



Economic Research Paper: 2009-08

A QUESTION OF SCALE: ETHNIC DIVERSITY AND THE PROVISION OF PUBLIC GOODS

KEN JACKSON*

JUNE 2009

* Department of Economics, School of Business and Economics, Wilfrid Laurier University.
Contact: kjackson@wlu.ca

A Question of Scale: Ethnic Diversity and the Provision of Public Goods*

Ken Jackson[†]

June 16, 2009

Abstract

Ethnic diversity has been shown to have negative consequences on the delivery of key public goods. It may be expected that the impact of ethnicity will be most critical in the diverse societies of Sub-Saharan Africa where the problems of development appear particularly intractable. However, given the centralized nature of African politics, the measurement of ethnic diversity is an important factor in the analysis. This paper demonstrates that ethnic diversity has a significant impact on the provision of basic public services such as water and electricity across 13 countries of Sub-Saharan Africa, when measured at local and regional scales, respectively.

Keywords - ethnic diversity, public goods, Africa JEL codes - H42, O17, H70

*Numerous people have provided helpful advice to thank Chris Barrington-Leigh, Chris Bidner, Alberto Alesina, Siwan Anderson, Patrick Francois, Nathan Nunn, and participants in seminars at the University of British Columbia and the NEUDC conference at Harvard University in 2008.

[†]*E-mail* - kjackson@wlu.ca, *Website* - <http://www.kjackson.net/>, <http://www.wlu.ca/sbe/kjackson/>

1 Introduction

“The international development community should speak of the Big Five development interventions that would spell the difference between hunger, disease and death and health and economic development.” (Sachs 2005, p.235)

The Big Five interventions discussed by Sachs are:

1. Agricultural improvements (related to poor credit markets)
2. Investments in basic health
3. Investments in education
4. Power, transport and communications services
5. Safe drinking water and sanitation

The striking fact about this list is that all five interventions relate to the provision of quasi-public goods. Further, there is little productive government spending in any country that does not fall into this list¹. This redefines the key issue in development economics. Why are the governments of developing countries unable to provide the basic services necessary to escape poverty?

One answer is social heterogeneity. Diverse societies, particularly with respect to ethnic diversity, have been the focus of extensive economic research over the past 10 years, starting with Easterly and Levine (1997). This research has demonstrated that ethnic diversity has a significant negative effect on the provision of public goods, and subsequently economic growth. Further research into sub-national governance, initiated by Alesina et al (1999), has further established that the effects of diversity occur at all levels of government. The impact of ethnicity could be critically important for Sub-Saharan Africa, where the greatest levels of diversity are found and the problems of underdevelopment are particularly intractable.

This paper analyzes the connection between ethnic diversity and public good provision in 13 Sub-Saharan African countries. In particular, I focus on the provision of piped drinking water and electricity as fundamental goods that are commonly provided or managed by government, and are critical to economic development. While neither good is a pure public good, as individuals can be excluded from accessing the service, there are significant economies of scale in their provision, thus creating an environment in which government regulation can create significant benefits.

¹The two main categories of government spending not on the list are military spending and transfers between groups - neither of which are usually considered productive.

Water and electricity are also both severely underprovided in Sub-Saharan Africa, particularly in rural areas. For Africa as a whole, less than 5% of people living in rural areas have access to piped water, and approximately 13% have access to electricity (Banerjee, Diallo, and Wodon 2007). This same report also finds that while access to electricity has increased slightly since the early 1990's, access to piped water has actually fallen. For improved water sources in general, 58% of the population Sub-Saharan Africa had access in 2006, as compared to a worldwide average of 86%². In fact, Sub-Saharan Africa is the only region in the world for which access to improved water is lower than the world average. Similarly, electricity consumption in Sub-Saharan Africa (excluding South Africa), lags far behind the rest of the world. While electricity consumption per capita has increased substantially throughout the world since 1990, it has barely increased in Sub-Saharan Africa.

However, while water and electricity are similar in this way, they are very different in their scale of operations. Water is a heavy, relatively low-value good, that is expensive to pipe across any significant distance. As reported in Clark and Stevie (1981), the efficient scale of piped water provision in the United States as of the late 1970's is on the order of 10miles. With the power needed for pumping electricity in much shorter supply in Sub-Saharan Africa, it seems safe to assume that the efficient scale in this survey is likely much smaller. Beyond this distance, the cost of piping water become extensive in comparison to the cost of establishing a new system. Electricity, in contrast, contains large economies of scale in production and the cost of transmitting electricity are relatively low. Filippini (1998) finds that electricity providers covering areas smaller than 2500km² would be more efficient at larger scales. These economies of scale disappear for the larger providers that cover territories of 5-10,000km². While electricity can be provided at smaller scales, particularly through the use of small-scale generators, this tends to be a very inefficient method of producing electricity and is not widespread throughout Sub-Saharan Africa³. In addition, the future of electricity generation in Sub-Saharan Africa likely includes far greater use of large-scale hydroelectric projects, as only 7% of this resource has been tapped(Karekezi and Kimani 2002).

This fact corresponds well to the economic literature, starting with Tiebout (1956)⁴, on the decentralization of government that has suggested that governments should decentralize responsibility to the lowest level of government that can capture the economies of scale necessary for provision. Panizza (1999) considers the theoretical and empirical implications of ethnic diversity for decentralization, and finds that decentralization should be, and is,

²All data drawn from the World Development Indicators.

³For example, in 1992 in Ethiopia, 9% of the population had access to hydroelectricity, while only 1% had access to diesel-generated electricity. This basic story has not changed significantly since(Jarso 2002).

⁴Or see Besley and Coate (2003) for a more recent theoretical analysis of the issue.

greater in more diverse countries.

For the most part, the literature on ethnic diversity and public good provision has therefore been able to sidestep the issue of how to measure ethnic diversity. In general the decision on how to measure ethnic diversity would be a critical step in the analysis, as one needs to decide the level of ethnic diversity that will affect the provision of the public good. Theoretically, the two geographic scales that could matter would be the scale of provision and the scale of governance. The scale of provision is the geographic area that encompasses the majority of the spillover benefits from a project, while the scale of governance is the geographic area that is represented by the government in question.

In the standard case, when the scale of provision and the scale of governance coincide, there appears to be no question as to the scale at which ethnic diversity might affect public good provision. As such, Alesina et al.(1999) demonstrates that the diversity of cities or counties affects the expenditures of city or county governments in the United States. Alternatively, Miguel and Gugerty (2005) measure the effect of community diversity on the provision of local education services in Western Kenya. In each case, the goods provided are essentially local, and diversity is measured over very narrow areas.

However, it is not always this clear. In particular, this paper is focused on Sub-Saharan Africa, where the decentralization of government has proceeded in a relatively haphazard fashion⁵. While many countries have devolved some responsibilities to local or district councils, these groups tend to have severely limited powers. Further, where national or district governments have de jure responsibility, these governments often abdicate their role, and leave the actual provision of local services to local governments⁶. As such, there is an underlying pressure for the scale of governance to shift toward the scale of provision. With this in mind, I hypothesize that ethnic diversity, when measured at the scale of provision, will affect the provision of public goods. It is anticipated that village level diversity will impact the provision of piped drinking water, and provincial level diversity will effect the development of electricity networks.

As mentioned, the scale of provision and the scale of government often coincide. Where they do not, the measurement of the impact of ethnic diversity is difficult. Banerjee et al

⁵See Olowu and Wunsch (2004) for a recent survey of selected countries' experiences with decentralization to local councils.

⁶As an example, in Chad, while the legal responsibility for education funding remained with the central government, over 60% of the countries teachers are paid by local community associations (Fass and Desloovere 2004). Similarly, as shown in Miguel and Gugerty (2005), local communities are critical in providing funding for education and water services, even though the official responsibility for the provision of these services remains at higher levels of government.

(2005) and Banerjee and Somanathan (2007) measure the impact of district levels of diversity in India on the provision of a variety of public goods that are provided at the town level, including schools, water systems or health units. In each case the provision of the public good in one town in a district can be expected to have little effect on the provision in a neighbouring town. However, the assumption of the paper is that the effect of diversity occurs through the motivation of the district level government official. In diverse communities, where the official will, on average, have fewer constituents of their own ethnic group, she will feel less pressure to bring projects to her home district. Measured in this way, these two papers show that the effects of ethnic diversity are economically small and not always statistically significant.

There are two possibilities. First, ethnic diversity in India may not play an important role in Indian politics. Alternatively, there may be effects of ethnic diversity as measured at a more local scale. Specifically, ethnic diversity, as measured at the village level could still have an impact on access to the public good but this effect would be severely underestimated in a district level analysis. While evidence for this effect is limited, there is a precedent for considering the differential impact of ethnic diversity as measured at different scales. Luttmer (2001) analyzes the effect of diversity on popular support for welfare policies in the United States. The results show that while local diversity reduces support for welfare, state level diversity has no effect.

Limited public good provision may be the result of inefficiencies in demand or supply. The theoretical literature on the provision of public goods in the context of ethnic diversity has considered both. Alesina et al. (1999) models the effects of ethnic diversity as a problem of demand. Individuals of different ethnic groups prefer the public good in different forms, thus limiting the utility derived from the single good provided by government. In contrast, Miguel and Gugerty (2005), considers the role of inefficiency in supply. If the government relies on mechanisms internal to each ethnic group to enforce contributions to the public good, then diverse communities will be less efficient at collecting funds, and will therefore provide lower levels of the public good.

If diversity creates differential demand, there is little question that the natural scale for diversity measurement is at the scale of provision, regardless of the scale of governance. A district government that oversees two communities that want different local public goods could, at least theoretically, provide those two communities with different goods. As an example, if two neighbouring villages are each comprised of a different ethnic group, speaking different languages, the problem of providing education is almost irrelevant if each of the villages has their own school. The district does not need to worry about compromising on the public good - simply set up schools in each community in the preferred language.

In contrast, when the challenges of diversity are created by an inefficient government, it is more difficult to determine the channel through which ethnic diversity affects public good provision. However, as discussed in section 2, an inefficient government that provides funding in response to local lobbying will create a similar pattern of public good provision that depends on local diversity. In contrast, where spillovers across a region are large, local diversity will not substantially impact the provision of the public good.

The main premise of this paper is that local diversity will affect the provision of piped water, while regional diversity will affect the provision of electricity. To test this idea, I use data on 104,000 households in 13 countries of Sub-Saharan Africa. The data comes from the most recent available wave of the Demographic and Health Surveys conducted by Measure DHS. The key to this data is the ability to collect households into communities based on their location, thus permitting the measurement of ethnic diversity and public good supply at the local scale that does not correspond with a well-defined governance body.

The results indicate that, on average, local diversity significantly affects the provision of piped water, yet has no effect on electricity. District level diversity has a negligible impact on either public good, while provincial diversity has a significant impact on the provision of electricity. To further emphasize this point, I demonstrate that the effect of ethnic diversity at local scales for piped water provision aggregates up to affect district level rates of piped drinking water. This demonstrates that the reduction in measured effects from altering the geographic unit of measurement to the district level is entirely the result of changing the definition of ethnic diversity.

2 Why Might the Scale of Diversity Matter?

Existing research has focused, primarily, on diversity as measured at the scale of governance. This is not surprising, as the mechanisms that underly our understanding of public good provision are at the scale of governance. A government is tasked with providing a public good, and they will be more or less effective in that undertaking. If diversity affects the ability of individuals in government to manage the economy effectively, then diversity at the scale of governance is critical.

However, it is far from certain that this is the path through which diversity matters. First, as discussed above, if the mechanism behind the effects of diversity is variation in demand, then the only critical scale is the scale over which the government can vary supply. If the government can efficiently vary the supply of public goods at a sufficiently small scale, then there is no problem with diversity of demand.

Where the effects of diversity are caused by the inability of government to efficiently supply a public good the relationship between different scales of diversity is more complicated. However, a simple extension to the model in Miguel and Gugerty (2005) demonstrates the role for diversity measured at a scale significantly less than the scale of governance. In Miguel and Gugerty, individuals contribute to the public good in response to social pressure from others in their ethnic group. In communities divided into small ethnic groups, the potential cost of social sanctions is insufficient to generate contributions, and for this reason, diverse communities are less likely to have access to the public good.

Now consider a situation in which local contributions do not directly affect the provision of the public good. Instead, the local community contributes to a lobbying effort to gain funds from either a central government or foreign organization⁷. These funds are then used to build the public goods in question. The contributions of individual citizens to the lobbying effort is now the public good that is driven by the expectation of social sanctions, and diverse communities that are unable to generate equivalent resources for lobbying are left with low levels of the public good. While the scale of governance has changed, the role for diversity remains at the scale of provision.

The argument advanced here is not that diversity at the scale of governance will have no effect, just that diversity at the scale of provision may have critical effects regardless. In addition, where the de facto scale of governance is unclear, measuring ethnic diversity at the scale of provision is a reasonable approximation of the impact of ethnic diversity. While existing research has focused entirely on the scale of governance, where the scale of provision and the scale of governance are different it is critical to take this into consideration in designing an empirical study.

3 The Governance of Water and Electricity in Sub-Saharan Africa

This paper considers the effects of ethnic diversity on the provision of public goods across 13 countries in Sub-Saharan Africa, as shown in figure 3. Incorporating a moderate number of relatively similar countries into the analysis generates benefits in terms of the generality of the results, yet poses problems as the institutional details of public good provision vary

⁷The source of funding is not critical to the analysis here. Of note is the fact that the World Bank has recently increased funding for community-driven development projects(Mansuri and Rao 2004). To the extent that local communities need to cooperate to access these funds, the involvement of the international community may actually exacerbate the problems of localized ethnic diversity.

across the sample.

(Insert figure 3 here.)

As expected given the nature of the two goods, the story is quite different for electricity and water. For electricity, the picture is relatively simple. While there has been attempts at reform, power utilities in Africa continue to hold monopolistic positions in most countries (Karekezi and Kimani 2002), with the majority run directly by government. Even in Ethiopia, a large country that is widely praised for implementing significant decentralization, the responsibility for providing electricity remains with the national power authority (Teferra 2002). Ghana and Cote D'Ivoire are two countries that have implemented significant reform of their power sector, yet production remains monopolized in a single national entity (Karekezi and Kimani 2002). In the remainder of the countries included in this study, the supply of electricity is managed by the national government. For the purposes of this paper, it is not possible to consider how electricity is affected by diversity at the scale of (de jure) governance.

The situation with regards to water is far more complicated. As noted previously, the provision of drinking water is usually best left in the hands of a local government. Historically, countries in Africa maintained the highly centralized governments that were left behind by the colonial powers. Starting in the mid-1980's, there was a significant push to decentralize the actions of government, particularly in some of the largely countries on the continent, such as Nigeria and Ethiopia (Olowu and Wunsch 2004). For example, in 1994, Ethiopia adopted a federalist constitution that devolved significant responsibility to provincial governments⁸.

While many countries devised systems that suggest a movement toward effective decentralization, the reality has been somewhat limited. In particular, it appears that even where local and district councils are officially developed, it is not clear that real power ever reached these councils. As was pointed out by Daniel Treisman:

In Sub-Saharan Africa, countries tend to have more tiers of government, but relatively low fiscal and personnel decentralization, and little local autonomy or local electoral accountability. (Treisman [p.35](2002))

Thus, for all the talk of decentralization, Africa has lagged behind the rest of the world in terms of the decentralization of government expenditures⁹.

⁸Throughout this paper, for clarity I refer to the largest sub-national units of a country as provinces, and the second-order administrative units as districts, regardless of differences in terms across countries.

⁹In the early 1990's, subnational governments accounted for just 11% of official government expenditures, with only the Middle East and North African region having more centralized governments (Treisman 2002).

For example, in Burkina Faso, the central government developed a series of laws through the 1990's that developed guidelines for decentralization, including both provincial and local governments. However, by 2003, while local government had been setup in urban areas, no such development had occurred in the rural parts of the country. Further, while temporary provincial governments had been developed, no elections had been held and their powers were extremely limited(Ouedraogo 2003). In other areas of Africa, such as the Central

In contrast, Ghana and Kenya have each developed district councils that, among other things, bear responsibility for the provision of drinking water. These districts, which match the districts present in the data in this paper, contain multiple villages and are in general much larger than the scale of provision for water. However, it is not clear that these district councils have sufficient funding or power to effectively govern the rural communities for which they are responsible(Brosio 2000). Further, Olowu and Wunsch (2004) characterize the effectiveness of district governance in Ghana and Kenya as low and very low, respectively. A broad overview of the 13 countries in the study indicates that while some have attempted to decentralize some services, this decentralization is not consistent either across countries or even necessarily within countries. Further, even where local and district governments have been developed, the power of these councils is extremely limited. Local governments often operate without sufficient funding and must therefore rely on the contributions of the local citizens.

Given the uncertainty regarding the division of governmental responsibility, the hypothesis tested here is that diversity will matter at the scale of provision. There are two reasons that this could be true. First, as in the optimal case, this could be the scale of governance. Alternatively, the governance or funding of the public good may rest at a higher level of government, but that government responds to the lobbying efforts from people in their communities. In either case, to the extent that ethnic diversity has an impact on public services, one logical place to start is by looking at the scale of provision.

4 Data and Empirical Specification

The household surveys conducted by Measure DHS in developing countries over the past 25 years provide an appealing source of data in which to consider this issue. In addition to household information, the geographic location of each survey cluster is recorded, allowing nearby clusters to be combined into single communities even in the absence of an official local governance structure. This paper draws on data from 13 countries of Sub-Saharan Africa and encompasses over 100,000 households. As discussed above, it is not immediately clear

how to analyze the impact of ethnic diversity on the provision of water and electricity. In particular, the scale of provision and the de jure scale of governance do not directly coincide in the vast majority of countries in the sample.

I focus on three distinct sub-national scales of governance. The first two, provincial and district, correspond to the first and second level administration districts that are defined for each country. The assignment of each surveyed household to a district is possible from the data provided by Measure DHS. Across the 13 countries in the sample, there are 114 provinces, and 1,200 districts. The third level, local governance, is more difficult to define, specifically because this level of administration is not legally defined in many countries.

4.1 Defining Local Communities

The data available from Measure DHS defines households by their survey cluster. One approach to this analysis would therefore measure ethnic diversity at the cluster level, and relate that to the level of public good access for households within that cluster. However, casual inspection reveals multiple clusters located at precisely the same geographic coordinates, particularly in urban areas. In these cases it seems unreasonable to define the two clusters to be in different communities with respect to the provision of piped water. I proceed with two approaches. First, I use the list of populated places, from the ESRI gazetteer, and assign all clusters that are located within 10 km of a populated place to the same community. For clusters outside this 10km radius, I combine clusters using a grid of roughly 10km x 10km squares. This procedure converts the 5,000 survey clusters into 2,800 arbitrarily defined communities. The 10km range is drawn directly from the previously cited research indicating that 10km is roughly the maximum efficient scale of provision for piped water.

Alternatively, I use a narrower geographic classification, where I simply group clusters according to an arbitrary grid with squares at 1km or 5km scales. The 1km grid creates a total of 4,300 'communities', thus creating a potentially important difference from the previous methodology. However, the variation is primarily found in the urban areas. Table 4.1 shows the number of artificial communities created under each method, and the urban/rural¹⁰ split.

(Insert table 4.1 here.)

As the results demonstrate, this differentiation is essentially meaningless when the sample is restricted to rural communities. When including urban communities, it probably makes sense to use the larger grid-size to accommodate the presence of cities that spread over

¹⁰Communities are classified as urban or rural based on the majority of residents in the community. Thus a community that contains 3 clusters, and a total of 150 urban individuals, and 200 rural individuals is classified as rural.

relatively large areas.

4.2 Public Goods

The two public goods that I focus on in this study are piped drinking water and electricity. A binary variable denoting access to electricity is available for all households in the Measure DHS surveys. Unfortunately it is not possible to assess quality of access, such as the fraction of the day or year during which electricity is available. If the quality of service is also affected by ethnic diversity, it is possible that the results here understate the true impact of heterogeneity.

For piped water, households are asked for the source of their drinking water and the time that it takes to get to their water source (if any). Possible answers include piped into the home/compound, a public tap, a well or borehole, or from a stream or lake. As piped drinking water systems generally offer significant economies of scale over local communities, it is my hypothesis that homogeneous communities will be more likely to use a piped water system. In keeping with the idea of community water systems, I code a household as accessing piped water if that water is piped into their home or compound, or if they access a public tap within 15 minutes of their home¹¹. As a secondary test, I consider the impact of ethnic diversity on access to piped drinking water that is available in the household or compound. Finally, I consider the effects of ethnic diversity on the availability of water to the general community. To do so, I construct a variable that measures, from the households that do not access water at home, the fraction of households that have access to a public tap.

There is no logical reason that the effects of ethnic diversity should fall equally on the very different goods that are provision of water at home and the provision of public taps. Generally, one might expect that there is less of a public good motivation involved in home delivery, and therefore ethnic diversity would have less of an effect in this case.

In keeping with the existing literature, I analyze the availability of each public good at the aggregate level, and therefore the ideal measure would be the existence and quality of the relevant infrastructure. As this information is unavailable for a sample as large as this, this information must be inferred from the choices of households. Here I make the assumption that the fraction of households that access piped drinking water, all else equal, reflects the quality and availability of the infrastructure. If in two otherwise equal communities, one has 20% of the population accessing piped drinking water, while in the other 80% do, then the interpretation of this analysis assumes that the second community has a more functional infrastructure.

¹¹The results are not sensitive to this distance.

4.3 Measuring Ethnic Diversity

The diversity measures contained in this study rely on the ethnicity data reported in the measure DHS surveys in each country. In the survey each respondent is asked to select their ethnicity from a list of potential ethnicities with 'Other' always as an included option. As shown in table 2 the number of ethnicities listed varies across countries from a low of 7 in Togo to a high of 66 in the Cote D'Ivoire. In contrast to the existing literature on ethnic identity and local public goods, this paper combines data from countries with widely varying ethnic structures. An example of this is the contrast between Burkina Faso where the Mossi are the dominant group in size and wealth as opposed to Niger where the Haoussa are the largest group but are relatively poor.

(Insert table 2 here.)

Choosing an ethnicity from a list may have a tendency to force individuals to make a choice that does not entirely fit their identity. To the extent that this biases the diversity results, and assuming that this error is unrelated to the provision of public goods, the measured effect of diversity will tend to be biased toward zero. This problem may have less significance in Africa as opposed to other regions due to the historical prevalence of ethnic identity as a tool of colonial rule. One form of evidence for this is the strong correlation between the results of the household ethnicity survey and recent alternative data sources of ethnic diversity in Africa. Not only do national measures of diversity created using this dataset agree closely with published data in other sources, such as Alesina et al. (2003) and Fearon (2003), table 3 uses Kenya as an example and provides striking evidence of the similarities between the individual components of the different measures¹². Where there is variation between the survey and published national level resources the survey tends to provide more options for survey respondents. This is important as local diversity may be significantly affected by the presence of groups that are locally concentrated and therefore have a relatively small presence nationally.

(Insert table 3 here.)

The independent variable of interest is a measure of ethnic diversity or homogeneity. The literature on the provision of public goods has focused primarily on a measure of fraction-

¹²Similar tables for other countries are available upon request. It seems likely that the consistency of how groups are defined across countries in Africa is the result of the colonial powers institutionalizing ethnicity to manage local areas. While there is evidence that this strategy was more commonly followed by the British it occurred to some extent in all colonial settings (Blanton, Mason, and Athow 2001).

alization¹³ that is drawn from the industrial organization literature. In keeping with this literature, I present the primary results here using fractionalization as the preferred measure of diversity. In Jackson (2008), I demonstrate that an alternative measure of ethnic diversity relevant for the provision of public goods is the size of the largest ethnic group. As such, the key results are repeated in the appendix using this as the measure of diversity. The results are not qualitatively different due to the very strong, negative, correlation between the two variables of < -0.95 , as measured at any geographic scale.

The issue of migration is critical to assessing the impact of diversity. To the extent that people move to a community in response to the provision of basic services, and migrants tend to come from a more diverse background than the local population, then this will tend to increase the diversity of communities with public goods. Ignoring this migration will therefore understate the true impact of ethnic diversity. For this reason, I could present results under three alternative measures. First, I use the diversity as measured over all current residents. Second, I measure the diversity using only residents that report living in their current location for either their entire life, or at least 20 years¹⁴. Third, I use the historical measure as an instrument for the current diversity. As is shown below, the results are not affected by this choice, primarily because the correlation between long-term diversity and current diversity is very high.

Theoretically, it is not clear which of the second two variables should be of greater importance. For example, as water systems may have existed for a significant time prior to the time of the survey, historical ethnic diversity may be precisely the relevant measure. Alternatively, recent construction or maintenance of the water systems could be the critical public goods problem, thus implying that current ethnic diversity is relevant, and that the correct solution is to instrument for ethnic diversity. As the results are not substantively affected¹⁵, I do not pursue this question further, though the persistent effects of poor governance are an

¹³Specifically, fractionalization is measured as

$$F = 1 - \sum_{e=1}^E p_e^2$$

where p_e is the population share of ethnic group, e . This measure is drawn from the Herfindahl index that is used in assessing industrial concentration, however in Jackson (2008) I also demonstrate that it is the optimal measure of diversity under specific conditions.

¹⁴This measure is therefore unaffected by in-migration in response to high levels of a public good. While this does not control for out-migration, possibly in response to low levels of a public good, it is unclear how out-migration would impact local diversity measures.

¹⁵For the key regressions in the paper, I include the instrumented results in the appendix. Similar results for any regression in the paper are available upon request.

interesting direction for future research.

4.4 Additional Explanatory Variables

Sub-national measures of ethnic diversity are strongly correlated with local wealth. Communities that are the center of trade, or home to significant resources, tend to draw migrants from all groups seeking wealth. It is therefore critical to control for the economic resources available to a community. In addition, diverse communities may be systematically different from homogeneous communities in other ways. For example, diverse communities may be the result of recent migration that is associated with ineffective local government. To account for these effects I control for economic factors such as the average household wealth, urbanization and population density, and average household tenure¹⁶.

The viability of piped drinking water or electricity are also directly related to the cost of installing infrastructure and the cost of alternatives. I therefore control for geographic factors including the distance to rivers, elevation and the proximity to the ocean. In particular, the distance to rivers will directly affect both the cost of installing piped drinking water and the cost of accessing alternative sources, such as water directly from a river¹⁷.

Specific regions may also have political advantages in accessing public funds. To control for this effect, I include a measure of the proximity to the national capital and the local share of the largest ethnic groups in the country. In addition, when considering geographic scales smaller than the province level, I control for provincial fixed effects, while in all regressions I control for national fixed effects.

Table 4 contains sample statistics for those measures that do not differ by the scale of aggregation. Notably, 27% of households have access piped drinking water, with over half of them having access being in their home or compound, while electricity is available in 20% of households.

(Insert table 4 here.)

Table 5 provides summary statistics on ethnic diversity as measured at local, district and provincial scales. Diversity is higher when measured over large regions, however there is significant ethnic diversity as measured at the local scale, with minority groups comprising 30% of the population.

(Insert table 5 here.)

¹⁶All but population density are drawn from the Measure DHS surveys. Population density is derived from the Gridded Population of the World, version 3.

¹⁷The data on the location of all geographic features are from ESRI data sources, and the distance to each cluster is based on the author's calculations.

The primary empirical specification throughout the paper is:

$$G = \beta_0 + \beta_1 \text{Homogeneity} + \beta_2 \text{Geog} + \beta_3 \text{Econ} + \beta_4 \text{Ethnic} + \beta_5 C + \beta_6 P \quad (4.1)$$

where C and P refer to country and provincial fixed effects respectively. At the local and district scales, there are sufficient observations to include the local share of the largest three ethnic groups in each country. At the provincial scale, in addition to removing provincial fixed effects, I only include a measure of the size of the largest ethnic group in each country.

G is the fraction of households in the community that have access to the specific public good in question, either water or electricity. The estimates reported throughout the paper are the result of OLS regressions. There are no substantive differences when using either a tobit or grouped logit specification and these results are available upon request.

5 Ethnic Diversity and the Provision of Piped Drinking Water

Access to clean drinking water is a fundamental need for a healthy population. From Jalan and Ravallion (2003)[p.153], "The World Health Organization estimates that four million children under the age of five die each year from diarrhea, mainly in developing countries. Unsafe drinking water is widely thought to be a major cause, and this has motivated public programs to expand piped water access." By 2006, 42% of the population in Sub-Saharan Africa still did not have access to an improved water source, while in rural areas, this increases to 54%. Expanding the network of piped drinking water is therefore a critical step in improving the quality of life in developing countries.

As discussed, the provision of clean drinking water is an inherently local problem. Water is a heavy commodity that is very costly to transport, thus the development of clean drinking water in one community provides very little benefit to neighbouring communities¹⁸. In this context, we should expect local diversity to affect provision. While piped water is not a classic public good, as there is no specific problem with restricting access to those that have paid for a service, the provision of piped water includes economies of scale that are much larger than, for example, providing wells, and therefore the efficient provision of piped water will require the resolution of public good issues. In addition, given the necessity of water, it may be morally difficult to limit access to those households that have sufficient ability to pay.

¹⁸It is possible that this benefit is provided in the sense that people are able to access public taps in neighbouring communities. For this reason, throughout the analysis that is run at the local level, I account for clustering of the standard errors by district.

A secondary issue related to piped water is the definition of household access. The DHS surveys record the primary source of drinking water for each household from a list that includes roughly 10 options. For my purposes, households are divided into three groups, those that access piped water in their home, those that access piped water at a public tap within 20 minutes of their home¹⁹, and those with any other source of drinking water.

The 'other' sources of drinking water are not, by definition, less suitable in all cases than piped water. Many communities, especially smaller ones, may efficiently use wells, boreholes, or local streams instead of piped water. The hypothesis underlying this analysis is simply that, all else equal, less diverse communities will find a piped water system viable in more circumstances than a more diverse community. Alternatively, less diverse communities will spend greater resources on the piped water system, thus encouraging a greater fraction of local residents to prefer piped water to alternative sources.

5.1 Piped Water and Local Measurement

Theory suggests that if ethnic diversity will affect the delivery of piped drinking water, it is likely to do so at the community level, as the efficient scale of provision of water is inherently local. As such, I start by estimating the impact of "local" diversity on the provision of piped drinking water.

As approximately one-third of the population that accesses piped water does so at a public tap, I measure the impact of ethnic diversity on three different dependent variables. The first two are simply the fraction of households that access piped drinking water at all, and the fraction that access piped water in their home or compound. Finally, I use a dependent variable that measures, out of the households that don't have access to piped water at home, what fraction access their piped water at a public tap.

(Insert table 6 here.)

The results of the primary local-level regressions are presented in table 6. Column (1) contains the primary specification, and the results demonstrate that households are significantly more likely to access piped water in homogeneous communities than in heterogeneous communities. In column (2), the dependent variable is the fraction of households that access piped water at home. The coefficient on ethnic homogeneity remains statistically significant. In both cases, the fraction of households that access piped water increases by approximately 30% as one moves from a very diverse community to a homogeneous community. The results presented in column (3) demonstrate that ethnic diversity affects the provision of water via

¹⁹Varying the cutoff length of time is inconsequential, the choice was made to be consistent with the notion of local provision.

public taps. The dependent variable in this regression is the ratio of households that access drinking water at a public tap to the number of households that do not access piped water at home. Controlling for the fraction of households that have access to water at home, homogeneous communities also increase the fraction of other households that have access to public taps by approximately 30%. The results in the first 3 columns thus suggest that the benefits to homogeneity are roughly equally split between increasing home access, and increasing public access for those households that don't receive drinking water at home.

In column (4), I restrict the sample to communities in the sample that are primarily rural, as this group is less likely to suffer from the endogeneity problems associated with migration. Further, rural areas in all countries in the sample do not have well-defined local governance structures and therefore most directly represent the notion that the scale of provision and the scale of governance are different. The results are essentially identical to the full sample, though the relative impact is far higher given that only 12% of all rural residents have access to piped drinking water.

The coefficients on the basic economic control variables in these regressions are consistent with existing theory. Wealthy, urban and long-standing communities are more likely to have piped water systems. In addition, and more surprisingly, communities that are further from the national capital are more likely to have a piped water system. One interpretation of this result would be that these communities have less of an expectation of help from the central government, and therefore build a piped water system without waiting for others. This phenomenon is consistent across specifications of the model, and may be an interesting direction for future research.

Households in communities close to the coast are more likely to access water in their home, yet less likely to access water overall, caused by a very low level of public taps. This may be related to variation in economic activity in coastal communities, but the issue is not explored further within the paper. The distance to a river is, occasionally, marginally statistically significant but not economically significant. As being closer to a river would reduce the costs of installing a piped water system, while also reducing the costs of the alternative source (water directly from the river), it seems that the increased costs of the alternative to a piped system are slightly more important than the reduced cost of implementing a piped system.

The results presented here are robust to a variety of alternative specifications. Measuring diversity by the size of the largest group, or using current diversity, with or without instrumental variables, has no effect on the results. Altering the definition of local, from 1km to 20km squares, is also inconsequential. Further, while there is the possibility that the wealth variable is affected by the presence of piped water, instrumenting for wealth does not have

a substantive effect on the results. The results from the key robustness checks are contained in the appendix.

5.2 Piped Water and District Measurement

The lowest level of administrative unit that is common across the countries included in the sample is what I have defined as the district. It is the 2nd administrative level in each of the 13 countries, and is drawn directly from the DHS database. The role of this governmental level varies significantly across countries, with some, such as Ghana, devolving significant responsibility, and others, such as Kenya giving little power to district councils.

(Insert table 7 here.)

Table 7 contains the results from the regressions conducted at the district level. The first three columns contain the estimates for piped water access, in general, at home, and at a public tap and district-level diversity is shown to have no significant effect on any of these outcomes. The correlation between local and district measures of diversity is 85%, and therefore the difference in the results could occur for a variety of reasons. In particular, in addition to the change in the measure of diversity, the observations in the previous results have been significantly aggregated. This aggregation may wipe out some of the variation in drinking water access that was at the heart of the previous result.

Column(4) demonstrates that the unit of observation is not important. In this specification, all variables included in column (1) are unchanged in column (4). However, the measure of diversity is changed from the district measure of fractionalization to a local measure of fractionalization, averaged across the district²⁰. The coefficient on this average measure of local diversity is therefore evidence that the significant change in column (1) is the measure of diversity, not the aggregation of observations.

The additional control variables do not show substantial differences from those reported in the previous section. As with all the results, these results are robust to the same variations reported earlier. In particular, the results are robust to limiting the sample to those districts that are primarily rural.

5.3 Piped Water and Provincial Measurement

For completeness and comparison with the results for electricity, table 8 contains the estimates using the province as the unit of measurement. The expectation is that provincial level

²⁰This measure is weighted by the sum of household weights in the DHS dataset, and can be taken as a population weighted measure of diversity.

diversity should have no impact on the provision of piped drinking water.

(Insert table 8 here.)

Diversity, measured at any of the three scales, has no effect on the extent of provision across provinces. Local fractionalization, having been shown to be important when aggregated to the district scale, may be expected to have a significant coefficient at the provincial scale as well. The insignificant result may be attributed to the significant level of aggregation involved in this specification. As before, the remaining control variables remain consistent, with wealth acting as a consistent determinant of piped water access.

Collectively, the results connecting ethnic diversity to the lack of piped drinking water support the hypothesis that local ethnic diversity has a significant impact on provision and that measurement at other scales acts to understate the true impact of diversity. This result suggests that local actions are critical to the provision of local goods, even where the legal responsibility for that provision rests at higher levels of government.

6 Ethnic Diversity and the Provision of Electricity

While clean drinking water is a fundamental human need, electricity is a key factor in enhancing productivity and promoting development. However, within Sub-Saharan Africa, access to a consistent source of electricity is relatively rare. In this sample, approximately 20% of households have access to electricity, and there is large regional variation with almost half the provinces having less than 10% of the population with access to electricity. The dependent variable throughout the analysis of electricity is the fraction of households that report having access to electricity. As mentioned, this dataset does not allow for the consideration of the quality of access, such as the reliability of electricity access.

In the context of this paper, electricity is a fundamentally different good than water. Expensive to generate, yet relatively easy to transmit, the economies of scale associated with electricity production are large. The minimum efficient scale for electricity production is apparently much larger than a community or a district, and we should therefore expect that, if ethnic diversity matters, it matters at the largest scale possible in this analysis, the province.

6.1 Electricity and Local Measurement

Given the economies of scale associated with the production of electricity, the hypothesis of this paper is that local diversity will not affect household access. In the first 3 columns of table 9, the results confirm this hypothesis. Local fractionalization does not impact household

access to electricity, in general, or specifically in rural or urban areas. Wealth, urbanization and the average tenure in the community all have logical and statistically significant impacts. In addition, as with piped drinking water, communities near the coast appear to be significantly different than other communities. For electricity, they are more likely than others to have electricity, though this seems to be only important for rural households. As with the variation in the case of water this is likely the result of differences in economic structure associated with coastal communities.

(Insert table 9 here.)

In column (4), I do not control for provincial fixed effects but instead include a measure of provincial diversity. The results indicate that communities in homogeneous provinces have a higher likelihood of electricity access than an otherwise identical community in a heterogeneous province. As with the results for piped water, this coincides with the hypothesis that regional diversity will affect electricity provision. This result is essentially unchanged if local diversity is controlled for in the same regression.

6.2 Electricity and District Measurement

Table 10 contains the results for regressions conducted with the district as the unit of observation. Columns (1) and (2) incorporate measures of local and district level homogeneity and show that diversity does not have a negative impact. In fact, measured at the district level, fractionalization is shown to have an unexpected statistically significant positive effect. I discuss this result further below.

In columns (3) and (4), I again remove the provincial controls, and measure the impact of ethnic diversity measured at the provincial scale. The results again indicate that ethnic diversity measured at the provincial scale has a negative effect on the provision of electricity. The results for the control variables are similar to the results obtained previously, with, in particular, wealth, tenure and proximity to the coast being important factors. Surprisingly, urbanization and population density are not consistently correlated with electrification. (Insert table 10 here.)

6.3 Electricity and Provincial Measurement

Table 11 contains the analysis conducted at the provincial level for electricity. The results indicate that ethnic homogeneity is associated with increased rates of electricity access. Moving from a very diverse province to an otherwise similar homogeneous province will reduce the fraction of households with access to electricity by approximately 40%. As shown in columns

(2) and (3), diversity measured at the district scale is also a significant factor. In column (4), we find that local homogeneity is not important in determining access to electricity.

Not surprisingly, wealth remains consistently important for the provision of electricity. The role of density is more difficult to assess, as the effects of urbanization and density have opposing signs. In addition, as for water, distance from the national capital appears to promote access to electricity. One explanation might be that communities or districts far from the centers of power tend to have a better developed system of governance that does not rely on the support of an unreliable central government. This suggests that further research into the economic effects of a weak, but legally centralized, central government might prove interesting. (Insert table 11 here.)

The results for electricity contained in tables 10 and 11 appear contradictory at first glance. In column (1) of table 10, district level fractionalization appears to have a positive effect on electrification, while in column (3) of table 11, the same variable appears to have a negative effect. However, these regressions are reporting the effects of diversity across very different samples. In the district specification, table 10, the variation being analyzed is between districts within the same province. The results therefore may indicate a secondary, positive, effect of district-level diversity that appears once the overall provincial diversity is controlled for.

In contrast, column (3) of table 11 considers the impact of average district diversity on the electricity access rates across provinces. In this regression, district-level diversity appears to have a negative impact. This reflects the negative, first-order effect of regional diversity, measured here at the district level.

7 Conclusion

Economic development depends on the effectiveness of governance structures, both formal and informal. This paper adds to the growing literature demonstrating that ethnic diversity has a negative effect on the provision of public goods. In addition, this paper demonstrates it is critical that measurement of diversity takes place at the correct scale. Unlike averaged measures such as wealth, the scale at which diversity is measured must be considered as a significant part of the measure and should be chosen after consideration of the underlying mechanism.

Regardless of the underlying mechanism, ethnic diversity at the scale of provision is a potentially limiting factor. The alternative scale that should be considered is the scale of governance. However, in many regimes in Africa, it is difficult to define the scale of

governance with sufficient accuracy to allow for accurate analysis. This is particularly true in the context of a multi-country study such as this one. It is therefore important that future research expands on these findings and considers the role of ethnic diversity at the scale of governance directly.

The two goods considered in the paper are water and electricity. In particular, the public good nature of water is inherently local, and therefore local diversity has an effect on the provision of piped water. For electricity, the minimum efficient scale is regional, and therefore ethnic diversity, measured regionally, has a negative impact on the provision of electricity. Further, if diversity is measured at the incorrect scale, this induces a form of measurement bias that is not fundamentally different, for example, than measuring income when wealth is the correct measure. While the two measures are likely correlated they are fundamentally different and the results are going to be strongly affected by the choice.

There are three general results that may be of interest. First, ethnic diversity has a statistically and economically significant effect on the provision of both piped water and electricity. Measured at the scale of provision, moving from a diverse community (province) to a homogeneous one increases access to piped water (electricity) by approximately 25-35%. Resolving the governance issues surrounding ethnic diversity would therefore have a significant impact on the development prospects in Sub-Saharan Africa. In addition, given the difficulty in assessing the key dimensions of diversity, the results reported here should be taken as a lower-bound on the impact of ethnic diversity.

Second, the consequences of measuring diversity at the wrong scale are significant. As demonstrated for water, the point estimate of the impact of district-level diversity roughly one-sixth of the impact of diversity measured at the community level. This strong downward bias associated with measurement error must be carefully considered in the context of ethnic diversity. The notion that aggregation errors may have significant consequences on economic measurement is not new, however it may be particularly important in the context of ethnic identity studies where the underlying mechanisms behind diversity effects are poorly understood.

Third, the measurement of diversity does not need to coincide with the unit of observation. Often, researchers will be limited in their access to all relevant data at a sufficiently local scale. As demonstrated for understanding the role of diversity in limiting access to piped water, it is sufficient to measure diversity at the local scale, and then aggregate that measure up to a larger scale where additional data may be available.

These results should not be taken as suggesting that the scale of governance is not important. There are two reasons for this caution. First, while most African countries are

heavily centralized, diversity at the national level could not be included within this study. Previous research has demonstrated that ethnic diversity, as measured at a national scale, is strongly linked to poor public good provision. In addition, when the official responsibility for governance falls to a disinterested central government, local governments may step in to ensure provision. In this sense, the official record likely overstates the degree of centralization in countries with ineffective national governments. As such, the distinction between scale of provision and scale of governance is difficult to determine.

References

- ALESINA, A., R. BAQIR, AND W. EASTERLY (1999): "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 114(4), 1243–1284.
- ALESINA, A., A. DEVLEESCHAUWER, W. EASTERLY, S. KURLAT, AND R. WACZIARG (2003): "Fractionalization," *Journal of Economic Growth*, 8, 155–194.
- BANERJEE, A., L. IYER, AND R. SOMANATHAN (2005): "History, Social Divisions, and Public Goods in Rural India," *Journal of the European Economic Association*, 3(2-3), 639–647.
- BANERJEE, A., AND R. SOMANATHAN (2007): "The Political Economy of Public Goods: Some Evidence from India," *Journal of Development Economics*, Forthcoming.
- BANERJEE, S., A. DIALLO, AND Q. WODON (2007): "Measuring trends in access to modern infrastructure in Sub-Saharan Africa: Results from Demographic and Health Surveys," Mimeo.
- BESLEY, T., AND S. COATE (2003): "Centralized versus decentralized provision of local public goods: a political economy approach," *Journal of Public Economics*, 87, 2611–2637.
- BLANTON, R., T. D. MASON, AND B. ATHOW (2001): "Colonial Style and Post-Colonial Ethnic Conflict in Africa," *Journal of Peace Research*, 38, 473–491.
- BROSIO, G. (2000): "Decentralization in Africa," Mimeo.
- CLARK, R. M., AND R. G. STEVIE (1981): "A Water Supply Cost Model Incorporating Spatial Variables," *Land Economics*, 57, 18–32.
- EASTERLY, W., AND R. LEVINE (1997): "Africa's Growth Tragedy: Policies and Ethnic Divisions," *The Quarterly Journal of Economics*, 112(4), 1203–1250.

- FASS, S. M., AND G. M. DESLOOVERE (2004): “Chad: Governance by the Grassroots,” in *Local Governance in Africa: The Challenges of Democratic Decentralization*, pp. 155–180. Lynne Rienner Publishers.
- FEARON, J. D. (2003): “Ethnic and Cultural Diversity by Country,” *Journal of Economic Growth*, 8, 195–222.
- FILIPPINI, M. (1998): “Are Municipal Electricity Distribution Utilities Natural Monopolies?,” *Annals of Public and Cooperative Economics*, 69, 157–174.
- JACKSON, K. (2008): “Why Does Diversity Matter? An Empirical Analysis of Water Provision in Africa?,” Mimeo.
- JALAN, J., AND M. RAVALLION (2003): “Does piped water reduce diarrhea for children in rural India?,” *Journal of Econometrics*, 112(1), 153–173.
- JARSSO, S. (2002): “Water Sector Development Programme,” .
- KAREKEZI, S., AND J. KIMANI (2002): “Status of power sector reform in Africa: impact on the poor,” *Energy Policy*, 30(11-12), 923–945.
- LUTTMER, E. (2001): “Group Loyalty and the Taste for Redistribution,” *Journal of Political Economy*, 109(3), 500–528.
- MANSURI, G., AND V. RAO (2004): “Community-based and-driven development: a critical review,” *The World Bank Research Observer*, 19(1), 1–39.
- MIGUEL, E., AND M. K. GUGERTY (2005): “Ethnic Diversity, Social Sanctions and Public Goods in Kenya,” *Journal of Public Economics*, 89, 2325–2368.
- OLOWU, D., AND J. WUNSCH (2004): *Local governance in Africa: The challenges of democratic decentralization*. Lynne Rienner Publishers.
- OUEDRAOGO, H. (2003): “Decentralisation and local governance: experiences from franco-phone West Africa,” *Public administration and development*, 23(1).
- PANIZZA, U. (1999): “On the determinants of fiscal centralization: Theory and Evidence,” *Journal of Public Economics*, 74, 97–139.
- SACHS, J. D. (2005): *The End of Poverty: Economic Possibilities For Our Time*. The Penguin Press.

- TEFERRA, M. (2002): “Power sector reforms in Ethiopia: options for promoting local investments in rural electrification,” *Energy Policy*, 30(11-12), 967–975.
- TIEBOUT, C. (1956): “A Pure Theory of Local Expenditures,” *The Journal of Political Economy*, 64(5), 416.
- TREISMAN, D. (2002): “Defining and measuring decentralization: A global perspective,” *unpublished paper. Available at <http://www.polisci.ucla.edu/faculty/treisman>.*

A Robustness Tests of Empirical Results

A.1 Piped Water

As mentioned previously, given the lack of direct data, there are a variety of methodologies for combining clusters into artificial local communities. Within the paper, I use a 10km distance from defined populated places, and a 10km x 10km grid for clusters further than 10km from a populated place. Especially for the rural communities on which the results related to piped water rely, 10km may be much further than is appropriate. As a robustness check, I present the results of the primary regressions using the 1km grid discussed in the paper, and an equivalent 5km grid. As can be clearly seen, the results are not substantively affected by the precise definition of “local”. (Insert table 12 here.)

Table 12 repeats the results of the first 3 columns of table 6, but restricts the analysis to narrowly defined rural communities²¹ As urban areas are characterized by a larger variety of economic activity and governance structures, which I am unable to properly control for, it is reasonable to assume that the effects of ethnic diversity would be most easily measured in rural communities. Further, the disconnect between scale of provision and de jure scale of governance is most apparent in rural communities. Perhaps not surprisingly, as rural communities make up approximately 80% of the total, the results are not substantively changed by this restriction.

Table 13 demonstrates that the choice of diversity measurement does not affect the measured relationship between diversity and the provision of piped drinking water. If anything, using the historical measure of diversity as a proxy for current diversity appears to result in an increase in the measured effects of ethnic diversity, as opposed to just using the historical

²¹The Measure DHS survey defines clusters as rural or urban. Where two or more clusters are collected into a single community, I code the community as rural if more than half the (weighted) households are characterized as rural.

measure. In addition, measuring diversity by fractionalization or the size of the largest ethnic group, as is done in columns (2) and (3), is irrelevant for the results. (Insert table 13 here.)

In each case, the first stage of the regression is highly significant, as the correlation between current and historical measures of diversity is over 90% in the sample.

A second potential source of endogeneity in this regression is the measurement of wealth, which may be affected by the presence of piped water for two reasons. First, the measure of wealth in the DHS survey is the result of a principal components analysis conducted on which objects are available to households. As some products are only valuable in the presence of piped water or electricity, this may naturally lead to a positive relationship between piped drinking water or electricity, and the measure of wealth. This form of simultaneity can be overcome by using other elements of household wealth as instruments. As such, I present results with the fraction of households choosing various materials for household flooring acting as an instrument for average wealth. The results presented here indicate that the type of household flooring is a very strong predictor of wealth within the sample.

A second source of endogeneity is the simple idea that access to piped water or electricity may make individuals more productive in their everyday life. This will also lead to a positive bias in the measured relationship between public good provision and community wealth. In this case, household assets will not serve as effective instruments for wealth, and therefore I use a variety of geographic controls. Specifically, I use the distance from large cities or towns, and the distance from the coast as instruments for community wealth. For this IV strategy to be acceptable, it must be that any direct effect of being close to a city or town is otherwise picked up in the existing specification. In this case, I assume that the measures of urbanization and population density assess the aspects of distance from cities that are relevant to the provision of public goods and that distance further impacts wealth by affecting trade. Since there is no strong reason for trade links to significantly affect public good provision other than through wealth effects, there is good reason for the instruments to be excluded from the regression.

While this set of instruments is not without problems, the results of a Hansen J-test indicate that the instruments are not incorrectly excluded from the base specification, and the first stage results suggest a strong relationship between the instruments and our measure of wealth. Further, it is not entirely clear how any bias related to the endogeneity of wealth would effect the key results relating ethnic diversity to public good provision. Given those caveats, the results of the IV-specifications are presented in table 14. (Insert table 14 here.)

The first two columns of table 14 present the results of specifications using the standard measure of access to piped water as the dependent variable, while the latter two columns

use access at one's home or compound as the dependent variable. The results indicate that the measurement of wealth does not appear to substantively affect the results presented in the main section of the paper, though, as would be predicted by theory, there is a slight reduction in the measured effect of wealth. The final two lines present the Cragg-Donald F-statistic related to weak identification and the Hansen-J p-statistic of overidentification, demonstrating that the instruments are both effective predictors of household wealth, and properly excluded from the original specification.

A.2 Electricity

The results for electricity provision are similarly robust to variations. Here I present two tables with alternative specifications for the effects of regional diversity on the provision of electricity. The first, table 15, repeats table 11 but replaces the fractionalization measure used previously with one that uses the population share of the largest group in the relevant province, district or community. The results suggest that the choice of diversity measure is not critical to the analysis of developed here. The results are also robust to using the current measure of diversity, while instrumenting using a historical measure. (Insert table 15 here.)

As with the specifications for water, there is the possibility that the wealth measure included in the regression is endogenous. Here I include the results using the same 2 sets of instruments that were used in the case of piped drinking water. While the results on either specification show no substantial change in terms of the measured effects of ethnic diversity, the geographic instruments are not particularly strong at the first stage. (Insert table 16 here.)

A Question of Scale - Countries in Dataset

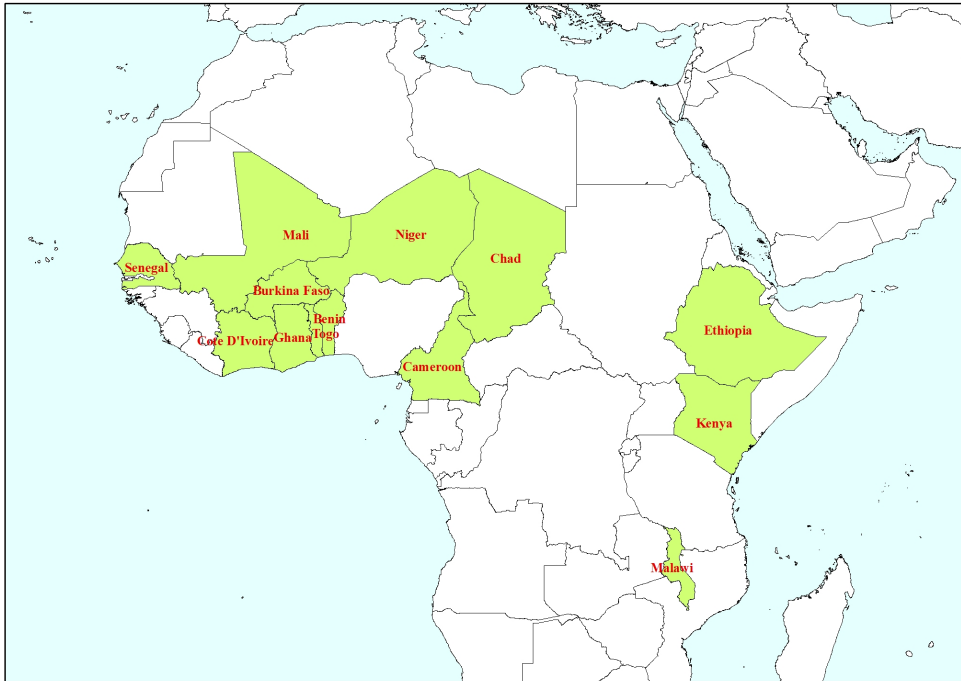


Figure 1: Question of Scale - Countries in Analysis

	1KM Grid	5KM Grid	10KM Grid, w/pop.places
Total Communities	4,321	3,579	2,866
Primarily Urban	1,375	749	444
Primarily Rural	2,946	2,830	2,422

Table 1: Local Communities

Country	# Ethnicities	Largest Ethnicity	Frac.Large Group	Wealth
Benin	10	Fon	45.5	Average
Burkina Faso	11	Mossi	53.6	Above Avg.
Cameroon	50	Bamileke	19.8	Above Avg.
Chad	14	Sara	26.4	Average
Cote D'Ivoire	66	???	15.6	Above Avg.
Ethiopia	56	Amharra	33.2	Above Avg.
Ghana	9	Akan	52.8	Above Avg.
Kenya	15	Kikuyu	23.3	Above Avg.
Malawi	10	Chewa	32.3	Below Avg.
Mali	10	Bambara	33.1	Below Avg.
Niger	10	Haoussa	60.6	Below Avg.
Senegal	12	Wolof/Lebou	36.5	Above Avg.
Togo	7	Adja/Ewe	47.2	Above Avg.

Table 2: Ethnic Structure

Table 3: Ethnicity in Kenya

Ethnic Group	% of Nat.Pop.(DHS)	Alesina et al. (2003)	Fearon(2003)
Kikuyu	23.4	22.0	28
Luhya	15.0	14.0	14
Luo	12.7	13.0	14
Kamba	10.8	11.0	11
Kalenjin	10.1	12.0	10
Meru	5.9	6.0	w/Kikuyu
Kisii	5.5	6.0	6.6
Mijikenda/Swahili	4.7	–	5
Somali	3.9	–	2
Masai	2.6	–	2
Embu	1.6	–	w/Kikuyu
Turkana	1.6	–	2
Taita/Taveta	1.2	–	
Kuria	0.6	–	–
Other	1.3	16.0	5

Table 4: Summary Statistics

Variable	Average	Stand. Dev.	Minimum	Maximum
Fraction w/ Piped Water	0.271	0.328	0	1
Fraction w/ Piped Water at Home	0.164	0.253	0	1
Fraction w/ Electricity	0.198	0.299	0	1
Distance to a River (km)	4.67	5.29	0.002	51.05
Elevation (m)	645.9	660.0	-1	3,690
Distance to Ocean (km)	501.9	352.5	0	1,698
Average Tenure	30.64	9.23	2.9	50*

Notes: *Tenure (years in current location) is top-coded at 50 years. Standard deviations are calculated at the local (or 10km) scale. When calculated at larger scales, standard deviations are somewhat smaller.

Table 5: Diversity Statistics for Communities, Districts and Provinces

Variable	Communities	Districts	Provinces
Share of Largest Group	0.702 (0.244)	0.660 (0.237)	0.541 (0.216)
Share of 2nd Largest Group	0.142 (0.112)	0.156 (0.107)	0.174 (0.080)
Ethnic Fractionalization	0.402 (0.285)	0.459 (0.267)	0.604 (0.224)
Number of Districts	2,866	1,208	114

Notes: Standard deviations in parentheses.

Table 6: Piped Water and Ethnic Diversity - Local Measurement

	Local	Home	Public	Rural
	(1)	(2)	(3)	(4)
Local Fractionalization	-.073*** (.020)	-.041*** (.014)	-.042** (.019)	-.063*** (.019)
Fraction w/ Piped Water at Home			-.061 (.040)	
Avg. Wealth	.333*** (.015)	.235*** (.012)	.197*** (.017)	.325*** (.020)
Urban	.094*** (.021)	.028* (.016)	.057*** (.021)	.178*** (.057)
Pop.Density	.008 (.006)	.017* (.010)	-.008 (.011)	-.025 (.059)
(ln)Distance to River	.006* (.003)	.003 (.002)	.005 (.003)	.008** (.003)
Elevation(km)	.001 (.014)	.010 (.011)	-.004 (.016)	.009 (.014)
(ln)Distance to Capital City	.037*** (.010)	.016* (.009)	.029*** (.010)	.033** (.013)
Coastal Community	-.058* (.032)	.043* (.026)	-.157*** (.032)	-.059 (.044)
Avg. Tenure	.001** (.0006)	.001*** (.0005)	.0003 (.0006)	.0009 (.0006)
Obs.	2828	2828	2810	2399
R^2	.753	.776	.618	.508

Notes: Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***, **, and * respectively. All regressions include country, province and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 7: Piped Water and Ethnic Diversity - District Measurement

	Base	Home	Public	Local
	(1)	(2)	(3)	(4)
District Fractionalization	-.009 (.025)	-.009 (.017)	-.013 (.023)	
Average Local Fractionalization				-.056** (.022)
Fraction w/ Piped Water at Home			-.114* (.059)	
Avg. Wealth	.289*** (.018)	.222*** (.014)	.179*** (.025)	.228*** (.014)
Urban	.148*** (.029)	.035* (.020)	.130*** (.032)	.040** (.020)
Pop.Density	.010** (.005)	.011 (.009)	.003 (.008)	.012 (.009)
(ln)Distance to River	.007 (.006)	.004 (.004)	.007 (.005)	.004 (.004)
Elevation(km)	-.002 (.021)	.014 (.017)	.0009 (.024)	.012 (.017)
(ln)Distance to Capital City	.036*** (.012)	.006 (.010)	.037*** (.012)	.007 (.010)
Coastal Community	-.075* (.040)	.076*** (.027)	-.207*** (.041)	.078*** (.027)
Avg. Tenure	.001 (.0009)	.001* (.0006)	.0004 (.001)	.001** (.0006)
Obs.	1196	1196	1189	1201
R^2	.842	.86	.728	.861

Notes: The dependent variable in columns (1) and (4) is the fraction of households that get drinking water from a piped source. The dependent variable in column (2) is the fraction of households that have access to a piped source of water in their home or compound, and the dependent variable in column (3) is the fraction of households that receive water at a public tap, of those households that don't access water at home. Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***, **, and * respectively. All regressions include country, province and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 8: Piped Water and Ethnic Diversity - Provincial Measurement

	Province	Home	District	Local
	(1)	(2)	(3)	(4)
Provincial Fractionalization	-.034 (.030)	-.003 (.023)		
Avg. District Fractionalization			-.003 (.032)	
Avg. Local Fractionalization				-.032 (.080)
Avg. Wealth	.351*** (.054)	.136*** (.044)	.347*** (.055)	.353*** (.054)
Urban	-.002 (.111)	.171* (.100)	.001 (.110)	.003 (.113)
Pop.Density	.028*** (.007)	.027*** (.007)	.028*** (.007)	.028*** (.007)
(ln)Distance to River	.058*** (.017)	.032** (.014)	.056*** (.017)	.057*** (.017)
Elevation(km)	-.013 (.033)	.021 (.017)	-.013 (.033)	-.016 (.032)
(ln)Distance to Capital City	.034*** (.011)	.008 (.012)	.034*** (.012)	.034*** (.012)
Coastal Community	-.103*** (.038)	.066 (.046)	-.100*** (.038)	-.104*** (.040)
Avg. Tenure	.002 (.002)	.002 (.001)	.002 (.002)	.002 (.002)
Obs.	114	114	114	114
R^2	.937	.933	.937	.937

Notes: The dependent variable in columns (1),(3) and (4) is the fraction of households that get drinking water from a piped source. The dependent variable in column (2) is the fraction of households that have access to a piped source of water in their home or compound. Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***,**, and * respectively. All regressions include country and a control for the size of the largest national ethnic group and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 9: Electricity and Ethnic Diversity - Local Measurement

	Local	Rural	Urban	Province
	(1)	(2)	(3)	(4)
Local Fractionalization	-.008 (.013)	-.011 (.012)	-.017 (.041)	
Provincial Fractionalization				-.037* (.019)
Avg. Wealth	.283*** (.014)	.263*** (.016)	.374*** (.021)	.260*** (.024)
Urban	.074*** (.016)	.048 (.030)	.053 (.047)	.076*** (.023)
Pop.Density	-.008* (.004)	.013 (.033)	-.015*** (.004)	.015*** (.006)
(ln)Distance to River	.001 (.002)	.001 (.002)	-.008 (.007)	-.002 (.002)
Elevation(km)	-.039*** (.013)	-.013* (.008)	-.063 (.047)	-.008 (.011)
(ln)Distance to Capital City	-.014 (.014)	-.010 (.008)	-.036** (.016)	-.004 (.010)
Coastal Community	.066*** (.017)	.047** (.021)	-.006 (.036)	.018 (.021)
Avg. Tenure	.002*** (.0004)	.0006* (.0003)	.004*** (.001)	.001** (.0005)
Obs.	2828	2399	420	2842
R^2	.869	.665	.915	.843

Notes: The dependent variable is the fraction of households that have access to electricity. Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district or province level) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***, **, and * respectively. All regressions include country and ethnic group controls, all but column (4) includes provincial controls, and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 10: Electricity and Ethnic Diversity - District Measurement

	District	Local	DistrictProvince	Province
	(1)	(2)	(3)	(4)
District Fractionalization	.039*** (.014)		.023 (.018)	
Average Local Fractionalization		.011 (.017)		
Provincial Fractionalization			-.073** (.030)	-.059* (.030)
Avg. Wealth	.272*** (.012)	.272*** (.011)	.276*** (.025)	.277*** (.025)
Urban	.022 (.019)	.027 (.019)	-.002 (.030)	.003 (.031)
Pop.Density	-.003 (.003)	-.002 (.003)	.019*** (.004)	.020*** (.004)
(ln)Distance to River	.007** (.003)	.007** (.003)	.003 (.005)	.003 (.005)
Elevation(km)	-.022 (.016)	-.022 (.016)	-.005 (.017)	-.007 (.017)
(ln)Distance to Capital City	-.013* (.007)	-.011 (.007)	.0009 (.010)	.001 (.010)
Coastal Community	.092*** (.014)	.091*** (.014)	.032* (.019)	.032* (.019)
Avg. Tenure	.003*** (.0005)	.003*** (.0005)	.002** (.0008)	.002** (.0008)
Obs.	1196	1201	1202	1208
R^2	.926	.925	.904	.903

Notes: The dependent variable is the fraction of households that have access to electricity. Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***, **, and * respectively. All regressions include country and ethnic group controls, columns (1) and (2) also include provincial controls and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 11: Electricity and Ethnic Diversity - Provincial Measurement

	Province	ProvDistrict	District	ProvLocal
	(1)	(2)	(3)	(4)
Provincial Fractionalization	-.084*** (.027)	-.058** (.029)		-.109*** (.032)
Avg. District Fractionalization		-.051* (.027)	-.081*** (.026)	
Avg. Local Fractionalization				.079 (.082)
Avg. Wealth	.362*** (.050)	.362*** (.049)	.356*** (.050)	.349*** (.052)
Urban	-.169* (.089)	-.144 (.089)	-.126 (.089)	-.177** (.090)
Pop.Density	.033*** (.005)	.034*** (.005)	.034*** (.005)	.032*** (.005)
(ln)Distance to River	.016 (.015)	.017 (.015)	.016 (.016)	.016 (.015)
Elevation(km)	-.009 (.019)	-.009 (.019)	-.010 (.019)	-.002 (.021)
(ln)Distance to Capital City	.020* (.012)	.022* (.012)	.023* (.012)	.020* (.011)
Coastal Community	-.021 (.030)	-.021 (.030)	-.017 (.029)	-.014 (.032)
Avg. Tenure	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)
Obs.	114	114	114	114
R^2	.966	.967	.966	.967

Notes: The dependent variable is the fraction of households that have access to electricity. Heteroskedasticity-robust standard errors are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***, **, and * respectively. All regressions include country and controls for the size of the largest national ethnic group and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 12: Piped Water and Local Diversity - 1KM Measurement, Rural Communities

	Local	Home	Public
	(1)	(2)	(3)
Local Fractionalization	-.054*** (.018)	-.024** (.011)	-.031** (.015)
Fraction w/ Piped Water at Home			.118*** (.044)
Avg. Wealth	.341*** (.020)	.202*** (.018)	.181*** (.023)
Urban	.030 (.093)	.144* (.080)	-.195*** (.040)
Pop.Density	-.049 (.032)	.026 (.025)	-.098*** (.032)
(ln)Distance to River	.007** (.003)	.003 (.002)	.004* (.002)
Elevation(km)	.012 (.016)	.017 (.012)	-.002 (.013)
(ln)Distance to Capital City	.027** (.013)	.017* (.009)	.014 (.009)
Coastal Community	-.060 (.042)	.029 (.028)	-.100*** (.037)
Avg. Tenure	.0008 (.0006)	.001** (.0004)	-.0004 (.0005)
Obs.	2903	2903	2887
R^2	.422	.45	.347

Notes: Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***, **, and * respectively. All regressions include country, province and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 13: Piped Water and Ethnic Diversity - Local Measurement - Diversity Measurement

	Instrument	Largest	IVLargest
	(1)	(2)	(3)
Local Homogeneity		.102*** (.024)	.119*** (.028)
Local Fractionalization	-.088*** (.024)		
Avg. Wealth	.337*** (.015)	.334*** (.015)	.338*** (.015)
Urban	.094*** (.021)	.096*** (.021)	.096*** (.021)
Pop.Density	.008 (.005)	.008 (.006)	.008 (.006)
(ln)Distance to River	.006* (.003)	.006* (.003)	.006* (.003)
Elevation(km)	-.0006 (.014)	.002 (.014)	.0005 (.014)
(ln)Distance to Capital City	.037*** (.010)	.037*** (.010)	.037*** (.010)
Coastal Community	-.059* (.032)	-.058* (.032)	-.058* (.032)
Avg. Tenure	.001* (.0006)	.001** (.0006)	.001* (.0006)
Obs.	2828	2828	2828
R^2	.753	.754	.754

Notes: Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***, **, and * respectively. All regressions include country, province and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 14: Piped Water and Ethnic Diversity - Local Measurement - Wealth Instruments

	AllInstrument	GeogInstrument	HomeAll	HomeGeog
	(1)	(2)	(3)	(4)
Local Fractionalization	-.061*** (.020)	-.054** (.023)	-.036** (.014)	-.059*** (.019)
Avg. Wealth	.285*** (.021)	.259*** (.061)	.212*** (.019)	.304*** (.059)
Urban	.141*** (.025)	.166*** (.060)	.050** (.021)	-.038 (.054)
Pop.Density	.012** (.006)	.014* (.007)	.019* (.010)	.012 (.011)
(ln)Distance to River	.006* (.003)	.006* (.003)	.003 (.002)	.003 (.002)
Elevation(km)	.006 (.014)	.008 (.015)	.013 (.011)	.004 (.013)
(ln)Distance to Capital City	.033*** (.010)	.030** (.012)	.014 (.009)	.023** (.011)
Coastal Community	-.049 (.032)	-.044 (.034)	.048* (.027)	.030 (.028)
Avg. Tenure	.0008 (.0007)	.0004 (.001)	.001** (.0005)	.002** (.001)
Obs.	2828	2828	2828	2828
R^2	.752	.749	.775	.77
e(jp)	.899	.663	.155	.422
e(cdf)	392.202	55.505	392.202	55.505

Notes: Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***, **, and * respectively. All regressions include country, province and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 15: Electricity and Provincial Diversity - Alternative Diversity Measure

	Province	ProvDistrict	District	ProvLocal
	(1)	(2)	(3)	(4)
Provincial Homogeneity	.104*** (.028)	.071** (.028)		.118*** (.030)
Avg. District Homogeneity		.076** (.033)	.120*** (.033)	
Avg. Local Homogeneity				-.051 (.085)
Avg. Wealth	.356*** (.047)	.355*** (.046)	.352*** (.048)	.349*** (.050)
Urban	-.159* (.082)	-.126 (.078)	-.110 (.082)	-.163** (.082)
Pop.Density	.033*** (.005)	.034*** (.005)	.034*** (.005)	.033*** (.005)
(ln)Distance to River	.018 (.014)	.019 (.015)	.018 (.015)	.018 (.015)
Elevation(km)	-.008 (.019)	-.008 (.019)	-.009 (.019)	-.005 (.021)
(ln)Distance to Capital City	.018 (.011)	.020* (.011)	.022* (.012)	.019* (.011)
Coastal Community	-.023 (.030)	-.023 (.030)	-.018 (.029)	-.020 (.032)
Avg. Tenure	-.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)
Obs.	114	114	114	114
R^2	.968	.969	.967	.968

Notes: Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level where applicable) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***, **, and * respectively. All regressions include country, province and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 16: Electricity and Provincial Diversity - Wealth Instruments

	AllInst1	AllInst2	GeogInst1	GeogInst2
	(1)	(2)	(3)	(4)
Provincial Fractionalization	-.090*** (.027)	-.064** (.030)	-.087*** (.030)	-.062* (.033)
Avg. District Fractionalization		-.050* (.030)		-.050* (.029)
Avg. Wealth	.435*** (.077)	.434*** (.078)	.405* (.210)	.414** (.209)
Urban	-.299** (.130)	-.273** (.133)	-.246 (.356)	-.238 (.359)
Pop.Density	.032*** (.006)	.032*** (.006)	.032*** (.007)	.033*** (.007)
(ln)Distance to River	.018 (.014)	.019 (.014)	.018 (.014)	.019 (.014)
Elevation(km)	-.009 (.019)	-.009 (.019)	-.009 (.019)	-.009 (.019)
(ln)Distance to Capital City	.027** (.013)	.029** (.014)	.024 (.025)	.027 (.025)
Coastal Community	-.025 (.032)	-.025 (.031)	-.023 (.033)	-.024 (.032)
Avg. Tenure	-.0009 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)
Obs.	114	114	114	114
R^2	.965	.966	.966	.966
e(cdf)	18.94	18.687	4.534	4.51
e(jp)	.828	.827	.411	.414

Notes: Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level where applicable) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***, **, and * respectively. All regressions include country, province and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.