Economic Research Paper: 2009-11

WHY DOES DIVERSITY MATTER? - AN EMPIRICAL ANALYSIS OF PIPED WATER PROVISION IN SUB-SAHARAN AFRICA

KEN JACKSON*

OCTOBER 2009

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

Why Does Diversity Matter? - An Empirical Analysis of Piped Water Provision in Sub-Saharan Africa *

Ken Jackson†

October 20, 2009

Abstract

Ethnic diversity has been shown to have significant, negative effects on the provision of basic public goods, in both developed and developing countries. However, the mechanism underlying this relationship is not fully understood. Two basic theories are drawn from the literature and incorporated within a single model, allowing for the derivation of key differences in their empirical predictions. The critical difference between models of differential demand and those of collective action problems lies in the distribution of public good provision across households. Using the DHS survey from 15 countries of Sub-Saharan Africa, covering over 100,000 households, I am able to test for the presence of aggregate effects of ethnic diversity and the distributional consequences. The results suggest that local ethnic diversity plays a critical role in limiting the provision of piped water in Sub-Saharan Africa, and that the mechanism behind this effect is ineffective local governance.

*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

A reliable supply of clean drinking water is a fundamental human need. The Millennium Development Goals recognize this need, and strive for halving the fraction of households that do not have access to an improved water source. Overall, the world is on pace to achieve this goal, with the fraction of households with an improved water source increasing from 76% to 86% between 1990 and 2007, and only needing to increase access to 88% by 2015. However, this masks critical variation around the world. In particular, the fraction of households in Sub-Saharan Africa only increased from 49% to 58% over the same period of time, indicating that it is unlikely that SSA will achieve this goal on a regional basis.

Why has the provision of a basic public service such as clean drinking water proven to be so difficult in Africa? One key factor may be ethnic diversity, which is very high in much of Sub-Saharan Africa. An extensive literature has developed over the past decade that focuses on the role of ethnic diversity in limiting effective governance. While this literature started at the national level, there has also been extensive research demonstrating the difficulty of providing public goods to diverse groups at sub-national scales.

This research raises an important question. If ethnic diversity is a problem, what is the solution? Not surprisingly, the answer will depend on the mechanism that causes public good provision to decline with diversity. If diversity has an effect through the demand for public goods, then the lack of public goods in diverse communities is not inefficient and therefore does not need nor permit a government solution. In contrast, if the difference is caused by the inability of diverse communities to coordinate the actions of their citizens, then changes in the governance structure responsible for providing piped water, particularly in diverse communities, may be warranted.

The literature on ethnicity and public good provision is primarily concerned with two main channels through which diversity would matter. First, in Alesina et al. (1999), different ethnic groups demand different goods, such as two groups that want their children educated in different languages. If education services involve significant economies of scale, the cost of provision in a diverse community will be higher than in a homogeneous one, generally resulting in lower rates of provision. This approach is also the basis of models in Alesina and La Ferrara (2000, 2005) along with Kimenyi (2006). With regard to a relatively neutral good such as water, segregated groups may disagree about the location of the main water supply to which households connect, or the location of a public tap. Or, as in Alesina and La Ferrara (2000), individuals may prefer to use a source that is not used by members of a different ethnic group. In any case, the underlying mechanism is that diversity affects the

aggregate demand for public goods.

In contrast to these demand-side effects, ethnic diversity could affect the ability of a group to act collectively. In Vigdor (2004) and Miguel and Gugerty (2005), collective action within an ethnic group is more efficient than collective action between groups. As such, individuals in diverse communities are less willing to contribute to the public good, thus reducing its level in equilibrium. While these two papers consider different specific mechanisms for this supply effect[1] the overall impacts are indistinguishable without knowledge of the motivation behind contributions to the collective good.

While the effective differences between supply and demand effects may be difficult to assess empirically, it is important to distinguish the relative importance of each mechanism. There is little that governments can, or possibly should, do about variation in public goods that is associated with ethnic diversity through demand effects. Improving efficiency in the face of a demand effect requires ethnic sorting, which is both morally difficult, and, if sufficiently important, will occur naturally. In contrast, if the effects of diversity demonstrate a vulnerability to collective action problems, then there may exist institutional changes that can improve the efficiency of public good provision.

This paper first develops a theoretical framework within which to consider the question of the relative importance of these two effects. This model incorporates the primary features of each of the demand or supply effect mechanisms within a single model in order to determine the testable implications of each effect. While the two mechanisms have similar aggregate effects, they differ sharply on the distribution of benefits from the public good. If demand effect models reflect reality, then the negative impact of ethnic diversity should fall disproportionately on the minority groups in the community. In contrast, the supply effect mechanism does not provide higher levels of the public good to any group.

This key difference provides the theoretical support for the empirical section that follows. Using household surveys conducted by the Demographic and Health Survey covering over 100,000 households across 15 countries of Sub-Saharan Africa, I demonstrate that ethnic diversity negatively impacts the supply of piped drinking water, yet is not provided preferentially to the locally dominant groups. This suggests that the effects of ethnic diversity are felt through their impact on the ability of diverse communities to resolve the collective action problem associated with the provision of public goods. The results indicate that the chance that a household in a very diverse community accesses piped water is reduced by approximately 30% as compared to a household in an otherwise identical homogeneous community,

[1] In Vigdor(2004), the effect is created by intra-ethnic altruism, while in Miguel and Gugerty (2005) it is the result of intra-ethnic sanctions.

regardless of the household's minority or majority status.

The existing literature on ethnic diversity and the provision of public goods has developed quickly over the past decade. Starting with Mauro (1995) and Easterly and Levine (1997), a large literature developed analyzing the effect of ethnic diversity on national outcomes. It found significant negative effects of ethnic diversity on economic growth, with the primary channel being the effect on government policy[2]. With regards to institutional outcomes such as property rights and democracy, Easterly (2001) and Collier (1999, 2000) demonstrated that effective institutions can eliminate the negative effects associated with ethnic diversity at the national level.

Focusing directly on the provision of public goods, La Porta et al. (1999) provided evidence that ethnic diversity negatively affects literacy and public health, while having less robust effects on educational attainment and infrastructure quality. Kuijs (2000) focus directly on the difference between measuring the effects on input measures as opposed to outcomes. His analysis finds that outcome variables in education and health are affected by ethnic diversity but that this is not entirely through reduced spending. For education he finds no significant effect of ethnic diversity on education spending[3].

At the local level, Alesina et al.(1999) provided the first extensive analysis of the effects of local diversity on the provision of public goods. The empirical results considered U.S. city and county expenditures on a variety of public goods and found significant variation in the effects of ethnic diversity on different goods. As an example, ethnic diversity was found to lower expenditures on roads and health-care, increase expenditures on policing and have no effect on fire services. In western Kenya, Miguel and Gugerty (2005) find significant negative effects of ethnic diversity on the provision of education and well water. In further research from developing countries, Bardhan and Dayton-Johnson (2002) summarizes the results of a series of papers on the impacts of diversity for agriculture projects and finds that most (but not all) report negative effects of ethnic diversity on local cooperation. Khawaja (2009) further documents significant negative effects of ethnic diversity on maintenance of public projects in Pakistan.

Caselli and Coleman (2006) provide a theoretical justification for focusing on forms of social heterogeneity that are relatively expensive to change such as ethnicity. However the model considered here does not exclude the possibility that groups may form around alter-

[2]Alesina et al (2003), Fearon (2003) and Posner (2004) refined the measurement of ethnic diversity - collectively demonstrating that the result is robust to variation in the way ethnicity is measured

[3]In the model developed below this outcome would be expected if inter-ethnic spillovers are relatively low in the case of education and relatively high in the case of health measures. I discuss this further at the end of section 2.

native social constructs such as religion. In addition to the results related to ethnic diversity above, Alesina et al. (2003) further estimates the impact of religious diversity on economic growth and finds no significant impact. In the empirical section I also consider the possible role for religious diversity and similarly find no evidence that indicates heterogeneity along this dimension is important to the provision of public goods in Sub-Saharan Africa.[4]

2 Model

The existing literature on ethnic diversity has considered a variety of mechanisms for how ethnic diversity could affect the provision of public goods. However these tend to fall into two categories. First, different ethnic groups could want different public goods. In the model of Alesina et al. (1999), this is the relevant source of conflict associated with diversity[5]. The provision of public goods in diverse communities involves lower average satisfaction for any level of funding and therefore the optimal level of funding is lower.

Alternatively, the papers of Vigdor (2004) and Miguel and Gugerty (2005) develop models in which the inter-ethnic free-riding problem is highlighted. In each case the intra-ethnic free-riding problem is at least partially mitigated by the presence of an additional feature that operates within ethnicities. In the case of Vigdor (2004) this is altruism whereas in Miguel and Gugerty (2005) it is an intra-ethnic sanctioning technology. This paper is not able to differentiate between these mechanisms though evidence provided in Humphreys et al. (2005) indicates that the presence of intra-ethnic sanctioning may play a vital role in this process. In either case, ethnic diversity is thus associated with lower provision of public goods because of the differential effectiveness of intra-ethnic institutions as opposed to inter-ethnic ones.

The model developed here embeds both mechanisms within a single reduced form model to investigate the effects of ethnic diversity at both the aggregate and individual levels. The model generates predictions related to aggregate provision of public goods and the distribution of public goods between majority and minority groups. In addition, the model provides direction as to how diversity affects the total level of spending on public goods and the efficiency of that spending.

[4] As the empirical section only considers countries in Sub-Saharan Africa it is not possible to conclude anything about the potential importance of religion in other areas of the world - or in countries of Africa not contained in the study. It seems likely that the relevant form of social heterogeneity will vary with local conditions.

[5] A similar model is developed in Alesina and La Ferrara (2000) where individuals from an ethnic group receive lower utility from the public good if the good is also accessed by members of a different ethnic group. Kimenyi (Kimenyi 2006) develops a model along similar lines with a specific focus on Africa.

2.1 Environment

A community provides a single public good funded by the contributions of local households.[6] The community has a population of mass N, divided among a total of E ethnic groups with p_e defining the population share of ethnic group e. Households receive utility from private consumption and the public good:

$$u_i = W_i - x_i + \mu \ln X_i \tag{2.1}$$

where W_i is household wealth[7], x_i is the household's contribution to the public good and X_i is the benefit the individual receives from the public good. The parameter μ captures demand and supply considerations including the local preference for the public good and the cost of installation and maintenance which could vary across communities[8] As a simplification[9], it is assumed that ethnic groups are able to fully overcome the internal free-riding problem and thus all individuals will make contributions to maximize group utility. Individuals belonging to the same ethnic group face identical decisions and therefore I focus on solutions that are symmetric within groups[10]. Therefore, define $x_i = x_e$.

The benefit received from the public good depends on the contributions of all other individuals but the contributions from one's own ethnic group may potentially have a greater effect.

$$X_e = p_e N x_e + \alpha \sum_{e' \neq e} p_{e'} N x_{e'} \tag{2.2}$$

The parameter $\alpha \in [0, 1]$ captures the extent to which the preferences for a public good vary across ethnicities with $\alpha = 1$ indicating no variation in preferences. This assumption is roughly consistent with the voting model developed in Alesina et al (1999) where the final allocation of funding on the public good is determined by majority vote. This parameterization provides a simplified reduced form while not detracting from the overall results.

[6] Alternatively funding from higher level governments could be available but responsive to local lobbying efforts. For the results below it would be sufficient that both the size and form of the public good provision are sensitive to local contributions.

[7] While households could potentially have different levels of wealth, the possible effects of wealth heterogeneity are excluded from the model through the use of quasi-linear preferences.

[8] In the model it is assumed to be constant for all individuals in the community, though the variation in demand across ethnic groups is considered below and controlled for directly in the empirical estimates.

[9] This is consistent within the model in MG and not relevant to the model of ABE in which institutions are essentially irrelevant. Alternatively, it could be assumed that the institution governing intra-ethnic coordination is less than perfect. To the extent that the ability of groups to overcome the within-group public good problem is not affected by group size the solutions provided here would not be changed in any qualitative way.

[10] Given quasi-linear utility there is no reason not to focus on symmetric solutions.

In the absence of institutions, each ethnic group chooses their contribution level to maximize the aggregate welfare of the group. The role of inter-ethnic institutions is to modify the incentives facing each ethnic group toward the investments that would maximize the aggregate social welfare of the community. The parameter $\tau \in [0,1]$ represents the effectiveness of local institutions in managing inter-ethnic coordination - with $\tau = 1$ the equilibrium is identical to that that would be chosen by a social planner. Each group takes the contributions of other groups as given and chooses $x_e \geq 0$ to solve:

$$U_e = \max_{x_e} \left[(1-\tau) p_e N (W - x_e + \mu \ln X_e) + \tau \sum_{j=1}^{E} p_j N (W - x_j + \mu \ln X_j) \right] \tag{2.3}$$

2.2 Extreme Cases

Before proceeding to the main propositions I present solutions to the extreme cases as these effectively demonstrate the different effects of the two primary mechanisms.

2.2.1 Divergent Preferences

First, I consider goods for which different ethnic groups have completely different preferences. The simplest example may be providing education where each group wants their children taught in a different language - with no room for compromise. One could define these as "ethnic" goods. In the event of completely divergent preferences ($\alpha = 0$) there is no effect of the institution, τ. Each group invests on their own behalf and each group invests until:

$$X_e = \mu p_e N \tag{2.4}$$

The average benefit is therefore:

$$\overline{X} = \mu (1 - F) N \tag{2.5}$$

where F is local ethnic fractionalization. This result is efficient regardless of institutional quality and in this case there is no practical policy implication in response to the 'inefficiency' caused by ethnic diversity. The reason for this is that contributions by one group have no impact on the welfare of any other group and hence inter-ethnic coordination has no effect.

Further, total spending is unrelated to the level of diversity[11]. The consequences of ethnic diversity for the provision of "ethnic" goods is entirely through the inefficiency of public spending and therefore it is critical that measurement of these effects occurs using outcome measures.

[11] And with alternative specifications of the concavity of preferences with regard to the public good it is not possible to even sign the effect that diversity will have on public good expenditures.

2.2.2 Complete Spillovers

The opposite of "ethnic" goods might be termed "community" goods, as all individuals benefit equally from their provision. A good that may be considered to be of this kind may be sanitation services in that all individuals benefit if others properly dispose of sewage[12]. In the event of complete spillovers ($\alpha = 1$) the results depend critically on the quality of inter-ethnic institutions. If inter-ethnic institutions are non-existent ($\tau = 0$) only the largest group will contribute and all other groups will choose to not contribute further. The intuition for this result is that the largest group is willing to contribute until the public good reaches a level higher than any other group's optimal level. In doing so they eliminate the incentive for any other group to contribute. The benefits to all groups are the same and are determined by the size of the largest group.

$$X_e = \overline{X} = \mu p_1 N \tag{2.6}$$

However if inter-ethnic institutions are complete $\tau = 1$ then all individuals contribute[13] and:

$$X_e = \overline{X} = \mu N \tag{2.7}$$

In contrast to the case with ethnic goods, all spending on community goods is highly effective. It is therefore not important whether analysis of public good provision is conducted using spending or outcome variables.

2.3 General Case

In a general solution we must consider the non-negativity constraint on public good investment by each group. However, lemma 1 significantly restricts the set of possible cases that require analysis.

Lemma 1 *A member of a larger group will never contribute less than a member of a smaller group. They will contribute strictly more unless $\alpha = 0$, $\tau\alpha = 1$, or they do not contribute at all. In addition, the largest group will always contribute.*

[12]The example of sewers as a community good and education as an ethnic good demonstrates the non-trivial nature of this question. In their papers, Miguel and Gugerty consider education in an "institutions" model while Alesina et al analyzes the provision of sewer services in the context of a "preferences" model.

[13]In the case of $\tau\alpha = 1$ the model faces further multiple equilibria as it is irrelevant which group(s) contribute. The predictions of the model do not depend on which individuals contribute and therefore this multiplicity of equilibria is irrelevant to this analysis.

Proof - see appendix. The intuition for this result is similar to that mentioned above in the case of "community" goods. A larger group always has a higher incentive to contribute than a smaller group and therefore will contribute more.

The effect of this lemma is that one can order groups by size and restrict analysis to cases where groups 1 through K contribute and groups K+1 through E do not. In all following results I assume that $p_i \geq p_{i+1}, \forall\ i < E-1$.

For the next result I define average public good provision as the weighted average[14] of group benefits:

$$\overline{X} = \sum_{e=1}^{E} p_e X_e \tag{2.8}$$

Proposition 1 *If $\tau\alpha < 1$ increasing the population share of the largest group (at the expense of any other group) will strictly increase average public good provision $\left(\frac{\partial \overline{X}}{\partial p_1} > 0\right)$.*

Proof - see appendix. In general there are two effects associated with a reduction in diversity. First, the individual(s) that move to the largest group receive a large direct benefit associated with improved access to public goods. Second, if the individual(s) were not previously contributing their contributions increase the total level of spending on the public good. The proof demonstrates that the result of these effects in unambiguously positive on the average level of public good provision.[15]

In general increasing the size of any group that was contributing at the expense of any smaller group will increase average public good provision. However, whether the second group is contributing depends on α. If $\alpha < \frac{p_2}{p_1}$ then the second largest group will also contribute. This leads to the following corollary:

Corollary 1 *If $\alpha < \frac{p_2}{p_1}$ then increasing the population share of the second largest group (at the expense of a smaller group) will strictly increase average public good provision.*

Proof - the proof of proposition 1 does not depend on the identity of the group increasing in size. It is sufficient that the increasing group is contributing to the public good. This is the case when $\alpha < \frac{p_2}{p_1}$.

Thus if variation in preferences is high enough, then holding the size of the largest group constant and increasing the size of the second largest group should increase aggregate provision. Further, as α is reduced, the number of groups that will contribute increases. For

[14] The proposition is unchanged if we define $\overline{X} = \sum_{e=1}^{E} p_e \ln X_e$.

[15] This proposition is qualitatively identical to the main propositions in the related papers and therefore provides no method to distinguish between the effects of different mechanisms.

$\alpha = 1$ the relevant measure of diversity (or homogeneity) is the size of the largest group. When $\alpha = 0$ the relevant measure is fractionalization. In between these extremes the optimal measure is the fractionalization between contributing groups. In addition, while the local literature has focused on fractionalization a significant portion of the cross-country literature on diversity considers the effects of polarization[16]. However, these results point toward potentially positive impacts associated with polarization. In general, holding the size of the largest group constant, increases in polarization are generally associated with increasing the size of the second largest group and thus possibly increasing public good provision.

Following directly from the proof of lemma 1 is proposition 2.

Proposition 2 *If spillovers are less than complete an individual in the largest group will receive greater value from the public good than an individual in any other group. For* $\alpha < 1$,*if* $p_1 > p_j$, $X_1 > X_j$. *For* $\alpha = 1$, $X_1 = X_j$.

Proof - see appendix.

If there is a single largest group they always contribute more than any other group and therefore the public good is more closely aligned with their preferences. This proposition provides the primary difference between the "preferences" approach and the "institutions" approach. In the preferences model $\alpha < 1$ and individuals in the largest group should receive a larger benefit from public goods than other individuals. Alternatively, if all individuals benefit equally from the public good then any effect of ethnic diversity on public good provision must be due to poor institutions. [17]

2.4 Quantity vs. Quality of Spending

As mentioned above in the extreme cases there is a definite advantage to measuring outcome variables in the case of public goods that have a strong ethnic component as the quantity of spending does not fully describe the impact of ethnic diversity. The empirical results in Kuijs (2000) indicates the variation in the effects of diversity on the quality of spending versus the quantity of spending, where, loosely speaking, quality is defined as the outcome measure divided by the spending measure. The model developed here provides one channel to explain why these effects would be different in specific cases. Diversity will therefore have

[16]See Montalvo and Reynal-Querol (2002, 2005b) and Alesina et al. (2003) for a further discussion of polarization.

[17]While this result is apparent from the model set out in ABE it is not discussed as the empirical focus on government spending prevents the authors from testing for this effect. In MG, while there is some discussion of anecdotal evidence regarding the applicability of spending on education services to all ethnic group, there is no further analysis of this result.

a larger effect on the quantity of spending when variation in preferences are low and a larger effect on the efficiency of spending when there is great variation in preferences. In Kuijs (2000) spending on education is not affected by diversity but the efficiency of spending is significantly affected. This is therefore consistent with education being a good where ethnic groups have large variation in preferences. Alternatively, both spending and the efficiency of spending on public health programs were reduced by diversity - indicating that health outcomes may face lower variation in preferences. Thus if the public good is associated with high spillovers between groups then a negative effect of ethnic diversity will be caused by a reduction in the quantity of spending with no expected effect on the quality of spending. Conversely, low spillovers between groups would indicate the problem lies with the quality of spending and we should not necessarily anticipate a correlation between diversity and spending.

2.5 Empirical Implications

The model provides two primary results. First, proposition 1 indicates that, if either mechanism has an effect, diverse communities will have lower rates of public good provision than homogeneous communities. This can be tested at both the community level and the household level. Tests are initially conducted at the community level to provide results that are directly comparable to the existing literature. Additionally, proposition 2 implies that, unless spillovers between groups are complete, individuals that are part of a local majority will have greater access to the public good than members of minority groups. Using household data regarding access to the public good I exploit this proposition to determine the relative importance of preference variation in the negative relationship between diversity and the provision of the public good in question.

In addition there are a pair of secondary results. First, the size of the second largest group should have a positive effect if the "preferences" model is correct. However, the starkness of this result is due to the quasi-linear nature of preferences and does not generalize. With a more complex specification, this result would not significantly differentiate between the two mechanisms. And second, measurement should occur over outcomes rather than spending as we are unsure as to what extent the "preferences" model is correct. To the extent that variation in preferences is the cause of poor provision of public goods, this could be associated with spending in diverse communities being either higher or lower than in homogeneous communities, depending on the precise nature of preferences.

3 Empirical Specification

The empirical section of this paper tests the implications of the model for the provision of piped water in Sub-Saharan Africa. While piped water is not a pure public good, as the use of a piped water system may be fairly characterized as non-rival[18] but excludable, it is generally provided through public investment[19]. The provision of water may therefore be best defined as a 'club good'(Cornes and Sandler 1996).

The excludability of piped water is further complicated by the difference between water that is piped into a person's home, and water that is accessed at a public tap. For this reason, I present results relating to both access to piped water in general, and for access to drinking water piped into ones home separately from drinking water accessed at a public tap.

A second concern is that piped water may not be an inclusive enough definition of public goods to capture variation in preferences across ethnic groups. Thus if individual ethnic groups place higher value on different types of water delivery or if ethnic groups place a different value on water as opposed to other public goods (such as education or electricity) then this variation will not be captured in these estimates. To the extent that this is true the results reported here should be considered a lower bound on the possible effects of ethnic diversity and in particular the effect of preference variation of this sort would not be captured in the results.

Finally, intuition may suggest that the "preferences" model of public good provision is not reasonable for the provision of piped water. As all people need clean drinking water, it is unlikely that the infrastructure to provide such a product would vary across ethnic groups. As such, the test suggested here is unnecessary. However there are reasons to believe that the test suggested here is worthwhile. First, while the basic goal of clean drinking water may be universal, there are a large variety of factors that could vary across groups. Location of public taps, location of the main water supply from which household access is drawn, the tradeoff between filtration cost and cleanliness or the use of specific chemicals to treat water are all issues that could arise in the context of developing a piped water supply. Further, as suggested in Alesina and La Ferrara (2000), ethnic groups may prefer to use a water supply that is not used by a neighbouring ethnic group. As such, while one's intuition might be that

[18] At least up to the point that water quality, or access time, is diminished by alternative uses.

[19] While private investment in the provision of piped water does occur it is a very small fraction of the total market. Budds and McGranahan (2003) find that approximately 1% of investment in water and sewerage projects comes from the private sector. In addition private provision of piped water appears to be primarily restricted to urban areas and thus the results below for the rural portion of the sample will be unaffected by private supply.

the "preferences" story is unlikely, there is sufficient reason to test this intuition. As clean drinking water is a critical good in the developing world, it is important to understand why access to drinking water may be limited in particular communities.

The existing empirical literature on ethnic diversity and public good provision has measured the impact of local diversity (usually measured as fractionalization) on the aggregate provision of public goods at the community level. The first results reported are similarly conducted at the aggregate level to provide estimates that are directly comparable to the existing literature. The model developed above indicates that if either preferences vary across ethnic groups, or local institutions are not powerful enough to overcome inter-group free-riding, or both, then the aggregate provision of public goods will be lower in communities that are more diverse. The ideal specification is therefore:

$$\overline{X}_c = \beta_0 + \beta_1 * D_c + \gamma_1 * G_c + \gamma_2 * E_c + \epsilon_c \tag{3.1}$$

where communities are indexed by c, D_c is the optimal measure of diversity as discussed above, G_c, E_c are vectors of geographic and economic controls, respectively, that capture the relevant cost-benefit factors that may affect the decision to invest in a piped water system. Two problems arise from this description. First, in the available data there is no direct measure of public good provision. To resolve this I calculate the fraction of households that access the public good, and thus the specification will be correct if this fraction is an increasing function of the aggregate level of the public good. Second, the model does not provide a single, optimal definition of diversity. For our purposes both the size of the largest group and fractionalization would be potential options and I provide results using each measure[20]. At the aggregate level, finding $\beta_1 < 0$ would indicate that ethnic diversity is having an effect, though the channel for that effect would be uncertain.

In order to differentiate between the two mechanisms I then test proposition 2 using a household-level specification.

$$X_i = \beta_0 + \beta_1 * D_c + \beta_2 * M_i + \gamma_1 * G_c + \gamma_2 * E_c + \gamma_3 * H_i + \epsilon_i \tag{3.2}$$

where in addition to the previous regressors, M_i is a measure for the household that takes the value 1 if individuals in the household are the same ethnicity as the largest local group, and H_i is a vector of household controls. If the variation in preferences between groups has an impact on the provision of piped water then individuals in the largest group in a community should have better access to the public good and therefore in addition to $\beta_1 < 0$ one should expect $\beta_2 > 0$. Only if $\alpha = 1$ does the model suggest that $\beta_2 = 0$. The relevance of the

[20] As noted below, these measures are highly correlated.

household results depend on the identifying assumption that a household is more likely to access piped drinking water if their own ethnic group has invested and thus likely influenced the resulting public good.

3.1 The Scale of Provision

It is critical to understanding the mechanism behind diversity effects that the geographic scale at which ethnic diversity is measured is related to the scale of provision. Clark and Stevie (1981) estimate the efficient scale of piped water projects as extremely local - average cost increases very quickly outside of approximately 10-15km. This is due to the relative inefficiency of piping water over long distances due to both pumping costs and water loss. Both of these problems are likely to be substantial in the context of Sub-Saharan Africa and therefore I aggregate diversity at the local scale of approximately 10km. The results are robust to minor variations to this scale. To demonstrate this robustness, the key results are replicated using a narrower definition of local ($<$ 1km between clusters) and included in the appendix.

4 Data

The primary data source for this analysis is the Demographic and Health Survey conducted by the Measure DHS project in 75 countries. This analysis is conducted on all countries of Sub-Saharan Africa for which sufficient data is available - resulting in 15 countries.[21]

For each country I use the last available survey that contains sufficient data, resulting in survey years that vary from 1994 in Cote D'Ivoire to 2004 in Cameroon. The complete dataset contains the results for approximately 180,000 individuals in 100,000 households across 5,700 survey clusters.

As mentioned above it is critical to measure diversity at a level appropriate to the public good under analysis. In the case of water, as discussed above, the appropriate scale is approximately 10km and therefore, as before, I define communities as including all clusters within 10km of a populated place[22]. Clusters that fall outside these boundaries are grouped according to a 10km x 10km grid. Alternatively, I group all clusters according to a grid that varies from 1km to 10km in size. As was demonstrated in the previous chapter, the results are not sensitive to this variation, particularly with regard to the rural communities that predominate in the sample.

[21] Excluded countries are those for which either ethnicity data or geographic data are not available.

[22] This is done using the gazetteer of populated places available from ESRI.

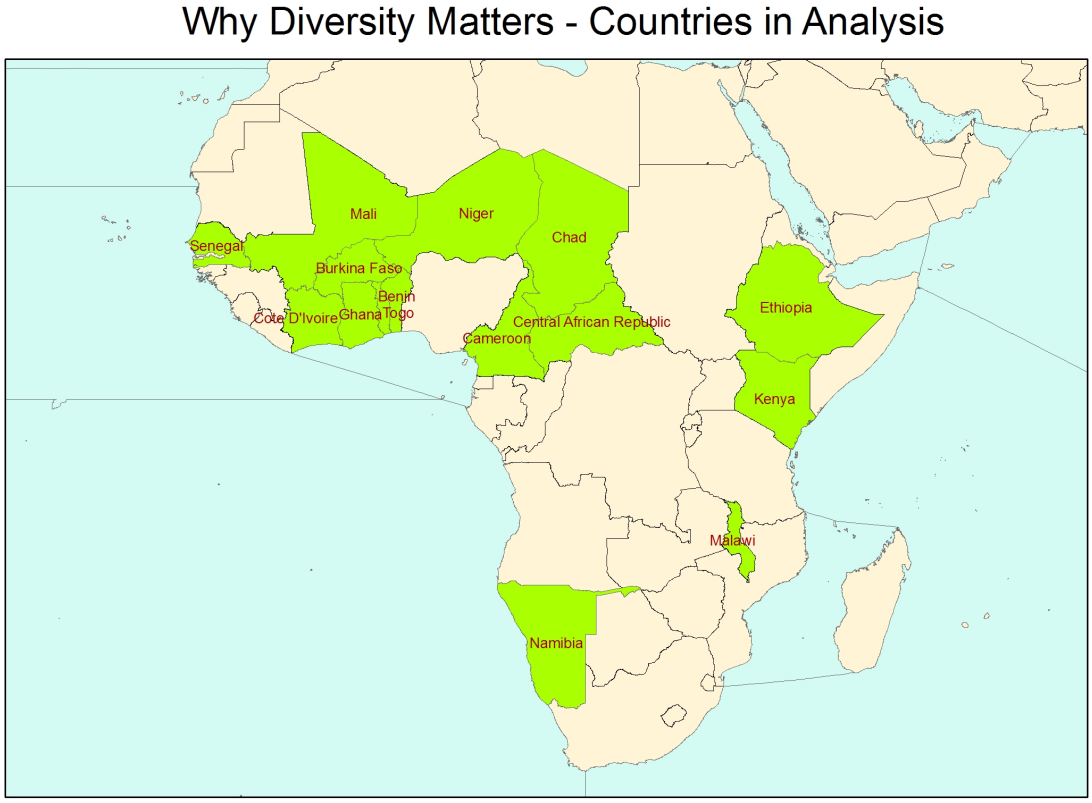

Figure 1: Why Diversity Matters - Countries in Analysis

This is partly done out of necessity as local government borders are not consistently available across countries. Further, even if these boundaries were available it is not clear that the intended influence of local governments would not be affected by the population living in the vicinity of the community. This is in part due to the dual nature of African politics with significant influence remaining for the traditional leadership structure in many countries. For example, in Ghana, Owusu-Sarpong (2003) points out that there are two sets of local political entities that affect development projects.

> 'the central government of Ghana,...,can safely carry out its development projects only by relying on a strong cooperation between the two complementary local political entities: the institutionalized local government structure and the perennial traditional authority structure; for the latter remains close to the heart of the

people.'(Owusu-Sarpong 2003, p.34)

4.1 Dependent Variables

For the aggregate specification, as before, the primary dependent variable is the fraction of households that report receiving their drinking water from a piped source. This includes both households with water piped into their house or compound along with those that access water at nearby[23] public taps. The alternatives include well water, boreholes, streams or rivers and rainwater. The necessary assumption is that the provision of a piped water system requires a more significant infrastructure cost than each of the alternatives and is thus more likely to be affected by the ability of communities to organize funding activities. For the household specification, the dependent variable is a binary variable indicating whether the household accesses water from a piped source.

In addition, in both the community and household specifications, I report the results with the dependent variable as the fraction of households that access piped drinking water in their home or compound, thus including those people that access piped water at a public tap with the group that access their water at alternative sources. Finally, access to water at a public tap is the closest that I come to a pure public good in this paper. Therefore, I construct a variable that measures the fraction of households that access water at a public tap of those households that do not have piped water in their home or compound.

4.2 A Measure of Diversity

The measure of diversity used throughout the empirical literature on public good provision is ethnic fractionalization. The model above predicts that the correct measure of diversity will depend on the extent of inter-ethnic spillovers associated with the public good, with measures varying from fractionalization to the population share of the largest group[24]. While I report key results using both measures, in practice all results are unaffected by the particular measure of diversity as the correlation between the size of the largest group and the level of ethnic fractionalization is 0.981.

An alternative measure of diversity that is commonly seen in the cross-country literature

[23]The measure of nearby is the length of time that it takes a household to collect water. If the household takes more than 20 minutes to collect water then calls into question both the efficiency of the water system and its local nature. However the results are not sensitive to this specification.

[24]The intermediate steps between these measures would involve a fractionalization measure that was restricted to groups above a specified relative size.

on ethnic diversity is polarization[25]. This attempts to measure the potential for inter-ethnic conflict and thus reaches a maximum with two equally sized groups. In this model a large second group will either be beneficial ($\alpha < 1$) or have no effect ($\alpha = 1$) and therefore we should not expect a negative effect associated with polarization once we take into account the size of the largest group. In table 4 I include a measure of the second largest group[26] and find no effect. Estimates (unreported) using the polarization measure are similar thus demonstrating that local ethnic conflict does not appear to be a significant factor in the provision of piped water.

Diversity is naturally affected by migration and this creates a potential endogeneity bias. In the simplest story migrants may be drawn to areas because of high levels of public good provision and to the extent that in-migration tends to increase diversity this would create a positive relationship between diversity and the public good. This would generally bias the results towards zero. However to control for in-migration I instrument for all diversity measures using the same statistic as calculated only on those households that have lived in their current area for at least 20 years[27]. Alternatively, the measure of diversity as calculated on long-term residents may be used directly in the regression. Given the very high correlations between the two measures this predictably has no effect. Further, this high correlation is the cause of very high F-statistics as reported in the results[28].

Alternatively it is possible that poor public good provision will lead to out-migration. Without historical data regarding ethnic diversity it is not possible to control for this factor as I do for in-migration. However it is also not clear what effect out-migration would have on local ethnic diversity. It is possible that individuals in minority groups will feel less tied to the community and would be more likely to emigrate. To the extent that this story is true the results will be biased toward zero and therefore should be taken as a lower bound. In general the average household reports living in their current area for over 30 years and

[25]See Montalvo and Reynol-Querol (2005a) for a discussion of the polarization measure and its effects on economic growth.

[26]I report estimates regarding the second largest group as the model provides a testable implication with regards to this measure as opposed to the measure of polarization.

[27]All individuals surveyed report the length of time they have lived in the area. I use all households for which the maximum tenure is either greater than 20 years or reported as "Always".

[28]It is not clear from theory which approach is more valid. If diversity now is affecting the provision of piped water through its affect on current infrastructure development and maintenance then using instrumental variables is correct. However, if previous diversity affected infrastructure development and this has a lasting impact then it would be better to use an older measure of diversity directly. Without taking a stand on when diversity mattered it is not possible to resolve this issue - however the results are consistent with either interpretation and there is likely an element of truth to both stories.

thus to the extent that diversity exists it is relatively historical[29] It seems likely that the endogeneity of diversity is likely to be less of a problem in rural areas as opposed to urban areas (which are generally more transient) and therefore I estimate effects both generally and over the rural portion of the sample without substantive differences.

4.3 Geographic Controls

The use of a piped water system may be strongly affected by geographic variables that we can control for in this study. In particular, the value of a piped water system is strongly affected by the quality and proximity of an alternative source of drinking water. A nearby stream would generally reduce the cost of developing a piped water system but would also reduce demand. The resulting effect is therefore an empirical question. Therefore, in all regressions I control for the distance from the community to the nearest river[30] (measured by the natural log of the distance in kilometres). There are many other geographic factors that may affect the demand for piped water as well as the cost of installation. These are partially captured by controlling for elevation (m), proximity to the ocean (a dummy variable taking the value 1 if the community is within 20km of the coast), suitability of land for agriculture, and distance to the capital city. In addition, I control for country and province fixed effects and thus further reduce the potential errors associated with uncontrolled geographic variation.

4.4 Ethnic Controls

African countries tend to be heavily centralized and funding for water projects may be strongly affected by political affiliation in the community. For this reason the presence of a large number of individuals of the dominant national ethnicity may affect the viability of development projects. To control for this effect I include the fraction of households in the community belonging to each of the three largest ethnic groups in the country.[31]

A second consideration is that in the preferences model the variation in demand could be between delivery methods of piped water or even the demand for public goods related to drinking water. Thus some ethnicities may prefer well water or to retrieve water from a stream while others prefer piped water. To the extent that the population share of these

[29]The historical nature of diversity over similar geographic scales is also imposed as an identifying assumption in Miguel and Gugerty (2005).

[30]It is inconsequential as to whether this is the nearest river or the nearest permanent river. The results reported in this paper use the nearest river of any sort.

[31]These are country specific thus adding 45 variables to the regressions. The results are robust to adding the local share of more ethnic groups.

specific ethnic groups is correlated with the size of the largest ethnic group this would bias the results. Controlling for the size of each ethnic group (and household ethnicity in the household specification) effectively controls for fluctuations in aggregate demand related to ethnic preferences.

4.5 Other Controls

The measure of wealth available within the DHS survey is constructed from household asset lists and then standardized for each country. Including a measure of wealth in the regression is critical, as wealth and diversity are strongly, and positively correlated. This is the predictable result of people migrating in response to economic opportunity.

Population density is likely to strongly affect the efficiency of a piped water system. I use 1990 measures of population density, measured at the very local scale, along with the urban or rural status of the community. A potential concern correlated with diversity is that diverse communities may be more likely to contain recent migrants or younger households and thus community organization may vary for reasons only indirectly associated with diversity. I therefore control for the length of residency reported by each household[32] and the average age of the household head. At the household level, previous research (Thomas 1990) has indicated that the sex of the household head affects household decision making and therefore I control for this, and the age of the household head, at both the household level and as a community average.

The summary statistics demonstrate the significant differences between rural and urban areas within the sample. Urban areas are far more likely to have piped water, and also are more diverse. In addition, households have shorter tenure, and are more likely to have a female head of household.

5 Provision of Piped Water - Community Access

Proposition 1 indicates that more diverse communities should have lower aggregate levels of public good provision regardless of the mechanism through which ethnic diversity has an effect. While the model does not provide a single measure of diversity there are two

[32] As mentioned above in the context of diversity instruments if more than one household member responded I use the longer reported tenure. Households also had the option to report "Always" and this was coded at either the maximum possible tenure (50 years) or the age of the individual responding. The results are robust to variations in the way this variable is calculated.

possible candidates related to the two extremes of the model. These are fractionalization[33] and the population share of the largest ethnic group[34]. In reality the (negative) correlation between these two measures across local communities is 98% and therefore this consideration is somewhat academic.

5.1 Primary Results

Table 2 contains the basic results of the community level specification. Columns (1) and (2) present the effects of diversity, measured as either the size of the largest group or fractionalization[35], on the fraction of the population that has access to piped water. As the results show, ethnic diversity, regardless of measure, has significant negative[36] effects on the provision of piped drinking water.

Columns (3) and (4) break down the supply of piped drinking water into households that access drinking water in their home or compound, and those that access drinking water at a public tap[37]. The results indicate that homogeneous communities provide a larger fraction of their households with home access to piped drinking water, and provide a larger fraction of the remaining community with access at a public tap.

Not surprisingly, wealth and urbanization have large and significant effects on access to piped water. In addition, wealth is a relatively more significant factor in determining access in one's home or compound, while urbanization is more important in the context of providing access to public taps. This seems natural, as home access is likely partially the result of private investment, while public taps will only be an effective community investment if demand is sufficiently high, which will occur more often in urban areas.

The other control variables offer further insight into the problem of providing piped drinking water to a population. Older communities, represented by higher values of average tenure and older household heads, are more likely to have higher rates of piped water, with this effect primarily the result of more households having direct access. In contrast, the distance to a river increases the likelihood that a community will provide public taps, which

[33]Fractionalization is identical to the Herfindahl of industry concentration commonly used in Industrial Organization. It is calculated as $F = 1 - \sum_{e=1}^{E} p_e^2$, where p_e is the population share of group e.

[34]As noted above fractionalization would be the correct measure if there were no spillovers between groups and the share of the largest group would be the correct measure if spillovers between groups are complete.

[35]In this table, I present the results using historical diversity, where diversity is measured only using households where a resident has been present in the community for more than 20 years, or has always been a resident.

[36]Or homogeneity has positive effects.

[37]As discussed previously, this variable is constructed as the ratio of the number of households that access water at a public tap to the number of households that do not have water piped into their home or compound.

is likely the result of increased demand as alternative sources are more costly to access. Surprisingly, coastal communities[38] are far less likely to provide access to public taps. The reason for this finding is unclear, though the magnitude is sufficient to potentially warrant further study.

These results are robust to a large variety of different specifications, including using current diversity measures, either with or without instruments, using instruments to account for the potential endogeneity of wealth, or varying the constructed size of a "community". In particular, moving to communities defined by clusters that are within 1km of each other, instead of 10km, does not affect the results.

5.2 Rural Access to Piped Water

Access to clean drinking water is of particular concern in rural communities[39] where only 13% of households have access to piped drinking water. Table 3 duplicates the previous table using only rural communities. The results show little to no variation from the table with all observations, though the point estimates are somewhat smaller.

In the appendix, I repeat the analysis using a narrower definition of "rural community", but the results show no substantive variation.

5.3 Alternative Measures of Diversity

As discussed in the theoretical section of the paper, the optimal measure of diversity depends on the mechanism that one expects to find. The existing literature on diversity overwhelming uses measures of ethnic fractionalization that are optimal in the case of specific models, but also act as a relatively general measure of ethnic diversity. The first two columns of table 2 demonstrate that there is little to choose between these two measures. In the first column of table 4, I include both measures at the same time. While this type of "horse-race" between variables is of limited validity, it is interesting to see that the size of the largest local group appears to be of greater relevance[40].

Alternatively, existing research at the national level has argued that competition between groups makes coordination of public projects difficult or costly. Competition between groups is generally measured by a polarization index that is maximized when a community has two

[38] Defined as communities that are within 20km of the coast. The result is robust to variations in this definition.

[39] Rural communities are those where a majority of households are listed as rural. Therefore the variable *urban* continues to contain information on the relative urbanization of the community.

[40] The sign on fractionalization is now reversed from the expected effects of diversity.

equally sized groups. As shown in column(2), polarization at the community-level does not significantly affect the provision of piped water.

In general, polarization is linked to the size of the second largest group in the community. As the theory above suggests, if there are spillovers between groups, there may be a positive effect from having a large second group. Column (3) presents a specification that tests this directly. As can be seen, after controlling for the size of the largest group, the size of the second largest group in the community has a marginally significant, positive effect. Within the context of the model, this could be interpreted as support for the preferences mechanism. However, the stark nature of the model, created by the quasi-linear formulation of preferences, does not generalize and this result should not be seen as particularly strong evidence in this regard.

While the empirical section of this paper is focused on the effects of ethnic diversity, the model does not differentiate between ethnicity and other features of personal identity such as religion. In column(4), I report the results of a regression including a measure of religious homogeneity. Religion does not appear to have a significant impact within the context of communities in Sub-Saharan Africa, and if anything, religious diversity is associated with greater provision of drinking water. This result is quite similar to the results found elsewhere in the literature on diversity, that suggest that reported religious diversity is associated with personal freedom and therefore might vary positively with a variety of economic outcomes. Further, it should be noted that including this measure does not affect the estimate of the effects of ethnic diversity.

6 Household Access

While the results at the community level are consistent with the existing evidence on the effects of ethnic diversity, in order to differentiate between the proposed mechanisms it is necessary to analyze the distribution of the public good in each community. The preferences mechanism suggests that the dominant group in the community should have higher rates of utility from the public good that is provided and therefore should have higher rates of access. Alternatively, if the local administration discriminates against minority groups in permitting access to the public good, we should again find that smaller groups in the community are relatively disadvantaged. In contrast, the institutions mechanism does not favour any group in the community, and where governance is the problem we should not expect the size of one's own ethnic group to matter.

To test this prediction at the household level, I include a measure of the share of the

local population in the same ethnic group as a household[41]. I also test to determine if it is important that one's own ethnic group is dominant in the town, relative to other groups. For this purpose, I test the effect of the size of one's own ethnic group relative to the largest group in town. Lastly, I include a dummy variable that takes the value 1 if one's own ethnic group is the largest in town.

Economic variables included in the community regression remain in the household regression, along with their household counterparts. Thus, I include average community wealth alongside household wealth, average tenure and household tenure, etc. The community variables are kept as the public good nature of the product suggests that the characteristics of other households in the community might affect a particular household's access. In addition, along with the share of the largest three ethnicities, I include a dummy variable for each ethnicity in each country.

A key to this analysis is that there is variation in household access within communities. Figure 2 shows a histogram of household access in communities where at least one household has access to piped water. As is clear from the figure, there is considerable variation in the fraction of households that access the piped source, with the majority of communities at a point where some households, but not all, access drinking water at a piped source.

6.1 Primary Results

Table 5 contains the results of the baseline specification[42]. Column (1) presents the baseline specification that corresponds to the primary specification at the community level.

The results indicate that the size of the largest ethnic group in the community is important, regardless of one's own minority status. When this result is broken down into home access in column (3), and public access[43] in column (4), the results suggest that ethnic homogeneity is a significant determinant of piped water access, but that minority status is not a significant detriment.

The other dependent variables provide reasonable outcomes that are not generally different from those found previously. The effects of wealth are split between the importance of household wealth and community wealth. Household wealth is relatively more important

[41] As there are households that report multiple ethnicities, the share is calculated by summing over the household share of each ethnicity multiplied by the community share of each ethnicity.

[42] To provide results that are directly comparable to the previous section, I present the results of a linear probability model using simple OLS. The results are not sensitive to this choice, and in particular, all results can be generated using a logit or probit specification.

[43] This regression is limited to households that do not access piped water at home. The results presented here are substantively the same as using an ordered probit or logit regression.

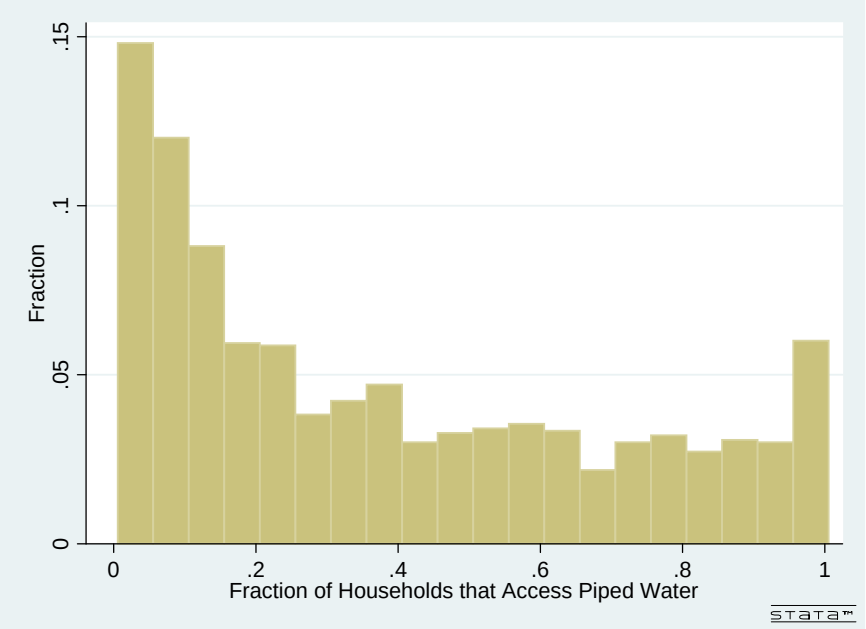

Figure 2: Household Access in Communities with Piped Water

in determining access to water in the home, which is reflective of the more private nature of this good. However, the fact that community wealth is a statistically and economically significant determinant of household access in all forms suggests that there is a strong public good aspect to the provision of piped water. In addition, in a result that mirrors previous findings, households with female heads are significantly more likely to access piped drinking water, either at home or at a public tap. However, this could be the result of a variety of factors and further analysis of this finding is beyond the scope of this paper.

6.2 Alternative Measures of Diversity and Dominance

The population share of one's own ethnic group is the measure that makes the most sense from the perspective of the theory outlined above. However, there are other measures that make more sense in the context of alternative models of political conflict. In particular, if local politics is driven by the relative size of individual groups, then the relative size of one's own group could be critical. Alternatively, in a pure median voter model as found in Alesina et al.(1999), being a member of the majority group would be the critical factor.

Table 6 presents the results using alternative measures of ethnic diversity and the relative importance of one's own group. Column (1) repeats the primary specification from table 5

for comparison. Column (2) incorporates a measure of the relative size of each ethnic group. Specifically, this variable is defined as the population share of one's own ethnic group, divided by the population share of the largest ethnic group, resulting in a variable bounded between 0 and 1. The results of this specification are unchanged from the previous analysis.

The key dependent variable in column (3) is the largest group variable. This takes the value of one if the relative share, defined above, is equal to one, and zero otherwise. If the political system provides power to the largest groups in society, then the coefficient on this variable should be positive. Alternatively, in column (4), the key variable is the majority variable. This takes the value of one if, and only if, the household is a member of a group that is larger than 50% of the local population. The results of both regressions indicate that local political factors appear to be minimal in determining access to piped water. Alternative specifications that look specifically at water piped into one's home or compound, or access at a public tap show similarly insignficant results.

6.3 Rural Communities

As discussed, the natural focus of this study is on rural communities that are the least likely to have access to piped drinking water. To this end, table 7 duplicates the results of table 5, but restricts the sample to households that live in primarily rural communities. This part of the sample makes up roughly two-thirds of the total, and the estimates provide further support to the idea that ethnic diversity affects access to piped drinking water, but that minority status is unimportant.

The results of this specification demonstrate that the effects encountered above are fully present in the rural communities that suffer from the most severe lack of critical water supplies. While diversity has a negative effect on the ability of all households in a community to access piped drinking water, being a member of a larger group in the community has no effect. The effects of the additional dependent variables are not significantly different from those discovered previously.

7 Conclusion

Achieving economic development requires that governance structures permit society to overcome the problems associated with the provision of public goods. As the aggregate results presented here demonstrate, ethnic diversity has a significant effect on the provision of piped water in Africa. This result is consistent with the existing evidence from around the world,

in both developed and developing countries. Knowing that diversity matters is only the first step. We must also understand the reasons behind the effect.

The model developed above demonstrates that the distributional pattern of benefits that individuals gain from the provision of public goods will vary depending on why diversity matters. Specifically, if ethnic groups vary in their preferences for a public good, then the majority group should be able to influence the resulting good in ways that their members prefer. In contrast, if the difficulty associated with the provision of public goods in a diverse community occurs because of an inter-ethnic collective action problem, then there is no reason to believe that the resulting public good will favour one group or another.

The community level results indicate that, within Sub-Saharan Africa, ethnic diversity severely limits the spread of piped drinking water, with possibly critical effects. The results of the household level regression indicate that minority groups are not significantly disadvantaged in terms of access to piped drinking water. The evidence is therefore supportive of the 'institutions' model - the effect of ethnic diversity is felt through ineffective governance rather than preference variation. This paper therefore presents a different model of "ethnic politics" than is sometimes suggested for diverse communities. Rather than demonstrating the effects of discrimination, where ethnic groups actively prevent the inclusion of others, the results suggest that ethnic politics leads to significant free-riding as groups wait for the contributions of others.

As mentioned previously, this same free-riding problem may be demonstrated by the fact that communities close to the national capital, including rural communities, are less likely to access piped water. Thus the idea of waiting for an "outsider" to accomplish the task of providing a public good appears to extend far beyond just an ethnic consideration.

It should be emphasized that, while the approach is general, the empirical result is specific to the context in question. Ethnic diversity may be related to variations in preferences for other goods, or in other countries. For example, it would not be surprising to find the household result reversed in the case of education services, where a majority group may be able to encourage instruction in the majority language. In the context of the model, investments in education may not have complete spillovers across ethnic groups. Alternatively, in countries where governance has greater structure, and the government does not rely on the voluntary actions of its citizens, the problems of ethnic diversity may not be as severe. Each of these issues is the subject of continuing research.

Further, the negative effect of ethnic diversity may simply be the visible sign of an ineffective governance structure. The benefits of homogeneity, whether they be through stronger social sanctions or intra-ethnic altruism, are unlikely to create an environment that provides

efficient levels of public goods. As such, it is likely that the results presented here represent a part of the inefficiency caused by the ineffective governance structures that persist in Sub-Saharan Africa.

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.

Alesina, A., and E. L. Ferrara (2000): "Participation in Heterogeneous Communities," *Quarterly Journal of Economics*, 115, 847–904.

——— (2005): "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 43(3), 762–800.

Bardhan, P., and J. Dayton-Johnson (2002): "Unequal Irrigators: Heterogeneity and Commons Management in Large-Scale Multivariate Research," in *The Drama of the Commons*. National Academy Press.

Budds, J., and G. McGranahan (2003): "Are the debates on water privatization missing the point? Experiences from Africa, Asia and Latin America," *Environment and Urbanization*, 15(2), 87.

Caselli, F., and W. J. C. II (2006): "On the Theory of Ethnic Conflict," Mimeo.

Clark, R. M., and R. G. Stevie (1981): "A Water Supply Cost Model Incorporating Spatial Variables," *Land Economics*, 57, 18–32.

Collier, P. (2000): "Doing Well Out of War," in *Greed and Grievance: Economic Agendas in Civil Wars*, ed. by M. Berdal, and D. Malone. Lynne Rienner Pub.

Collier, P., and J. W. Gunning (1999): "Explaining African Economic Performance," *Journal of Economic Literature*, 37, 64–111.

Cornes, R., and T. Sandler (1996): *The theory of externalities, public goods, and club goods*. Cambridge Univ Pr.

Easterly, W. (2001): "Can Institutions Resolve Conflict," *Economic Development and Cultural Change*, 49(4), 687–706.

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

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

Garcia-Montalvo, J., and M. Reynal-Querol (2002): "Why Ethnic Fractionalization? Polarization, Ethnic Conflict and Growth," Mimeo.

Humphreys, M., D. Posner, and J. Weinstein (2005): "Ethnic Identity, Collective Action, and Conflict: An Experimental Approach," Mimeo.

Khwaja, A. (2009): "Can Good Projects Succeed in Bad Communities?," *Journal of Public Economics*.

Kimenyi, M. (2006): "Ethnicity, Governance and the Provision of Public Goods," *Journal of African Economies*, 15, 62–99.

Luijs, L. (2000): "The Impact of Ethnic Heterogeneity on the Quantity and Quality of Public Spending," IMF Working Paper 00-49.

Mauro, P. (1995): "Corruption and Growth," *The Quarterly Journal of Economics*, 110(3), 681–712.

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

Montalvo, J. G., and M. Reynal-Querol (2005a): "Ethnic Diversity and Economic Development," *Journal of Development Economics*, 76, 293–323.

——— (2005b): "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review*, 95, 796–816.

Owusu-Sarpong, C. (2003): "Setting the Ghanian Context of Rural Local Government: Traditional Authority Values," in *Grassroots Governance: Chiefs in Africa and the Afro-Caribbean*, chap. 2, pp. 31–67. University of Calgary Press.

Porta, R. L., F. L. de Silanes, A. Shleifer, and R. W. Vishny (1999): "The Quality of Government," *Journal of Law, Economics and Organization*, 15(1), 222–279.

Posner, D. N. (2004): "Measuring Ethnic Fractionalization in Africa," *American Journal of Political Science*, 48(4), 849–863.

Thomas, D. (1990): "Intra-Household Resource Allocation: An Inferential Approach," *Journal of Human Resources*, 25, 635–664.

Vigdor, J. L. (2004): "Community Composition and Collective Action: Analyzing Initial Mail Response to the 2000 Census," *The Review of Economics and Statistics*, 86, 303–312.

Appendix

A Proofs

A.1 Lemma 1

- Proof of Lemma 1
  The FOC facing each group can be written as:

$$1 = \frac{\mu p_e N}{X_e} + \tau\alpha\mu N \sum_{j\neq e} \frac{p_j}{X_j} \tag{A.1}$$

$$1 - Q \geq (1-\tau\alpha)\mu \frac{p_e N}{X_e} \tag{A.2}$$

$$Q = \tau\alpha\mu \sum_{j=1}^{E} \frac{p_j N}{X_j} \tag{A.3}$$

  and Q is constant for all groups. If two groups contribute positive amounts the FOC holds with equality and thus:

$$\frac{p_e N}{X_e} = \frac{p_j N}{X_j} \tag{A.4}$$

$$p_e((1-\alpha)p_j x_j + \alpha S) = p_j((1-\alpha)p_e x_e + \alpha S) \tag{A.5}$$

$$\text{where S equals total spending} \tag{A.6}$$

$$S = \sum_{k=1}^{E} p_k x_k \tag{A.7}$$

  Rearranging equation A.5 defines the difference between the contributions of two groups:

$$x_e - x_j = \frac{(p_e - p_j)\alpha S}{(1-\alpha)p_e p_j} \tag{A.8}$$

This difference is positive if $p_e > p_j$ unless $\alpha = 0$[44]. If $\tau\alpha = 1$ then both groups face identical FOCs ($Q = 1$) and their contributions are equal. Given the first part of the lemma it is obvious that the largest group must contribute. If no one else contributes the marginal value of their first dollar invested is infinite and if anyone else contributes then the largest group will as well.

A.2 Proposition 2

- Proof of Proposition 2
  Rearranging the equation defining the public goods for each of two groups we have:

$$X_e = (1-\alpha)p_e x_e + \alpha S > (1-\alpha)p_j x_j + \alpha S = X_j \tag{A.9}$$

For $\alpha < 1$ this inequality holds as $p_e x_e > p_j x_j$.

A.3 Proposition 1

- Proof of Proposition 1
  For any equilibrium define C as the number of ethnic groups that contribute positively and therefore $J = E - K$ groups do not contribute. Further define the total share of all groups that contribute:

$$p_c = \sum_{k=1}^{K} p_e \tag{A.10}$$

$$\tag{A.11}$$

For any group that contributes, $\frac{p_e}{X_e} = \frac{p_c}{X_c} = \beta$. We therefore have:

$$\begin{aligned} \sum_{k=1}^{K} X_k &= \frac{p_c}{\beta} && \text{(A.12)} \\ &= (1+\alpha(K-1)) \sum_{k=1}^{K} p_k x_k && \text{(A.13)} \\ &= (1+\alpha(K-1)) * S && \text{(A.14)} \end{aligned}$$

where S is total spending on the public good as defined above. For any group that does not contribute their public good is determined by the contributions of the other

[44]Note that total spending is always positive as at least these two groups have made contributions.

groups:

$$X_j = \alpha \sum_{k=1}^{K} p_k x_k \tag{A.15}$$

$$= \frac{\alpha \sum_{k=1}^{K} X_k}{(1+\alpha(K-1))} \tag{A.16}$$

$$= \frac{\alpha p_c}{(1+\alpha(K-1))\beta} \tag{A.17}$$

An ethnic group for whom the FOC holds with equality when they make precisely zero contribution can be considered contributing or not. Ethnic group K+1 contributes if:

$$X_{K+1} = \frac{\alpha p_c}{(1+\alpha(K-1))\beta} = \frac{p_e}{\beta} = X_e \tag{A.18}$$

where p_c is the total share of all ethnic groups from 1 to K. Rearranging a group will not contribute if:

$$\frac{\alpha}{(1+\alpha(K-1))} p_c \geq p_e \tag{A.19}$$

The largest group always contributes and therefore we can write their FOC as:

$$1 = \mu N \frac{p_1}{X_1} + N\tau\alpha\mu \sum_{k=2}^{K} \frac{p_k}{X_k} + N\tau\alpha\mu \sum_{j=K+1}^{E} \frac{p_k}{X_k} \tag{A.20}$$

$$= \mu N\beta + N\tau\alpha\mu(K-1)\beta + N\tau\mu(1+\alpha(K-1))\beta\frac{(1-p_c)}{p_c} \tag{A.21}$$

Rearranging this equation results in:

$$\beta = \frac{p_c}{N\mu\left[p_c\left(1-\tau\right)+\tau\left(1+\alpha(K-1)\right)\right]} \tag{A.22}$$

$$= \frac{p_c}{N\mu\left[p_c\left(1-\tau\right)+\tau Q\right]} \tag{A.23}$$

Having defined β in terms of parameters we are now able to measure the marginal effect of increasing the size of the largest group at the expense of any other single group.[45]

[45] Obviously any change that involves more than one smaller group can be broken into multiple changes.

Average public good provision can be written as:

$$\overline{X} = \sum_{e=1}^{E} p_e X_e \tag{A.24}$$

$$= \sum_{k=1}^{K} \frac{p_k^2}{\beta} + \sum_{j=K+1}^{E} \frac{p_j p_c \alpha}{Q\beta} \tag{A.25}$$

$$= \frac{1}{\beta}\left[\sum_{k=1}^{K} p_k^2 + \frac{\alpha}{Q}(1-p_c)p_c\right] \tag{A.26}$$

with $Q = 1 + \alpha(S-1)$. Increasing the size of the largest group has three consequences for group size. The largest group increases, some other group decreases and the share of the population that contributes may or may not increase.

- Case 1 - $p_1 \uparrow, p_e \downarrow, p_c$ unchanged - the mass moving to the largest group comes from another group that was already contributing a positive amount.
  In this case there is no change in β and therefore:
  $$\frac{\partial \overline{X}}{\partial p_1} = \frac{1}{\beta}\left[2p_1 - 2p_e\right] \tag{A.27}$$
  which is positive as $p_1 > p_e$ by construction and $\beta > 0$.
- Case 2 - $p_1 \uparrow, p_j \downarrow, p_c \uparrow$ - the group moving to the largest group comes from another group that was not contributing.
  The key in this case is that the change in p_c alters β. Therefore:
  $$\frac{\partial \overline{X}}{\partial p_1} = \frac{\beta I_1 - I\beta_1}{\beta^2} \tag{A.28}$$
  where:
  $$I = \left[\sum_{k=1}^{K} p_k^2 + \frac{\alpha}{Q}(1-p_c)p_c\right] \tag{A.29}$$
  $$I_1 = \frac{\partial I}{\partial p_1} \tag{A.30}$$
  $$= 2p_1 - p_c Q + (1-p_c)Q \tag{A.31}$$
  $$\tag{A.32}$$

while β_1 is defined in a similar way:

$$\beta_1 = \frac{\partial \beta}{\partial p_c} \tag{A.33}$$

$$= \frac{\tau Q}{\mu N \left[p_c \left(1-\tau\right) + \tau Q\right]^2} \tag{A.34}$$

$$= \frac{\beta^2 \tau \mu Q N}{p_c^2} \tag{A.35}$$

Substituting equations A.23 and A.35 into equation A.28 results in:

$$\frac{\partial \overline{X}}{\partial p_1} = \frac{\mu N}{p_c}\left[p_c\left(1-\tau\right) + \tau Q\right] I_1 - I\frac{\tau \mu Q N}{p_c^2} \tag{A.36}$$

$$= p_c\left(1-\tau\right) I_1 + \tau Q I_1 - \frac{I \tau Q}{p_c} \tag{A.37}$$

The first term is positive if $\tau < 1$. I will therefore show that the second term is greater than the third. Factoring τQ leaves:

$$I_1 \geq \frac{I}{p_c} \tag{A.38}$$

$$2p_1 - 2\frac{\alpha}{Q}p_c + \frac{\alpha}{Q} \geq \frac{1}{p_c}\sum_{k=1}^{K} p_k^2 + \frac{\alpha}{Q}(1-p_c) \tag{A.39}$$

Collecting terms results in:

$$p_1 - \frac{1}{p_c}\sum_{k=1}^{K} p_k^2 + p_1 - \frac{\alpha}{Q}p_c \geq 0 \tag{A.40}$$

As:

$$p_1 \geq \frac{1}{p_c}\sum_{k=1}^{K} p_k^2 \tag{A.41}$$

$$\geq \frac{\alpha}{Q}p_c \tag{A.42}$$

while at least one of these is a strict relation if $p_c < 1$ which is necessary to even consider this case.

Thus for $\alpha\tau < 1$ increasing the size of the largest group will strictly increase average public good provision.

B Robustness Check

The results of this paper are driven, primarily, by the distribution of piped water in rural communities. However, while urban communities that stretch over distances of up to 10km

are regularly observed, this is not necessarily the case with most rural communities. In this section, I redo this exercise looking at only rural communities, where clusters are combined into communities based on a narrow, 1km x 1km grid. Doing so increases the number of rural communities slightly over the previous analysis, as clusters that are, for example, 5km apart, are now considered to be distinct villages.

B.1 Rural Communities - Narrow Definition

Table 8 presents the results of the community-level regressions, corresponding to table 3 in the main section of the paper. The results confirm the earlier findings, and demonstrate that the link between ethnic diversity and poor public good provision is not a byproduct of the choice of village aggregation. Many of the point estimates are reduced in this specification, which could be the result of too little aggregation. Incorrect aggregation would likely generate a strong downward bias in the estimates. Clusters that are measured to be 1-2km apart may be very likely to coordinate the provision of piped water, and the effects of ethnic diversity on this type of coordination is lost here.

At the household level, table 9 duplicate the regressions from table 7 over the narrower gridsize. Again, I find no evidence that the size of one's own group affects access to piped water, though the reduction in grid-size does reduce the measured effects of diversity on access to piped water through a public tap. There are two caveats to this result. First, diversity has a significant effect on the fraction that have access at home, and this has secondary effects on access at a public tap. Second, if there is any measure that is likely to be incorrectly specified over a narrow geographic region, it is the number of households that access piped water at a public tap.

Table 1: Summary Statistics for 10km Communities

Variable	All Communities	Rural	Urban
Fraction w/ Piped Water	0.287	0.127	0.646
	(0.346)	(0.225)	(0.298)
Fraction w/ Piped Water at home	0.181	0.073	0.430
	(0.271)	(0.174)	(0.289)
Share of Largest Group	0.720	0.804	0.530
	(0.245)	(0.198)	(0.236)
Share of 2nd Largest Group	0.137	0.120	0.175
	(0.117)	(0.123)	(0.090)
Ethnic Fractionalization	0.378	0.601	0.278
	(0.285)	(0.242)	(0.245)
Distance to a River (km)	4.97	4.82	5.31
	(6.27)	(4.83)	(6.37)
Elevation (m)	663.9	683.1	620.6
	(643.93)	(637.73)	(656.41)
Distance to Ocean (km)	511.7	555.9	412.0
	(358.4)	(350.15)	(357.05)
Distance to Large City (km)	115.3	133.3	74.8
	(110.4)	(105.7)	(109.9)
Average Tenure	30.99	33.07	26.31
	(9.35)	(9.05)	(8.27)
Average Age of HH Head	44.4	45.4	42.1
	(5.13)	(5.18)	(4.24)
Prop.w/ Female Head	0.22	0.21	0.25
	(0.14)	(0.15)	(0.12)
Number of Communities	3,349	2,724	624

Notes: Standard deviations in parentheses.

Table 2: Community Specification - Base Results

	Base	Fractionalization	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.098***		.052***	.061**
	(.024)		(.017)	(.024)
Fractionalization		-.069***		
		(.019)		
Fraction w/ Piped Water at Home				-.048
				(.038)
Avg. Wealth	.338***	.338***	.251***	.196***
	(.015)	(.015)	(.012)	(.017)
Urban	.092***	.090***	.026*	.063***
	(.020)	(.020)	(.015)	(.021)
Pop.Density(1990)	.007	.008	.015	-.010
	(.006)	(.006)	(.010)	(.010)
Dist.to Cap.City (LogKM)	.035***	.036***	.016	.030***
	(.010)	(.010)	(.010)	(.009)
Dist.to River (LogKM)	.009***	.009***	.004	.007**
	(.003)	(.003)	(.002)	(.003)
Coast.Comm.	-.057*	-.057*	.033	-.151***
	(.030)	(.030)	(.025)	(.035)
Elevation(km)	-.006	-.006	.007	-.010
	(.015)	(.015)	(.011)	(.015)
Avg. Tenure	.001*	.001*	.001***	.0002
	(.0006)	(.0006)	(.0004)	(.0006)
Frac.Fem. HH Head	.024	.025	-.045*	.027
	(.033)	(.033)	(.026)	(.036)
Avg. Age of Head	.001	.001	.002***	-.0007
	(.0009)	(.0009)	(.0007)	(.0009)
Obs.	3062	3062	3062	3031
R^2	.778	.778	.806	.639

Notes: The dependent variable is the fraction of households that get drinking water from a piped source. Heteroskedasticity-robust standard errors (adjusted for intra-district correlation) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***,**, and * respectively. All regressions include country and province fixed effects and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 3: Rural Communities

	Base	Fractionalization	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.081***		.030**	.049**
	(.023)		(.014)	(.022)
Fractionalization		-.057***		
		(.019)		
Fraction w/ Piped Water at Home				.034
				(.035)
Avg. Wealth	.333***	.333***	.219***	.177***
	(.020)	(.020)	(.017)	(.024)
Urban	.184***	.183***	.100***	.073
	(.046)	(.047)	(.028)	(.050)
Pop.Density(1990)	-.033	-.031	.113***	-.199***
	(.056)	(.056)	(.042)	(.063)
Dist.to Cap.City (LogKM)	.029**	.030**	.021**	.015*
	(.012)	(.012)	(.010)	(.008)
Dist.to River (LogKM)	.009***	.010***	.003	.007**
	(.003)	(.003)	(.002)	(.003)
Coast.Comm.	-.061	-.061	.019	-.092***
	(.038)	(.039)	(.031)	(.033)
Elevation(km)	.002	.002	.014	-.005
	(.015)	(.015)	(.011)	(.013)
Avg. Tenure	.0008	.0008	.0008**	.0002
	(.0005)	(.0005)	(.0004)	(.0005)
Frac.Fem. HH Head	-.012	-.012	-.033	-.0007
	(.034)	(.034)	(.026)	(.032)
Avg. Age of Head	.001	.001	.001**	-.0002
	(.0009)	(.0009)	(.0006)	(.0007)
Obs.	2601	2601	2601	2588
R^2	.579	.578	.634	.431

Notes: Heteroskedasticity-robust standard errors (adjusted for intra-group correlation over 50km districts) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***,**, and * respectively. All regressions include country and province fixed effects and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 4: Community-Level Specification - Diversity Measures

	Both	Polarization	TwoGroups	Religion
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.215***		.131***	.081***
	(.081)		(.035)	(.018)
Fractionalization	.123*			
	(.067)			
Eth.Polar.		-.015		
		(.010)		
2nd Large. Eth.Group			.104**	
			(.048)	
Large.Relig.Grp				-.050**
				(.020)
Avg. Wealth	.342***	.339***	.346***	.343***
	(.012)	(.012)	(.012)	(.012)
Urban	.093***	.087***	.097***	.093***
	(.016)	(.016)	(.016)	(.016)
Obs.	4611	4611	4611	4611
R^2	.781	.779	.781	.781
e(jp)				

Notes: The dependent variable is the fraction of households that get drinking water from a piped source. Heteroskedasticity-robust standard errors (adjusted for intra-district correlation) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***,**, and * respectively. All regressions include country and province fixed effects and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS. In addition, the coefficients on the majority of the control variables have been suppressed.

Table 5: Household Specification - Base Regression

	Base	Fractionalization	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.143***		.045**	.119***
	(.028)		(.020)	(.025)
Fractionalization		-.097***		
		(.022)		
Hist.Share Own Eth.	-.026***	-.017**	-.007	-.015*
	(.009)	(.009)	(.008)	(.008)
Fraction w/ Piped Water at Home				.109**
				(.045)
HH Wealth	.146***	.147***	.172***	.087***
	(.007)	(.007)	(.009)	(.007)
Avg. Wealth	.171***	.170***	.071***	.100***
	(.016)	(.016)	(.013)	(.019)
Urban	.103***	.102***	.024	.069***
	(.023)	(.023)	(.015)	(.024)
Pop.Density(1990)	.018***	.018***	.023***	.025**
	(.006)	(.006)	(.007)	(.011)
Tenure	-.0001*	-.0002**	-2.91e-07	-.00006
	(.00007)	(.00007)	(.00006)	(.00007)
Avg. Tenure	.001**	.001**	.002***	-.0002
	(.0006)	(.0006)	(.0005)	(.0005)
Female Head	.016***	.016***	.009***	.014***
	(.003)	(.003)	(.003)	(.003)
Frac.Fem. HH Head	.043	.043	-.010	.023
	(.038)	(.038)	(.027)	(.037)
Dist.to Cap.City (LogKM)	.028***	.028***	.010	.022**
	(.009)	(.009)	(.007)	(.009)
Dist.to River (LogKM)	.010***	.010***	.004*	.007**
	(.003)	(.003)	(.002)	(.003)
Coast.Comm.	-.076**	-.076**	.016	-.122***
	(.034)	(.034)	(.023)	(.044)
Obs.	96859	96859	96859	79983
R^2	.503	.503	.518	.304

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

Table 6: Household Specification - Measuring Dominance

	GroupShare	Relative	Largest	Majority
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.143***	.124***	.124***	.123***
	(.028)	(.026)	(.026)	(.026)
Hist.Share Own Eth.	-.026***			
	(.009)			
Relative Share Own Eth.		-.021***		
		(.007)		
Largest Eth.Grp			-.008	
			(.005)	
Majority Eth.Grp.				-.005
				(.005)
HH Wealth	.146***	.146***	.147***	.147***
	(.007)	(.007)	(.007)	(.007)
Avg. Wealth	.171***	.171***	.171***	.171***
	(.016)	(.016)	(.016)	(.016)
Urban	.103***	.103***	.103***	.103***
	(.023)	(.023)	(.023)	(.023)
Obs.	96859	96859	96859	96859
R^2	.503	.503	.503	.503

Notes: Heteroskedasticity-robust standard errors (adjusted for intra-group correlation over 10km communities) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***,**, and * respectively. All regressions include country and province fixed effects and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS. Additional coefficients are unchanged by the change in specification and are therefore not displayed.

Table 7: Households in Rural Communities

	Base	Fractionalization	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.104***		.040**	.076***
	(.026)		(.019)	(.021)
Fractionalization		-.067***		
		(.021)		
Hist.Share Own Eth.	-.017**	-.010	-.015**	-.005
	(.008)	(.008)	(.007)	(.006)
Fraction w/ Piped Water at Home				.134***
				(.045)
HH Wealth	.152***	.152***	.122***	.072***
	(.007)	(.007)	(.007)	(.006)
Avg. Wealth	.168***	.167***	.092***	.101***
	(.019)	(.019)	(.016)	(.021)
Urban	.181***	.179***	.105***	.066
	(.056)	(.056)	(.033)	(.063)
Pop.Density(1990)	-.004	-.002	.099***	-.090**
	(.045)	(.045)	(.032)	(.044)
Tenure	.00006	.00005	-9.84e-06	.00008
	(.00007)	(.00007)	(.00005)	(.00006)
Avg. Tenure	.0007	.0007	.0009**	-.0003
	(.0006)	(.0006)	(.0004)	(.0005)
Female Head	.008**	.008**	.002	.007**
	(.003)	(.003)	(.002)	(.003)
Frac.Fem. HH Head	.004	.003	-.003	-.008
	(.040)	(.040)	(.025)	(.037)
Dist.to Cap.City (LogKM)	.026*	.027*	.021**	.011
	(.014)	(.014)	(.010)	(.010)
Dist.to River (LogKM)	.010***	.010***	.004**	.006**
	(.003)	(.003)	(.002)	(.003)
Coast.Comm.	-.077	-.078	.011	-.102*
	(.051)	(.052)	(.030)	(.053)
Obs.	69915	69915	69915	64543
R^2	.305	.305	.323	.191

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

Table 8: Rural Community Specification - 1KM Grid-size

	Base	Fractionalization	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.070***		.034**	.036*
	(.023)		(.014)	(.020)
Fractionalization		-.051***		
		(.018)		
Fraction w/ Piped Water at Home				.102***
				(.038)
Avg. Wealth	.345***	.344***	.216***	.187***
	(.022)	(.022)	(.018)	(.028)
Urban	.075	.077	.132**	-.138***
	(.061)	(.061)	(.053)	(.034)
Pop.Density(1990)	-.046	-.046	.023	-.095***
	(.032)	(.032)	(.024)	(.037)
Dist.to Cap.City (LogKM)	.025**	.026**	.015*	.015**
	(.010)	(.011)	(.009)	(.007)
Dist.to River (LogKM)	.009***	.009***	.003	.006**
	(.003)	(.003)	(.002)	(.003)
Coast.Comm.	-.061	-.061	.027	-.100***
	(.038)	(.038)	(.025)	(.032)
Elevation(km)	.004	.004	.008	-.0002
	(.016)	(.016)	(.012)	(.013)
Avg. Tenure	.0006	.0007	.001***	-.0003
	(.0005)	(.0005)	(.0004)	(.0004)
Frac.Fem. HH Head	-.014	-.013	-.019	-.006
	(.037)	(.037)	(.027)	(.033)
Avg. Age of Head	.002**	.002**	.002***	1.00e-05
	(.0009)	(.0009)	(.0006)	(.0007)
Obs.	3133	3133	3133	3112
R^2	.483	.483	.536	.364

Notes: The dependent variable is a binary measure indicating whether the households acquires drinking water from each type of piped source. Heteroskedasticity-robust standard errors (adjusted for intra-group correlation over 1km districts) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***,**, and * respectively. All regressions include country and province fixed effects and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.

Table 9: Rural Household Specification - 1KM Grid-Size

	Base	Fractionalization	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.080***		.043**	.041**
	(.024)		(.017)	(.020)
Fractionalization		-.049***		
		(.019)		
Hist.Share Own Eth.	-.003	.003	-.010	.009
	(.008)	(.008)	(.007)	(.007)
Fraction w/ Piped Water at Home				.246***
				(.046)
HH Wealth	.109***	.110***	.091***	.043***
	(.006)	(.006)	(.006)	(.004)
Avg. Wealth	.227***	.226***	.121***	.141***
	(.018)	(.018)	(.015)	(.022)
Urban	.093	.093	.147***	-.167***
	(.060)	(.060)	(.053)	(.030)
Pop.Density(1990)	-.047	-.046	.025	-.074***
	(.034)	(.034)	(.028)	(.028)
Tenure	.0001**	.0001**	1.00e-05	.0001**
	(.00006)	(.00006)	(.00004)	(.00005)
Avg. Tenure	.0004	.0004	.001***	-.0007
	(.0006)	(.0006)	(.0004)	(.0005)
Female Head	.004	.004	.0001	.004*
	(.003)	(.003)	(.002)	(.003)
Frac.Fem. HH Head	.004	.005	.0005	-.009
	(.037)	(.037)	(.024)	(.034)
Dist.to Cap.City (LogKM)	.023**	.023**	.013*	.016**
	(.010)	(.010)	(.008)	(.007)
Dist.to River (LogKM)	.010***	.010***	.003*	.007***
	(.003)	(.003)	(.002)	(.002)
Coast.Comm.	-.057	-.057	.023	-.090*
	(.050)	(.050)	(.029)	(.049)
Obs.	69376	69376	69376	64879
R^2	.273	.273	.294	.181

Notes: The dependent variable is a binary measure indicating whether the households acquires drinking water from each type of piped source. Heteroskedasticity-robust standard errors (adjusted for intra-group correlation over 1km districts) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***,**, and * respectively. All regressions include country and province fixed effects and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS.