Culture and Government

An Analysis of the Interaction Between Formal and Informal Institutions

by

Kenneth Jackson

B.Sc., The University of British Columbia, 1998

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Economics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July, 2010

Abstract

Institutions are the key to economic development. While the community of development economists appear to have converged on this assessment, our understanding of the specific role of institutions remains extremely limited. In particular, economists have avoided the question of how government institutions interact with cultural factors that vary around the world. The three papers in this thesis address this question with respect to ethnic identity and social norms.

Chapters 2 and 3 focus on the role of ethnic diversity in determining the effectiveness of governance in Sub-Saharan Africa. The existing literature has demonstrated a tenuous link between diversity and ineffective governance, but many questions remain. Chapter 2 demonstrates that ethnic diversity has a significant negative effect on the provision of piped drinking water and electricity across a large sample of countries from Sub-Saharan Africa. In addition, it is demonstrated that the geographic scale of diversity measurement is critical to the analysis. Local diversity is negatively associated with an inherently local good, piped water, whereas regional diversity is negatively associated with a good provided regionally, electricity. This link between the scale of provision and the measurement of diversity is a key finding that should inform future work on diversity.

Chapter 3 extends this analysis to consider why ethnic diversity matters for the provision of public goods. Including the two mechanisms suggested in the literature within a single model demonstrates that the key difference between them is the distribution of household access across ethnic groups. The empirical section of the chapter then provides evidence that ineffective governing institutions are the critical factor connecting local ethnic diversity to low levels of piped drinking water across Sub-Saharan Africa.

Finally, chapter 4 focuses on the dynamic interaction between government institutions and social norms of honesty. While government institutions can be effective in promoting social capital development, there exists a social capital trap. For countries with social capital below this level, advances in government institutions will not improve the social environment. Further, social divisions caused by ethnic heterogeneity or economic inequality may increase the minimum level of social capital necessary to escape this institutional trap.

Table of Contents

Appendices

List of Tables

List of Figures

Acknowledgements

First, I need to thank Kristie for her support over the years, both personal and financial. We have been through a lot, and I couldn't have done it without you. In addition, Reid and Claire have offered special incentives that were necessary to get this project completed.

I would also like to thank the faculty, staff and students at UBC. In particular, my thesis committee, Patrick Francois, Nathan Nunn and Mukesh Eswaran each provided critical guidance at various stages of my education. In addition, my colleagues in the PhD program, notably Chris Barrington-Leigh, Chris Bidner, Kelly Foley and Ben Sand, helped to create an academic and social environment in which I could thrive.

The individual chapters in the thesis have been greatly improved by comments of participants at the Canadian Economics Association meetings in Montreal (2006), the Northeast Universities Development Consortium Committee Conferences at Harvard University (2007), Boston University (2008) and Tufts University (2009), the Centre for the Study of African Economies Conference (2010) and numerous presentations at the University of British Columbia and Wilfrid Laurier University. Finally, the advice, support, and patience, of my new colleagues at Wilfrid Laurier University has been very much appreciated.

For Kristie, Reid and Claire.

Chapter 1

Introduction

An extensive literature has developed over the past 20 years, evaluating the importance of institutional quality in determining patterns of economic development. While this research has successfully demonstrated that institutions are critical, it is also clear that institutions mean different things to different people. The literature generally follows Douglass North in using a very general definition:

> Institutions are the humanly devised constraints that structure political, economic and social interaction. (North 1991, p.97)

However, given this definition, it is not clear how the different political, economic and social structures in the economy interact to create a coherent set of institutions. This thesis is focused on the analysis of the interaction between, on one side, two types of social structures, ethnic identity and social norms, and on the other side, government institutions.

The first two chapters are focused on the role that ethnic diversity plays in determining the ability of government structures to efficiently provide key public goods. Starting with Easterly and Levine (1997) and Alesina et. al. (1999), a growing literature has developed to estimate the impact of diversity on the supply of goods such as piped water or schooling. The general consensus from this literature is that diversity has a negative impact, but the robustness of these results are questionable. In particular, there is relatively little research directly focused on the relevant scale of ethnic diversity. In particular, is diversity within communities as important as diversity across communities.

Identifying the negative impacts of diversity leads directly into the second key question of this thesis. Why does diversity matter? If ethnic diversity matters, as suggested by the results in chapter 2, what mechanisms could generate such an outcome. Two very different primary mechanisms have been proposed in the literature, and while their implications at an aggregate level are very similar, their distributional impacts are quite different. In addition, the two mechanisms suggest very different policy responses, thus making it critical that we have a greater understanding of this problem.

In chapter 2, I evaluate the effect of ethnic diversity on the provision of two key publicly provided goods in Sub-Saharan Africa. A key element of this analysis is a focus on the geographic scale at which ethnic diversity is important. The existing literature,

primarily drawn from either developed countries, or very specific examples in the developing world, has been able to ignore the question of how to measure ethnic diversity. In this study, given the variation in political systems across the sample, it is necessary to measure diversity at a scale that is appropriate for the public good in question. I therefore propose that diversity should be measured at the scale of provision, as this scale is expected to be relevant in a large variety of circumstances.

In the empirical section of the paper, I measure the effect of ethnic diversity on the provision of water and electricity. Water is a critical local good, in that supply in one community is expected to have a relatively minor impact on the cost of providing the same good in a neighbouring community. In contrast, electricity is a regional good, and once provided to one area in a region, the cost of extending provision to neighbouring communities or districts falls quickly.

The results confirm the results in the existing literature that ethnic diversity has a significant impact on the delivery of publicly provided goods. This paper is the first in the literature to extend this result to a large set of developing countries, and demonstrates that resolving the inefficiencies surrounding public provision in a diverse environment would significantly improve the access to African communities to these critical goods. Further analysis demonstrates that the results are subject to strong non-linearities, in which small amounts of diversity are not important, but once communities are more diverse than average, further diversity has strong negative effects. One interpretation of this is that existing social institutions are designed for diversity levels ranging from homogeneity to moderate diversity, but further diversity places excessive stress on the existing institutions.

Further, I demonstrate that the scale of diversity measurement is critical to the analysis. When measured at scales inappropriate to the analysis, either regional diversity for water provision, or local diversity for electricity provision, the results would indicate that ethnic diversity has no effect on the provision of these goods. This study therefore provides evidence to support the notion that ethnic diversity is important, but that measurement of diversity must occur within the context of a well-founded theoretical model.

In chapter 3, I extend the study of ethnic diversity and public good provision to consider the question of how diversity affects the provision of public goods. In this paper, I develop a simple theoretical model that embeds both candidate mechanisms and derive a key empirically testable difference between them. Specifically, the household distribution of the public good will vary according to the relevance of each theory. Using data from 15 countries of Sub-Saharan Africa, I test the applicability of each mechanism in affecting the supply of piped water. By using household level data, I am able to evaluate the importance of local minority or majority status on access to piped drinking water.

The results indicate that while, as before, ethnic diversity has a significant impact on the provision of piped water systems, there is no consistent preference for local minority groups. In contrast with the underlying interpretation of some previous work, this result suggests that the underprovision of piped water to diverse communities in Africa represents a significant inefficiency that could be partially resolved through improvements to the governance system.

Chapter 4 focuses on an alternative informal institution, the social norms related to honesty and trust. In particular, I consider the interaction between the development of government enforced contracting institutions and the evolution of social norms of honesty. These different institutions, while having similar effects, follow very different processes of development. In addition, both types of institutions have been noted for their persistence over time. I follow a growing literature on the evolution of preferences and assume that social norms of honesty are subject to a process of cultural evolution. In contrast, government institutions adapt in response to a conscious process of optimization.

The interaction between these two institutions takes on three forms. Government institutions create incentives that generate economic activity that is necessary for the development of social norms. In addition, the enforcement of contracts reduces the advantage that accrues to opportunistic individuals. Each of these factors is complementary to the development of social capital, and thus higher levels of contracting institutions generates positive cultural evolution. However, government institutions can crowd out the civic engagement that drives the development of positive social norms, and therefore lead to their decline.

The combination of these three effects implies that government institutions will only be effective over an intermediate range of social capital levels. For countries with sufficiently weak social norms related to honesty, there is no contracting institution that can generate sufficient incentives for individuals to undertake risky production without crowding out civic engagement. In addition, there is an upper bound on the level of social capital, beyond which government institutions cannot encourage its development.

In addition, I consider the role of social divisions caused by ethnic diversity or economic inequality. To the extent that these elements weaken the effectiveness of civic engagement in generating information, the ability of government to support social capital development will be weakened further.

Collectively, this thesis demonstrates that, while formal institutions may be critical to long-term development, we cannot ignore their interaction with the informal institutions that have developed around the world. It is therefore critical to understand the landscape of informal institutions in either aiding or hindering the performance of all political and economic systems.

Chapter 2

A Question of Scale

2.1 Introduction

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

The Big Five interventions discussed by Sachs are:

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

The striking fact about this list is that all five interventions relate to the provision of quasi-public goods. Further, there is little productive government spending in any country that does not fall into this list[1]. This redefines the key issue in development economics. Why are the governments of developing countries unable to provide the basic services necessary to escape poverty?

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

This chapter analyzes the connection between ethnic diversity and public good provision in 13 Sub-Saharan African countries. In particular, I focus on the provision of

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

piped drinking water and electricity as fundamental goods that are commonly provided or managed by government, and are critical to economic development. While neither good is a pure public good, as individuals could be excluded from accessing the service, there are significant economies of scale in their provision, thus creating an environment in which government regulation can create significant benefits.

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

However, while water and electricity are similar in terms of under-provision, they are very different in their scale of operations. Water is a heavy, relatively low-value good, that is expensive to pipe across any significant distance. As demonstrated in Clark and Stevie (1981), the efficient scale of piped water provision in the United States as of the late 1970's is on the order of 10 miles. With the power needed for pumping electricity in much shorter supply in Sub-Saharan Africa, it seems safe to assume that the efficient scale in this survey is likely much smaller. Beyond this distance, the cost of piping water become extensive in comparison to the cost of establishing a new system. Electricity, in contrast, contains large economies of scale in production and the cost of transmitting electricity are relatively low. Filippini (1998) finds that electricity providers covering areas smaller than 2,500km^2 would be more efficient at larger scales. These economies of scale disappear for the larger providers that cover territories of 5-10,000km^2. While electricity can be provided at smaller scales, particularly through the use of small-scale generators, this tends to be a very inefficient method of producing electricity and is not widespread throughout Sub-Saharan Africa[2]. In addition, the future of electricity generation in Sub-Saharan Africa likely includes far greater use of large-scale hydroelectric projects, as only 7% of this resource has been tapped(Karekezi and Kimani 2002).

One can think of the "scale of provision" as the geographic area that encompasses the majority of the spillover benefits from a project, and thus water has a much smaller

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

scale of provision. In contrast, one can think of the "scale of governance" as the the geographic area that is represented by the government responsible for provision. The "scale of governance" can therefore vary substantially across countries, or even regions of countries. Further, the scale of governance is complicated by de jure and de facto governance. For example, a central government may retain de jure responsiblity for the provision of piped water, while the local community acquires de facto control when they build themselves a water system in the absence of government activity. In this way, the scale of governance is a far more complicated notion than the scale of provision. Regardless, the literature on ethnic diversity and public good provision has been able to sidestep the issue of how to measure ethnic diversity, as the two are generally coincident.

This fact corresponds well to the economic literature, starting with Tiebout (1956)[3], on the decentralization of government that has suggested that governments should decentralize responsibility to the lowest level of government that can capture the economies of scale necessary for provision. Panizza (1999) considers the theoretical and empirical implications of ethnic diversity for decentralization, and finds that decentralization should be, and is, greater in more diverse countries.

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

However, it is not always this clear. In particular, this paper is focused on Sub-Saharan Africa, where the decentralization of government has proceeded in a relatively haphazard fashion[4]. While many countries have devolved some responsibilities to local or district councils, these groups tend to have severely limited powers. Further, where national or district governments have de jure responsibility, these governments often abdicate their role, and leave the actual provision of local services to local governments[5]. As such, there may be an underlying pressure for the scale of governance to shift toward the scale of provision. With this in mind, I hypothesize that ethnic diversity, when measured at the scale of provision, will affect the provision of public goods. It is antici-

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

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

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

pated that village level diversity will impact the provision of piped drinking water, and provincial level diversity will effect the development of electricity networks.

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

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

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

If diversity creates differential demand, there is little question that the natural scale for diversity measurement is at the scale of provision, regardless of the scale of governance. A district government that oversees two communities that want different local public goods could, at least theoretically, provide those two communities with different goods. As an example, if two neighbouring villages are each comprised of a different ethnic group, speaking different languages, the problem of providing education is almost

irrelevant if each of the villages has their own school. The district does not need to worry about compromising on the public good - simply set up schools in each community in the preferred language.

In contrast, when the challenges of diversity are created by an inefficient government, it is more difficult to determine the channel through which ethnic diversity affects public good provision. However, as discussed in section 2.2, an inefficient government that provides funding in response to local lobbying will create a similar pattern of public good provision that depends on local diversity. In contrast, where spillovers across a region are large, local diversity will not substantially impact the provision of the public good.

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

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

2.2 Why Might the Scale of Diversity Matter?

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

However, it is far from certain that this is the path through which diversity matters. First, as discussed above, if the mechanism behind the effects of diversity is variation in demand, then the only critical scale is the scale over which the government can vary

supply. If the government can efficiently vary the supply of public goods at a sufficiently small scale, then there is no problem with diversity of demand.

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

Now consider a situation in which local contributions do not directly affect the provision of the public good. Instead, the local community contributes to a lobbying effort to gain funds from either a central government or foreign organization[6]. These funds are then used to build the public goods in question. The contributions of individual citizens to the lobbying effort is now the public good that is driven by the expectation of social sanctions, and diverse communities that are unable to generate equivalent resources for lobbying are left with low levels of the public good. While the scale of governance has changed, the role for diversity remains at the scale of provision.

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

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

This paper considers the effects of ethnic diversity on the provision of public goods across 13 countries in Sub-Saharan Africa, as shown in figure 2.3. Incorporating a moderate number of relatively similar countries into the analysis generates benefits in terms of the generality of the results, yet poses problems as the institutional details of public good provision vary across the sample.

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

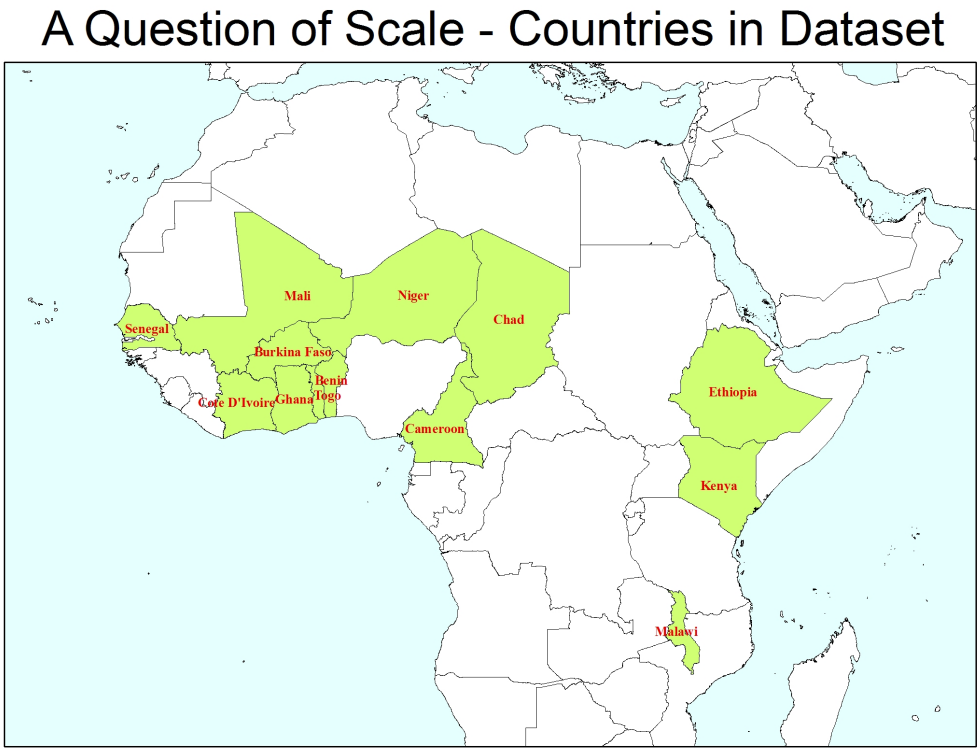


Figure 2.1: Question of Scale - Countries in Analysis

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

The situation with regards to water is far more complicated. As noted previously, the provision of drinking water is usually best left in the hands of a local government. Historically, countries in Africa maintained the highly centralized governments that were left behind by the colonial powers. Starting in the mid-1980's, there was a significant push to

decentralize the actions of government, particularly in some of the largely countries on the continent, such as Nigeria and Ethiopia(Olowu and Wunsch 2004). For example, in 1994, Ethiopia adopted a federalist constitution that devolved significant responsibility to provincial governments[7].

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

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

Thus, for all the talk of decentralization, Africa has lagged behind the rest of the world in terms of the decentralization of government expenditures[8].

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

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

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

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

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

In the literature on public good provision in Africa, there is essentially no discussion of local diversity and its effects on public goods provided by higher levels of government. Either of the two stories suggested above, either variation in demand, or variation in lobbying effort, may provide the reason for this. In each case, the connection between local diversity and public good provision would be secondary in nature. Either model suggests why diversity at a scale different from the scale of governance would not be obvious without careful statistical analysis. The fact that the results are not the

2.4 Data and Empirical Specification

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

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

2.4.1 Defining Local Communities

The data available from Measure DHS defines households by their survey cluster. One approach to this analysis would therefore measure ethnic diversity at the cluster level,

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

Alternatively, I use a narrower geographic classification, where I simply group clusters according to an arbitrary grid with squares at 1km or 5km scales. The 1km grid creates a total of 4,300 'communities', thus creating a potentially important difference from the previous methodology. However, the variation is primarily found in the urban areas. Table 2.4.1 shows the number of artificial communities created under each method, and the urban/rural[9] split.

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

Table 2.1: Local Communities

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

2.4.2 Public Goods

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

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

For drinking water, households are asked for the source of their drinking water and the time that it takes to get to their water source. Possible answers include piped into the home/compound, a public tap, a well or borehole, or from a stream or lake. I then code these responses into three broad categories. First, people that receive their drinking water from a piped source, either in their home or compound, or from a nearby public tap[10]. Second, households that receive their water from any improved source (ie. piped water, well or borehole). And third, those that collect water from an unimproved source, generally a lake, stream or rainwater. Finally, within the group that has access to piped water, I differentiate between those that access piped water at home, or at a public tap, as the delivery method may have different implications in regards to the relevance of ethnic diversity. There is no logical reason that the effects of ethnic diversity should fall equally on the very different goods that are provision of water at home and the provision of public taps. Generally, one might expect that there is less of a public good motivation involved in home delivery, and therefore ethnic diversity would have less of an effect in this case.

In keeping with the existing literature, I analyze the availability of each public good at the aggregate level. In chapter 3 I focus on access by individual households. At the community level, the ideal measure would be the existence and quality of the relevant infrastructure[11] . As this information is unavailable for a sample as large as this, this information must be inferred from the choices of households. Here I make the assumption that the fraction of households that access drinking water from different sources, all else equal, reflects the quality and availability of the infrastructure. If in two otherwise equal communities, one has 20% of the population accessing piped drinking water, while in the other 80% do, then the interpretation of this analysis assumes that the second community has a more functional infrastructure.

2.4.3 Measuring Ethnic Diversity

The diversity measures contained in this study rely on the ethnicity data reported in the measure DHS surveys in each country. In the survey each respondent is asked to select their ethnicity from a list of potential ethnicities with 'Other' always as an included option. As shown in table 2.2 the number of ethnicities listed varies across countries from a low of 7 in Togo to a high of 66 in the Cote D'Ivoire. In contrast to the existing literature on ethnic identity and local public goods, this paper combines data from countries with

[10]The definition of nearby is that the public tap is within 15 minutes of their home. The results are not sensitive to this distance

[11]If price data were available, this could also be incorporated into the dependent variable. In the key question, "How does ethnic diversity affect the availability of piped drinking water?", price data would be an alternative source of information on the state of water provision.

widely varying ethnic structures. An example of this is the contrast between Burkina Faso where the Mossi are the dominant group in size and wealth as opposed to Niger where the Haoussa are the largest group but are relatively poor.

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

Table 2.2: Ethnic Structure

Choosing an ethnicity from a list may have a tendency to force individuals to make a choice that does not entirely fit their identity. To the extent that this biases the diversity results, and assuming that this error is unrelated to the provision of public goods, the measured effect of diversity will tend to be biased toward zero. This problem may have less significance in Africa as opposed to other regions due to the historical prevalence of ethnic identity as a tool of colonial rule. One form of evidence for this is the strong correlation between the results of the household ethnicity survey and recent alternative data sources of ethnic diversity in Africa. Not only do national measures of diversity created using this dataset agree closely with published data in other sources, such as Alesina et al. (2003) and Fearon (2003), table 2.3 uses Kenya as an example and provides striking evidence of the similarities between the individual components of the different measures[12]. Where there is variation between the survey and published national level resources the survey tends to provide more options for survey respondents. This is important as local diversity may be significantly affected by the presence of groups that are locally concentrated and therefore have a relatively small presence nationally.

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

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

Table 2.3: Ethnicity in Kenya

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

fractionalization[13] that is drawn from the industrial organization literature. In keeping with this literature, I present the primary results here using fractionalization as the preferred measure of diversity. In chapter 3, I demonstrate that an alternative measure of ethnic diversity relevant for the provision of public goods is the size of the largest ethnic group. As such, the key results are repeated in the appendix using this as the measure of diversity. The results are not qualitatively different due to the very strong, negative, correlation between the two variables of < -0.95, as measured at any geographic scale[14].

The issue of migration is critical to assessing the impact of diversity. To the extent that people move to a community in response to the provision of basic services, and migrants tend to come from a more diverse background than the local population, then this will tend to increase the diversity of communities with public goods. Ignoring this migration will therefore understate the true impact of ethnic diversity. For this reason, there are three reasonable measures of diversity. First, diversity as measured over all

[13]Specifically, fractionalization is measured as

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

where p_e is the population share of ethnic group, e. This measure is drawn from the Herfindahl index that is used in assessing industrial concentration, however in chapter 3 I also demonstrate that it is the optimal measure of diversity under specific conditions.

[14]An alternative measure of diversity that is critical in the literature on ethnic diversity and conflict is the notion of polarization. The results of specifications containing a measure of polarization indicate no negative effects. As discussed in chapter 3, as polarization is linked primarily to the size of the second largest group in society, it is possible that polarization is linked to positive outcomes for public goods in some circumstances.

current residents. Second, diversity using only residents that report living in their current location for either their entire life, or at least 20 years[15]. Third, one could use the historical measure as an instrument for the current diversity. As is shown below, the results are not affected by this choice, primarily because the correlation between long-term diversity and current diversity is very high.

Theoretically, it is not clear which of the second two variables should be of greater importance. For example, as water systems may have existed for a significant time prior to the time of the survey, historical ethnic diversity may be precisely the relevant measure. Alternatively, recent construction or maintenance of the water systems could be the critical public goods problem, thus implying that current ethnic diversity is relevant, and that the correct solution is to instrument for ethnic diversity. As the results are not substantively affected[16], I do not pursue this question further, though the persistent effects of poor governance are an interesting direction for future research.

2.4.4 Additional Explanatory Variables

Sub-national measures of ethnic diversity are strongly correlated with local wealth. Communities that are the center of trade, or home to significant resources, tend to draw migrants from all groups seeking wealth. It is therefore critical to control for the economic resources available to a community. In addition, diverse communities may be systematically different from homogeneous communities in other ways. For example, diverse communities may be the result of recent migration that is associated with ineffective local government. To account for these effects I control for economic factors such as the average household wealth, urbanization and population density, and average household tenure[17].

The viability of piped drinking water or electricity are also directly related to the cost of installing infrastructure and the cost of alternatives. I therefore control for geographic factors including the distance to rivers, precipitation, elevation and the proximity to the ocean[18]. In particular, the distance to rivers and the level of precipitation may directly affect both the cost of installing piped drinking water and the cost of accessing alterna-

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

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

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

[18]A country is considered close to the coast if the distance to the coast is less than 20km. This is intended to capture variation in demand for water that arises from differences in economic structure, such as when a coastal town is more likely to be focused on fishing than farming. The results are not sensitive to this inclusion, or to changing the definition in various ways.

tive sources, such as rainwater or water directly from a river[19].

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

Table 2.4 contains sample statistics for those measures that do not differ by the scale of aggregation. Notably, 27% of households have access piped drinking water, with over half of them having access being in their home or compound, while electricity is available in 20% of households.

Table 2.4: Summary Statistics

Variable	Average	Stand. Dev.	Minimum	Maximum
Fraction w/ Piped Water	0.271	0.328	0	1
Fraction w/ Piped Water at Home	0.164	0.253	0	1
Fraction w/ Imp. Source	0.773	0.317	0	1
Fraction w/ Electricity	0.198	0.299	0	1
Distance to a River (km)	4.67	5.29	0.002	51.05
Elevation (m)	645.9	660.0	-1	3,690
Distance to Ocean (km)	501.9	352.5	0	1,698
Average Tenure	30.64	9.23	2.9	50*

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

Table 2.5 provides summary statistics on ethnic diversity as measured at local, district and provincial scales. Diversity is higher when measured over large regions, however there is significant ethnic diversity as measured at the local scale, with minority groups comprising 30% of the population.

Table 2.5: Diversity Statistics for Communities, Districts and Provinces

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

Notes: Standard deviations in parentheses.

[19]The data on the location of all geographic features are from ESRI data sources, and distances are based on the author's calculations.

The primary empirical specification throughout the paper is:

$$G = \beta_0 + \beta_1 \text{Homogeneity} + \beta_2 \text{Geog} + \beta_3 \text{Econ} + \beta_4 Ethnic + \beta_5 C + \beta_6 P \tag{2.1}$$

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

G is the fraction of households in the community that have access to the specific public good in question, either water or electricity. The estimates reported throughout the paper are the result of OLS regressions. There are no substantive differences when using either a tobit or grouped logit specification and these results are available upon request. Given the spatial nature of public good provision, it is also reasonable to be concerned about the degree of spatial correlation across the sample. For this reason, I also show that the results are not changed by accounting for spatial correlations.

2.5 Ethnic Diversity and the Provision of Piped Drinking Water

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

As discussed, the provision of clean drinking water is an inherently local problem. Water is a heavy commodity that is very costly to transport, thus the development of clean drinking water in one community provides very little benefit to neighbouring communities[20]. In this context, we should expect local diversity to affect provision. While piped water is not a classic public good, as there is no specific problem with restricting access to those that have paid for the service, the provision of piped water includes economies of scale that are much larger than, for example, providing wells, and therefore

[20]It is possible that this benefit is provided in the sense that people are able to access public taps in neighbouring communities. For this reason, throughout the analysis that is run at the local level, I account for clustering of the standard errors by district. In addition, spatial correlations are directly considered in the appendix.

the efficient provision of piped water will require the resolution of public good issues. In addition, given the necessity of water, it may be morally difficult to limit access to those households that have sufficient ability to pay.

A secondary issue related to piped water is the definition of household access. The DHS surveys record the primary source of drinking water for each household from a list that includes roughly 10 options. For my purposes, households are divided into four groups, those that access piped water in their home, those that access piped water at a public tap within 20 minutes of their home[21], those with any improved source of water (including piped water) and those with any other source of drinking water.

2.5.1 Piped Water and Local Measurement

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

As approximately one-third of the population that accesses piped water does so at a public tap, I measure the impact of ethnic diversity on four different dependent variables. The first is the fraction of households that have access to an improved source of drinking water. Next, I consider the fraction of households that access piped drinking water, either at home or a public tap. Finally, I differentiate between those households that access water at home, versus those that access water at a public tap. In column (4), the dependent variable is the fraction of households that access public taps, out of the population that do not have access to piped water at home.

The results of the primary local-level regressions are presented in table 2.6. The results demonstrate that ethnic diversity is not significantly correlated with access to improved sources of water, but is significantly correlated with access to piped water, both at home (column 3), and at public taps (column (4). The results suggest that the impact of moving from a very diverse community to a homogeneous community is on the same order of magnitude as moving from an urban area to a rural area, and this is true for both access at home, or at public taps. In both cases, the difference of being in a homogeneous community, versus a diverse community, is correlated with a 30% increase in the fraction of households that access a piped source. In particular, the elasticity of piped water access with respect to ethnic diversity is 0.12.

While the distinct difference in the results depending on the type of water source is striking, there are multiple reasons this could be true. First, it may be that the provision

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

Table 2.6: Piped Water and Ethnic Diversity - Local Measurement

	Improved	Piped	Home	Public
	(1)	(2)	(3)	(4)
Ethnic Fractionalization	.013	-.074***	-.042***	-.041**
	(.024)	(.020)	(.014)	(.019)
Fraction w/ Piped Water at Home				-.061
				(.040)
Avg. Wealth	.151***	.333***	.235***	.197***
	(.016)	(.015)	(.012)	(.017)
Urban	.021	.096***	.029*	.058***
	(.019)	(.021)	(.017)	(.021)
Pop.Density	-.013**	.008	.017*	-.008
	(.007)	(.006)	(.010)	(.011)
(ln)Distance to River	.023***	.006*	.003	.005
	(.005)	(.003)	(.002)	(.003)
Elevation(km)	.023	.0004	.011	-.004
	(.030)	(.015)	(.011)	(.017)
(ln)Distance to Capital City	.032**	.037***	.016*	.030***
	(.014)	(.010)	(.009)	(.010)
Coastal Community	.020	-.059*	.045*	-.159***
	(.026)	(.032)	(.026)	(.032)
Precipitation(mm/year)	.00002	9.87e-06	-6.66e-07	1.27e-06
	(.00004)	(.00003)	(.00002)	(.00002)
Avg. Tenure	.0005	.001**	.001***	.0004
	(.0007)	(.0006)	(.0005)	(.0006)
Obs.	2824	2824	2824	2806
R^2	.464	.754	.776	.619

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

of wells or boreholes have such low economies of scale, that their provision is unaffected by the ability of local groups to cooperate. An alternative, and consistent with the results in chapter 3, is that the more basic improved sources are the result of local actions, while piped sources of drinking water are the result of involvement by higher levels of government. In this case, it may be that diversity negatively affects the ability of communities to lobby higher levels of government, while having little effect on other forms of local cooperation.

The coefficients on the basic economic control variables in these regressions are consistent with existing theory. Wealthy, urban and long-standing communities are more likely to have piped water systems. Somewhat surprisingly communities closer to the national capital appear to be less likely to have access to piped water, but the statistical

significance of this result disappears if we exclude the communities within 20km of each capital[22]. It therefore appears to be picking up a very specific result, possibly related to the provision of piped water in large metropolitan areas.

Households in communities close to the coast are more likely to access water in their home, yet less likely to access water overall, due to a very low level of public taps. This may be related to variation in economic activity in coastal communities, but the issue is not explored further within the paper. The distance to a river is, occasionally, marginally statistically significant but not generally economically significant. As being closer to a river would reduce the costs of installing a piped water system, while also reducing the costs of the alternative source (water directly from the river), it seems that the increased costs of the alternative to a piped system are slightly more important than the reduced cost of implementing a piped system. There appears to be no correlation between the source of drinking water and precipitation levels. This could again be the result of competing factors, as rainfall would reduce the demand for piped water, while a wetter climate might make it cheaper to build a piped water system. Various alternative specifications of the geographic variables, including testing for non-linear effects, do not generate any different results.

The results presented here are robust to a variety of alternative specifications. Measuring diversity by the size of the largest group, or using current diversity, with or without instrumental variables, has no effect on the results. Neither does including alternative dimensions of diversity, such as wealth diversity or religious diversity. Altering the definition of local, from 1km to 20km squares, or focusing solely on rural communities is also inconsequential. Further, while there is the possibility that the wealth variable is affected by the presence of piped water, instrumenting for wealth does not have a substantive effect on the results. The results from the key robustness checks are contained in the appendix.

2.5.2 Piped Water and District Measurement

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

Table 2.7 contains the results from the regressions conducted at the district level. The first three columns contain the estimates for piped water access, in general, at home,

[22]Regression not shown, but available upon request.

Table 2.7: Piped Water and Ethnic Diversity - District Measurement

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

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

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

Column(4) demonstrates that the unit of observation is not important. In this specification, all variables included in column (1) are unchanged in column (4). However, the measure of diversity is changed from the district measure of fractionalization to a

local measure of fractionalization, averaged across the district[23]. The coefficient on this average measure of local diversity is therefore evidence that the significant change in column (1) is the measure of diversity, not the aggregation of observations.

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

2.5.3 Piped Water and Provincial Measurement

For completeness and comparison with the results for electricity, table 2.8 contains the estimates using the province as the unit of measurement. The expectation is that provincial level diversity should have no impact on the provision of piped drinking water.

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

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

2.6 Ethnic Diversity and the Provision of Electricity

While clean drinking water is a fundamental human need, electricity is a key factor in enhancing productivity and promoting development. However, within Sub-Saharan Africa, access to a consistent source of electricity is relatively rare. In this sample, approximately 20% of households have access to electricity, and there is large regional variation with almost half the provinces having less than 10% of the population with access to electricity. The dependent variable throughout the analysis of electricity is the fraction of households that report having access to electricity. As mentioned, this dataset

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

Table 2.8: Piped Water and Ethnic Diversity - Provincial Measurement

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

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

does not allow for the consideration of the quality of access, such as the reliability of electricity access.

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

2.6.1 Electricity and Local Measurement

Given the economies of scale associated with the production of electricity, the hypothesis of this paper is that local diversity will not affect household access. In the first 3 columns of table 2.9, the results confirm this hypothesis. Local fractionalization does not impact household access to electricity, in general, or specifically in rural or urban areas. Wealth, urbanization and the average tenure in the community all have logical and statistically significant impacts. In addition, as with piped drinking water, communities near the coast appear to be significantly different than other communities. For electricity, they are more likely than others to have electricity, though this is seems to be only important for rural households. As with the variation in the case of water this is likely the result of differences in economic structure associated with coastal communities.

Table 2.9: Electricity and Ethnic Diversity - Local Measurement

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

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

In column (4), I do not control for provincial fixed effects but instead include a mea-

sure of provincial diversity. The results indicate that communities in homogeneous provinces have a higher likelihood of electricity access than an otherwise identical community in a heterogeneous province. As with the results for piped water, this coincides with the hypothesis that regional diversity will affect electricity provision. This result is essentially unchanged if local diversity is controlled for in the same regression.

2.6.2 Electricity and District Measurement

Table 2.10 contains the results for regressions conducted with the district as the unit of observation. Columns (1) and (2) incorporate measures of local and district level homogeneity and show that diversity does not have a negative impact. In fact, measured at the district level, fractionalization is shown to have an unexpected statistically significant positive effect. I discuss this result further below.

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

2.6.3 Electricity and Provincial Measurement

Table 2.11 contains the analysis conducted at the provincial level for electricity. The results indicate that ethnic homogeneity is associated with increased rates of electricity access. Moving from a very diverse province to an otherwise similar homogeneous province will reduce the fraction of households with access to electricity by approximately 40%. The elasticity of electricity with respect to ethnic fractionalization is 0.23, slightly larger than the elasticity reported for piped water. As shown in columns (2) and (3), diversity measured at the district scale is also a significant factor. In column (4), we find that local homogeneity is not important in determining access to electricity.

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

Table 2.10: Electricity and Ethnic Diversity - District Measurement

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

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

The results for electricity contained in tables 2.10 and 2.11 appear contradictory at first glance. In column (1) of table 2.10, district level fractionalization appears to have a positive effect on electrification, while in column (3) of table 2.11, the same variable appears to have a negative effect. However, these regressions are reporting the effects of diversity across very different samples. In the district specification, table 2.10, the variation being analyzed is between districts within the same province. The results therefore may indicate a secondary, positive, effect of district-level diversity that appears once the overall provincial diversity is controlled for.

In contrast, column (3) of table 2.11 considers the impact of average district diversity on the electricity access rates across provinces. In this regression, district-level diversity appears to have a negative impact. This reflects the negative, first-order effect of regional

Table 2.11: Electricity and Ethnic Diversity - Provincial Measurement

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

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

diversity, measured here at the district level.

2.7 Spatial Correlation

The fact that the provision of piped water and electricity involve significant economies of scale suggest that there may be a significant issue with spatial correlation in the sample. If the choice of the sampling grid, where sampling groups are collected into "communities" is correct for the public good in question, this problem should be minimized. However, it is difficult to be certain if the sampling level is perfectly chosen. At the local level, the OLS results further attempt to partially account for this effect by clustering standard errors at the district level, however this may not be completely ef-

fective. Second, it is possible that there are additional variables, particularly geographic factors, that are not included in the regression and this will lead to spatial correlations in the error terms, and therefore bias the results. In this section, I show that the key results for both piped water and electricity are not substantively affected by moving the analysis to a spatial regression model.

In the analysis of piped water, to reduce the computational load in the analysis, I sample half the rural communities[24]. Table 2.12 shows the results of the analysis of this reduced sample. Column (1) shows the OLS result on the reduced sample, which shows no difference from the previous analysis. Column (2) controls for spatial lags, which would be important if the provision of piped water in one community has an effect on nearby communities. As expected under the assumption that the provision of piped drinking water has only local effects, the results indicate that this form of spatial correlation is not critical to our finding that ethnic diversity has a negative effect on the provision of piped water. In contrast, column (3) considers the role of a spatial error model, which would be more relevant if there were omitted variables that were important, and spatially correlated. Again, the results indicate no substantive change to the measured effect of diversity.

The one result that has changed is the effect of distance from the capital city. In previous regressions, this variable had been shown to be positively correlated with access to piped water, suggesting that being further from the capital city may have beneficial effects. Here, controlling for the spatial correlations, we see the opposite, and more intuitive result, that communities close to the national capital are more likely to have access to piped drinking water.

Similarly, table 2.13 shows the results for electricity at the provincial level. As with the effects of local diversity on piped water provision, the analysis suggests that spatial correlations are not critical to the measurement of the effects of provincial diversity on the provision of electricity.

2.8 Extensions

The two key results suggest that local diversity is negatively associated with the provision of piped water, while provincial level diversity is negatively associated with the provision of electricity. In this section, I explore these key results in more depth, with a focus on the possibility of non-linear relationships between diversity and public good provision, alternative forms of diversity or variation across countries or groups of countries. I start by analyzing the potential for different effects in water provision in

[24]Resampling does not change the result, nor does keeping urban households within the sample.

Table 2.12: Spatial Regression and Water

	Base	Lag	Error
	(1)	(2)	(3)
Fractionalization	-.066**	-.068**	-.054**
	(.026)	(.027)	(.028)
Avg. Wealth	.326***	.322***	.319***
	(.022)	(.017)	(.016)
Urban			
Pop.Density(1990)	-.036	-.064	-.064
	(.069)	(.060)	(.065)
Dist.to River (LogKM)	.008	.006	.005
	(.005)	(.005)	(.005)
Elevation(km)	.0003	.005	.022
	(.016)	(.015)	(.018)
Dist.to Cap.City (LogKM)	-.007	.011	-.037***
	(.010)	(.011)	(.009)
Coast.Comm.	.051	.067*	.032
	(.056)	(.035)	(.038)
Precipitation (mm/year)s	-9.75e-06	-.00002	-1.35e-06
	(.00002)	(.00002)	(.00002)
Avg. Tenure			
Obs.	1186	1186	1186
R^2	.365		

Notes: Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***,**, and * respectively. All regressions include country and ethnic group controls.

2.8.1 Rural-Urban Differences and Piped Water

The regression controls for differences between urban and rural communities, however it is possible that the effects of ethnic diversity are different. In particular, the forms of governance that have developed in urban communities may be different from the primarily traditional forms of governance that prevail in rural areas. To test for differences between the effects of ethnic diversity in rural areas vs. urban areas, I include an interaction effect that takes the value of diversity if the community is primarily urban. This regression is shown in column (1) of table 2.14.

As shown, there is no evidence for a statistically significant difference between the effects of ethnic diversity in urban and rural areas. In addition, the effect of diversity in urban areas is shown to be significant and negative, as a Wald test rejects the notion that the sum of the two fractionalization measures is 0, with a p-value of 0.027. Though there are no apparent differences between urban and rural areas, it is reasonable to as-

Table 2.13: Spatial Regression and Electricity

	Base	Lag	Error
	(1)	(2)	(3)
Fractionalization	-.090***	-.070**	-.093***
	(.028)	(.028)	(.028)
Avg. Wealth	.349***	.327***	.347***
	(.056)	(.035)	(.036)
Urban	-.141	-.078	-.122*
	(.093)	(.067)	(.067)
Pop.Density(1990)	.034***	.033***	.030***
	(.006)	(.006)	(.006)
Dist.to River (LogKM)	.018	.005	.006
	(.015)	(.010)	(.010)
Elevation(km)	-.011	-.009	-.014
	(.019)	(.018)	(.018)
Dist.to Cap.City (LogKM)	.018	.019**	.012
	(.013)	(.008)	(.008)
Precipitation (mm/year)	.00002	7.85e-06	.00002
	(.00002)	(.00002)	(.00002)
Coast.Comm.	-.026	-.010	-.022
	(.026)	(.029)	(.030)
Avg. Tenure	-.0008	.0006	.001
	(.001)	(.001)	(.001)
Obs.	114	114	114
R^2	.967		

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

sume that the rural districts are a more homogeneous subgroup than urban areas, and the diversity in average wealth is much higher between urban communities than rural communities. In addition, the problems of in-migration are expected to be much lower in rural areas, as the average household has lived in the community far longer. As such, columns (2)-(4) repeat the analysis from the related columns in table 2.6, but restricting the analysis to primarily rural communities. As shown, there are no substantive differences from the previous analysis[25].

2.8.2 Non-Linear Specification

There is no particular reason to believe that the effects of ethnic diversity are linear, or affect all communities equally. In particular, it is conceivable that introducing small

[25]Communities are classified as primarily rural if over half the households are listed as rural. The results are unchanged if we limit the analysis to communities that are completely rural.

Table 2.14: Rural-Urban Differences and Piped Water

	UrbanRural	Piped	Home	Public
	(1)	(2)	(3)	(4)
Ethnic Fractionalization	-.070***	-.063***	-.023**	-.041**
	(.019)	(.019)	(.012)	(.017)
Fractionalization (Urban)	-.020			
	(.038)			
Fraction w/ Piped Water at Home				.054
				(.040)
Avg. Wealth	.334***	.325***	.202***	.164***
	(.015)	(.020)	(.017)	(.022)
Urban	.103***	.178***	.102***	.073
	(.026)	(.057)	(.033)	(.059)
Pop.Density	.008	-.026	.127**	-.195***
	(.006)	(.059)	(.051)	(.059)
(ln)Distance to River	.006*	.008**	.003	.005*
	(.003)	(.003)	(.002)	(.003)
Elevation(km)	.0008	.008	.023**	-.009
	(.015)	(.014)	(.010)	(.013)
Precipitation(mm/year)	1.00e-05	5.51e-06	-7.32e-06	8.78e-06
	(.00002)	(.00002)	(.00002)	(.00002)
(ln)Distance to Capital City	.037***	.033**	.023**	.016
	(.010)	(.013)	(.010)	(.010)
Coastal Community	-.059*	-.062	.022	-.094***
	(.032)	(.045)	(.032)	(.035)
Avg. Tenure	.002**	.0009	.0008*	.00006
	(.0006)	(.0006)	(.0004)	(.0005)
Obs.	2824	2396	2396	2388
R^2	.754	.509	.554	.393

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

amounts of ethnic diversity to homogeneous communities fundamentally changes the governance process and has important effects. Alternatively, it could be that the local governance structures in African communities are designed to accommodate relatively small amounts of diversity but large amounts of diversity would require significant changes to the system of governance. To test for this effect, I include higher-order polynomials or piecewise linear versions of the diversity measure into the regressions in tables 2.15 and 2.16.

Column (1) includes a quadratic term of fractionalization. The results show that the effects of ethnic diversity follow an inverted-U shaped pattern, with small amounts of diversity possibly being beneficial, while larger amounts of diversity have significant

Table 2.15: Non-Linear Specification - Local Analysis

	Quadratic	Piecewise	Improved	Home	Public
	(1)	(2)	(3)	(4)	(5)
Local Fractionalization	.099*	.079*	.163**	.003	.082*
	(.052)	(.044)	(.063)	(.034)	(.043)
Fract. Squared	-.249***				
	(.079)				
Fract. Above Avg.		-.240***	-.235***	-.069	-.195***
		(.072)	(.089)	(.055)	(.069)
Fraction w/ Piped Water at Home					-.063
					(.040)
Avg. Wealth	.334***	.333***	.151***	.235***	.198***
	(.015)	(.015)	(.016)	(.012)	(.017)
Urban	.101***	.099***	.024	.030*	.060***
	(.021)	(.021)	(.019)	(.016)	(.021)
Pop.Density	.009	.009	-.012*	.017*	-.008
	(.006)	(.006)	(.007)	(.010)	(.011)
(ln)Distance to River	.006*	.006*	.023***	.003	.005
	(.003)	(.003)	(.005)	(.002)	(.003)
Elevation(km)	.003	.003	.025	.011	-.002
	(.015)	(.015)	(.030)	(.011)	(.017)
(ln)Distance to Capital City	.036***	.036***	.031**	.016*	.029***
	(.010)	(.010)	(.013)	(.009)	(.010)
Coastal Community	-.057*	-.059*	.020	.045*	-.159***
	(.033)	(.032)	(.026)	(.026)	(.032)
Precipitation (mm/year)	1.00e-05	1.00e-05	.00002	4.00e-08	3.35e-06
	(.00002)	(.00003)	(.00004)	(.00002)	(.00002)
Avg. Tenure	.002**	.002**	.0005	.001***	.0004
	(.0006)	(.0006)	(.0007)	(.0005)	(.0006)
Obs.	2824	2824	2824	2824	2806
R^2	.755	.755	.465	.776	.621

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

negative effects. From the point estimates, the turning point is at a fractionalization that is very close to the average diversity in society, 0.267. The marginal effects are more clearly seen in column (2), which estimates the differential linear effect of diversity greater than average. Increasing fractionalization up to the average is associated with increases in piped drinking water, while increases beyond the average are associated with reductions in piped water[26]. Columns (4) and (5) show a similar pattern for access

[26]A Wald test rejects the hypothesis that the sum of the two fractionalization coefficients is zero.

to water at home, or at a public tap, though the results with respect to water at home are not significant[27].

In addition, column (3) demonstrates that the previous result showing no relationship between diversity and improved sources of piped water may be the result of the linear specification. In particular, it appears that there is a strong positive relationship between small amounts of diversity and access to improved sources of piped water, while increasing diversity past the average reduces access. The results in this section indicating that small amounts of diversity are positively associated with public good access is deserving of future study.

Including additional non-linear controls does not offer further enlightIn addition, the previous analysis with respect to electricity is unchanged, and the inclusion of various non-linear terms does not indicate any statistically, or economically, significant relationship between local ethnic diversity and electricity provision.

The results for electricity mirror those for piped water. The overall negative relationship between ethnic diversity and electricity provision is the result of a large, and significant, effect for provinces that are more diverse than average, while small amounts of diversity are actually associated with increases in public good provision.

2.8.3 Other Forms of Diversity

The focus to this point has been on the effects of ethnic diversity. However, the mechanisms suggested for why diversity might matter are not specific to the form of diversity. While previous research[28] has indicated that less durable forms of diversity, such as religious diversity, may be less important in economic contexts, it is important to consider their effects. Separately, it is unclear how diversity in economic outcomes, measured here as the standard deviation in wealth in the community, will affect public good provision. It may have identical effects to those considered for ethnic diversity, as individuals are not strongly connected, and therefore coordination is weaker. Alternatively, as suggested by Mancur Olson (1965), a diversity of wealth may imply that a small number of people have a heightened incentive to provide public goods for themselves, and once provided these become available to others in the community. For our purposes, this would imply that if someone is rich enough to build themselves a piped water system, the benefits would flow to other members of the community. In addition, ethnic diversity is correlated with wealth diversity and religious diversity, and ignoring their effects may

[27] We can however reject the hypothesis that the two measures sum to zero, with a p-value on the Wald test of 0.01.

[28] At the cross-country level, Alesina et al. (2003) demonstrates that religious diversity has no effect on economic growth. In addition, Caselli and Coleman (2010) provide a theoretical justification for why religious diversity may not be a significant factor driving conflict.

Table 2.16: Non-Linear Specification - Provincial Analysis

	WaterQuad	WaterPiece	ElecQuad	ElecPiece
	(1)	(2)	(3)	(4)
Local Fractionalization	.030	-.034	.206*	-.040
	(.118)	(.125)	(.114)	(.111)
Fract. Squared	-.065		-.301***	
	(.125)		(.116)	
Fract. Above Avg.		-.001		-.059
		(.153)		(.135)
Avg. Wealth	.344***	.350***	.321***	.347***
	(.058)	(.058)	(.047)	(.054)
Urban	.013	.001	-.087	-.138
	(.123)	(.122)	(.079)	(.090)
Pop.Density	.028***	.028***	.033***	.034***
	(.007)	(.007)	(.005)	(.006)
(ln)Distance to River	.059***	.058***	.022	.018
	(.017)	(.017)	(.014)	(.015)
Elevation(km)	-.012	-.013	-.009	-.010
	(.032)	(.033)	(.021)	(.020)
(ln)Distance to Capital City	.033***	.034***	.014	.018
	(.011)	(.011)	(.011)	(.012)
Coastal Community	-.104***	-.104***	-.024	-.025
	(.040)	(.039)	(.025)	(.026)
Precipitation(mm/year)	3.40e-06	2.09e-06	.00003	.00002
	(.00003)	(.00003)	(.00002)	(.00002)
Avg. Tenure	.002	.002	-.0004	-.0007
	(.002)	(.002)	(.001)	(.001)
Obs.	114	114	114	114
R^2	.937	.937	.969	.967

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

therefore introduce omitted variable bias. The results of these tests are shown in tables 2.17 and 2.18.

As demonstrated in table 2.17, the results do not change with the inclusion of alternative diversity measures. Columns (1) and (2) control for religious fractionalization, where the definition of a religious group is based on major religion (ie. Christianity, Muslim, Traditional), or denominational groups (ie. Catholic and Protestant are separate groups). Religious fractionalization has no impact, regardless of the measure. Column (3) includes the standard deviation of wealth, as measured at the community level, and does not have a statistically significant impact, and doesn't affect the measured corre-

Table 2.17: Alternate Forms of Diversity and Piped Water

	ReligionA	ReligionB	Wealth	Interaction
	(1)	(2)	(3)	(4)
Ethnic Fractionalization	-.077***	-.079***	-.076***	-.039
	(.020)	(.020)	(.020)	(.027)
Relig. Fract. (Major)	.018			
	(.024)			
Relig. Fract. (Denom)		.033		
		(.022)		
Wealth Diversity			.031	.058*
			(.028)	(.033)
Diversity Interaction				-.077
				(.054)
Avg. Wealth	.334***	.333***	.322***	.324***
	(.015)	(.015)	(.018)	(.018)
Urban	.095***	.095***	.091***	.094***
	(.021)	(.021)	(.021)	(.021)
Pop.Density	.008	.008	.009	.010*
	(.006)	(.006)	(.006)	(.006)
(ln)Distance to River	.006*	.006*	.006*	.006*
	(.003)	(.003)	(.003)	(.003)
Elevation(km)	.0007	.0004	.0009	.002
	(.015)	(.015)	(.015)	(.015)
(ln)Distance to Capital City	.037***	.037***	.037***	.037***
	(.010)	(.010)	(.010)	(.010)
Coastal Community	-.059*	-.060*	-.058*	-.058*
	(.032)	(.032)	(.033)	(.033)
Precipitation(mm/year)	1.00e-05	9.48e-06	1.00e-05	9.23e-06
	(.00003)	(.00002)	(.00003)	(.00002)
Avg. Tenure	.001**	.001**	.001**	.001**
	(.0006)	(.0006)	(.0006)	(.0006)
Obs.	2824	2824	2824	2824
R^2	.754	.754	.754	.754

Notes: The dependent variable in all regressions is access to piped water. Heteroskedasticity-robust standard errors (adjusted for intra-group correlation at the district level) are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***,**, and * respectively. All regressions include country, province and ethnic group controls and observations are weighted by the sum of local household weights provided by Measure DHS. In addition, I include religious group controls in the regressions that include religious diversity.

lation with ethnic fractionalization. Column (4) includes an interaction effect between ethnic diversity and wealth diversity. The results suggest that ethnic diversity may be more critical in the presence of wealth diversity, but the results are not statistically significant.

Table 2.18 extends the analysis of electricity provision to consider the role of other forms of diversity. The results in columns (1) and (2) indicate that there is no significant

Table 2.18: Alternate Forms of Diversity and Electricity

	ReligionA	ReligionB	Wealth	Interaction
	(1)	(2)	(3)	(4)
Ethnic Fractionalization	-.053**	-.055**	-.053**	.062
	(.022)	(.023)	(.024)	(.085)
Relig. Fract.(Major)	-.044			
	(.043)			
Relig. Fract.(Denom)		-.015		
		(.047)		
Wealth Diversity			-.103***	.016
			(.038)	(.087)
Diversity Interaction				-.193
				(.138)
Avg. Wealth	.381***	.383***	.410***	.399***
	(.052)	(.053)	(.051)	(.047)
Urban	-.202**	-.204**	-.160**	-.138*
	(.095)	(.096)	(.075)	(.070)
Pop.Density	.037***	.037***	.029***	.027***
	(.005)	(.005)	(.005)	(.006)
(ln)Distance to River	.017	.016	.011	.010
	(.014)	(.015)	(.012)	(.011)
Elevation(km)	-.009	-.010	.0007	.003
	(.031)	(.031)	(.029)	(.029)
(ln)Distance to Capital City	.028**	.027**	.034***	.028**
	(.012)	(.012)	(.012)	(.011)
Precipitation(mm/year)	-1.00e-05	-1.00e-05	-.00002	-9.79e-06
	(.00002)	(.00002)	(.00002)	(.00002)
Coastal Community	-.039	-.037	-.026	-.022
	(.028)	(.028)	(.025)	(.025)
Avg. Tenure	.001	.002	.002	.002
	(.001)	(.001)	(.001)	(.001)
Obs.	111	111	111	111
R^2	.974	.974	.976	.977

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

role for religious diversity within our sample of countries. In contrast, column (3) shows a strong negative correlation between the diversity of wealth and electricity provision. The exact causal links are difficult to assess, and are worthy of further research, but it does not affect the measured relationship between ethnic diversity and electricity provision. In column (4), and similar to the analysis of piped water, ethnic diversity may be more strongly correlated with poor outcomes in the presence of greater diversity of wealth, but

the results are again not statistically significant.

Both results suggest that the interplay between ethnic diversity and wealth diversity is an important area for future study. One issue that has arisen in the literature on social capital is the presence of cross-cutting links. One direction for further work is to consider the specific links between economic circumstances and ethnic diversity. In particular, if employment choices creates the opportunity for social interaction across ethnic groups then it might be expected that diversity would have larger effects where wealth diversity and ethnic diversity move together.

2.8.4 Cross-Country Variation

The results presented here represent the average effect of ethnic diversity across a wide range of countries. It is reasonable that this hides significant variation in the effects of diversity across different countries, or within more specific subgroups. Two natural subgroupings would be either geographic (ie. West Africa vs. non-West Africa) or colonial (ie. French vs. non-French). In the sample at hand, these two groups are almost coincident, with Ghana being the lone non-French country in West Africa, while Chad and Cameroon were French colonies not part of West Africa.

An alternative possibility is that the underlying governance structures of the countries are important. As suggested in Easterly (2001), institutions may be able to reduce the effects of ethnic differences. To explore this, I use the rule of law measure from the 1996 World Bank Governance Indicators, and measure the relative effects of diversity in countries with a high or low rule of law. To simplify the analysis, I normalize the rule of law measure to a 0-1 scale. The weakest rule of law in the sample is in Cameroon, measured at 0.177, and the strongest is in Benin, measured at 0.452.

Table 2.19 contains the results for piped water measured at the local scale. In column (1), one can see that there are differences in the measured effects of diversity, but the point estimates are negative in 10 out of 13 countries, and significantly negative (at least at the 10% level) in 6. Using a Wald-test we can reject, at the 5% level, the hypothesis that the effects of ethnic diversity are identical in all 13 countries. In columns (2), (3), and (4), the point estimates suggest that diversity is more of a problem for water in French Africa, West Africa and countries with poor governance, but there is little power to actually resolve the issue. Further analysis of the differences across countries is beyond the scope of this study, and in particular would require observations across a wider group of countries.

Table 2.20 repeats the identical exercise for electricity. Columns (1), (2), and (3) collectively suggest that diversity is more of a problem in formerly French countries that are not in West Africa, which is the result of the measured effects in Cameroon and Chad

shown in column (5). Column (4) is very similar to the regression in the previous table, where diversity has a negative effect in countries with a weak rule of law, and that effect might diminish as the rule of law improves. With only 13 countries in the sample, these results should only be taken as suggestive evidence that is consistent with previous findings, and as a direction for future study. In column (5), again the results vary across countries, though the country specific effect is negative in 10 out of 13 countries, and we find significant negative effects in 4 individual countries. Again, we can reject the hypothesis that the effects are the same in all countries.

The results in this subsection reinforce the idea that the results measured in the main section of the paper represent an average effect across countries. The important result presented here is that the average effect varies significantly depending on the scale of measurement, rather than the absolute magnitude of the effect in any single country. Further analysis, not reported, shows no significant differences in the main results based on the year of observation.

2.9 Conclusion

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

Regardless of the underlying mechanism, ethnic diversity at the scale of provision is a potentially limiting factor. The alternative scale that should be considered is the scale of governance. However, in many regimes in Africa, it is difficult to define the scale of governance with sufficient accuracy to allow for accurate analysis. This is particularly true in the context of a multi-country study such as this one. It is therefore important that future research expands on these findings and considers the role of ethnic diversity at the scale of governance directly.

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

measures are likely correlated they are fundamentally different and the results are going to be strongly affected by the choice.

There are four general results from this paper that should be of particular interest. First, ethnic diversity has a statistically and economically significant effect on the provision of both piped water and electricity. Measured at the scale of provision, moving from a diverse community (province) to a homogeneous one is associated with a 25-35% increase in access to piped water (electricity) %. With regard to the provision of water, this magnitude is similar to moving from a rural community to an urban one. Resolving the governance issues surrounding ethnic diversity would therefore have a significant impact on the development prospects in Sub-Saharan Africa. In addition, given the difficulty in assessing the key dimensions of diversity, the results reported here should be taken as a lower-bound on the impact of ethnic diversity.

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

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

Finally, the results provide suggestive evidence that the effects of ethnic diversity are non-linear, and do not appear to be uniform across Africa. Small amounts of diversity, up to roughly the average diversity level in society is generally insignificant, while increasing diversity beyond this point is associated with large reductions in the supply of public goods. In addition, while the results are not statistically significant, there is reason to believe that the effects of ethnic diversity are exacerbated by poor governance or wealth diversity, and these results deserve further study.

These results should not be taken as suggesting that the scale of governance is not important. There are two reasons for this caution. First, while most African countries are heavily centralized, diversity at the national level could not be included within this study. Previous research has demonstrated that ethnic diversity, as measured at a national scale, is strongly linked to poor public good provision. In addition, when the official responsibility for governance falls to a disinterested central government, local govern-

ments may step in to ensure provision. In this sense, the official record likely overstates the degree of centralization in countries with ineffective national governments. As such, the distinction between scale of provision and scale of governance is difficult to determine.

Table 2.19: Cross-Country Differences - Water

	All	French	West	Rule
	(1)	(2)	(3)	(4)
Fract. Benin	.038			
	(.092)			
Fract. Burkina Faso	.072*			
	(.038)			
Fract. Cote D'Ivoire	-.198**			
	(.095)			
Fract. Cameroon	-.093*			
	(.056)			
Fract. Chad	-.111**			
	(.051)			
Fract. Ghana	-.100**			
	(.050)			
Fract. Kenya	-.019			
	(.055)			
Fract. Mali	-.126**			
	(.061)			
Fract. Malawi	.003			
	(.038)			
Fract. Ethiopia	-.051			
	(.034)			
Fract. Togo	-.177			
	(.114)			
Fract. Niger	-.102			
	(.070)			
Fract. Senegal	-.156*			
	(.093)			
Fractionalization		-.034	-.046*	-.169**
		(.023)	(.024)	(.076)
Diversity (French)		-.065*		
		(.035)		
Diversity (West Afr.)			-.054	
			(.037)	
Diversity*Rule of Law				.275
				(.211)
Obs.	2824	2824	2824	2824
R^2	.756	.754	.754	.754

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

Table 2.20: Cross-Country Differences - Electricity

	French	West	FrenchWest	RuleofLaw	Countries
	(1)	(2)	(3)	(4)	(5)
Local Fractionalization	.010	-.111***	-.055	-.238*	
	(.035)	(.040)	(.038)	(.128)	
Diversity (French)	-.093		-.214***		
	(.062)		(.074)		
Diversity (West Afr.)		.061	.195***		
		(.050)	(.070)		
Diversity*Rule of Law				.441	
				(.340)	
Fract. Benin					.063
					(.047)
Fract. Burkina Faso					-.044
					(.050)
Fract. Cote D'Ivoire					-.173*
					(.091)
Fract. Cameroon					-.448***
					(.147)
Fract. Chad					-.191**
					(.095)
Fract. Ghana					-.117
					(.120)
Fract. Kenya					-.048
					(.050)
Fract. Mali					-.054
					(.123)
Fract. Malawi					-.402***
					(.107)
Fract. Ethiopia					-.057
					(.043)
Fract. Togo					.003
					(.149)
Fract. Niger					.095
					(.088)
Fract. Senegal					-.120
					(.184)
Obs.	114	114	114	114	114
R^2	.938	.967	.97	.967	.964

Notes: The dependent variable in all regressions is the fraction of the population with access to electricity. Heteroskedasticity-robust standard errors are included in parentheses. Variables significant at the 1%, 5% and 10% levels are noted by ***,**, and * respectively. All regressions include country dummies and observations are weighted by the sum of local household weights provided by Measure DHS.

Chapter 3

Why Does Diversity Matter?

3.1 Introduction

A reliable supply of clean drinking water is a fundamental human need. The Millennium Development Goals (MDG) recognize this need, and strive for halving the fraction of households that do not have access to an improved water source. Overall, the fraction of households around the world with an improved water source increasing from 76% to 86% between 1990 and 2007, and only needing to increase access to 88% by 2015 to meet the MDG. 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[29].

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. As was demonstrated in chapter 2, the problems created by ethnic diversity do have significant impacts on the provision of piped water within Sub-Saharan Africa.

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),

[29]The fraction in our sample that access water at an improved source is 82%, which reflects differences in the definition of "improved" across different datasets.

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[30] 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 an expanded version of the dataset from chapter 2 containing over 100,000

[30] 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.

households across 15 countries[31] of Sub-Saharan Africa, I demonstrate that households in diverse communities are less likely to have access to piped water, yet being a member of a locally dominant group is not associated with preferential access. 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, regardless of the household's minority or majority status. The reduction in access to water in diverse communities, in comparison to homogeneous ones, is approximately the same as moving from an urban area to a rural area.

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[32]. 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[33].

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

[31]The Central African Republic and Namibia are added to the dataset used in chapter 2.

[32]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

[33]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 3.2.

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.

While most published papers on the issue of ethnic diversity demonstrate significant effects, Banerjee et al.(2005) and Banerjee and Somanathan (2007) show the opposite. In chapter 2, I demonstrate the importance of measuring ethnic diversity at the appropriate scale for the public good in question. The papers by Banerjee et al. are restricted to measurements of district level diversity, rather than a more local measure, and as shown in the previous chapter, their results may be significantly biased downwards.

Caselli and Coleman (2010) 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 alternative 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.[34]

3.2 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[35]. 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

[34] 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. It is also likely that the measurement of religion is not sufficiently accurate within the sample. In particular, the fraction of people following traditional belief structures is far below that normally reported in such studies.

[35] 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.

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 Habyarimana et al. (2007) 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.

3.2.1 Environment

A community provides a single public good funded by the contributions of local households.[36] 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:

$$u_i = W_i - x_i + \mu \ln X_i \tag{3.1}$$

where W_i is household wealth[37], 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[38] As a simplification[39], 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

[36] 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.

[37] 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.

[38] 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.

[39] 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.

group utility. Individuals belonging to the same ethnic group face identical decisions and therefore I focus on solutions that are symmetric within groups[40]. 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.

$$X_e = p_e N x_e + \alpha \sum_{e' \neq e} p_{e'} N x_{e'} \tag{3.2}$$

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.

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:

$$U_e = \max_{x_e} \left[(1-\tau) p_e N (W - x_e + \mu \ln X_e) + \tau \sum_{j=1}^{E} p_j N (W - x_j + \mu \ln X_j) \right] \tag{3.3}$$

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.

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

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

is no effect of the institution, τ. Each group invests on their own behalf and each group invests until:

$$X_e = \mu p_e N \tag{3.4}$$

The average benefit is therefore:

$$\overline{X} = \mu(1 - F)N \tag{3.5}$$

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[41]. 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.

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[42]. 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.

$$X_e = \overline{X} = \mu p_1 N \tag{3.6}$$

However if inter-ethnic institutions are complete $\tau = 1$ then all individuals con-

[41] 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.

[42] 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.

tribute[43] and:

$$X_e = \overline{X} = \mu N \tag{3.7}$$

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.

3.2.3 General Case

In a general solution we must consider the non-negativity constraint on public good investment by each group. However, lemma 3.1 significantly restricts the set of possible cases that require analysis.

Lemma 3.1. *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.*

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[44] of group benefits:

$$\overline{X} = \sum_{e=1}^{E} p_e X_e \tag{3.8}$$

Proposition 3.1. *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

[43] 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.

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

public good. The proof demonstrates that the result of these effects in unambiguously positive on the average level of public good provision.[45]

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 3.1. *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 3.1 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 $\alpha = 1$ 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[46]. 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 3.1 is proposition 3.2.

Proposition 3.2. *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,

[45]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.

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

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. [47]

3.2.4 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 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.

3.2.5 Empirical Implications

The model provides two primary results. First, proposition 3.1 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 3.2 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

[47] 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.

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.

3.3 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[48] but excludable, it is generally provided through public investment[49]. 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.

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

[49] 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.

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 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:

$$\overline{X}_c = \beta_0 + \beta_1 * D_c + \gamma_1 * G_c + \gamma_2 * E_c + \epsilon_c \tag{3.9}$$

where communities[50] 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[51]. At the aggregate level, finding $\beta_1 < 0$ would indi-

[50] As discussed below in section 3.3.1, I use a variety of definitions of community and the results are robust to variations in which the measure of diversity remains consistent with the model.

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

cate 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 3.2 using a household-level specification.

$$X_i = \beta_0 + \beta_1 * D_c + \beta_2 * M_i + \gamma_1 * G_c + \gamma_2 * E_c + \gamma_3 * H_i + \epsilon_i \tag{3.10}$$

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 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.

3.3.1 The Scale of Provision

As demonstrated in chapter 2, it is critical 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.

3.4 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.[52]

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

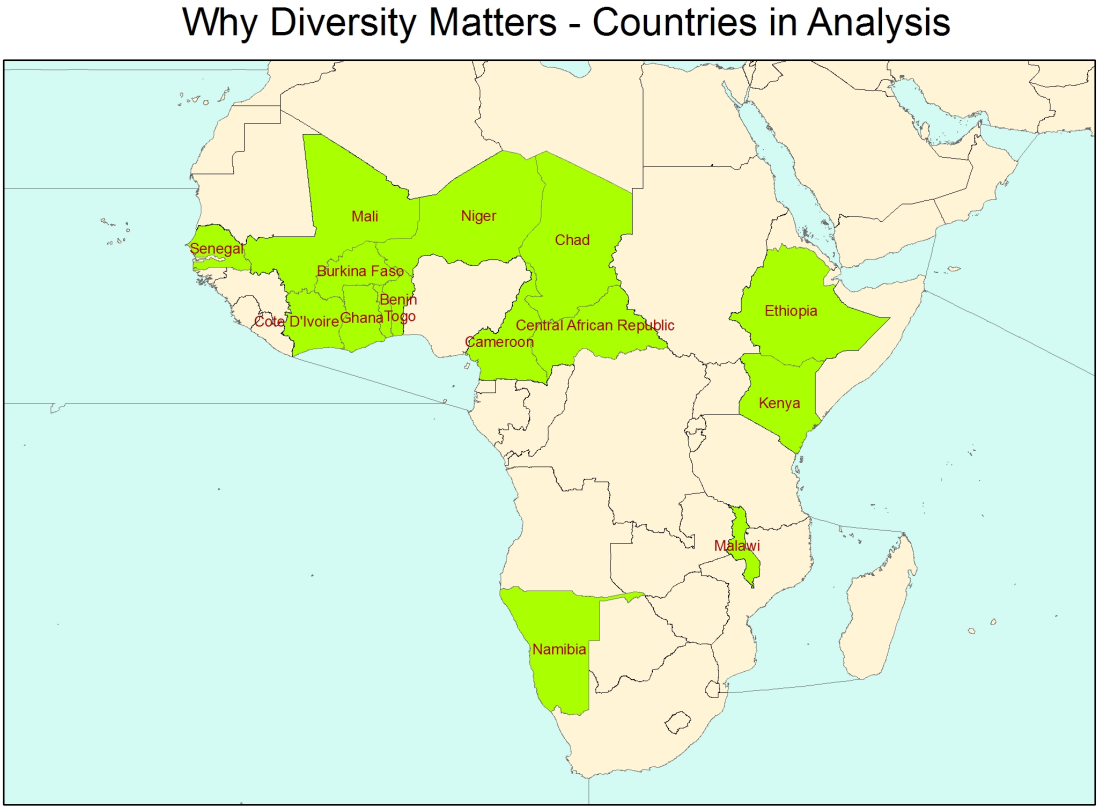


Figure 3.1: Why Diversity Matters - Countries in Analysis

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[53]. 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.

This is partly done out of necessity as local government borders are not consistently

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

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)

3.4.1 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[54] 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.

3.4.2 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

[54]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.

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[55]. 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 on ethnic diversity is polarization[56]. 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 3.4 I include a measure of the second largest group[57] 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[58]. 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[59].

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

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

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

[57]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.

[58]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".

[59]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.

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 thus to the extent that diversity exists it is relatively historical[60] 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.

3.4.3 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, or significant rainfall, 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[61] (measured by the natural log of the distance in kilometres) and the yearly precipitation (in mm). 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), 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.

3.4.4 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.[62]

A second consideration is that in the preferences model the variation in demand could

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

[61]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.

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

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 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.

3.4.5 Other Controls

The measure of wealth available within the DHS survey is constructed from household asset lists and then standardized for each country. It is necessary to include a measure of wealth in the regression as wealth and diversity are strongly, and positively, correlated. As demonstrated in the appendix to chapter 2, the potential endogeneity of wealth does not have a significant impact on the connection between ethnic diversity and public good provision at the community level. Further, within the dataset there is little that can act as a convincing instrument for wealth at the household level. Therefore I have little choice but to assume that the potential endogeneity of wealth does not have a significant effect on the relationship between diversity, minority status and access to piped water.

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[63] 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, along with the age of the household head, for each household, 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.

[63] 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.

Table 3.1: 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.

3.5 Provision of Piped Water - Community Access

The dataset for this analysis is slightly larger than for chapter 2, and therefore I duplicate the primary community-level regressions from that dataset here. Proposition 3.1 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 possible candidates related to the two extremes of the model. These are fractionalization[64] and the population share of the largest ethnic group[65]. In reality the (negative) correlation between these two measures across local communities is 98% and therefore this

[64]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.

[65]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.

consideration is somewhat academic.

3.5.1 Primary Results

Table 3.2 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[66], on the fraction of the population that has access to piped water. As the results show, ethnic diversity, regardless of measure, has significant negative[67] effects on the provision of piped drinking water.

Following the results of chapter 2, 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[68]. 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 is likely the result of increased demand as alternative sources are more costly to access. Surprisingly, coastal communities are far less likely to provide access to public taps. A coastal community is defined in the data as one that is within 20km of the ocean, though the result is not sensitive to this definition. The reason for this finding is unclear, though the magnitude is sufficient to potentially warrant further study.

As was shown in chapter 2, these results are robust to a large variety of different specifications, including using current diversity measures, either with or without instru-

[66] 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.

[67] Or homogeneity has positive effects.

[68] 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.

Table 3.2: Community Specification - Base Results

	Improved	Piped	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	-.009	.099***	.052***	.061**
	(.026)	(.024)	(.017)	(.024)
Fraction w/ Piped Water at Home				-.048
				(.038)
Avg. Wealth	.137***	.338***	.251***	.196***
	(.016)	(.015)	(.012)	(.017)
Urban	.027	.093***	.027*	.063***
	(.018)	(.020)	(.015)	(.021)
Pop.Density(1990)	-.012*	.007	.015	-.011
	(.007)	(.006)	(.010)	(.010)
Dist.to Cap.City (LogKM)	.034**	.035***	.015	.030***
	(.013)	(.010)	(.010)	(.009)
Dist.to River (LogKM)	.022***	.009***	.004	.007**
	(.004)	(.003)	(.002)	(.003)
Avg. Precipitation	.00002	.00002	5.57e-06	6.48e-06
	(.00004)	(.00003)	(.00002)	(.00002)
Coast.Comm.	.025	-.059*	.033	-.153***
	(.028)	(.030)	(.025)	(.036)
Elevation(km)	.024	-.008	.006	-.011
	(.025)	(.016)	(.011)	(.016)
Avg. Tenure	.0004	.001*	.001***	.0003
	(.0007)	(.0006)	(.0004)	(.0006)
Frac.Fem. HH Head	.022	.025	-.044*	.025
	(.044)	(.033)	(.026)	(.036)
Avg. Age of Head	-.003**	.001	.002***	-.0008
	(.001)	(.0009)	(.0007)	(.0009)
Obs.	3058	3058	3058	3027
R^2	.496	.779	.806	.64

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.

ments, 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.

3.5.2 Rural Access to Piped Water

Access to clean drinking water is of particular concern in rural communities[69] where only 13% of households have access to piped drinking water. Table 3.3 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.

Table 3.3: Rural Communities

	Improved	Piped	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	-.042	.081***	.030**	.050**
	(.031)	(.023)	(.014)	(.022)
Fraction w/ Piped Water at Home				.034
				(.035)
Avg. Wealth	.194***	.333***	.219***	.177***
	(.021)	(.020)	(.016)	(.024)
Urban	.022	.186***	.100***	.074
	(.045)	(.046)	(.028)	(.050)
Pop.Density(1990)	-.035	-.037	.113***	-.202***
	(.065)	(.055)	(.042)	(.063)
Dist.to Cap.City (LogKM)	.034**	.029**	.021**	.015*
	(.016)	(.012)	(.010)	(.008)
Dist.to River (LogKM)	.025***	.009***	.003	.007**
	(.004)	(.003)	(.002)	(.003)
Avg. Precipitation	-2.53e-06	.00002	-5.14e-07	1.00e-05
	(.00004)	(.00003)	(.00002)	(.00002)
Coast.Comm.	.045	-.063	.020	-.095***
	(.042)	(.038)	(.031)	(.032)
Elevation(km)	.042*	.00005	.013	-.006
	(.025)	(.015)	(.011)	(.014)
Avg. Tenure	.0008	.0008	.0008**	.0003
	(.0008)	(.0005)	(.0004)	(.0005)
Frac.Fem. HH Head	-.010	-.013	-.033	-.002
	(.047)	(.034)	(.026)	(.032)
Avg. Age of Head	-.001	.001	.001**	-.0003
	(.001)	(.0009)	(.0006)	(.0007)
Obs.	2599	2599	2599	2586
R^2	.502	.58	.634	.433

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.

In the appendix, I repeat the analysis using a narrower definition of "rural commu-

[69] 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.

nity", but the results show no substantive variation.

3.5.3 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 column of table 3.2 shows that measuring diversity by ethnic fractionalization shows the same general picture as using the size of the largest group[70]. 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 equally sized groups. As shown in column(2), polarization at the community-level does not significantly affect the provision of piped water.

Table 3.4: Community-Level Specification - Diversity Measures

	Fractionalization	Polarization	TwoGroups	NonLinearity
	(1)	(2)	(3)	(4)
Fractionalization	-.070***			
	(.019)			
Eth.Polar.		-.019		
		(.012)		
Largest Ethnic Group			.160***	.437***
			(.045)	(.162)
2nd Large. Eth.Group			.114*	.031
			(.060)	(.068)
Largest Group Squared				-.232*
				(.122)
Avg. Wealth	.337***	.332***	.339***	.337***
	(.015)	(.015)	(.015)	(.015)
Urban	.092***	.084***	.097***	.095***
	(.020)	(.020)	(.020)	(.020)
Obs.	3058	3058	3058	3058
R^2	.778	.777	.779	.78

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.

In general, polarization is linked to the size of the second largest group in the com-

[70]Note that the standard deviation of fractionalization is 25% larger than the standard deviation of the size of the largest group, and thus the relevant magnitude of the two coefficients is very similar.

munity. 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.

Further, the specification reported in column (4) includes a quadratic term in the size of the largest group, as suggested by the evidence for non-linearities in the previous chapter. Doing so suggests that the size of the second largest group is not related to public good provision. As shown in the previous chapter, including alternative measures of diversity, does not affect the underlying relationship showing that high levels of ethnic diversity are associated with poor provision of piped drinking water.

3.6 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[71]. 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

[71]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.

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 3.2 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.

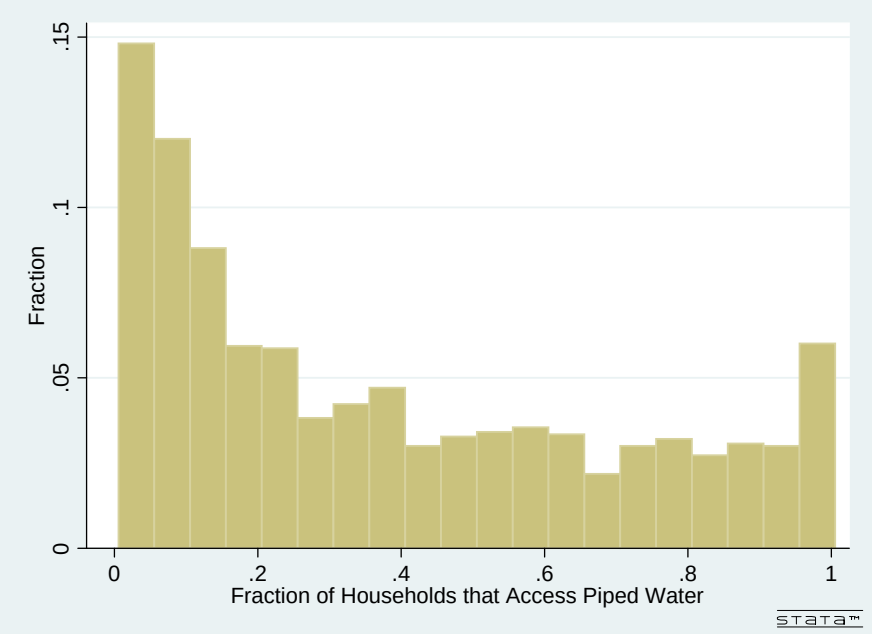


Figure 3.2: Household Access in Communities with Piped Water

3.6.1 Primary Results

Table 3.5 contains the results of the baseline specification[72]. 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

[72]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.

Table 3.5: Household Specification - Base Regression

	Improved	Piped	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.015	.138***	.046**	.114***
	(.029)	(.028)	(.020)	(.025)
Hist.Share Own Eth.	-.032***	-.017*	-.005	-.009
	(.010)	(.009)	(.008)	(.008)
Fraction w/ Piped Water at Home				.100**
				(.044)
HH Wealth	.049***	.136***	.168***	.077***
	(.005)	(.007)	(.009)	(.007)
Avg. Wealth	.076***	.174***	.066***	.106***
	(.013)	(.013)	(.012)	(.016)
Urban	.046***	.131***	.048***	.092***
	(.013)	(.015)	(.012)	(.014)
Pop.Density(1990)	-.014***	.017***	.024***	.024**
	(.005)	(.006)	(.007)	(.011)
Tenure	-1.00e-05	-.00009	.00002	-.00003
	(.00006)	(.00007)	(.00006)	(.00007)
Avg. Tenure	.0008	.001**	.002***	-.0002
	(.0008)	(.0006)	(.0005)	(.0005)
Female Head	.011***	.013***	.007***	.012***
	(.003)	(.003)	(.003)	(.003)
Frac.Fem. HH Head	.020	.030	-.016	.011
	(.050)	(.038)	(.027)	(.037)
Dist.to Cap.City (LogKM)	.029***	.025***	.009	.021**
	(.011)	(.009)	(.007)	(.009)
Dist.to River (LogKM)	.018***	.009***	.004*	.007**
	(.004)	(.003)	(.002)	(.003)
Avg. Precipitation	-1.00e-05	-.00003	-3.30e-06	-.00004*
	(.00003)	(.00002)	(.00002)	(.00002)
Coast.Comm.	.035	-.076**	.014	-.122***
	(.027)	(.035)	(.023)	(.044)
Obs.	95498	95498	95498	79472
R^2	.264	.5	.512	.306

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.

into home access in column (3), and public access[73] 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.

[73]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.

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 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.

3.6.2 Non-Linear Effects of Diversity

As shown in the previous chapter, and in the community-level regressions, the relationship between diversity and public good provision may be non-linear. Table 3.6 repeats the analysis of table 3.5 but includes a quadratic term in the size of the largest group[74]. Both the linear and quadratic terms are significant in all the regressions except the one focused on home access. Even there, we can reject at the 5% level the test that both coefficients are equal to zero. Further analysis of the coefficients suggests that having a largest group that encompasses at least 80% of the community population is associated with the highest levels of public good provision, and that further increases have relatively little effect. In addition, the negative association found in table 3.5 between the size of one's own group and access to the public good falls, both in the size of the coefficient and in its statistical significance.

In terms of magnitude, these results suggest that moving from a community that has a largest group comprising 40% of the population, to one with a largest group comprising 80% of the population, would correspond to roughly a 7% increase in the likelihood that a household has access to piped water. As

3.6.3 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

[74]Further terms (cubics/quartics) are statistically, but not economically significant. The overall picture does not change, and if anything the effects of diversity may be underrepresented here.

Table 3.6: Household Specification - Quadratic Diversity Term

	Improved	Piped	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.456***	.514***	.106	.636***
	(.166)	(.166)	(.131)	(.145)
Largest Group Squared	-.326***	-.278**	-.044	-.379***
	(.119)	(.112)	(.089)	(.097)
Hist.Share Own Eth.	-.021**	-.008	-.003	.002
	(.010)	(.009)	(.007)	(.008)
Fraction w/ Piped Water at Home				.104**
				(.043)
HH Wealth	.049***	.136***	.168***	.077***
	(.005)	(.007)	(.009)	(.007)
Avg. Wealth	.077***	.175***	.066***	.107***
	(.013)	(.013)	(.012)	(.016)
Urban	.048***	.132***	.048***	.094***
	(.013)	(.015)	(.012)	(.014)
Pop.Density(1990)	-.013***	.018***	.024***	.023**
	(.005)	(.006)	(.007)	(.011)
Tenure	-.00003	-.0001	.00002	-.00005
	(.00006)	(.00007)	(.00006)	(.00007)
Avg. Tenure	.0008	.001**	.002***	-.0002
	(.0007)	(.0006)	(.0005)	(.0005)
Female Head	.011***	.013***	.007***	.012***
	(.003)	(.003)	(.003)	(.003)
Frac.Fem. HH Head	.024	.032	-.015	.014
	(.050)	(.038)	(.027)	(.036)
Dist.to Cap.City (LogKM)	.027**	.024***	.009	.018**
	(.012)	(.009)	(.007)	(.008)
Dist.to River (LogKM)	.018***	.009***	.004*	.007**
	(.004)	(.003)	(.002)	(.003)
Avg. Precipitation	-9.98e-06	-.00003	-2.98e-06	-.00004*
	(.00003)	(.00002)	(.00002)	(.00002)
Coast.Comm.	.037	-.074**	.015	-.120***
	(.027)	(.035)	(.024)	(.045)
Obs.	95498	95498	95498	79472
R^2	.265	.5	.512	.308

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.

model as found in Alesina et al.(1999), being a member of the majority group would be the critical factor.

Table 3.7 presents the results using alternative measures of ethnic diversity and the

Table 3.7: Household Specification - Measuring Dominance

	GroupShare	Relative	Largest	Majority
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.514***	.511***	.517***	.519***
	(.166)	(.166)	(.166)	(.166)
Largest Group Squared	-.278**	-.279**	-.285**	-.286**
	(.112)	(.111)	(.111)	(.111)
Hist.Share Own Eth.	-.008			
	(.009)			
Relative Share Own Eth.		-.011		
		(.007)		
Largest Eth.Grp			-.002	
			(.005)	
Majority Eth.Grp.				-.002
				(.005)
HH Wealth	.136***	.135***	.136***	.136***
	(.007)	(.007)	(.007)	(.007)
Avg. Wealth	.175***	.175***	.176***	.176***
	(.013)	(.013)	(.013)	(.013)
Urban	.132***	.132***	.133***	.133***
	(.015)	(.015)	(.015)	(.015)
Obs.	95498	95498	95498	95498
R^2	.5	.5	.5	.5

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.

relative importance of one's own group. Column (1) repeats the primary specification from table 3.5 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 household is identified as a member of the largest ethnic group in the community. Assuming 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.

3.6.4 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 3.8 duplicates the results of table 3.5, 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.

3.7 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 is associated with lower levels of piped water provision in Africa. This result is consistent with the growing body of evidence from around the world, in both developed and developing countries. However, 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 signifi-

Table 3.8: Households in Rural Communities

	Improved	Piped	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.482**	.641***	.256**	.554***
	(.217)	(.174)	(.127)	(.117)
Largest Group Squared	-.359**	-.379***	-.147*	-.338***
	(.152)	(.116)	(.084)	(.078)
Hist.Share Own Eth.	-.012	-.002	-.007	.004
	(.012)	(.008)	(.006)	(.006)
Fraction w/ Piped Water at Home				.084**
				(.042)
HH Wealth	.088***	.142***	.106***	.074***
	(.007)	(.008)	(.006)	(.007)
Avg. Wealth	.085***	.142***	.082***	.070***
	(.018)	(.016)	(.012)	(.017)
Pop.Density(1990)	.028	.003	.017	-.013
	(.024)	(.024)	(.020)	(.023)
Tenure	6.20e-06	.00008	.00004	.00005
	(.00008)	(.00007)	(.00004)	(.00006)
Avg. Tenure	.001	.0005	.0008**	-.0005
	(.0009)	(.0006)	(.0004)	(.0005)
Female Head	.010**	.005	.0002	.005*
	(.004)	(.003)	(.002)	(.003)
Frac.Fem. HH Head	.018	.004	-.007	-.003
	(.054)	(.040)	(.025)	(.036)
Dist.to Cap.City (LogKM)	.037**	.022**	.006	.022***
	(.016)	(.010)	(.008)	(.007)
Dist.to River (LogKM)	.019***	.008***	.002	.006***
	(.004)	(.003)	(.002)	(.002)
Avg. Precipitation	-.00002	-.00005**	-.00003**	-.00003*
	(.00004)	(.00002)	(1.00e-05)	(.00002)
Coast.Comm.	.030	-.082	.016	-.115**
	(.042)	(.052)	(.031)	(.051)
Obs.	67429	67429	67429	63237
R^2	.257	.284	.309	.179

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.

cantly 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.

This also suggests why the presence of ethnic effects is largely hidden from view. If all members of a diverse community are equally affected, there is no "other" group that can be identified. It is therefore not obvious that ethnic diversity is having an effect. In contrast, if the effects of diversity were felt primarily by local minorities, that would far more obvious and would likely have been a subject for comment before now.

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 small part of the inefficiency caused by the ineffective governance structures that persist in Sub-Saharan Africa.

Chapter 4

Adaptive Institutions

4.1 Introduction

The prevailing wisdom within economics is that institutions are the key to development. If governments enact policies that promote investment and encourage trade, prosperity follows. Yet the stark reality in many countries of the world is ineffective government, lax rule of law and consequently a dire lack of economic activity. Are the poor institutions in the countries of the "bottom billion"[75] the result of a lack of political will, or are there additional constraints on the development of institutions?

While it may be clear that institutions, broadly defined, are important, it is not clear which specific institutions are critical to development, or how those institutions change over time. The two broad categories of economic institutions, formal constraints including government structures and laws, and informal constraints such as norms of behaviour (North 1994)[76] have each been shown to be relevant. Papers by Acemoglu et al. (2001, 2002, 2005) provide significant support to the notion that formal, government institutions are important, while Knack and Keefer (1997), Tabellini (2005) and Guiso et al. (2006) demonstrate that informal institutions, such as cultural norms, also have significant long-term effects.

Of further importance are questions related to the long-term persistence of institutional quality. As demonstrated in Acemoglu et al. (2001, 2002, 2005), and in Guiso et al. (2008), the quality of institutions, both formal and informal, appear to be strongly influenced by historical factors[77]. This paper develops the idea that low levels of social capital may generate an institutional trap in which the economic activity necessary to develop social capital does not occur, and where formal institutions are powerless to change this.

I develop a theoretical framework within which to consider the interaction between formal and informal institutions. An evolutionary process of social capital development

[75]The notion of a "bottom billion" was presented in(Collier 2007) and the associated literature and refers to the approximately 1 billion people in the world that live in very poor circumstances, primarily as a result of poor governance.

[76]As an alternative to the terms formal and informal constraints, the terms public-order and private-order have been used in the institutional literature with a very similar meaning (Platteau 1994, Greif 2005).

[77]Settler mortality rates or historical density patterns affect current formal institutions in the papers by Acemoglu et al., while a historical period of independence does the same for social norms in the case of Guiso et al.

is matched to a simple economic model in which agents choose between producing alone, or entering risky joint production. Legal institutions, governing the enforcement of contracts, are critical to determining the equilibrium that results in each period of the game, and therefore the resulting evolutionary development of social capital.

The evolutionary model presented here is standard, with a fraction of agents incurring a cost for dishonest actions, while the remainder simply make choices that maximize fitness. The evolutionary model then exhibits societal level changes in social capital in the direction of the group with higher average fitness. In contrast, the production model differs from previous work on this subject by assuming that the behaviour type of a particular individual is observable, though only imperfectly and at a cost. This is in keeping with an extensive literature in psychology and behavioural economics[78] that demonstrates that people are able to imperfectly determine the trustworthiness of individuals after minimal interaction. Agents are thus able to use this information to decide whether or not to enter into risky joint production, and signal acquisition becomes critical to the evolution of social capital. In terms of social capital development, signal acquisition acts as a screening mechanism that causes the honest agents a higher probability of entering joint production, and therefore increasing their average income (or fitness).

The results of the model generate three critical interactions between contracting institutions enacted by government, and the development of social capital. First, institutions may be necessary to promote joint production by reducing the associated risks, and the risks associated with joint production are the incentive to engage in the screening process. Second, the enforcement of contractual obligations reduces the benefit to opportunistic activity, and therefore encourages the development of honesty within society. Finally, strong government institutions may crowd out the benefits associated with signal acquisition, and without signal acquisition the honest individuals are at a disadvantage, leading to the downward evolution of social capital.

These interactions generate two critical barriers in the effectiveness of government institutions in the development of social capital. First, at low levels of social capital, the contracting institutions necessary to encourage joint production are also sufficient to crowd out the screening activity that promotes social capital development. This has the effect of generating a social capital trap, in which a society endowed with poor social capital is unable to use contracting institutions to escape the effects of a poor endowment. Second, the crowding out effect generates a maximum level of social capital, at less than the perfectly honest society, beyond which social capital will not increase. The final result in the paper is that reducing the cost of screening, or increasing its efficiency, has the effect of lowering the bar to social capital development, and increasing the maximum level of social capital in society. Thus it may be that government actions to promote the

[78]This literature is summarized below.

exchange of information in society, rather than in improving the environment for formal contracting, would be more effective in building generalized trust.

Starting with Coleman (1988) and Putnam (1993), an extensive literature has developed to understand how social structures affect economic outcomes. The initial research focused on a definition of social capital as the content of interpersonal relationships and civic involvement. However, the subsequent literature has also fixed on a notion of social capital that is far larger than the interpersonal relationships first envisioned. In an empirical analysis by Knack and Keefer (1997), the authors define social capital further to include the level of interpersonal trust within society. That is, do people believe that "'most people can be trusted'" within the society[79].

The primary focus of the paper is on the trustworthiness of a society. A fraction of the population is unwilling to cheat once they have committed to a particular action, regardless of the possibility of external enforcement. The paper therefore deals directly with the second form of social capital, generalized trust. Additionally, as mentioned above, individuals are able to incur a cost to identify those individuals that have trustworthy preferences. This is representative of the civic engagement form of social capital, with individuals making decisions as to their engagement with society according to private incentives. Individuals that incur the cost associated with civic engagement are then able to acquire information through their networks about a prospective match. By modeling both extensive social capital, or trust, and intensive social capital, or relationships, this paper provides a connection for the two key ideas behind the social capital literature.

The primary motivation for this project is to understand the causes of the apparently large cross-country variations in the degree of interpersonal trust held by individuals. The World Values Survey has repeatedly included the following question about generalized trust in their surveys.

> "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?"

At the individual level, the answers to this question can be interpreted to indicate either the trust an individual has in their fellow citizens, or alternatively as a reflective measure of trustworthiness (Glaeser, Laibson, Scheinkman, and Soutter 2000). In the aggregate, this distinction is lost and the fraction of individuals reporting that most people can be trusted may be viewed as a measure of the level of trust and trustworthiness in society[80]. This fraction varies from a high of roughly two-thirds of the population in

[79]Their results suggest that both interpersonal trust and civic engagement have important consequences for economic development.

[80]Note that while trust and trustworthiness are very different concepts, the common assumption is that there are no systematic errors in people's equilibrium beliefs. Therefore, the level of trust, and the opposite notion, the level of trustworthiness, are assumed to be strongly correlated.

the Scandinavian countries, to a low in the single digits in a variety of developing and underdeveloped countries. A related question included in the World Values Survey asks:

"Do you think most people try to take advantage of you?"

At the national level, the results from this question are strongly correlated with the fraction of people that say most people can be trusted. These results, and other related findings, suggest that there are real differences in trust levels around the world[81].

A significant literature has focused on explaining how a social norm such as trust has developed such variation across societies. The primary mechanism for analyzing the adaptation of social norms for trust has been evolutionary theory. Guth and Kliemt (1994, 1998) are early examples of this, while Francois and Zabojnik (2005), Tabellini (2008), Ahn and Esarey (2008) and Francois (2008) are more recent adaptations. All start from the premise that there is a subset of the population that is inherently trustworthy, and incorporate evolutionary mechanisms to describe the development of trust in society. In particular, the general assumption throughout these papers is that there is a fraction of the population that is fundamentally honest, and therefore will not act dishonestly even if it were advantageous to do so.

For social norms to develop, there must be a mechanism allowing the honest to have different utility from the opportunistic. As the opportunistic have the ability to gain from dishonest actions when they have value, the ability to generate downward evolution of social norms is simple. In contrast, generating conditions under which social norms of honesty are subject to positive selection is less clear, though there have been two basic assumptions employed. Models either alter the evolutionary environment or introduce the idea that preferences are observable. Building on the model of Bisin and Verdier (2001), Francois and Zabojnik (2005) and Tabellini (2008) generate evolutionary dynamics by altering the evolutionary process to allow for a 'warm glow' effect, in which the effective fitness of trustworthy individuals is increased through a non-pecuniary benefit when they act honestly. Specifically, parents may choose to teach their children to be honest, if the parent sees those beliefs as beneficial. In doing so, the parent takes the "warm glow" benefits into account, and as such the evolutionary environment is affected by the otherwise non-pecuniary benefit[82].

Alternatively, and following the more general game theory literature on preference evolution, non-standard preferences may develop if those preferences are observable

[81] In addition, Knack and Keefer (1997) demonstrate that there is a strong correlation between the World Values Survey measure of trust, and the fraction of lost wallets that are returned in a "wallet drop" experiment. This suggests that the survey measure is capturing a real difference across countries.

[82] In a slightly different treatment, Bidner and Francois (2009) adopts a different evolutionary type space in which there are honest individuals that never cheat, dishonest individuals that always cheat, and opportunistic individuals that play to maximize fitness. In this setup, the evolutionary dynamics are played between the honest and the dishonest individuals, and there is a constant fraction of opportunistic agents.

(Heifetz, Shannon, and Spiegel 2007, Dekel, Ely, and Yilankaya 2007). Guth and Kliemt (1994, 1998) and Ahn and Esarey (2008), each generate evolutionary dynamics with observability as a key factor. This paper develops further on the idea of observable preferences, by allowing for that observation to be costly and imperfect. Doing so generates the possibility of additional interactions between social norms and formal institutions.

Each of these basic ideas, parental indoctrination of values and observable preferences, has empirical support. Of particular importance for this paper is the notion that preference types, or an indication of honesty, are at least partially observable[83]. This has broad support within anthropology, biology, psychology and economics. Frank (1987) and Frank et al. (1993), first demonstrated the theoretical justification for an evolved marker of trustworthiness, and then showed that people are able to differentiate between real and faked enjoyment of a situation. More directly, Brown and Moore (2000), Brown et al. (2003) and Verplaetse et al. (2007) provide experimental evidence that people are able to detect altruism and honesty in prospective partners. The idea that preferences are observable has been defended in, among others, Ellingsen (1997), and is also incorporated extensively in the literature on evolutionary preferences[84] The model developed here allows agents to, at a cost, acquire a signal of the preference type of a prospective match. This cost could be interpreted as the cost of conducting an extensive interview, or could be the cost associated with civic engagement that provides an individual with additional information from other sources.

As mentioned, in addition to the idea of generalized trust, the notion of social capital has also been used to understand the role of civic engagement, or social connections, in economic situations. Building on the ideas proposed by Coleman (1988) and Putnam (1993, 1995b, 1995a), a large literature has developed to analyze this alternate form of social capital.[85] However, the connections between the two forms of social capital are not well understood. Fafchamps (2006) models the role of social connections, and the resulting improvements in trust. The empirical predictions of this model are complicated due to both positive and negative externalities related to social capital, possibly accounting for earlier findings showing little relationship between the two forms of social capital (Knack and Keefer 1997). In the model developed here, civic engagement, in the form of the costly observation of the preference type of other individuals, is key to the evolution of trustworthiness in society, but not beneficial from the perspective of current income.

Carlin et al. (2009) develop a related model in which social capital and government institutions interact. However, in this model, agents consciously choose their trustworthiness, or social capital, in response to economic incentives. The results of that paper

[83] For a literature review on the importance of family transmission of values, see Bisin and Verdier (2001).
[84] See Eaton and Eswaran (2003) for one such example.
[85] See Sobel (2002) for a summary of the early work in this area.

are similar in form to the basic results here, as regulation is found to act as a complement to trust at low levels of social capital, and as a substitute for trust at high levels of social capital. However, given the nature of social capital, the persistent impact of poor social norms, or of government institutions cannot be examined in their model.

The focus of this paper on the connection between social capital and legal institutions builds on a growing recent body of work. Platteau (Platteau 1994) and Greif (2005) each discuss the relationship between public-order and private-order institutions and their ability to coordinate complex production. More recently, Tabellini (2008) and Francois (2008) develop theoretical models of norm development in the presence of varying legal institutions. In both papers, the effect of legal institutions is to promote the development of social capital by creating negative consequences for dishonest behaviour. In Tabellini, as the beneficial effects of cheating are reduced, the warm glow effect for honest individuals of acting honestly dominates, thus promoting the evolutionary development of trust. In Francois, the dishonest are particularly hurt by the increase in legal institutions as they are unable to not try to cheat.

The results in this paper are more closely linked to Bohnet et al. (2001), in which legal institutions may either encourage or discourage the evolution of trustworthiness. Similarly to the model developed here, at low levels of legal institutions the incentives create a scenario in which agents only match with honest partners, thus providing an evolutionary advantage for these agents. With stronger legal institutions, the dis-incentive to matching with opportunistic partners is reduced, thus creating an environment in which the opportunistic succeed. However, that paper assumes that honesty is a freely observable trait, and this eliminates the positive complementarity between contracting institutions and social norms that is found in this paper. Guth and Ockenfels (2005) derive a similar model in which the interaction between legal institutions can crowd out trustworthiness when types are observable, but crowd-in trustworthiness when preference types are unobservable.

The notion that government institutions can crowd out social norms is not specific to the theoretical literature on the development of trust. Papers by Kumlin and Rothstein (2005) and van Oorschot and Arts (2005) demonstrate empirically that government programs can affect social capital, and in particular the development of generalized trust. This builds on previous theoretical work by Frey (1997) that overly strict legal systems crowd out civic virtue.

Fukuyama (2001) provides an interesting discussion of the critical tradeoff that is behind the results in this paper. First, he states that:

> ...states indirectly foster the creation of social capital by efficiently providing necessary public goods, particularly property rights and public safety.[p.18]

but also that:

> ...states can have a serious negative impact on social capital when they start to undertake activities that are better left to the private sector or to civil society.[p.18]

These two statements reflect the opposing effects of formal institutions present in the model developed here. First, government policy is supportive of the development of social capital by fostering an environment in which risky, complex production can function. However, if that support is excessive it renders civic engagement unnecessary and thus causes the erosion of social norms.

In considering the interaction between formal and informal institutions, the process of institutional change is critical to the analysis. In particular, the results derived here depend on the idea that formal, government institutions, can respond to the existence of strong or weak social norms of honesty. This requires that, as defined by Roland (2004), government institutions are "'fast-moving'", whereas social norms tend to change gradually. The institutional model in this paper follows this basic format, with government institutions chosen in any period in response to the existence of social norms that evolve slowly over time. While there is significant evidence that political institutions change relatively infrequently, it is also apparent that significant shifts are possible in short periods of time.

A key result of the paper is that at low levels of social norms, there exists the possibility of a dynamic complementarity between social norms and legal institutions. As such, government institutions must be sufficiently effective to encourage the complex production methods that provide incentives for civic engagement that is necessary for the evolutionary development of social norms. However, government institutions that are too strong, eliminate the incentive for social interaction even while promoting complex production, thus eliminating the upward pressure on social norms. This result suggests that the effectiveness of government in promoting social norms is strictly limited, which leads to the possibility of an institutional trap where social norms remain low, regardless of government action.

The notion that institutional traps may exist has been explored at length in the development literature[86], and is also a key result in Tabellini (2008). However, the poverty trap in that paper is specific to the political process of decision making. Effectively, when social norms are weak, the large body of opportunistic individuals know that they would be worse off with stronger government institutions and therefore block reform that would have the effect of raising average welfare. In contrast, the institutional trap

[86] See Bowles et al. (2006) for a full-length treatment of the subject.

developed here is structural. A government with weak social norms is unable to implement a contracting institution that promotes social development, regardless of its preferences. This generates a critical, and at least plausibly testable, difference between the two models. In Tabellini (2008), if a country with low social capital did introduce slight improvements in contract enforcing institutions, social capital would increase. In contrast, the model developed here suggests that changes to the contracting environment below a cutoff will have no effect on social capital.

Further, the model developed here links the existing research on social diversity and institutional development. If diversity causes an increase in the cost of civic engagement, or a reduction in the accuracy of information provided by social networks, that diversity increases the level of social capital necessary to escape the institutional trap.

The next section presents the basic economic model with fixed rates of social capital and contracting institutions. Section 4.3 derives the equilibrium solution to this model and presents the resulting outcomes in graphical form. Section 4.5 presents the evolutionary model that defines the development of social capital. Sections 4.6 and 4.7 discuss the interaction effects between contracting institutions and social capital, and how those interactions combine to generate four distinct social capital regimes, including an institutional trap at low levels of social capital. Section 4.8 discusses the effects of social division on the development of social capital, and section 4.9 concludes.

4.2 Model

A model of cultural evolution is an inherently dynamic problem. With this in mind, I first develop a simple single-period model for exogenous levels of social capital and institutional development. Agents are paired randomly and, after possibly acquiring information about their partner, may choose to work together or apart. This model generates the behaviour and outcomes that then permit the discussion of cultural evolution and institutional development that follows.

There are three key elements to the model. First, agents may choose to work in joint production or remain on their own, and the moral hazard associated with joint production is the key factor in this decision. Second, workers are heterogeneous. A fraction of all individuals in society have an inherent preference to act honestly, and will do so even in the presence of gains to opportunism. Third, agents may acquire information regarding the behaviour type (honest or opportunistic) of their prospective partner. However, this information is both costly and imperfect. These three elements provide the basic structure necessary to proceed to the discussion of the evolution of social norms of honesty, in section 4.5.

4.2.1 The Population and Production Technology

There is a population of measure 1, of which a fraction $\beta \in (0,1)$ incur a non-pecuniary cost, ψ of behaving dishonestly. In the model, this occurs if an honest person accepts a contract and then shirks their responsibility. However, this non-pecuniary cost does not affect the evolutionary model that follows[87]. Specifically, 2 people, one honest and the other dis-honest, that take identical actions, would have identical fitness.

Production occurs either alone, or in 2-person teams. Production alone generates a net payoff normalized to 0, while production in teams generates an amount $\alpha \in (0,1)$ for each participant. Team production, while inherently more productive than producing alone, creates a 1-sided moral hazard problem. At the beginning of the period, all individuals, regardless of type, are randomly paired, and then one is defined as Player 1(referred to by feminine pronouns), while the other is defined as Player 2(referred to by masculine pronouns). Production creates a moral hazard problem for Agent 2, and thus it will be necessary for Player 1 to determine whether she wants to undertake joint production. If joint production is undertaken, but Player 2 decides to shirk, the direct payoffs from production are (-1) for player 1 and (+1) for player 2[88].

4.2.2 Signals,Offers and Acceptance

Prior to the decision to enter joint production, player 1 has the option to pay a cost, $\theta > 0$, to acquire a signal of the behaviour type of player 2[89]. With a probability $\kappa \in (\frac{1}{2},1)$, the signal will correctly match the type of player 2, with a good signal if he is honest, and a bad signal if he is opportunistic. Conversely, with a probability $(1-\kappa)$, the signal is incorrect, with good signals attached to opportunistic types and bad signals associated with honest players. I restrict focus to situations in which signals are sufficiently cheap and informative as to be acquired in at least certain cases. The specific assumption that is necessary and sufficient[90] is that:

[87] And would not even if it occurred in equilibrium. In the model, this is simply the sufficient incentive for honest individuals to not shirk.

[88] The net utility after shirking is therefore identical to the net production from producing alone. An alternative would be to have the total shirking outcome be lower than solo production, though this would not affect the results presented here. To see this, simply define $\alpha < 1$, and player 2 receives an amount μ where $\alpha < \mu < 1$. As player 2's utility from shirking is only relevant in this game to the extent that it is greater than α, there is no effect on the equilibrium that results. An extension of the paper to consider optimal government choice would need to consider this issue further.

[89] As discussed in the introduction, this could be the cost of civic engagement. With greater civic engagement, a player is more likely to have access to reports from other members of society as to the personality of a randomly matched individual.

[90] I discuss the logic behind this assumption further below.

Assumption 1.

$$\theta < \frac{\alpha(2\kappa - 1)}{1 + \alpha} \tag{4.1}$$

After acquiring the signal, Agent 1 can choose to offer Agent 2 the opportunity to enter joint production, or produce on her own and receive a payoff normalized to zero. If agent 1 chooses to offer joint production, agent 2 may choose to accept or reject that offer.

4.2.3 Institutions, Effort and Outcomes

If an offer of joint production is extended, and accepted, the two agents draw up a contract that covers as many eventualities as possible. Once the contract is written, and if it is accepted, both agents begin work. On starting production, agent 2 learns whether the contract will actually be enforceable in court, as determined by the quality of a country's institutions. With a probability, i, the contract is enforceable, and shirking will be punished by an amount P. Conversely, with a probability $1 - i$, the contract will not be enforceable, and the worker knows that there will be no punishment for shirking.

A key difference from the existing literature is that the agent knows the institutional effects prior to making his choice regarding shirking. While this approach to modeling institutions is non-standard, there is a strong reason for doing so. The natural alternative is to have the effectiveness of the contract determined after the worker makes the decision to shirk or work. In the context of this model, that would generate a binary shift in behaviour at a particular institution, below which no dishonest individual would work, while above that point all dishonest individuals work. This result would be even less realistic than the modeling choice made above. Optimally, a combination of the two concepts could be incorporated, with the agent imperfectly observing the consequences of shirking, and this may be explored in a future extension. As long as the effects of increasing institutions leads to a continuous increase in the probability of opportunistic agents working, the results will not change from those presented here.

The punishment that results from shirking in the face of an enforceable contract, P, is assumed to be greater than the benefits of shirking, $1-\alpha$, and therefore when contracts are enforceable, both honest and opportunistic workers have a dominant strategy to not shirk. Similarly, the non-pecuniary cost of dishonest behaviour incurred by an honest individual, ψ, is also assumed to be large enough such that working dominates shirking.

4.2.4 Timing

1. Random matching
2. Optional signal acquisition by player 1, at a cost θ, of quality κ
3. Agree to joint production or not
4. Contract enforceability determined, probability i
5. Player 2 chooses to work or shirk
6. Production is realized

The key variables in the model are fraction of the population that is honest, β, and the quality of contracting institutions, i. There are three parameters of interest, θ and κ represent the cost and quality of signals, while α is the value of joint production.

4.3 Equilibrium

4.3.1 Defining an Equilibrium

The relevant equilibrium concept in a game with incomplete information is a sequential equilibrium, which requires that each player's strategy incorporates optimal choices at any information set, while also placing restrictions on player's beliefs. However, given the simplicity of the game modeled here, it is sufficient to focus on Nash equilibria that result from the elimination of dominated strategies, and then confirm that these equilibria meet the requirements of a sequential equilibrium.

Proposition 4.1. *For any set of parameter values, social capital, β, and contracting institutions, i, a sequential equilibrium exists. Further, for almost any set of parameter values, social capital and contracting institutions the equilibrium is unique.*

The exact characterization of the strategy along the equilibrium path is defined here[91].Each behaviour type of Player 2 has a dominant strategy. If he is honest, he accepts any contract and works to complete the contract. If he is opportunistic, he accepts any contract, and works if and only if the contract is enforceable. The institutional

[91]Off equilibrium path strategies are effectively irrelevant in this game, due to the information structure of the game, and as such, I do not focus on them. They are however defined within the context of the proof.

environment determines the optimal choice for player 1. First, I define 3 cut-off values:

$$I_{AB}(\beta) = \frac{\theta + (1-\beta)(1-\kappa) - \beta\kappa\alpha}{(1-\beta)(1-\kappa)(1+\alpha)} \tag{4.2}$$

$$I_{AC}(\beta) = \frac{1-\beta-\beta\alpha}{(1-\beta)(1+\alpha)} \tag{4.3}$$

$$I_{BC}(\beta) = \frac{(1-\beta)\kappa - \beta\alpha(1-\kappa) - \theta}{\kappa(1-\beta)(1+\alpha)} \tag{4.4}$$

The strategy choice along the equilibrium path for player 1 is:

1. $i \leq I_{AB}(\beta)$ & $i \leq I_{AC}(\beta)$
    - Player 1 does not screen prospective partners and will not offer to enter into any contracts.
2. $i \geq I_{AB}(\beta)$ & $i \leq I_{BC}(\beta)$
    - Player 1 acquires a signal of the type of player 2, and offers to enter joint production if and only if she receives a good signal.
3. $i \geq I_{AC}(\beta)$ & $i \geq I_{BC}(\beta)$
    - Player 1 does not acquire a signal of type, and offers to enter joint production with player 2.
4. At specific contracting institutions where more than one equilibrium condition is satisfied, there are multiple equilibria.

4.3.2 Proving Proposition 4.1

There are four decisions in the game, the first two by player 1, and the last two by player 2, and the game can be solved by backward induction. The next subsection defines the dominant strategy for player 2, which must form his strategy in any equilibrium. I then define the optimal choice for player 1 in deciding whether to offer contracts, followed by the optimal choice in determining whether to acquire a signal of player 2's behaviour type. Finally, I collect the results of these three subsections to demonstrate that a sequential equilibrium exists for all values of parameters, social capital and institutions, and that it is almost everywhere unique.

Work or Shirk

The last move of the game occurs when Player 2 chooses to either work or shirk.

Lemma 4.1. *If player 2 is honest, his dominant strategy in this sub-game is to work, rather than shirk. If player 2 is opportunistic and his contract is enforceable, then his dominant strategy is to work. If he is opportunistic, and his contract is unenforceable, then his dominant strategy is to shirk.*

At this stage, he knows whether his contract is enforceable, and as he is player 2, he knows his own type. If he is honest, the payoff to working is α, and the payoff to shirking is $1-\psi$ is less than α by assumption. If he is opportunistic, and his contract is enforceable, then the payoff to working is higher than the payoff to shirking, $\alpha > 1-P$, and he will work. If the contract is not enforceable, the payoff to shirking is higher than the payoff to working, $1 > \alpha$ and he will shirk.

Before the enforceability of institutions is known, player 2 must decide if he will accept the contract offer.

Lemma 4.2. *Player 2 will accept a contract offer under any circumstances.*

Accepting the contract offer offers a payoff of α if he works, and a payoff of 1 if he is opportunistic, his contract is unenforceable, and he shirks. Either payoff is greater than the alternative of not accepting the offer and receiving a payoff of 0.

Corollary 4.1. *Player 2 has a dominant strategy to accept any offer, and then work unless he is opportunistic and his contract is unenforceable, in which case he will shirk.*

Any equilibrium will necessarily be based on player 1 having accurate beliefs about the expected play of player 2, and this corollary defines what those beliefs must be.

Offering a Contract

The decision to accept an offer by player 2, is preceded by the decision to make an offer by player 1. At the time that she does this, she is at one of 3 information sets. Either she has not acquired a signal, or she has acquired a signal and it was either good or bad[92]. The next 3 lemmas define the optimal strategy at each information set. At each information set, she is in one of two states of nature - either player 2 is honest or he is opportunistic.

First, I define the critical value:

$$I_{AC}(\beta) = \frac{1-\beta-\beta\alpha}{(1-\beta)(1+\alpha)}. \tag{4.5}$$

[92] I use the terminology good/bad as the signals may be incorrect and are therefore not interchangeable with the notation honest/dishonest. Also note that good/bad is not interchangeable with correct/incorrect - which I avoid as that is information unavailable to the player.

Lemma 4.3. *If no signal has been acquired, it is optimal for player 1 to offer player 2 a contract if $I \geq I_{AC}(\beta)$. It is optimal for player 1 to not offer player 2 a contract if $I \leq I_{AC}(\beta)$. At $I = I_{AC}(\beta)$, either strategy is optimal.*

Proof - see appendix. If institutions are higher than the cutoff value, the expected payoff to offering a contract is greater than zero, the payoff to not offering a contract.

In the second information set under consideration, player 1 has received a signal of player 2's type, and that signal is good. Define the critical value:

$$I_G(\beta) = \frac{(1-\beta)(1-\kappa) - \beta\kappa\alpha}{(1-\beta)(1-\kappa)(1+\alpha)} \tag{4.6}$$

Lemma 4.4. *After receiving a good signal, it is optimal for player 1 to offer a contract if $I \geq I_G(\beta)$. If is optimal to not offer a contract if $I \leq I_G(\beta)$. Either strategy is optimal at the critical value of institutions.*

Proof - see appendix. Again, if institutions are higher than this cutoff value, the payoff to offering a contract in this situation is higher than not offering a contract.

Similarly, we can define a cutoff value after a bad signal. Let:

$$I_B(\beta) = \frac{(1-\beta)\kappa - \beta\alpha(1-\kappa)}{\kappa(1-\beta)(1+\alpha)} \tag{4.7}$$

Lemma 4.5. *After receiving a bad signal, it is optimal for player 1 to offer a contract if $I \geq I_B(\beta)$. If is optimal to not offer a contract if $I \leq I_B(\beta)$. Either strategy is optimal at the critical value of institutions.*

Proof - see appendix.

Corollary 4.2. *If it is optimal to offer a contract after a bad signal, it is also optimal to offer a contract after a good signal.*

This is easily shown as $I_G(\beta) < I_B(\beta)$ when $\kappa > \frac{1}{2}$, which is true by assumption. Given this corollary, there remain three possible strategies after acquiring a signal. Either never offer a contract, always offer a contract, or offer a contract if and only if the signal was good. In the next lemma I demonstrate that the first two options will never be part of the equilibrium path.

Choosing to Screen

The first decision in the game occurs when player 1 chooses to pay to acquire a signal of player 2's type.

Lemma 4.6. *It is never optimal to acquire a signal, and then ignore it.*

Proof - see appendix. Acquiring the signal is costly and if it is ignored, then there is no possibility that it provides value. We can therefore eliminate these strategies from consideration.

On the equilibrium path, player 1 was originally choosing from 6 possible strategies[93]. Three of these (2 ways of acquiring and ignoring signals, or acquiring a signal and then offering a contract if and only if the signal is bad) have been eliminated. The game is therefore reduced to player 1 choosing from one of 3 effective strategies.

A - Don't acquire signals and never enter joint production.

B - Acquire a signal, and enter joint production if and only if the signal is good. (Separating equilibrium)

C - Don't acquire signals, and enter joint production. (Pooling equilibrium)

The payoff to each strategy for player 1 is:

$$U_{1A}(i,\beta) = 0 \tag{4.8}$$

$$U_{1B}(i,\beta) = \beta\kappa\alpha + (1-\beta)(1-\kappa)i\alpha + (1-\beta)(1-\kappa)(1-i)(-1) - \theta \tag{4.9}$$

$$U_{1C}(i,\beta) = \beta\alpha + (1-\beta)I\alpha + (1-\beta)(1-i)(-1) \tag{4.10}$$

Strategy B will be preferred to strategy A if:

$$\beta\kappa\alpha + (1-\beta)(1-\kappa)i\alpha + (1-\beta)(1-\kappa)(1-i)(-1) - \theta \geq 0 \tag{4.11}$$

while strategy A would naturally be preferred to strategy B if the converse were true. If the payoffs from the 2 strategies are equal, then it is possible that either will be selected. Rearranging this equation indicates that strategy B will be preferred to strategy A if:

$$i \geq \frac{\theta + (1-\beta)(1-\kappa) - \beta\kappa\alpha}{(1-\beta)(1-\kappa)(1+\alpha)} \tag{4.12}$$

Define this as the cutoff value, $I_{AB}(\beta)$, listed with proposition 4.1.

Similarly, Strategy C will be preferred to strategy A if:

$$\beta\alpha + (1-\beta)I\alpha + (1-\beta)(1-i)(-1) \geq 0 \tag{4.13}$$

[93] As mentioned, the off-equilibrium path elements of player 1's strategy are inconsequential for play, as these choices are completed before player 2 makes any choices. I therefore do not discuss them in detail. However, if player 1 chooses to acquire signals, then the off-equilibrium element of his strategy must be determined according to lemma 4.3. Similarly, if player 1 chooses to not acquire signals, his choices if he were to acquire signals are determined according to lemmas 4.4 and 4.5.

Rearranging results in the cut-off value:

$$I_{AC}(\beta) = \frac{1-\beta-\beta\alpha}{(1-\beta)(1+\alpha)} \tag{4.14}$$

Finally, strategy C will be preferred to strategy B if:

$$\beta\kappa\alpha + (1-\beta)(1-\kappa)i\alpha + (1-\beta)(1-\kappa)(1-i)(-1) - \theta \geq \tag{4.15}$$

$$\beta\alpha + (1-\beta)I\alpha + (1-\beta)(1-i)(-1) \tag{4.16}$$

and the associated cutoff value is:

$$I_{BC}(\beta) = \frac{(1-\beta)\kappa - \beta\alpha(1-\kappa) - \theta}{\kappa(1-\beta)(1+\alpha)} \tag{4.17}$$

These three cutoff values determine the optimal strategy for player 1, as defined in proposition 4.1.

Alternatively, one can rearrange these equations to solve for the critical level of social capital at which each strategy is optimal for a given institutional quality, by defining $\beta_{AB}(i), \beta_{AC}(i)$ and $\beta_{BC}(i)$. In particular, I refer to $\beta_{ii}(0)$ below to denote the level of social capital at which there is a change in strategy in the absence of contracting institutions. Assumption 1 regarding the cost of signal acquisition (or civic engagement) is necessary for $\beta_{AB}(0) < \beta_{BC}(0)$ - when contracting institutions are non-existent, there exists a range of social capital levels at which it is optimal for agents to acquire signals. If this were not true, then there would be no institutional value for which signal acquisition would be an optimal strategy.

A Sequential Equilibrium

As mentioned, a sequential equilibrium requires that all players strategies contain optimal choices at any information set (on or off the equilibrium path), given their beliefs, and that those beliefs are consistent with the strategies chosen by the respective players. The beliefs of player 1 and 2 are the easiest to deal with. Player 1 believes that player 2 will follow his dominant strategies. Player 2's beliefs about player 1 do not affect the equilibrium in any way as has full information regarding player 1's strategy when he makes his choice.

The requirement that all players make optimal choices at every information set only has the potential to affect strategy B, in which player 2's strategy calls for the acquisition of signals, and then offering a contract if and only if the signal was good. The next lemma shows that strategy is still optimal after the signal is acquired.

Lemma 4.7. *For institutional values at which each strategy of player 1 is preferred, that strategy is optimal at every information set.*

Proof - see appendix. On the institutional range in which strategy B is optimal ex ante, it is optimal to offer contracts after good signals, and not offer contracts after bad signals.

The payoff functions for player 1 are well-defined and finite-valued, and therefore at least one must be weakly higher than the other two. Define the optimal strategy for player 1 as the strategy A,B, or C associated with the highest payoff.

$$U_1(i,\beta) = \underset{A,B,C}{\operatorname{argmax}}\{U_{1A}(i,\beta), U_{1B}(i,\beta), U_{1C}(i,\beta)\} \tag{4.18}$$

If player 1 chooses the strategy (A,B,C) that maximizes his payoff, and player 2 follows his dominant strategy, and both believe that the other will do so, then the equilibrium is both rational and consistent, and forms a sequential equilibrium. Therefore an equilibrium exists, proving the first part of proposition 4.1.

Lemma 4.8. *If $i \notin \{I_{AB}(\beta), I_{AC}(\beta), I_{BC}(\beta)\}$, the equilibrium is unique along the equilibrium path.*

There are three strategies under consideration, and if any two result in the identical payoff, then $i \in \{I_{AB}(\beta), I_{AC}(\beta), I_{BC}(\beta)\}$. Therefore, under the restriction noted, the 3 strategies for player 1 must provide different payoffs. Further, all 3 payoff functions are well-defined. Therefore, one must provide a strictly higher payoff to player 1, and therefore forms that player's strategy when combined with optimal play off the equilibrium path. If player 1 or player 2 chose a different strategy, that would not be sequentially rational. And if either believed that the other would not choose their optimal strategy, that would not be consistent. Therefore, there cannot exist any other sequential equilibrium. This proves the second part of the proposition, concerning uniqueness[94].

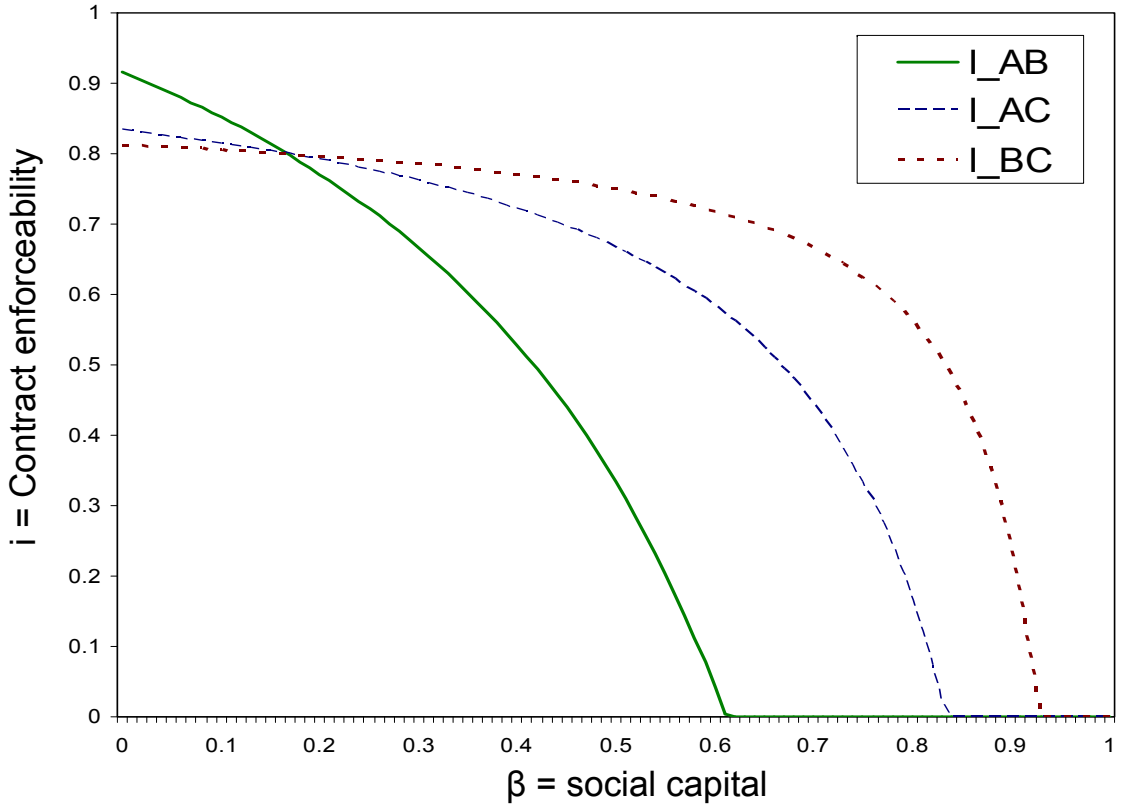
At the points with multiple equilibria, player 1 has multiple strategies that provide an identical payoff, and therefore also has a range of mixed strategies that are also optimal. These are the only points at which mixed strategies can be optimal for either player.

[94] If one is concerned about the strategies on the off-equilibrium path having a unique solution, then there are three more specific institutional levels that must be excluded from the lemma. Either way, there is a small number of cases in which there could be multiple equilibria.

4.4 Characterization of the Equilibrium

The strategy space is easily represented graphically in terms of the cutoff values for each strategy[95].

Figure 4.1: Critical Values for Player 1's Choice



The key feature of this graph is the fact that the 3 cut-off lines cross at a single point. At this institutional pair,

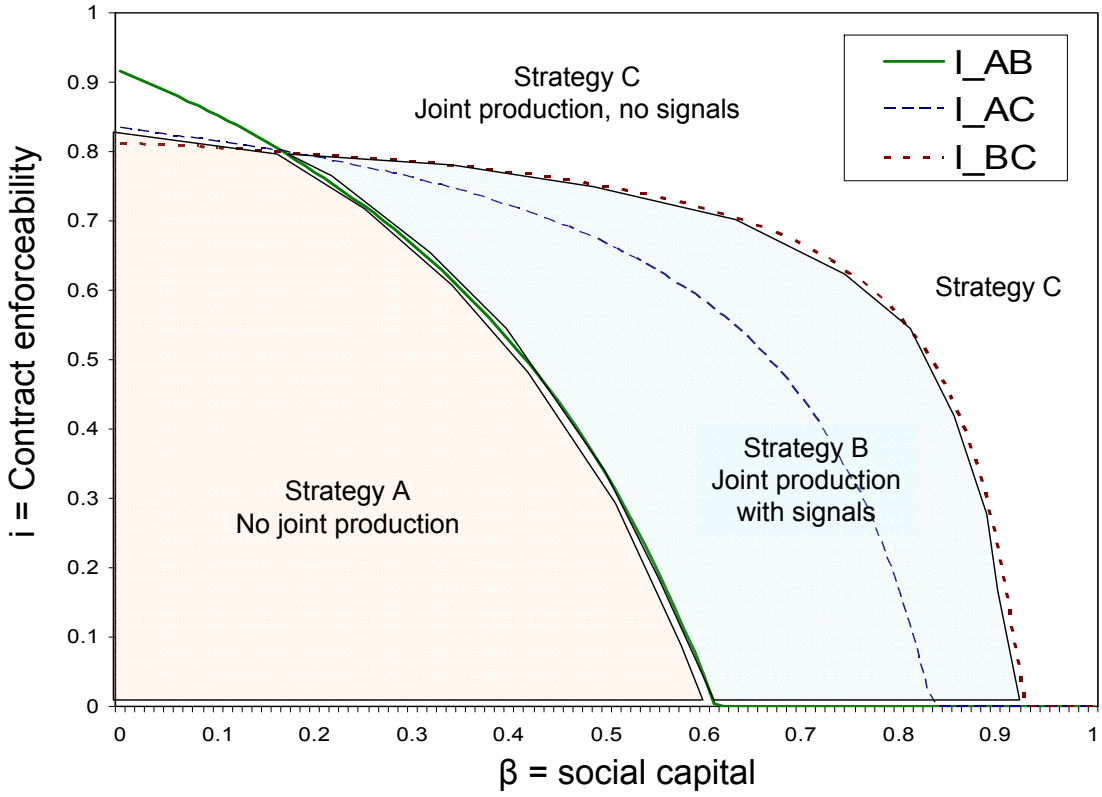
$$(I^*, \beta^*) = \left(\frac{2\alpha\kappa - \alpha - \theta - \theta\alpha}{(1+\alpha)(2\alpha\kappa - \alpha - \theta)}, \frac{\theta}{\alpha(2\kappa - 1)} \right) \tag{4.19}$$

the expected utility of each strategy for player 1 is equal to zero. Further, it is easy to show that with zero social capital, $\beta = 0$, the ordering of critical values is as pictured, $I_{BC}(0) < I_{AC}(0) < I_{AB}(0)$. Similarly, given assumption 1, it is easy to show that $\beta_{AB}(0) < \beta_{AC}(0) < \beta_{BC}(0)$.

As shown in figure 4.2, there are three relevant regions of strategies. In the lower

[95] All figures in the paper are calculated with parameter values $\alpha = 0.20, \kappa = 0.8$, and $\theta = 0.02$)

Figure 4.2: Player 1 Strategy as a Function of Institutions



left corner, where both contracting institutions and social capital are low, players choose to avoid joint production. When either institutions or social capital are high (or both), all players enter into joint production. In the skewed wedge-shaped region in the middle, the equilibrium strategy calls for the acquisition of signals, and then continuing to joint production if and only if there is a good signal.

4.5 The Evolution of Preferences

Whereas contracting institutions are a conscious decision of the government, social norms are assumed to adapt through a process of cultural evolution. While the underlying process is naturally complex, the model developed here is exceedingly simple. The primary reason for this is that the level of detail that must be added to the structure to derive complete analytical solutions would necessarily require specific assumptions on this process that would have little analytical value. The goal here is to provide an analysis of the interaction between government institutions and social norms that is general to a

variety of processes of cultural evolution.

The path of cultural evolution is defined by the success of honest individuals within a society, relative to the opportunistic individuals. When trustworthy individuals are seen to be rewarded, it is assumed that there is an underlying shift in preferences toward honesty. In contrast, when trustworthy individuals are at a disadvantage there is a shift toward opportunism.

The relative success of honest and opportunistic individuals is determined by the decisions to acquire signals of behaviour type or not, and to enter joint production or not. In all cases, half the agents, those assigned by nature to be player 1, will have identical expected utility regardless of behaviour type. The key for the evolutionary argument is the relative utility of honest and opportunistic agents when assigned by nature to be player 2. The following 3 lemmas define the relative utility of individuals assigned to be player 2, under each of the 3 types of equilibria that exist.

Lemma 4.9. *When joint production does not occur (equilibrium A), honest and opportunistic agents have identical utility* $(U_{hA} = U_{oA})$.

When no one is willing to enter joint production, all agents have an expected utility of 0.

Lemma 4.10. *If joint production occurs only after the revelation of a positive signal (equilibrium B), honest agents have strictly higher expected utility than opportunistic agents if and only if institutions are sufficiently strong.* $\left(U_{hB} > U_{oB} \text{ if and only if } i > \frac{1-\kappa-\alpha\kappa}{(1-\alpha)(1-\kappa)}\right)$

Proof - see appendix. There is contrasting evolutionary pressure in equilibrium B, as honest agents are more likely to be hired than dishonest agents, but dishonest agents receive a higher expected utility when they are hired as they have the possibility to gain from shirking. If contracts are sufficiently enforceable, then the gains from shirking are limited, and therefore the increased rate of hiring for honest agents dominates this comparison. Further, if signals are sufficiently informative, $\kappa > \frac{1}{1+\alpha}$, then honest agents have higher utility in this equilibrium, regardless of institutional quality. For future reference, I define this critical institutional level as:

$$i^e = \max\left\{0, \frac{1-\kappa-\alpha\kappa}{(1-\alpha)(1-\kappa)}\right\} \tag{4.20}$$

Lemma 4.11. *When agents enter joint production without acquiring signals (equilibrium C), and contracting institutions are less than perfect ($i < 1$), honest agents have strictly lower expected utility than opportunistic agents* $(U_{hB} < U_{oB})$.

Proof - see appendix. In this equilibrium, honest agents gain no hiring advantage, but the advantages of shirking remain for dishonest agents. The only caveat is that if

contract enforceability were perfect, then there would be no advantage from shirking and all agents receive the same utility.

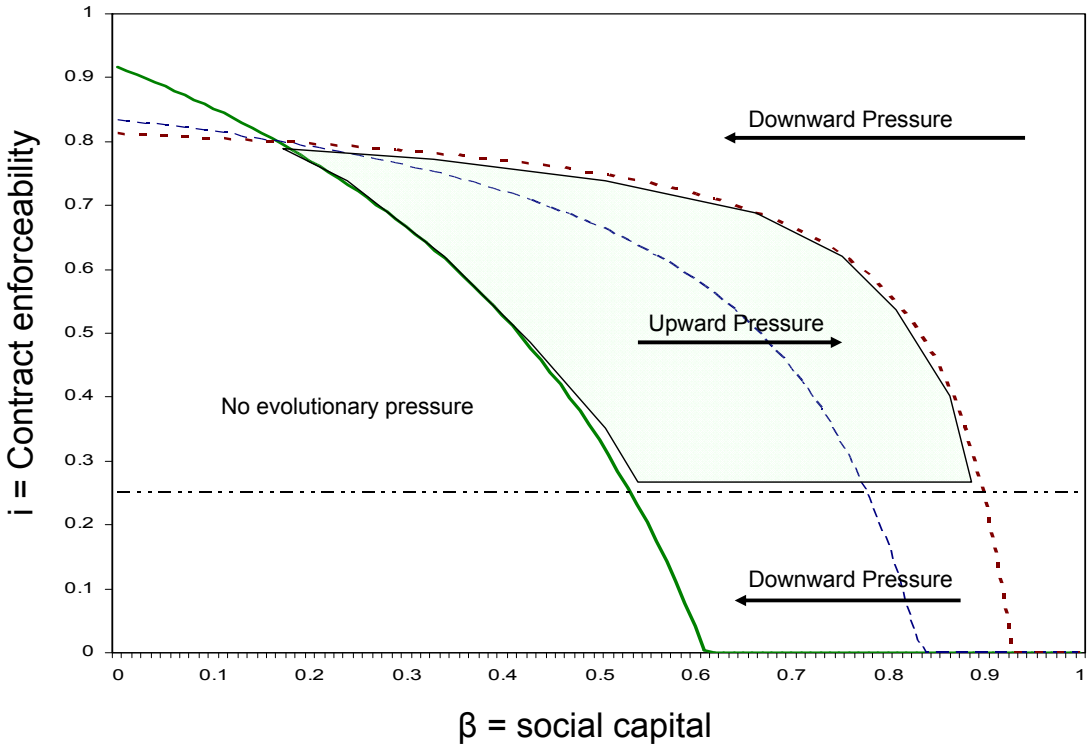
4.5.1 The Evolutionary Process

The evolutionary process assumed to be at work here is extremely simple. Assuming interior values of β, if honest types have higher utility than opportunistic types then the proportion of honest types will increase. Define F_h and F_o, as the average fitness of honest and opportunistic individuals, respectively. The evolutionary process then follows the functional form:

$$\beta_{t+1} - \beta_t = \beta_t(1 - \beta_t) * G[F_h - F_o) \tag{4.21}$$

where $G(a, a) = 1$ and $G'() > 0$. Perfect honesty ($\beta = 1$) and perfect opportunism ($\beta = 0$) are therefore absorbing states, while at intermediate levels, evolutionary pressure will push in the direction of the behaviour type with higher average fitness.

Figure 4.3: The Evolution of Preferences



This evolutionary process, with the results of Lemmas 4.9 to 4.11, creates the simple

dynamics indicated in figure 4.3. Where agents avoid joint production, there is no evolutionary pressure. When joint production occurs, evolutionary pressure will increase social capital when signals are acquired and institutions are sufficiently strong, and reduce social capital otherwise. Where multiple equilibria exist, the dynamics depend on the specific equilibrium selected. For example, along the I_{BC} cutoff, the equilibrium must result in a fraction γ of agents following equilibrium B, and acquiring signals of type, and a fraction $1-\gamma$ choose not to acquire signals and simply enter joint production with anyone. The fraction γ will determine whether honest or opportunistic agents do better under the selected equilibrium.

4.6 Institutional Interaction

From this model, it is possible to identify three separate interaction effects between the two institutional forms. First, there is an **economic activity** effect, where contracting institutions promote the joint production that is partially dependent on trust. Second, there is an **enforcement** effect, in which contracting institutions are necessary to limit the gains from opportunism. Third, there is a **crowding out** effect, where contracting institutions replace the screening mechanism that preferentially hires the honest types. The combination of the three effects varies significantly accordingly to the level of generalized trust in society, as discussed in section 4.7.

4.6.1 An Economic Activity Effect

For countries with relatively low trust levels, $\beta < \beta_{AB}(0)$, in the absence of contracting institutions, the economy would be unable to undertake joint production. If contracting institutions are sufficiently strong, then the risk associated with joint production is mitigated. Thus contracting institutions can generate the economic activity that is necessary before social capital can develop.

When screening is imperfect, $\kappa < 1$, this will entail a non-empty set of social capital levels[96], at which there exists an institutional level that will promote joint production. As joint production is necessary for social capital development, institutions can have a complementary effect on social capital. Effectively, contracting institutions reduce the cost of mistakes in screening, and therefore encourage the use of joint production.

[96]When screening is perfect, $\kappa = 1$, the I_{AB} line is undefined, and lies vertically at the point $\beta = \frac{\theta}{\alpha}$. For social capital levels below that, strategy A is preferred to strategy B, and for social capital levels above that, strategy B is preferred to strategy A. Critically, institutions have no effect when screening is perfect.

4.6.2 A Crowding-Out Effect

However, institutions also crowd out the screening process. When an agent screens for type, institutions only matter when her partner is opportunistic *and signals are incorrect*. In contrast, when agents do not engage in screening, the institution matters whenever ones partner is opportunistic. For this reason, at any level of social capital improvements in institutions will have a larger increase in the expected utility of agents that choose strategy C (no signals), versus agents that acquire signals and condition on their outcome, strategy B.

4.6.3 An Enforcement Effect

The two previous effects generate the strategy space shown in figure 4.2, where there is the potential for a range of social capital levels at which joint production occurs after a screening process. However, agents following such a strategy is not necessarily sufficient to generate positive evolution of social capital. In addition, the benefits that accrue to opportunistic agents that are lucky enough to be hired must be relatively small. As was mentioned above, the evolutionary pressure is determined by the trade-off between a greater chance of being hired for the honest, against a higher benefit if they are hired for the opportunistic. As the extra benefit to the opportunistic only occurs when contracts are unenforceable, increasing institutional effectiveness limits these gains. If signals are sufficiently strong, then the enforcement effect is unnecessary.

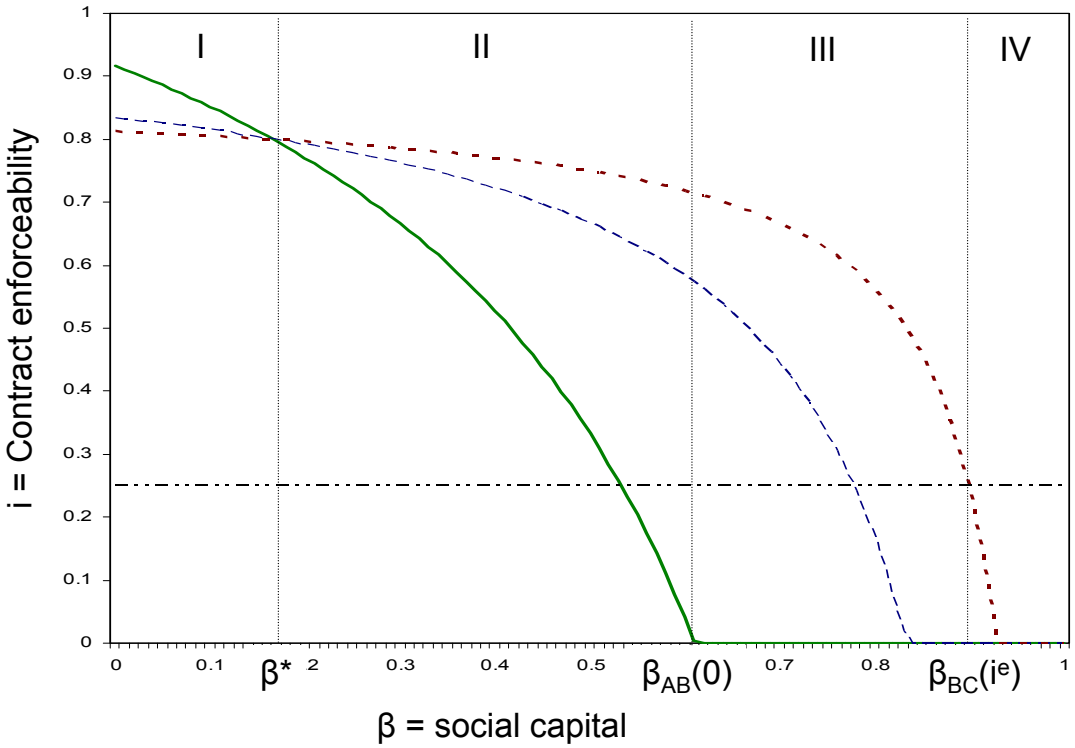
4.7 Social Capital Regimes

When can government be an effective force in promoting social capital development? Tabellini (2008) suggests that increases in the formal contracting environment will lead to higher levels of social capital. This is driven by the fact that the model in that paper is primarily based on the enforcement effect seen here, and critically, there is no observation of preferences, and therefore no crowding out effect. In this section I show that at low and high levels of social capital, government institutions cannot promote its development, while at intermediate levels of social capital, there are government institutions that could be effective. This result is driven by the relative importance of the three interaction effects defined in section 4.6.

Proposition 4.2. *If $\beta < \beta^* = \frac{\theta}{\alpha(2\kappa-1)}$, or $\beta \geq \beta_{BC}(i^e)$ there is no contracting institution that can cause the social capital in the economy to increase, regardless of the intentions of government. Conversely, if $\beta^* \leq \beta < \beta_{BC}(i^e)$, there exists a level of contracting institutions that will promote the development of social capital.*

Proof - see appendix. Figure 4.4 shows graphically the critical values associated with the proposition, and the discussion that follows describes the relative importance of each of the interaction effects. To do so, I further divide the intermediate region of social capital into two parts, with the split occurring at $\beta_{AB}(0)$.

Figure 4.4: Social Capital Regimes



4.7.1 An Institutional Trap

β^* corresponds to the social capital level associated with the intersection of the 3 cutoff values defined above. Intuitively, when social capital is below this level, there are two possibilities. First, contracting institutions could be relatively weak, and in this case it is not worthwhile to pursue joint production. Alternatively, contracting institutions are strong enough to encourage joint production, however they are also strong enough to crowd out the use of a screening strategy. In this case, by lemma 4.11, social capital will decrease. As such, any society that finds itself with social capital lower than β^*, social capital will not increase, though.

This result is in contrast to the outcome of Tabellini (2008), in which contracting

institutions would promote social capital in any society, but political economy considerations prevent the development of effective institutions, and therefore generate an institutional trap. Here, it is simply not possible to use contracting institutions to promote social capital if social capital is too low.

However, it should be added that this social capital trap does not necessarily create an institutional trap that affects contracting institutions. If formal institutions are cheap enough, it may be efficient to have strong contracting institutions and allow social capital to decay toward zero. In particular, if institutions are free, then an equilibrium with perfect contracting is perfectly efficient, though social capital would tend toward zero[97].

4.7.2 Institutions to Promote Production

For countries with $\beta \in [\beta^*, \beta_{AB}(0)$, there exists a contracting institution that will generate the development of social capital. Contracting institutions are necessary to promote joint production, and also possibly to ensure that honest agents gain disproportionately from that joint production. In the absence of sufficiently strong contract enforceability, there is insufficient protection for agents to enter joint production, and social capital does not increase or decline. If institutions are overly strong, joint production will occur without screening, and therefore social capital will decline. In a middle range, institutions are strong enough to promote joint production without crowding out the use of social signals, and provide sufficient enforcement to ensure that honest agents achieve higher fitness than opportunistic agents.

The importance of contracting institutions in generating the incentives supporting the acquisition of signals is critically tied to the imperfections in the signals. As signals improve, $(\kappa \to 1)$, the $I_{AB}(\beta)$ line approaches vertical, and this social capital regime no longer exists $(\beta^* = \beta_{AB}(0))$. When signals are perfect, agents that choose the strategy of acquiring signals derive no benefit from improving contracts, as they only continue to joint production with agents that are honest.

4.7.3 Institutions as Enforcement

Similarly, for countries with $\beta \in [\beta_{AB}(0), \beta_{BC}(i^e))$, there also exists a contracting institution that will generate the development of social capital. However, in contrast with the previous range, contracting institutions are not necessary to promote joint production, as generalized trust is sufficient to allow joint production. However, in the absence of sufficient contracting institutions, social capital would decline over time as opportunistic

[97] If institutions are actually perfect, then there is no decay of social capital, however if contracting institutions are slightly less than perfect $(i = 1 - \epsilon)$, then social capital will decline slowly.

individuals are the advantaged group. In addition, there is a maximum level of contracting institutions, beyond which no screening would take place and social capital would decline. If the screening process is sufficiently accurate, $\kappa \geq \frac{1}{1+\alpha}$, institutions have no role in enforcement, and this regime is characterized only by a maximum level of institutions that is consistent with stable or increasing levels of social capital.

4.7.4 Maximum Social Capital

For a country with $\beta > \beta_{BC}(i^e)$, it is impossible to maintain the existing levels of social capital. The economy runs on joint production, and whether or not screening is conducted, the opportunistic in society have the advantage. The starkness of this result is softened by the fact that there is no mechanism in the model by which any society would reach such a level of social capital in the first place. This regime is therefore effectively defined by the maximum level of stable social capital, $\beta_{BC}(i^e)$.

4.8 Social Diversity and the Effectiveness of Institutions

Social divisions, such as ethnic or religious diversity or wealth inequality, have been linked to lower levels of social capital (Knack and Keefer 1997). Further, ethnic diversity has been linked consistently with low levels of institutional development (Easterly 2001). A possible effect of social divisions is to either increase the cost of acquiring information about a prospective partner, or to reduce the accuracy of the information obtained. Within the context of the model developed here, this would mean increasing θ or reducing κ. Either change has the effect of increasing range of the institutional trap associated with low social capital.

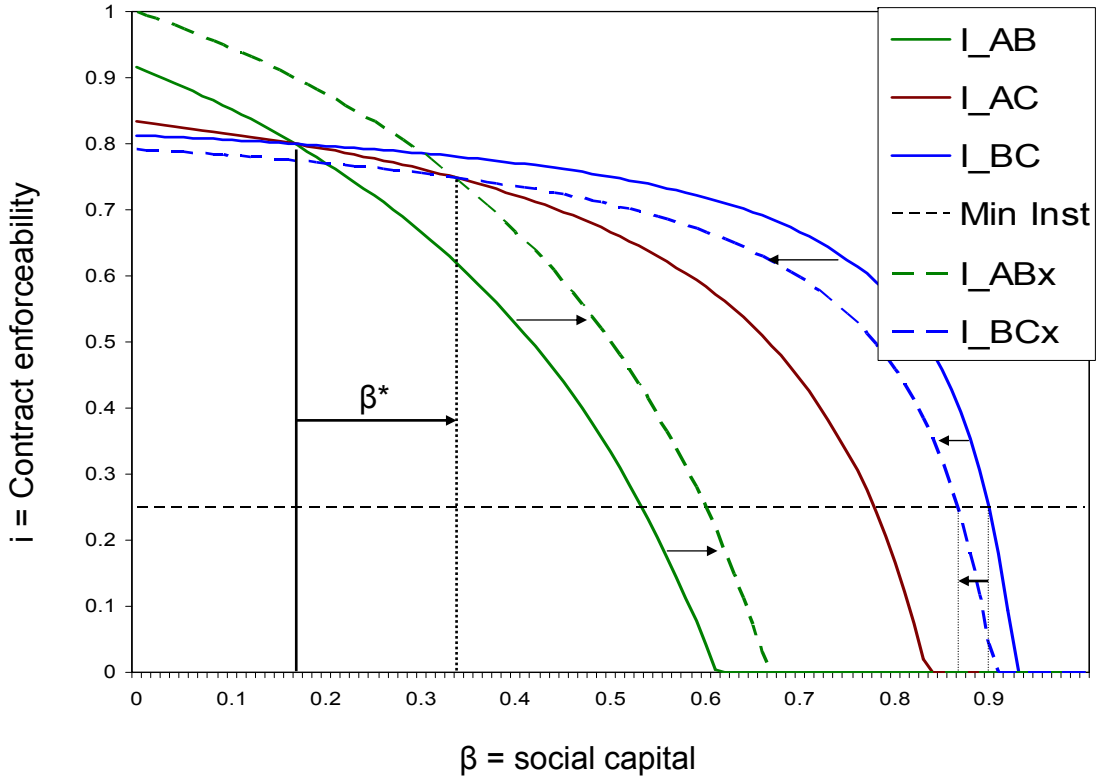
Proposition 4.3. *Social divisions that reduce the effectiveness of signals, or increase their cost, strictly increase the critical value of social capital below which a society is caught in an institutional trap with low social capital, and strictly decrease the maximum level of social capital. (*$\frac{\partial \beta^*}{\partial \theta} > 0, \frac{\partial \beta^*}{\partial \kappa} < 0, \frac{\partial \beta_{BC}(i^e)}{\partial \theta} < 0, \frac{\partial \beta_{BC}(i^e)}{\partial \kappa} > 0$*)*

Proof - see appendix. Increasing the cost of signals, or reducing their effectiveness, reduces the institutional space in which the acquisition of signals is optimal, thus increasing the level of social capital necessary to escape a social capital trap, and limiting the maximum development of social capital.

Increasing θ, or reducing κ, both have the effect of reducing the profitability of acquiring signals, thus shifting out the I_{AB} curve, and shifting in the I_{BC} curve, thus shifting their intersection to the right. This is shown graphically in figure 4.5[98]. The other two

[98]The solid lines in figure 4.5 have not changed from the previous graphs, while the dotted lines are the result of increasing the cost of signal acquisition from $\theta = 0.02$ to $\theta = 0.04$.

Figure 4.5: An Increase in the Cost of Signals



lines from the figure, the I_{AC} curve, and the line for a necessary set of institutions for social capital development, are not affected by a change in the cost of signal acquisition[99]. As in the proposition, and as shown in the figure, the range of social capital levels in regime 1, the institutional trap, and in regime 4, above the maximum social capital level, increase with social divisions. This result is consistent with the data demonstrating a negative link between social divisions and social capital.

4.9 Conclusion

There are two key approaches to modeling the evolution of generalized trust in response to changes in the forml institutional environment. Tabellini (2008) follows the work of Bisin and Verdier (2001) and models the evolutionary process as one of parental socialization. Modeling trust in this way generates a particular set of results, notably that

[99]If instead the figure demonstrated a reduction in the effectiveness of signals, the i^e line would also shift up, thus further reducing the maximum level of social capital.

improvements in the formal contracting environment will always lead to higher levels of social capital.

While there is extensive empirical support for the parental socialization approach, there is also considerable support for the idea that people are able to, at least imperfectly, identify whether a potential partner will cheat if given the opportunity. This paper models the interaction between legal institutions and the cultural evolution of trust that can occur if agents are able to screen potential partners for trustworthiness at a cost.

When that screening is imperfect, then sufficiently strong government institutions can promote the development of social capital by reducing the risk associated with joint production in the face of a moral hazard problem, and reducing the potential opportunities for dishonest behaviour. However, the use of institutions to promote social capital is limited, as at high levels of government involvement, the civic engagement that is critical to the development of social capital is crowded out.

In countries with low levels of trust and trustworthiness, government institutions are unable to promote risky production without crowding out civic engagement. This results in the presence of a critical level of social capital, below which a society will find themselves in an institutional trap. A secondary feature of the model is the presence of a maximum level of social capital, beyond which an economy cannot move.

The advent of an institutional trap defined here is quite different from the institutional trap present in related work by Tabellini (2008), where the trap is created primarily by political economy concerns, where the group of opportunistic individuals gains from poor institutions and therefore blocks the development of better institutions. In contrast, the model developed here does not depend on a political economy consideration - a benevolent dictator would be similarly unable to extract the economy from the trap through the use of government institutions. Of particular importance is the empirical question of how social capital responds to improvements in legal institutions when all institutions are weak. The model of Tabellini (2008) would suggest that social capital would increase in response to changes in legal institutions, whereas the model developed here would suggest that at very low initial levels of social capital there would be no change.

I then consider the role of social divisions on the development of social capital and contracting institutions. If social divisions increase the cost or reduce the effectiveness of social signaling, then those divisions will tend to exacerbate the problems of an institutional trap. This result is consistent with the extensive literature demonstrating the negative effects of ethnic diversity on the development of both formal and informal institutions. In addition, these results provide a positive message regarding the development of social capital. If government choices can promote civic engagement, by making investments in social signals cheaper, or more informative, then those choices may be

more effective in promoting social capital than any investment in contracting institutions. In addition, to escape a social capital trap, it may be necessary to strongly promote civic engagement while developing an effective, intermediate level of legal institutions to promote economic activity. This suggests the possibility that economic development requires the interaction of various institutional elements, rather than a simple focus on one element of institutional quality.

Chapter 5

Conclusion

There is no question that cultural values, such as social norms of honesty or the relevance of ethnic identity, are difficult concepts to encapsulate accurately within an economic context. For this reason, the economics literature has primarily focused on tangible forms of institutions such as government structure or the rule of law. The argument outlined in this thesis is that ignoring the effect of cultural institutions limits our understanding of the process of economic development. In particular, the effectiveness of various government institutions depends on their interaction with specific social aspects of a society. In this type of environment, it is important that the development of government institutions is only done after careful consideration of the cultural environment. The developing literature on growth diagnostics[100] is representative of the notion that institutional development occurs in the context of an existing set of formal and informal institutions.

In chapters 2 and 3, I focus on the role of ethnic diversity in limiting government effectiveness. In particular, ethnic diversity stands as a barrier to the resolution of basic collective action problems that are necessary for development. The results indicate that ethnic diversity, at both local and regional scales, have significantly limited the provision of basic goods such as piped drinking water and electricity. Further, as the evidence in chapter 3 suggests, this limitation is not the result of unresolvable preference differences across groups. Rather, the effects of ethnic diversity, specifically in the case of piped drinking water, result from the inability of diverse ethnic groups to coordinate appropriate actions.

Further research is necessary to understand the policy choices that governments can make to overcome these difficulties. A particularly promising avenue of research is related to government decentralization. If, as demonstrated in chapter 2, diversity matters at the scale of provision, it might be critical to devolve power over decision-making to an equivalent level of government. Specifically, this might be most important in countries with high rates of local and national diversity. The empowerment of lower levels of government, especially within Sub-Saharan Africa, may be an important step toward development.

The third essay focuses more directly on institutional change. Social norms of hon-

[100] See Hausmann et al. (2008) for a summary.

esty reflect underlying cultural values that have proven very persistent over time. This may create the appearance of persistence in government institutions, even when those institutions are relatively flexible. The existing literature within economics on institutional change is extremely limited, in part, I believe, because the social aspects of the institutional framework have been ignored. This paper contributes to the literature by identifying how government institutions affecting contract enforcement may have a limited effect on the promotion of social capital. This limitation is particularly true in the presence of social divisions caused by ethnic heterogeneity or economic inequality. Further research in this area will focus on the interaction between government choices and changes in the social framework of society. A paper currently under development connects the two streams of research presented here as it considers the effect that formal legal institutions and globalization have on the evolution of inter-ethnic relations.

Bibliography

ACEMOGLU, D., AND S. JOHNSON (2005): "Unbundling Institutions," *Journal of Political Economy*, 113, 949–995.

ACEMOGLU, D., S. JOHNSON, AND J. A. ROBINSON (2001): "The Colonial Origins of Comparative Development:An Empirical Investigation," *The American Economic Review*, 91, 1369–1401.

——— (2002): "Reversal of Fortune:Geography and Institutions in the Making of the Modern World Income Distribution," *The Quarterly Journal of Economics*, 117, 1231–1294.

AHN, T., AND J. ESAREY (2008): "A Dynamic Model of Generalized Social Trust," *Journal of Theoretical Politics*, 20(2), 151.

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.

BANERJEE, A., L. IYER, AND R. SOMANATHAN (2005): "History, Social Divisions, and Public Goods in Rural India," *Journal of the European Economic Association*, 3(2-3), 639–647.

BANERJEE, A., AND R. SOMANATHAN (2007): "The Political Economy of Public Goods: Some Evidence from India," *Journal of Development Economics*, Forthcoming.

BANERJEE, S., A. DIALLO, AND Q. WODON (2007): "Measuring trends in access to modern infrastructure in Sub-Saharan Africa: Results from Demographic and Health Surveys," Mimeo.

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.

BESLEY, T., AND S. COATE (2003): "Centralized versus decentralized provision of local public goods: a political economy approach," *Journal of Public Economics*, 87, 2611–2637.

BIDNER, C., AND P. FRANCOIS (2009): "Cultivating Trust:Norms, Institutions and the Implications of Scale," Mimeo.

BISIN, A., AND T. VERDIER (2001): "The economics of cultural transmission and the dynamics of preferences," *Journal of Economic Theory*, 97(2), 298–319.

BLANTON, R., T. D. MASON, AND B. ATHOW (2001): "Colonial Style and Post-Colonial Ethnic Conflict in Africa," *Journal of Peace Research*, 38, 473–491.

BOHNET, I., B. FREY, AND S. HUCK (2001): "More Order with Less Law: On Contract Enforcement, Trust, and Crowding," *American Political Science Review*, 95(01), 131–144.

BOWLES, S., S. DURLAUF, AND K. HOFF (2006): *Poverty Traps*. Princeton University Press.

BROSIO, G. (2000): "Decentralization in Africa," Mimeo.

BROWN, W., AND C. MOORE (2000): "Is prospective altruist-detection an evolved solution to the adaptive problem of subtle cheating in cooperative ventures? Supportive evidence using the Wason selection task," *Evolution and Human Behavior*, 21(1), 25–37.

BROWN, W., B. PALAMETA, AND C. MOORE (2003): "Are there nonverbal cues to commitment? An exploratory study using the zero-acquaintance video presentation paradigm," *Evolutionary Psychology*, 1, 43.

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.

CARLIN, B., F. DOROBANTU, AND S. VISWANATHAN (2009): "Public trust, the law, and financial investment," *Journal of Financial Economics*.

CASELLI, F., AND W. J. C. II (2010): "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.

COLEMAN, J. (1988): "Social Capital in the Creation of Human Capital," *American Journal of Sociology*, 94(S1), 95.

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.

——— (2007): *The Bottom Billion: Why the Poorest Countries Are Failing and What Can Be Done About It*. Oxford University Press.

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.

DEKEL, E., J. C. ELY, AND O. YILANKAYA (2007): "Evolution of Preferences," *Review of Economic Studies*, 74, 685–704.

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.

EATON, B., AND M. ESWARAN (2003): "The evolution of preferences and competition: a rationalization of Veblen's theory of invidious comparisons," *Canadian Journal of Economics*, pp. 832–859.

ELLINGSEN, T. (1997): "The Evolution of Bargaining Behavior*," *Quarterly Journal of Economics*, 112(2), 581–602.

FAFCHAMPS, M. (2006): "Development and social capital," *Journal of Development Studies*, 42(7), 1180–1198.

FASS, S. M., AND G. M. DESLOOVERE (2004): "Chad: Governance by the Grassroots," in *Local Governance in Africa: The Challenges of Democratic Decentralization*, pp. 155–180. Lynne Rienner Publishers.

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

FILIPPINI, M. (1998): "Are Municipal Electricity Distribution Utilities Natural Monopolies?," *Annals of Public and Cooperative Economics*, 69, 157–174.

FRANCOIS, P. (2008): "Norms and Institution Formation," Mimeo.

FRANCOIS, P., AND J. ZABOJNIK (2005): "Trust, Social Capital and Economic Development," *Journal of the European Economics Association*, 3, 51–94.

FRANK, M. G., P. EKMAN, AND W. V. FRIESEN (1993): "Behavioral markers and recognizability of the smile of enjoyment," *Journal of personality and social psychology*, 64(1), 83–93.

FRANK, R. H. (1987): "If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience?," *American Economic Review*, 77, 593–604.

FREY, B. (1997): "A constitution for knaves crowds out civic virtues," *The Economic Journal*, pp. 1043–1053.

FUKUYAMA, F. (2001): "Social capital, civil society and development," *Third World Quarterly*, 22(1), 7–20.

GARCIA-MONTALVO, J., AND M. REYNAL-QUEROL (2002): "Why Ethnic Fractionalization? Polarization, Ethnic Conflict and Growth," Mimeo.

GLAESER, E., D. LAIBSON, J. SCHEINKMAN, AND C. SOUTTER (2000): "Measuring Trust*," *Quarterly Journal of Economics*, 115(3), 811–846.

GREIF, A. (2005): "Commitment, Coercion, and Markets: The Nature and Dynamics of Institutions Supporting Exchange," *Handbook of New Institutional Economics*, pp. 727–786.

GUISO, L., P. SAPIENZA, AND L. ZINGALES (2006): "Does Culture Affect Economic Outcomes?," *The Journal of Economic Perspectives*, 20(2), 23–48.

——— (2008): "Long Term Persistence," NBER Working Papers 14278.

GÜTH, W., AND H. KLIEMT (1994): "Competition or Co-operation: On the Evolutionary Economics of Trust, Exploitation and Moral Attitudes," *Metroeconomica*, 45(2), 155–187.

——— (1998): "The Indirect Evolutionary Approach: Bridging the Gap Between Rationality and Adaptation," *Rationality and Society*, 10(3), 377.

GÜTH, W., AND A. OCKENFELS (2005): "The Coevolution of Morality and Legal Institutions: An Indirect Evolutionary Approach," *Journal of Institutional Economics*, 1(02), 155–174.

HABYARIMANA, J., M. HUMPHREYS, D. POSNER, AND J. WEINSTEIN (2007): "Why does ethnic diversity undermine public goods provision?," *American Political Science Review*, 101(04), 709–725.

HAUSMANN, R., D. RODRIK, AND A. VELASCO (2008): "Growth diagnostics," *The Washington Consensus Reconsidered: Towards a New Global Governance*, p. 324.

HEIFETZ, A., C. SHANNON, AND Y. SPIEGEL (2007): "The Dynamic Evolution of Preferences," *Economic Theory*, 32(2), 251–286.

JALAN, J., AND M. RAVALLION (2003): "Does piped water reduce diarrhea for children in rural India?," *Journal of Econometrics*, 112(1), 153–173.

JARSSO, S. (2002): "Water Sector Development Programme," *The Federal Democratic Government of Ethiopia - Ministry of Water Resources*.

KAREKEZI, S., AND J. KIMANI (2002): "Status of power sector reform in Africa: impact on the poor," *Energy Policy*, 30(11-12), 923–945.

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.

KNACK, S., AND P. KEEFER (1997): "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *The Quarterly Journal of Economics*, 112, 1251–1288.

KUMLIN, S., AND B. ROTHSTEIN (2005): "Making and Breaking Social Capital: The Impact of Welfare-State Institutions," *Comparative Political Studies*, 38(4), 339.

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

LUTTMER, E. (2001): "Group Loyalty and the Taste for Redistribution," *Journal of Political Economy*, 109(3), 500–528.

MANSURI, G., AND V. RAO (2004): "Community-based and-driven development: a critical review," *The World Bank Research Observer*, 19(1), 1–39.

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.

NORTH, D. (1994): "Economic performance through time," *The American Economic Review*, pp. 359–368.

NORTH, D. C. (1991): "Institutions," *Journal of Economic Perspectives*, 5, 97–112.

OLOWU, D., AND J. WUNSCH (2004): *Local governance in Africa: The challenges of democratic decentralization*. Lynne Rienner Publishers.

OLSON, M. (1965): *The logic of collective action: Public goods and the theory of groups*. Harvard University Press.

OUEDRAOGO, H. (2003): "Decentralisation and local governance: experiences from francophone West Africa," *Public administration and development*, 23(1).

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.

PANIZZA, U. (1999): "On the determinants of fiscal centralization: Theory and Evidence," *Journal of Public Economics*, 74, 97–139.

PLATTEAU, J. (1994): "Behind the market stage where real societies exist-part I: The role of public and private order institutions," *Journal of Development Studies*, 30(3), 533–577.

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.

PUTNAM, R. (1995a): "Bowling Alone: America's Declining Social Capital," *Journal of Democracy*, 6, 65–65.

PUTNAM, R. D. (1993): *Making Democracy Work:Civic Traditions in Modern Italy*. Princeton University Press.

——— (1995b): "Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America," *Political Science and Politics*, 28(4), 664–683.

ROLAND, G. (2004): "Understanding institutional change: Fast-moving and slow-moving institutions," *Studies in Comparative International Development (SCID)*, 38(4), 109–131.

SACHS, J. D. (2005): *The End of Poverty: Economic Possibilities For Our Time*. The Penguin Press.

SOBEL, J. (2002): "Can We Trust Social Capital," *Journal of Economic Literature*, 40, 139–154.

TABELLINI, G. (2005): "Culture and Institutions: Economic Development in the Regions of Europe," Mimeo.

——— (2008): "The Scope of Cooperation: Values and Incentives," *Quarterly Journal of Economics*, 123(3), 905–950.

TEFERRA, M. (2002): "Power sector reforms in Ethiopia: options for promoting local investments in rural electrification," *Energy Policy*, 30(11-12), 967–975.

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

TIEBOUT, C. (1956): "A Pure Theory of Local Expenditures," *The Journal of Political Economy*, 64(5), 416.

TREISMAN, D. (2002): "Defining and measuring decentralization: A global perspective," *unpublished paper. Available at http://www. polisci. ucla. edu/faculty/treisman*.

VAN OORSCHOT, W., AND W. ARTS (2005): "The social capital of European welfare states: the crowding out hypothesis revisited," *Journal of European Social Policy*, 15(1), 5.

VERPLAETSE, J., S. VANNESTE, AND J. BRAECKMAN (2007): "You can judge a book by its cover: the sequel. A kernel of truth in predictive cheating detection," *Evolution and Human Behavior*, 28(4), 260–271.

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

Additional Material for Chapter 2

A.1 Robustness Checks for Piped Water

A.1.1 Piped Water

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

Table A.1 repeats the results of the first 3 columns of table 2.6, but restricts the analysis to narrowly defined rural communities[101] As urban areas are characterized by a larger variety of economic activity and governance structures, which I am unable to properly control for, it is reasonable to assume that the effects of ethnic diversity would be most easily measured in rural communities. Further, the disconnect between scale of provision and de jure scale of governance is most apparent in rural communities across the 13 countries. Perhaps not surprisingly, as rural communities make up approximately 80% of the total, the results are not substantively changed by this restriction. The results are further robust to altering the grid specification, either to a shifted 1km grid, or to a 5km grid.

Table A.2 demonstrates that the choice of diversity measurement does not affect the measured relationship between diversity and the provision of piped drinking water. If anything, using the historical measure of diversity as a proxy for current diversity appears to result in an increase in the measured effects of ethnic diversity, as opposed to just using the historical measure. In addition, measuring diversity by fractionalization or the size of the largest ethnic group, as is done in columns (2) and (3), is irrelevant for the results.

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

Table A.1: Piped Water and Local Diversity - Rural Communities with a Smaller Grid

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

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

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

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

Table A.2: Piped Water and Local Diversity - Alternative Diversity Measures

	Instrument	Largest	IVLargest
	(1)	(2)	(3)
Large.Eth.Grp		.102***	.119***
		(.024)	(.028)
Eth.Fract.	-.088***		
	(.024)		
Avg. Wealth	.337***	.334***	.338***
	(.015)	(.015)	(.015)
Urban	.094***	.096***	.096***
	(.021)	(.021)	(.021)
Pop.Density(1990)	.008	.008	.008
	(.005)	(.006)	(.006)
Dist.to River (LogKM)	.006*	.006*	.006*
	(.003)	(.003)	(.003)
Elevation(km)	-.0006	.002	.0005
	(.014)	(.014)	(.014)
Dist.to Cap.City (LogKM)	.037***	.037***	.037***
	(.010)	(.010)	(.010)
Coast.Comm.	-.059*	-.058*	-.058*
	(.032)	(.032)	(.032)
Avg. Tenure	.001*	.001**	.001*
	(.0006)	(.0006)	(.0006)
Obs.	2828	2828	2828
R^2	.753	.754	.754

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

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

While neither set of instruments are ideal, the results of a Hansen J-test indicate that the instruments are not incorrectly excluded from either specification, and the first stage results suggest a strong relationship between the instruments and our measure of wealth. Further, it is not entirely clear how any bias related to the endogeneity of wealth would effect the key results relating ethnic diversity to public good provision. Given those caveats, the results of the IV-specifications are presented in table A.3.

Table A.3: Piped Water and Local Diversity - Wealth Instruments

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

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

The first two columns of table A.3 present the results of specifications using the standard measure of access to piped water as the dependent variable, while the latter two columns use access at one's home or compound as the dependent variable. The results indicate that the measurement of wealth does not appear to substantively affect the results presented in the main section of the paper, though, as would be predicted by theory, there is a slight reduction in the measured effect of wealth. A Hansen-J test shows that the instruments are properly excluded from the second stage of the regression (p-values

are all insignificant), while a Cragg-Donald F-test shows that the instruments are effective at identifying household wealth in the first stage.

A.1.2 Electricity

Table A.4 demonstrates that the choice of diversity measure is not critical to the analysis. Whether we use ethnic fractionalization, as in the main section, or the size of the largest group (column 2), the results are not substantively changed. Further, using current measures and instrumenting using historical measures does not affect the results either. This is primarily the result of very high correlations between different diversity measures, and a population that faces relatively little migration [102].

As with the specifications for water, there is the possibility that the wealth measure included in the regression is endogenous. Here I include the results using the same 2 sets of instruments that were used in the case of piped drinking water. Again, the results show no significant change under the instrumental variables strategy. The Hansen-J test p-values indicate that the instruments are properly excluded from the regression, however it should be noted that the geographic instruments are not particularly strong in the first stage. As such, the Cragg-Donald statistics show that this limited set of instruments may suffer from a weak first stage. Experimenting with slightly different mixes of the instruments (not reported) does not resolve this problem, but does show that the overall results are not particularly sensitive to the choice of instruments.

[102] As evidence for this statement, note that the average household has lived in the same community for 30 years.

Table A.4: Electricity and Provincial Diversity - Alternative Diversity Measure

	Instrument	LargestGrp	LargestIV
	(1)	(2)	(3)
Ethnic Fractionalization	-.093***		
	(.030)		
Share of Largest Ethnic Group		.111***	.116***
		(.029)	(.031)
Avg. Wealth	.350***	.340***	.342***
	(.056)	(.054)	(.054)
Urban	-.142	-.126	-.127
	(.094)	(.087)	(.088)
Pop.Density	.034***	.034***	.034***
	(.006)	(.005)	(.006)
(ln)Distance to River	.017	.020	.019
	(.015)	(.014)	(.015)
Elevation(km)	-.013	-.010	-.012
	(.019)	(.019)	(.019)
(ln)Distance to Capital City	.019	.017	.017
	(.012)	(.012)	(.012)
Precipitation(mm/year)	.00002	.00002	.00002
	(.00002)	(.00002)	(.00002)
Coastal Community	-.022	-.029	-.025
	(.026)	(.025)	(.025)
Avg. Tenure	-.001	-.0007	-.001
	(.001)	(.001)	(.001)
Obs.	114	114	114
R^2	.967	.968	.968

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

Table A.5: Electricity and Provincial Diversity - Wealth Instruments

	AllInst1	AllInst2	GeogInst1	GeogInst2
	(1)	(2)	(3)	(4)
Ethnic Fractionalization	-.092***	-.066**	-.091***	-.065**
	(.027)	(.031)	(.027)	(.030)
Dist. Fract.		-.049		-.048*
		(.030)		(.029)
Avg. Wealth	.430***	.433***	.394*	.408*
	(.092)	(.092)	(.229)	(.228)
Urban	-.289*	-.271*	-.223	-.225
	(.155)	(.155)	(.396)	(.393)
Pop.Density	.032***	.032***	.033***	.033***
	(.006)	(.006)	(.008)	(.008)
(ln)Distance to River	.019	.020	.018	.019
	(.014)	(.014)	(.014)	(.014)
Elevation(km)	-.010	-.009	-.010	-.009
	(.020)	(.019)	(.020)	(.020)
(ln)Distance to Capital City	.026*	.029*	.023	.026
	(.015)	(.015)	(.027)	(.027)
Precipitation(mm/year)	7.17e-06	3.50e-06	1.00e-05	7.41e-06
	(.00003)	(.00003)	(.00005)	(.00005)
Coastal Community	-.027	-.026	-.026	-.026
	(.031)	(.029)	(.028)	(.028)
Avg. Tenure	-.0008	-.001	-.0008	-.001
	(.001)	(.001)	(.001)	(.001)
Obs.	114	114	114	114
R^2	.965	.966	.966	.967
Cragg-Donald F-Stat	15.873	15.629	5.08	4.964
Hansen J-test (p-value)	.879	.86	.487	.429

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

Appendix B

Additional Material for Chapter 3

B.1 Proofs

B.1.1 Lemma 3.1

- Proof of Lemma 3.1
  The FOC facing each group can be written as:

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

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

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

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

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

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

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

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

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

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

  This difference is positive if $p_e > p_j$ unless $\alpha = 0$[103]. If $\tau\alpha = 1$ then both groups face identical FOCs ($Q = 1$) and their contributions are equal. Given the first part

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

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.

B.1.2 Proposition 3.2

- Proof of Proposition 3.2
  Rearranging the equation defining the public goods for each of two groups we have:

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

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

B.1.3 Proposition 3.1

- Proof of Proposition 3.1
  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:

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

$$\tag{B.11}$$

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

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

  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

groups:

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

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

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

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:

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

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

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

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

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

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

Rearranging this equation results in:

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

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

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.[104] Average public good provision can be written as:

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

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

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

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:

$$\frac{\partial \overline{X}}{\partial p_1} = \frac{1}{\beta}\left[2p_1 - 2p_e\right] \tag{B.27}$$

  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:

$$\frac{\partial \overline{X}}{\partial p_1} = \frac{\beta I_1 - I\beta_1}{\beta^2} \tag{B.28}$$

  where:

$$I = \left[\sum_{k=1}^{K} p_k^2 + \frac{\alpha}{Q}(1-p_c)p_c\right] \tag{B.29}$$

$$I_1 = \frac{\partial I}{\partial p_1} \tag{B.30}$$

$$= 2p_1 - p_c Q + (1-p_c)Q \tag{B.31}$$

$$\tag{B.32}$$

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

while β_1 is defined in a similar way:

$$\begin{aligned} \beta_1 &= \frac{\partial \beta}{\partial p_c} && \text{(B.33)} \\ &= \frac{\tau Q}{\mu N \left[p_c \left(1-\tau\right)+\tau Q\right]^2} && \text{(B.34)} \\ &= \frac{\beta^2 \tau \mu Q N}{p_c^2} && \text{(B.35)} \end{aligned}$$

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

$$\begin{aligned} \frac{\partial \overline{X}}{\partial p_1} &= \frac{\mu N}{p_c}\left[p_c\left(1-\tau\right)+\tau Q\right] I_1 - I\frac{\tau \mu Q N}{p_c^2} && \text{(B.36)} \\ &= p_c\left(1-\tau\right) I_1 + \tau Q I_1 - \frac{I \tau Q}{p_c} && \text{(B.37)} \end{aligned}$$

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

$$\begin{aligned} I_1 &\geq \frac{I}{p_c} && \text{(B.38)} \\ 2p_1 - 2\frac{\alpha}{Q}p_c + \frac{\alpha}{Q} &\geq \frac{1}{p_c}\sum_{k=1}^{K} p_k^2 + \frac{\alpha}{Q}(1-p_c) && \text{(B.39)} \end{aligned}$$

Collecting terms results in:

$$p_1 - \frac{1}{p_c}\sum_{k=1}^{K} p_k^2 + p_1 - \frac{\alpha}{Q}p_c \geq 0 \qquad \text{(B.40)}$$

As:

$$\begin{aligned} p_1 &\geq \frac{1}{p_c}\sum_{k=1}^{K} p_k^2 && \text{(B.41)} \\ &\geq \frac{\alpha}{Q}p_c && \text{(B.42)} \end{aligned}$$

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.2 Robustness Check - Smaller Rural Communities

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.2.1 Rural Communities - Narrow Definition

Table B.1 presents the results of the community-level regressions, corresponding to table 3.3 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. As was shown in chapter 2, incorrect aggregation can 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 B.2 duplicate the regressions from table 3.8 over the narrower gridsize. The results for access to improved sources, piped water or piped water at home show little to no change. The one change of significance is that there is a statistically significant correlation between being in a larger group and accessing piped water at a public tap. This may indicate that there is some ethnic consideration in either the placement of public taps, or in access to them, or it may be a statistical outlier. In terms of economic significance, the magnitude of the coefficient is tiny. In terms of access to water, a household in the largest group in a divided rural area, would have less access to water than a minority household in a relatively homogeneous area.

As a further robustness check, table B.3 repeats the analysis in table B.2 using a probit regression model[105] The results are not substantively different, though there is no longer a statistically significant association between the size of one's own group and accessing piped water at a public tap.

[105]As with this one, the choice of specification does not substantively affect any of the results reported within the paper.

Table B.1: Rural Community Specification - 1KM Grid-size

	Base	Fractionalization	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.401*	.381**	.310***	.044
	(.235)	(.167)	(.117)	(.156)
Largest Group Squared	-.298*	-.219**	-.191**	-.010
	(.160)	(.111)	(.078)	(.103)
Fraction w/ Piped Water at Home				.104***
				(.040)
Avg. Wealth	.204***	.341***	.216***	.185***
	(.021)	(.020)	(.018)	(.025)
Urban	.260*	.073	.124**	-.142***
	(.141)	(.064)	(.055)	(.035)
Pop.Density(1990)	-.005	-.043	.023	-.092**
	(.041)	(.032)	(.024)	(.037)
Dist.to Cap.City (LogKM)	.028	.023**	.013	.015**
	(.019)	(.010)	(.009)	(.007)
Dist.to River (LogKM)	.021***	.009***	.003*	.006**
	(.004)	(.003)	(.002)	(.003)
Avg. Precipitation	-4.85e-07	1.00e-05	1.00e-05	-7.31e-06
	(.00004)	(.00003)	(.00002)	(.00002)
Coast.Comm.	.032	-.062*	.023	-.098***
	(.039)	(.036)	(.027)	(.031)
Elevation(km)	.044*	-.002	.004	-.001
	(.027)	(.016)	(.010)	(.014)
Avg. Tenure	.0007	.0007	.001***	-.0001
	(.0007)	(.0005)	(.0004)	(.0004)
Frac.Fem. HH Head	.012	-.006	-.018	.001
	(.043)	(.034)	(.025)	(.031)
Avg. Age of Head	-.002	.002*	.002***	-.0001
	(.001)	(.0008)	(.0006)	(.0007)
Obs.	3128	3128	3128	3107
R^2	.46	.494	.554	.357

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 B.2: Rural Household Specification - 1KM Grid-Size

	Improved	Piped	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	.569**	.520***	.252**	.371**
	(.235)	(.164)	(.113)	(.144)
Largest Group Squared	-.425***	-.311***	-.148**	-.228**
	(.162)	(.110)	(.075)	(.096)
Hist.Share Own Eth.	-.003	.009	-.0004	.013**
	(.012)	(.008)	(.006)	(.007)
Fraction w/ Piped Water at Home				.216***
				(.047)
HH Wealth	.062***	.101***	.079***	.050***
	(.006)	(.005)	(.005)	(.004)
Avg. Wealth	.138***	.228***	.135***	.127***
	(.018)	(.016)	(.014)	(.021)
Pop.Density(1990)	.018	-.014	.002	-.022
	(.023)	(.021)	(.018)	(.020)
Tenure	.00006	.0002**	.00005	.0001**
	(.00008)	(.00006)	(.00004)	(.00006)
Avg. Tenure	.0008	.0001	.0009**	-.0009*
	(.0008)	(.0006)	(.0003)	(.0005)
Female Head	.008*	.002	-.002	.004
	(.004)	(.003)	(.002)	(.003)
Frac.Fem. HH Head	.022	-.006	-.003	-.015
	(.048)	(.037)	(.024)	(.034)
Dist.to Cap.City (LogKM)	.035**	.020**	.005	.022***
	(.016)	(.010)	(.008)	(.007)
Dist.to River (LogKM)	.020***	.009***	.003	.007***
	(.004)	(.003)	(.002)	(.002)
Avg. Precipitation	-.00003	-.00006***	-.00003**	-.00004**
	(.00004)	(.00002)	(1.00e-05)	(.00002)
Coast.Comm.	.030	-.077	.021	-.113**
	(.042)	(.051)	(.030)	(.050)
Obs.	67253	67253	67253	63131
R^2	.261	.29	.31	.19

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 B.3: Rural Household Specification - 1KM Grid-Size - Probit Specification

	Improved	Piped	Home	Public
	(1)	(2)	(3)	(4)
Largest Ethnic Group	1.869**	3.240***	3.101**	2.562*
	(.942)	(1.113)	(1.245)	(1.465)
Largest Group Squared	-1.372**	-2.016***	-1.962**	-1.611*
	(.639)	(.754)	(.862)	(.979)
Hist.Share Own Eth.	-.003	.067	.078	.068
	(.055)	(.049)	(.064)	(.060)
Fraction w/ Piped Water at Home				1.419***
				(.270)
HH Wealth	.347***	.559***	.812***	.391***
	(.026)	(.023)	(.034)	(.025)
Avg. Wealth	.798***	.984***	.751***	.720***
	(.088)	(.078)	(.087)	(.105)
Pop.Density(1990)	.148	-.156**	-.158*	-.187***
	(.100)	(.069)	(.082)	(.071)
Tenure	4.17e-09	.0009**	.0003	.001**
	(.0003)	(.0004)	(.0005)	(.0005)
Avg. Tenure	.004	.003	.009**	-.003
	(.003)	(.003)	(.004)	(.004)
Female Head	.033**	.025	-.009	.051**
	(.015)	(.019)	(.027)	(.022)
Frac.Fem. HH Head	.099	.129	-.012	.196
	(.188)	(.194)	(.233)	(.239)
Dist.to Cap.City (LogKM)	.137**	.131*	.051	.203***
	(.060)	(.069)	(.080)	(.078)
Dist.to River (LogKM)	.088***	.065***	.050**	.062***
	(.017)	(.018)	(.021)	(.024)
Avg. Precipitation	-1.00e-05	-.00004	-.00005	-.00007
	(.0001)	(.0002)	(.0002)	(.0002)
Coast.Comm.	.083	-.390**	-.081	-.697***
	(.183)	(.162)	(.142)	(.222)
Obs.	66718	64883	49642	48676
R^2				

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.

Appendix C

Additional Material for Chapter 4

C.1 Proving Lemma 4.3

The cutoff value comes from a comparison of the payoffs associated with offering or not offering a contract. The expected utility to player 1 of choosing to not offer a contract is zero. The expected utility of offering a contract, in the absence of signals, is:

$$U_{1C}(\beta, i) = \beta\alpha + (1-\beta)I\alpha + (1-\beta)(1-i)(-1) \tag{C.1}$$

There are two nodes in the sequential game that are within this information set - either player 2 is honest or dishonest. If player 2 is honest, with probability β, or is dishonest but the contract is enforceable, then player 1 will receive a payoff of α. If he is dishonest and contracts are unenforceable, then player 1 will receive a payoff of (-1).

Player 1 will optimally offer a contract if the payoff from doing so is greater than zero, or:

$$U_{1C}(\beta, i) = \beta\alpha + (1-\beta)I\alpha + (1-\beta)(1-i)(-1) > 0 \tag{C.2}$$

Rearranging equation C.2, results in the cutoff value defined in lemma 4.3.

C.2 Proving Lemma 4.4

If player 1 does not offer player 2 a contract, her total payoff in the game is $(-\theta)$, as she has paid to receive a signal. If player 1 does offer player 2 a contract, her total payoff is:

$$U_{1|G}(\beta, i) = -\theta + \frac{\beta\kappa\alpha + (1-\beta)(1-\kappa)i\alpha + (1-\beta)(1-\kappa)(1-i)(-1)}{\beta\kappa + (1-\beta)(1-\kappa)} \tag{C.3}$$

Either the partner is honest and the signal was correct, or the player was dishonest and the signal was incorrect. In the latter case, the expected outcome is dependent on the quality of institutions. The denominator is simply the ex ante probability of being in this

state. Making an offer will be optimal if and only if:

$$U_{1|G}(\beta, i) = -\theta + \frac{\beta\kappa\alpha + (1-\beta)(1-\kappa)i\alpha + (1-\beta)(1-\kappa)(1-i)(-1)}{\beta\kappa + (1-\beta)(1-\kappa)} > -\theta \tag{C.4}$$

Solving this for I results in the defined cutoff value.

C.3 Proving Lemma 4.5

As before, the payoff to not offering a contract is $-\theta$. The payoff to offering a contract is:

$$U_{1|B}(\beta, i) = -\theta + \frac{\beta(1-\kappa)\alpha + (1-\beta)\kappa i\alpha + (1-\beta)\kappa(1-i)(-1)}{\beta(1-\kappa) + (1-\beta)\kappa} \tag{C.5}$$

Setting $U_{1|B}(\beta, i) = -\theta$, results in the bad signal cutoff value.

C.4 Proving Lemma 4.6

As shown after lemma 4.3, the expected payoff to player 1 after not acquiring a signal is:

$$U_{1N} = \max(0, \beta\alpha + (1-\beta)I\alpha + (1-\beta)(1-i)(-1)) \tag{C.6}$$

The first element is the payoff if no contract is offered, while the second is the payoff if a contract is offered. The payoff to acquiring the signal, and then not offering a contract, regardless of the signal is $-\theta$. Alternatively, the payoff to acquiring the signal, and then offering a contract, regardless of the signal is:

$$U_{1SY} = -\theta + \beta\alpha + (1-\beta)I\alpha + (1-\beta)(1-i)(-1) \tag{C.7}$$

Neither of these can be larger than the payoff associated with not acquiring a signal, and therefore neither can be part of an optimal strategy.

C.5 Proving Lemma 4.7

This is only relevant in the case where player 1 chooses to acquire signals of player 2's type[106]. That is, for institutional values where the overall payoff is highest to signal acquisition, it is optimal to offer contracts if a signal has been good, and not to offer a contract if the signal has been bad. Combining equations 4.6, 4.7, 4.2 and 4.4, the

[106]Sequential rationality of strategies A or C require that optimal off-equilibrium strategies are chosen, but as has been pointed out, this does not affect the equilibrium. The strategies in this case are drawn directly from 4.4 and 4.5 and are therefore sequentially optimal.

institutional cutoffs can be ordered as $I_G(\beta) < I_{AB}(\beta) < I_{BC}(\beta) < I_B(\beta)$. In words, if institutions are sufficient for strategy B to be optimal, then $I_{AB}(\beta) \leq i \leq I_{BC}(\beta)$. If so, then it must be that $I_G(\beta) < i < I_B(\beta)$, the institutions are sufficient to offer contracts if a signal has been good, but not offer a contract if the signal was bad. Therefore the strategy is sequentially rational.

C.6 Proving Lemma 4.10

Expected utility of honest agents that are assigned to be Player 2 in the equilibrium where Player 1 chooses strategy B is:

$$E[U_{h2B}] = \kappa\alpha \tag{C.8}$$

With a probability κ, the honest individual will be correctly identified, joint production will occur and player 2 receives a payoff of α. The remainder of the time, the player is incorrectly identified as opportunistic and joint production does not occur.

For the opportunistic agent assigned to be player 2 in the same equilibrium, the expected utility is:

$$E[U_{o2B}] = (1-\kappa)(i\alpha + (1-\alpha) * 1) \tag{C.9}$$

The opportunistic agent is hired if they are incorrectly identified, and once hired receive a payoff of α when institutions are effective, and 1 when contracts are unenforceable.

The expected utility of honest individuals is greater than opportunistic individuals if:

$$\kappa\alpha > (1-\kappa)(i\alpha + (1-\alpha) * 1) \tag{C.10}$$

$$i_m > \frac{1-\kappa-\alpha\kappa}{(1-\alpha)(1-\kappa)} \tag{C.11}$$

Further, when $\kappa > \frac{1}{1+\alpha}$, the critical institution, $i_m < 0$, and thus there will always be increasing evolutionary pressure on social capital in a type B equilibrium.

C.7 Proving Lemma 4.11

The expected utility of honest agents that are assigned to be player 2 in a match where player 1 chooses strategy C is α - he will always enter joint production and work productively.

For the opportunistic agent assigned to be player 2 in the same equilibrium, the

expected utility is:

$$E[U_{o2C}] = i\alpha + (1 - \alpha) \tag{C.12}$$

If institutions are enforceable, the opportunistic agent does as well as the honest agent, but when institutions are not enforceable, he is able to shirk and receives a higher payoff. As long as contracting is not perfect, $i < 1$, opportunistic agents have the advantage in this equilibrium.

C.8 Proving Proposition 4.2

There are three parts to the proposition, and the proof deals with each in sequence.

C.8.1 The Existence of a Social Capital Trap

The proposition states that if social capital is less than a critical level, $\beta^* < \frac{\theta}{\alpha*(2\kappa-1)}$, then there cannot be an equilibrium in which there is evolutionary pressure for social capital to increase. As shown in lemmas 4.9 to 4.11, evolutionary pressure for social capital to increase only occurs if strategy B is chosen at equilibrium - agents acquire signals and use them to determine entry into joint production. In addition, contracting institutions must be strong enough to ensure that honest agents have the advantage over opportunistic agents.

Strategy B is only chosen if $i \geq I_{AB}(\beta)$ and $i \leq I_{BC}(\beta)$. However, this option cannot exist if $I_{AB}(\beta) > I_{BC}(\beta)$.

$$I_{AB}(\beta) = \frac{\theta + (1-\beta)(1-\kappa) - \beta\kappa\alpha}{(1-\beta)(1-\kappa)(1+\alpha)} > \frac{(1-\beta)\kappa - \beta\alpha(1-\kappa) - \theta}{\kappa(1-\beta)(1+\alpha)} = I_{BC}(\beta) \tag{C.13}$$

Simplifying and solving this expression for β results in:

$$\beta^* < \frac{\theta}{\alpha(2\kappa - 1)} \tag{C.14}$$

which is the critical value of social capital below which countries find themselves in the social capital trap.

C.8.2 A Cap on Social Capital

As shown in lemma 4.10, social capital development requires two elements. First, strategy B must be chosen, and second, institutions must provide sufficient enforcement that the honest agents in society receive higher utility than the opportunistic.

At $\beta > \beta_{BC}(i^e)$, there cannot exist an institution that satisfies both requirements. First, if $i \leq i^e$, then honest agents receive no advantage. This is shown in 4.10.

Second, at $\beta_{BC}(i^e)$, the cutoff institution that is the maximum for employing strategy B is $I_{BC}(\beta_{BC}(i^e) = i^e$, as that is how the function was defined. Any institution greater than i^e would result in strategy C being chosen. Given that $I_{BC}(\beta)$ is declining in β, this is also true for any level of social capital higher than $\beta_{BC}(i^e)$.

As institutions must either be less than or equal to the cutoff value, or greater than it, it is impossible to choose an institution that promotes social capital development.

C.8.3 The Effective of Institutions in the Intermediate Range

If $\beta^* \leq \beta < \beta_{BC}(i^e)$, then there exists a contracting institution that will promote the development of social capital. This will occur if $I_{BC}(\beta) \geq I_{AB}(\beta)$ and $I_{BC}(\beta) > 0$ (there is an institution for which strategy B is chosen) and $I_{BC}(\beta) > i^e$ (there is an institutional option that generates sufficient enforcement), and $I_{BC}(\beta) > 0$. The second two requirements are effectively identical, as $i^e \geq 0$ by definition. If both are satisfied, choosing $i = I_{BC}(\beta) - \epsilon$ will promote the development of social capital, for small enough ϵ.

As shown in the earlier part of the proof, $I_{BC}(\beta) \geq I_{AB}(\beta)$, if and only if $\beta \geq \beta^*$. Additionally, as shown previously, $I_{BC}(\beta_{BC}(i^e) = i^e$. With $I_{BC}(\beta)$ declining in β, this implies that $I_{BC}(\beta) > i^e \forall \beta < \beta_{BC}(i^e)$. As such, all intermediate levels of social capital meet both requirements.

Finally, we have not assured that this intermediate range exists. An additional assumption would be necessary to ensure that $\beta^* < \beta_{BC}(i^e)$. After rearranging the relevant definitions, the necessary assumption is that:

$$\kappa > \frac{2\alpha + \alpha\theta + \theta}{\alpha(3 + \alpha)} \tag{C.15}$$

or rearranging:

$$\theta < \frac{\alpha(\kappa(3 + \alpha) - 2)}{1 + \alpha} \tag{C.16}$$

If signals are sufficiently informative relative to cost, or cheap enough relative to accuracy, then the intermediate range exists.

C.9 Proving Proposition 4.3

The critical value of social capital is $\beta = \frac{\theta}{\alpha(2\kappa-1)}$. The relevant derivatives as related to the poverty trap are therefore:

$$\frac{\partial \beta^*}{\partial \theta} = \frac{1}{\alpha(2\kappa-1)} > 0 \tag{C.17}$$

$$\frac{\partial \beta^*}{\partial \kappa} = \frac{-2\alpha\theta}{(\alpha(2\kappa-1))^2} < 0 \tag{C.18}$$

For the maximum level of social capital, $\frac{\partial \beta_{BC}(i^e)}{\partial \theta}$, and $\frac{\partial \beta_{BC}(i^e)}{\partial \kappa}$ must similarly be shown to be negative and positive, respectively.

With regards to the cost of signals, θ, the level of institutions necessary to ensure enforceability is unaffected by the change. Therefore, it is sufficient to show that $\frac{\partial \beta_{BC}(i)}{\partial \theta} < 0$, or equivalently, that $\frac{\partial I_{BC}(\beta)}{\partial \theta} < 0$. This last point is easily demonstrated as:

$$I_{BC}(\beta) = \frac{(1-\beta)\kappa - \beta\alpha(1-\kappa) - \theta}{\kappa(1-\beta)(1+\alpha)} \tag{C.19}$$

$$\frac{\partial I_{BC}(\beta)}{\partial \theta} = \frac{-1}{\kappa(1-\beta)(1+\alpha)} < 0 \tag{C.20}$$

For the effectiveness of signals, if the enforcement level of institutions is relevant, $i^e = \frac{1-\kappa-\alpha\kappa}{(1-\alpha)(1-\kappa)} \in [0,1)$, then this level increases with the reduction in signal quality , $\frac{\partial i^e}{\partial \kappa}$ if $i^e < 1$.

As with increasing the cost of signals, the maximum cutoff institution necessary to encourage the use of a screening strategy also falls with the decrease signal effectiveness, $\frac{\partial I_{BC}(\beta)}{\partial \kappa} > 0$. Shifting the $I_{BC}(\beta)$ curve down, while shifting the i^e line up, necessarily results in their intersection at a lower level of social capital, β. Thus, the maximum level of social capital achievable in society is reduced when the effectiveness of signals falls.