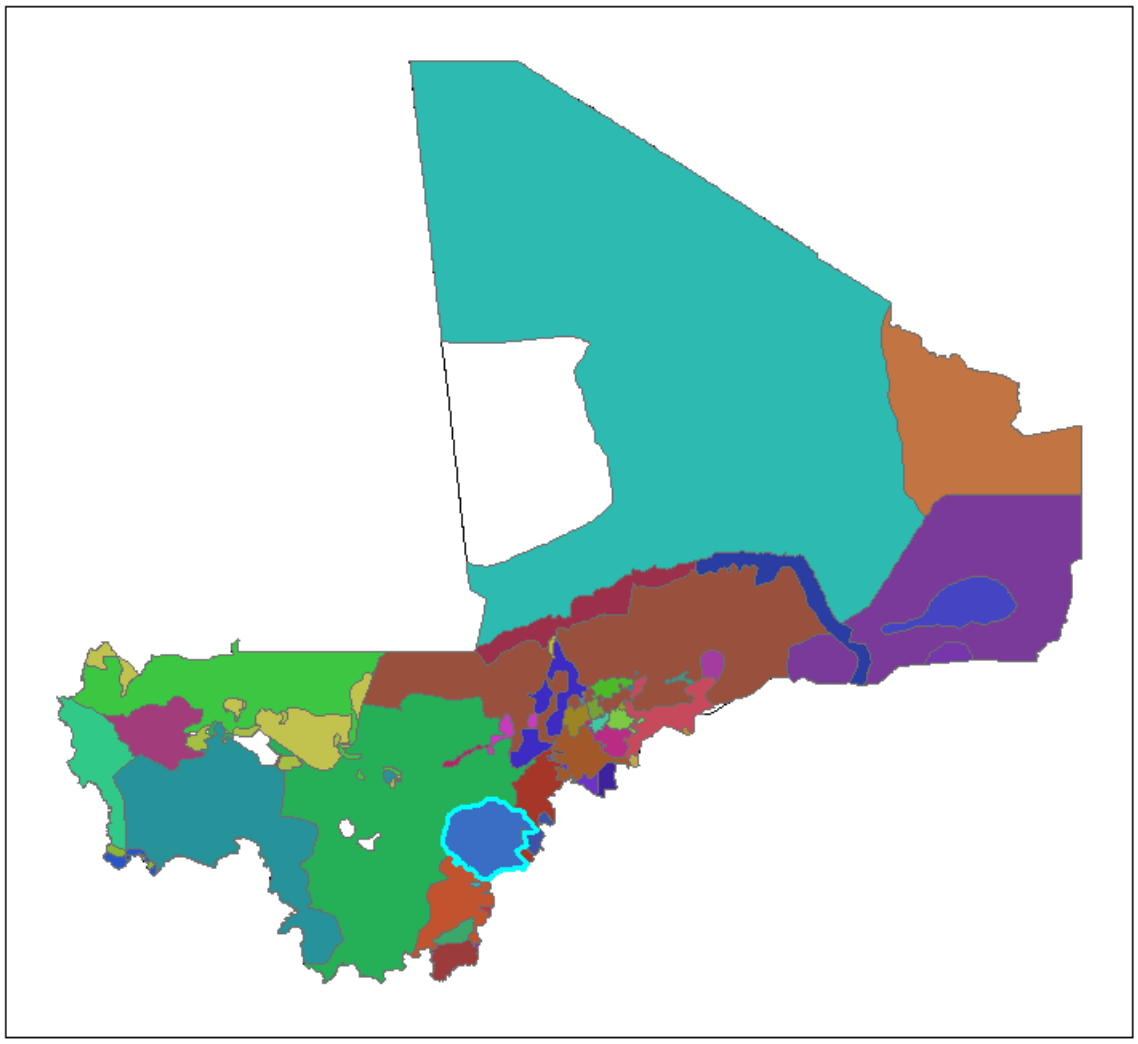
Notes: This map plots the fourteen countries used in the study: Benin, Burkina Faso, Ethiopia, Ghana, Guinea, Kenya, Malawi, Mali, Namibia, Niger, Senegal, Sierra Leone, Uganda, and Zambia.

Figure C.2: Languages used



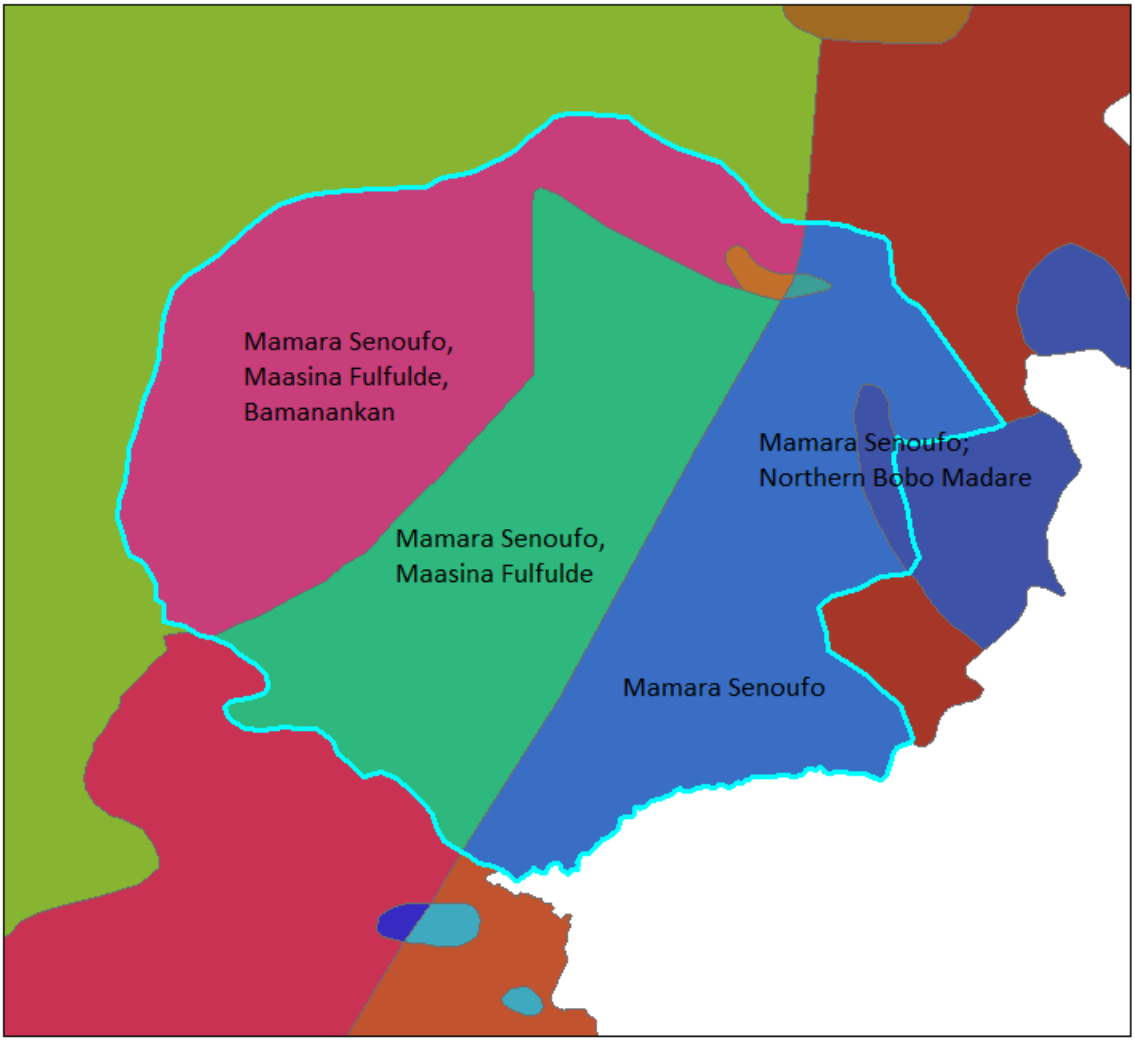
NOTES: This map plots the linguistic groups for the fourteen countries used in the study from the Ethnologue database. Polygons of different colours represent the different language groups. Areas where multiple languages are spoken are represented by overlapping polygons, which are not distinguishable in this map.

Figure C.3: The Languages of Mali



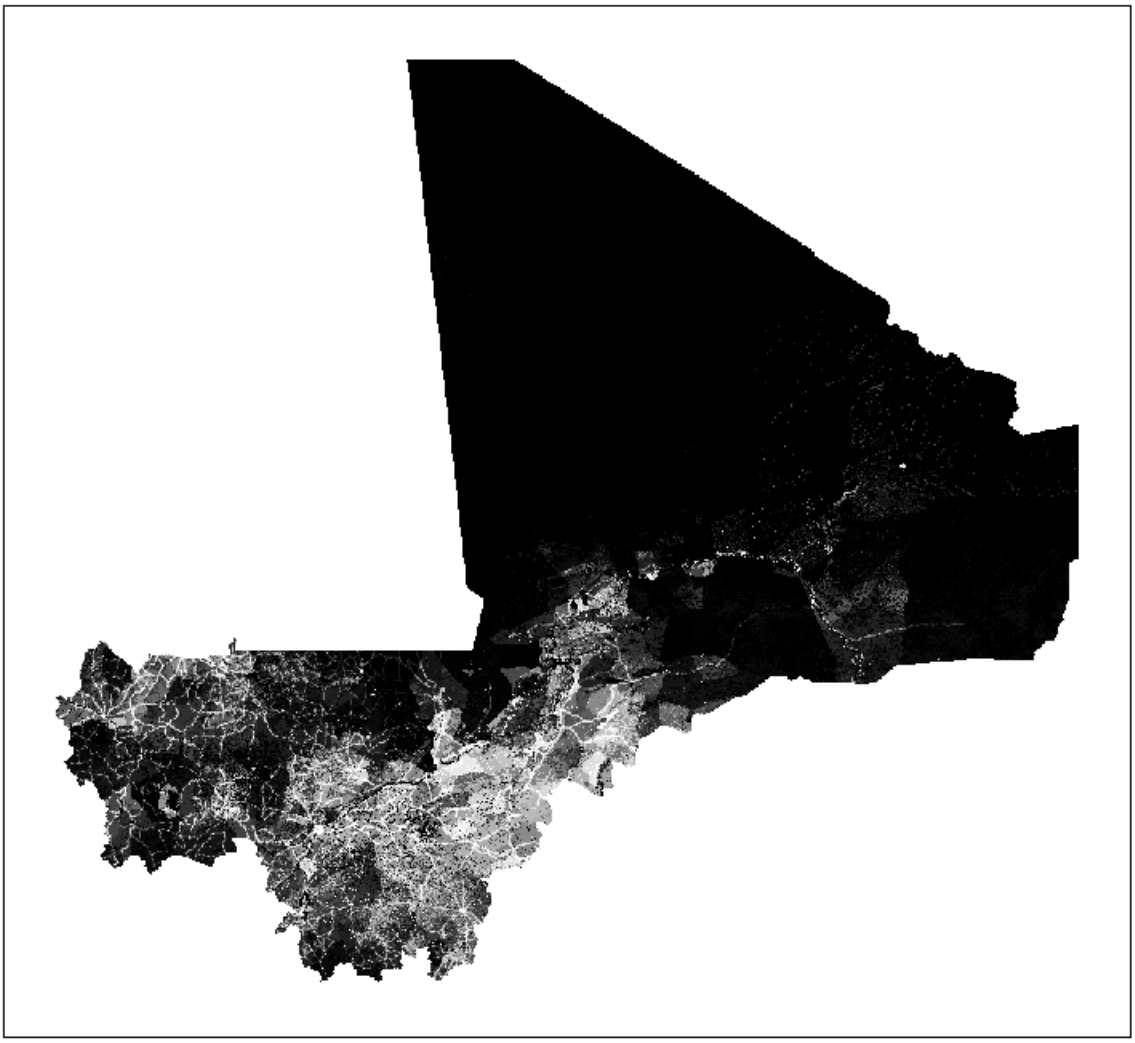
NOTES: This map plots the linguistic groups of Mali from the Ethnologue database. Polygons of different colours represent the different language groups. Areas where multiple languages are spoken are represented by overlapping polygons, which are not distinguishable in this map. The polygon highlighted in blue in the south-eastern corner of the map demarcates the linguistic homeland of the Mamara Senoufo language speakers.

Figure C.4: Example of Overlapping Language Polygons



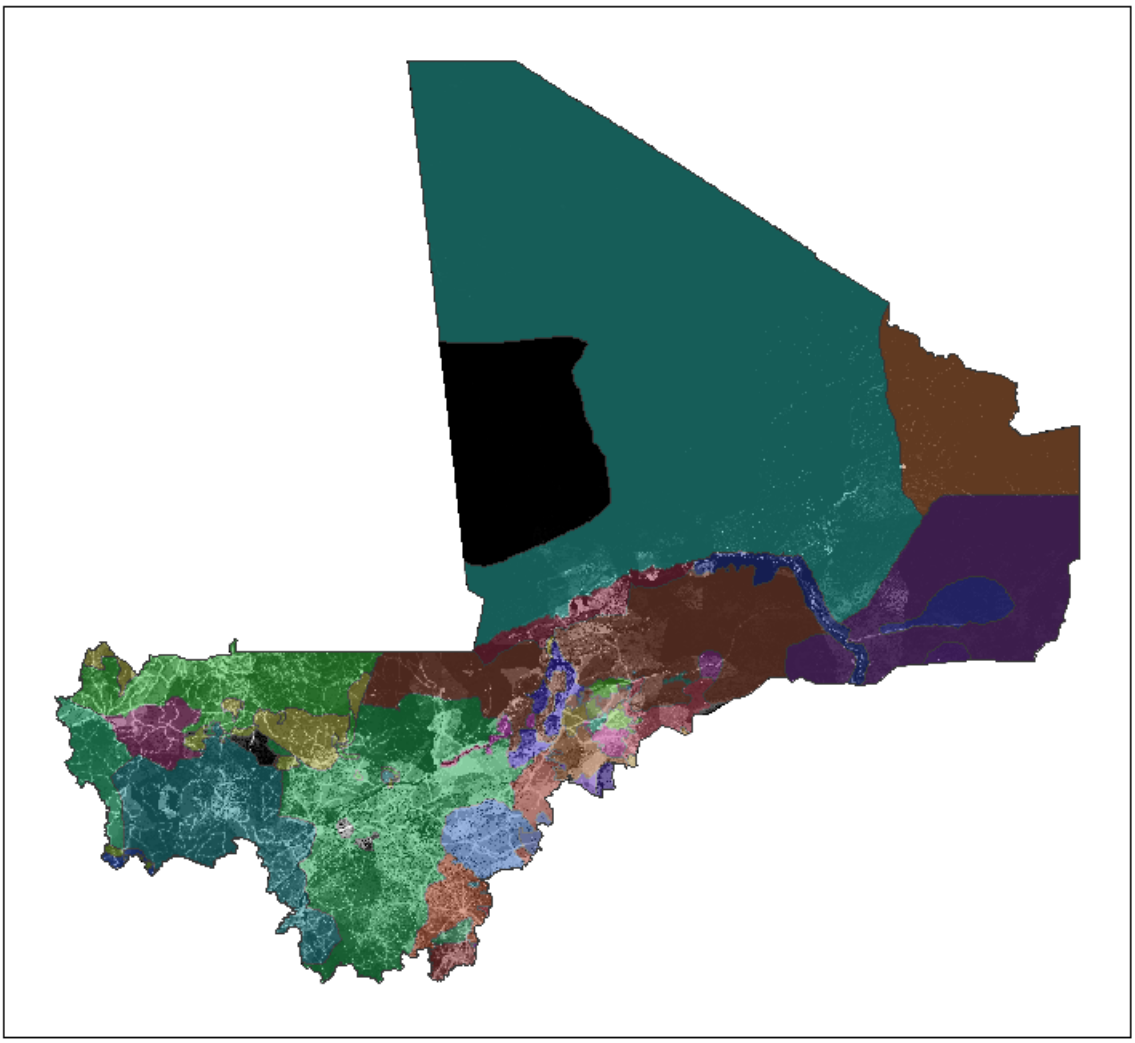
NOTES: This map plots the linguistic groups in the south-eastern region of Mali from the Ethnologue database. Polygons of different colours represent the different linguistic areas. The polygon highlighted in blue demarcates the linguistic homeland of the Mamara Senoufo language speakers. In the light blue shaded polygon in the south-east corner of the map, there are no other languages spoken apart from Mamara Senoufo. In the polygon with a darker shade of blue, just north of this area, both Mamara Senoufo and Northern Bobo Madare are spoken. In the green shaded polygon in the centre of the map, Mamara Senoufo and Maasina Fulfulde are spoken. Finally, in the pink shaded polygon in the west, Mamara Senoufo is spoken with two other languages viz. Maasina Fulfulde and Bamanankan.

Figure C.5: The Population of Mali



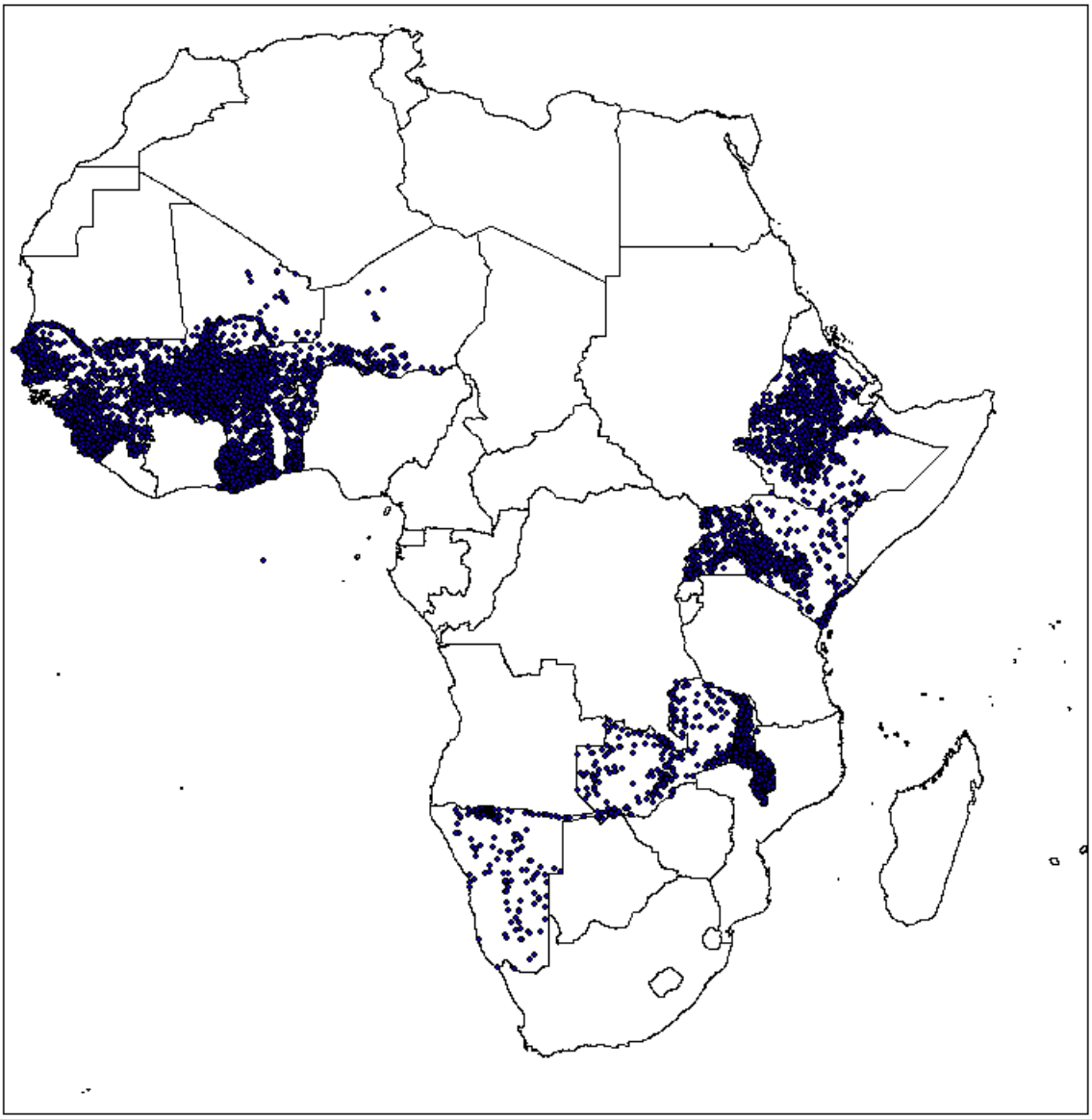
NOTES: This map plots the population distribution of Mali from the LandScan database at the 30 arc seconds x 30 arc seconds (roughly 1 x 1 sq. *km* at the equator) resolution. The brighter (darker) pixels within the geographic boundaries of Mali represent more (less) populated areas.

Figure C.6: Mali Overlay



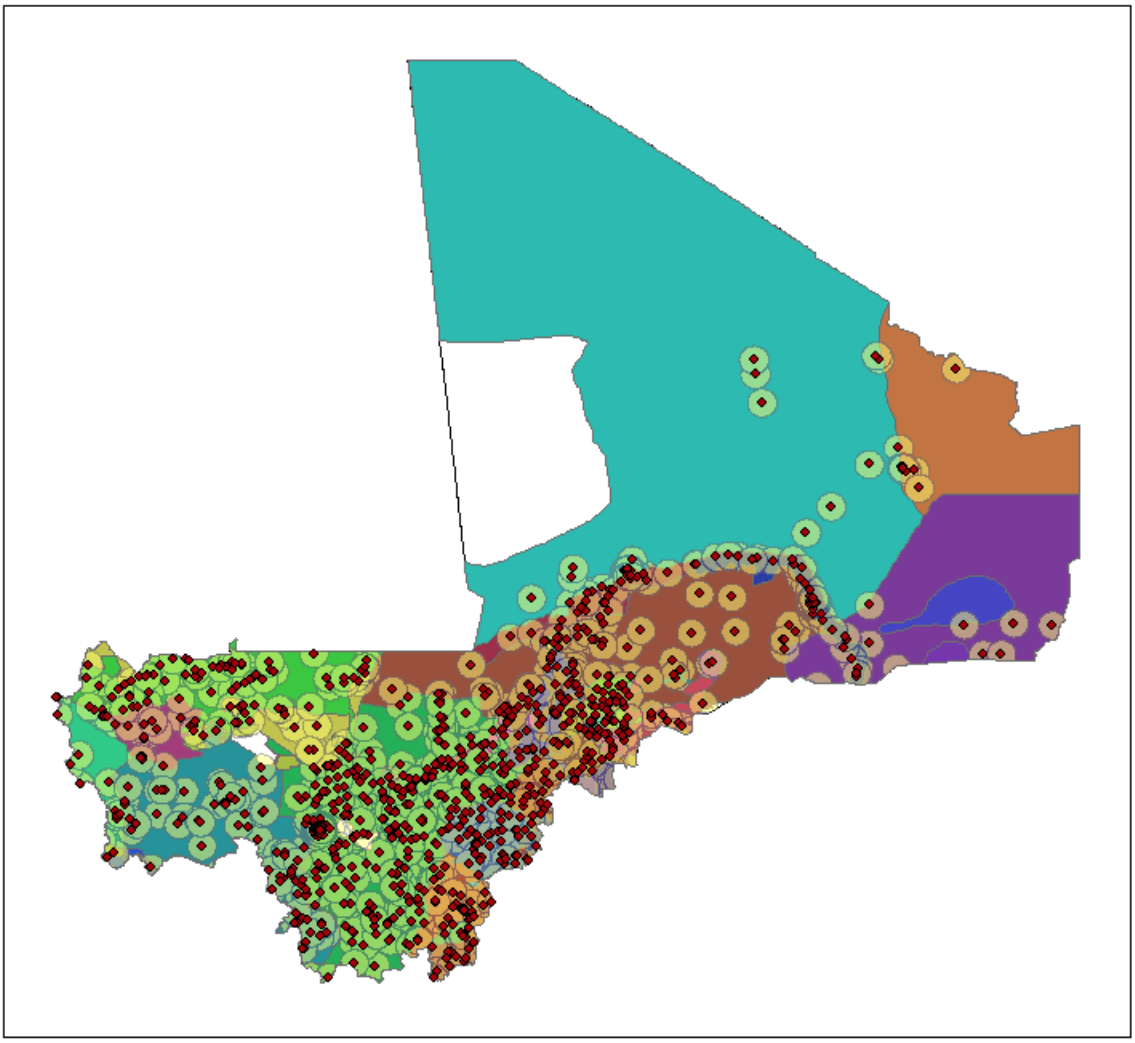
NOTES: In this map we overlay the language polygons for Mali from the Ethnologue database (see Figure C.3) on the population distribution of Mali from the LandScan database (See Figure C.5). Polygons of different colours represent the different language groups in the language group map. Areas where multiple languages are spoken are represented by overlapping polygons, which are not distinguishable in this map. The brighter (darker) pixels within the geographic boundaries of Mali in the population map represent more (less) populated areas.

Figure C.7: Mothers' Locations



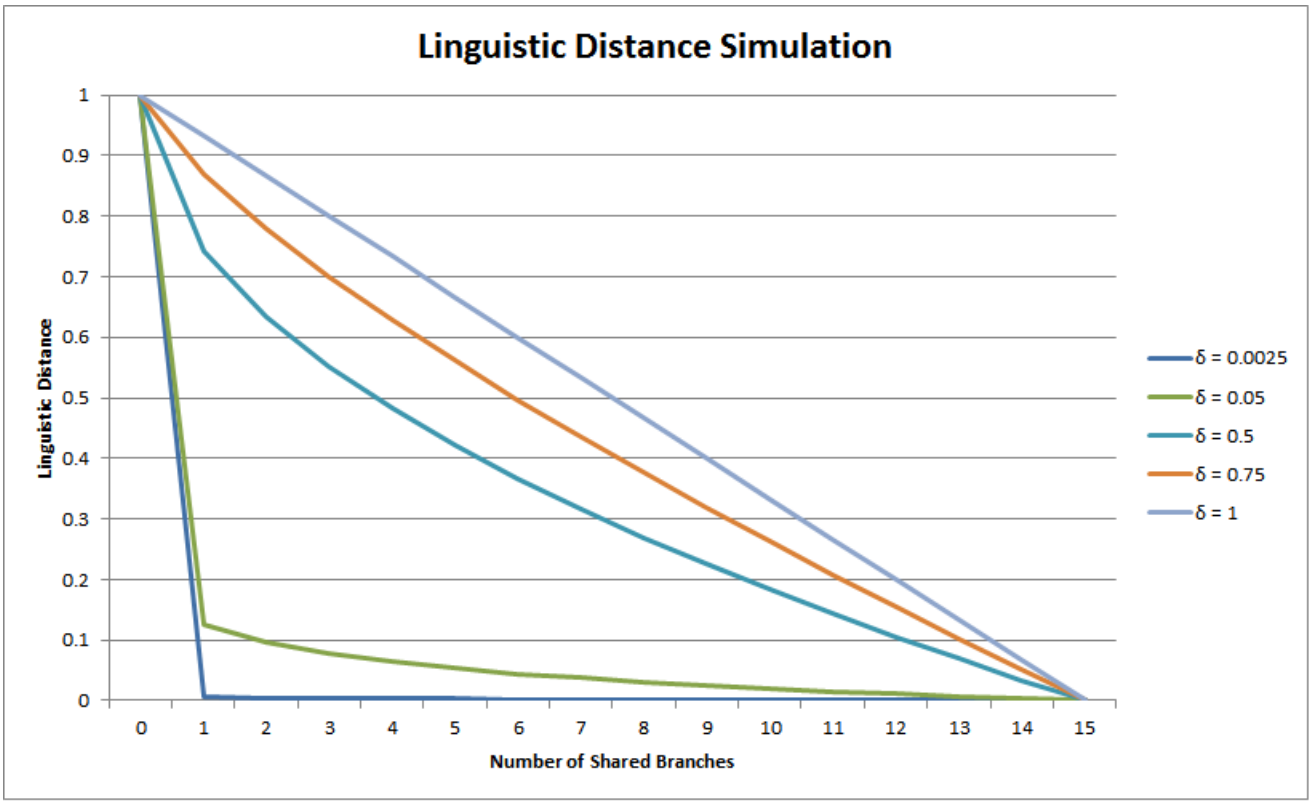
NOTES: This map plots the locations of the 28,839 DHS clusters where the 208,898 individual mothers, that comprise our sample, are located.

Figure C.8: Mali DHS Clusters



NOTES: This map plots the linguistic groups of Mali from the Ethnologue database (See Figure C.3). The red dots represent the locations of the mother's (DHS clusters) for Mali and the circles around them represent 25 km circles around the mothers.

Figure C.9: Linguistic Distance for alternative values of the decay factor δ



NOTES: In this graph we simulate how linguistic distance changes for alternative values of the decay factor δ. The x-axis gives how many branches any two languages share and the y-axis gives the corresponding values of linguistic distance for different values of δ.