



ELSEVIER

Journal of Public Economics 89 (2005) 2325–2368

JOURNAL OF
PUBLIC
ECONOMICS

www.elsevier.com/locate/econbase

Ethnic diversity, social sanctions, and public goods in Kenya

Edward Miguel^{a,*}, Mary Kay Gugerty^b

^a549 Evans Hall, #3880, Department of Economics, University of California,
Berkeley, CA 94720-3880, United States

^bDaniel J. Evans School of Public Affairs, University of Washington,
WA, United States

Received 6 January 2003; received in revised form 12 August 2004; accepted 9 September 2004
Available online 6 November 2004

Abstract

This paper examines ethnic diversity and local public goods in rural western Kenya. The identification strategy relies on the stable historically determined patterns of ethnic land settlement. Ethnic diversity is associated with lower primary school funding and worse school facilities, and there is suggestive evidence that it leads to poor water well maintenance. The theoretical model illustrates how inability to impose social sanctions in diverse communities leads to collective action failures, and we find that school committees in diverse areas do impose fewer sanctions on defaulting parents. We relate these results to the literature on social capital and economic development and discuss implications for decentralization in less developed countries.

© 2004 Elsevier B.V. All rights reserved.

JEL classification: D71; H41; O12

Keywords: Local public goods; Ethnic diversity; Collective action; Africa

1. Introduction

Well-known cross-country empirical research indicates that ethnically diverse societies have slower economic growth and are more prone to corruption and political instability

* Corresponding author. Tel.: +1 510 642 7162; fax: +1 510 642 6615.

E-mail address: emiguel@econ.berkeley.edu (E. Miguel).

than ethnically homogeneous societies as a result of political conflict and lack of cooperation across ethnic groups.¹ Easterly and Levine (1997) argue that ethnic diversity has had a particularly negative impact on economic outcomes in sub-Saharan Africa, which has suffered from a series of destructive ethnic conflicts in recent years and is the most ethnically diverse and the poorest region in the world.² Yet, the impact of ethnic diversity on local collective action in sub-Saharan Africa remains largely unexplored empirically.³

This is the first microeconomic study to our knowledge that examines the impact of ethnic diversity on local public good provision in sub-Saharan Africa. We examine the relationship between ethnic diversity and local funding of primary schools and community water wells in western Kenya using a unique data set⁴ and show that ethnic diversity is negatively related to local public goods provision in this rural African setting. This paper contains two main innovations. First, we employ an innovative empirical strategy to identify the relationship between diversity and collective action, relying on historically determined ethnic settlement patterns. Second, we present the first quantitative evidence (we are aware of) on the role that social sanctions play in sustaining local public goods provision.

To address potential biases caused by endogenous household sorting among local schools and wells, we identify the relationship between ethnic diversity and collective action in western Kenya using local residential composition in the surrounding area as the key measure of ethnic diversity. Historical evidence indicates that ethnic land claims in western Kenya were established in the 1800s and have remained largely stable during the past century. Ethnically diverse areas in rural western Kenya are also similar to homogeneous areas along a range of socioeconomic and school quality measures, ameliorating concerns that the estimated ethnic diversity effects are driven by omitted variables. The use of historically determined ethnic settlement patterns to estimate the impact of local ethnic diversity on public good provision constitutes a considerable improvement over the recent studies from the United States inasmuch as high rates of residential mobility in the US increase the likelihood that unobserved household characteristics are correlated with local ethnic composition.

¹ In cross-country work, Mauro (1995) finds that ethnic diversity is significantly related to poor bureaucratic performance and political instability, and Knack and Keefer (1997) find that low social capital reduces growth.

² The 1994 Rwandan genocide is the most tragic example (Des Forges, 1999).

³ Other recent work has suggested that ethnic divisions are related to local public goods funding and social capital across United States municipalities. Goldin and Katz (1997) argue that public secondary schooling expanded slowly in ethnically diverse US school districts from 1910 to 1940. Poterba (1997) finds that an increase in the share of ethnic minorities among the school age child population is associated with lower local school spending per child in US cities. Alesina et al. (1999) find that high levels of ethnic diversity are associated with lower funding for schools and other local public goods in US municipalities, and Alesina and LaFerrara (2000) show that ethnically diverse regions in the US have lower levels of social capital as measured by participation in community groups. Refer to Coleman (1989) and Putnam (1993) for seminal work on social capital and to Costa and Khan (2002) for a recent review of the growing economics literature on diversity and collective action.

⁴ The authors are grateful to Michael Kremer for sharing the primary school data with us.

We find that local ethnic diversity is associated with sharply lower local school funding and lower quality school facilities in 84 primary schools. This effect operates through lower contributions at public fundraising events. The drop in school funding associated with the change from complete ethnic homogeneity to average school ethnic diversity is approximately 20% of mean local school funding per pupil, and this relationship is robust to socioeconomic, geographic, and demographic controls.

There is additional evidence that these collective action problems may extend beyond the school setting. Data from 667 community water wells in rural western Kenya indicate that local ethnic diversity may also be associated with poor well maintenance; areas with average levels of ethnic diversity are 6% points less likely to have a functioning water well than homogenous areas. While the well results are not definitive on their own and are not as robust as the primary school results, this evidence does suggest that ethnic diversity has implications for collective action beyond school funding.

We develop a simple stylized model that proposes a specific channel through which ethnic diversity affects local collective action outcomes: social sanctions. The model highlights the role that social sanctions play in overcoming free-rider problems in collective action, following Besley et al. (1993), Besley and Coate (1995), Wydick (1999), and Fehr and Gächter (2000),⁵ and as such is complementary to existing theories that rely on variation in the level (Wade, 1994) or distribution of public good benefits (Khwaja, 2000) or on taste differences (Alesina et al., 1999; Alesina and LaFerrara, 2000; Vigdor, 2004) to explain why diversity leads to failed collective action. We make the key assumption that social sanctions are imposed more effectively within ethnic groups than between groups, drawing on the recent social capital literature, as well as on a wealth of anthropological evidence from rural Africa on the importance of kinship in governing access to resources. This assumption implies that public good contributions will generally be lower in ethnically diverse areas because of free-riding in the absence of effective community sanctions.

We find empirical support for the idea that social sanctions play an important role in public goods provision in the primary school setting, using a unique data set of primary school committee records describing the use of sanctions against noncontributing parents. School committees in ethnically diverse areas threaten fewer sanctions and use less verbal pressure against parents who do not contribute at public fundraisings, pay school fees, or contribute in other ways to the school. We also find that the frequency of ordinary administrative (i.e., nonsanction) items discussed in the records is not significantly related to local ethnic diversity, indicating that the sanctions result does not simply reflect poor record keeping in diverse areas. We cannot completely rule-out all other competing hypotheses for these patterns; for instance, weaker support for primary schooling could drive both lower funding and fewer sanctions without a causal link between the two, and a community that intends to hold fewer fundraising events may also choose to use fewer social sanctions because they will be less effective. Yet even in light of these caveats, we

⁵ In a related study, Wydick (1999) finds that a willingness to apply informal social pressure reduced the extent of risky borrower behavior in Guatemalan microfinance groups. Fehr and Gächter (2000) find that the ability to impose punishments on free-riders in laboratory experiments (conducted among Swiss university students) leads to dramatically improved cooperation in public goods games.

argue that the cumulative evidence from the school committee record results, interviews with school headmasters, field worker observations, and anthropological evidence from this area makes a strong case that the inability to impose sanctions across ethnic groups is a crucial determinant of local public goods provision.

This research project makes at least three contributions. First, we contribute to the growing literature on social capital and economic development by presenting empirical evidence for a specific channel through which ethnic diversity affects local collective action—the inability of ethnically diverse communities to sanction free-riders. The finding that social sanctions are weaker in diverse communities has potentially far-reaching implications for less developed countries inasmuch as a variety of informal collective action, contracting, and credit market outcomes are thought to rely on effective sanctions.

Second, the empirical result linking ethnic diversity to low public good provision contributes to the current debate on the sources of poor African economic performance because as local contributions play an increasingly critical role in the provision of education and health care in Africa, both of which are important determinants of human capital accumulation and, therefore, economic growth.⁶ Inasmuch as the Kenyan region we study is fairly typical for Africa in terms of income levels and ethnic diversity, the results are likely to have implications for local collective action in other contexts. Third, this study illustrates a potential downside of decentralized public finance reforms currently in favor in many less developed countries, a theme that is taken up again in the conclusion. Finally, we believe that our results point to the need for further research on policies and institutions that build cooperation or “social capital” in societies with pronounced ethnic divisions—although a full discussion of these policies is beyond the scope of this paper.

The remainder of the paper is organized as follows. Section 2 surveys alternative theories of diversity and public goods provision and presents our theory of collective action in ethnically diverse communities. Section 3 discusses the ethnicity measures and identification strategy, and Section 4 describes the data. Section 5 presents the empirical results for primary schools and water well maintenance. The final section summarizes and discusses implications of the results.

2. A theory of ethnic diversity, social sanctions, and local public good provision

2.1. Alternative theories of ethnic diversity and collective action

Alesina and LaFerrara (2000) present a theory of ethnic diversity and community group formation in which individuals dislike mixing across ethnic lines, and this taste for homogeneity drives their theoretical prediction that diverse areas exhibit lower participation in community activities. Their theory implies that all individuals would opt to sort into ethnically homogeneous organizations to avoid the costs of mixing with individuals from other ethnic groups. Alesina and LaFerrara’s result is particularly

⁶ Collier and Gunning (1999) discuss the debate on African underdevelopment. Refer to Krueger and Lindahl (2000) for evidence on the cross-country relationship between education and growth.

pronounced when it is costly for distinct ethnic groups to self-segregate into separate organizations.

Ethnic diversity could also lead to lower public goods funding because different ethnic groups have divergent preferences over the type of public good to be funded and are therefore less willing to contribute toward compromise types. This could explain [Alesina et al. \(1999\)](#)'s finding that higher levels of ethnic diversity are empirically associated with lower local public goods funding across US municipalities. In their model, heterogeneous tastes across ethnic groups are once again the channel through which diversity affects collective action, and as in [Alesina and LaFerrara \(2000\)](#), the model in [Alesina et al. \(1999\)](#) implies that individuals would sort into homogeneous communities if moving costs were not too high. This theory predicts that funding for ethnically contentious public goods, such as school funding if the preferred language of school instruction divides ethnic groups, should be lower in ethnically diverse areas. The model, however, does not explain why funding for local public goods that are not ethnically contentious should be related to ethnic diversity. In fact, the empirical results indicate that funding is significantly lower in diverse US municipalities for many local public goods with no obvious ethnic dimension, including libraries, roads, and garbage collection.

A third model relying on ethnic taste differences assumes that individuals prefer to fund public goods if the beneficiaries are from their own ethnic group but care less about individuals from other ethnic groups. [Vigdor \(2004\)](#) develops this framework to examine the effect of county-level ethnic composition on 2000 US census postal response rates and finds that response rates, which affect subsequent Congressional representation and Federal transfers to the county and hence can be thought of as a public good, are in fact significantly lower in more diverse counties. As in the other “taste” models, individuals in this framework would prefer to sort into homogeneous communities to maximize local public good funding levels. However, unlike [Alesina et al. \(1999\)](#), the Vigdor model can explain lower funding for public goods without a clear ethnic dimension like roads and libraries.

An additional factor linking social heterogeneity to public goods is the level of benefits from collective action. [Wade \(1994\)](#) provides evidence that socially diverse farming communities in India that face greater crop risks due to lack of irrigation are more likely to develop effective collective action mechanisms to deal with this problem than communities with similar levels of heterogeneity but less crop risk. [Khwaja \(2002\)](#) shows that social heterogeneity has a negative affect on the maintenance of community projects in northern Pakistan, arising from a more unequal distribution of project returns in diverse communities.

A final possible theoretical channel between ethnic diversity and public goods provision is violent conflict among ethnic groups, which could disrupt local fundraising and service delivery ([Easterly and Levine, 1997](#)). Although ethnic divisions are a seemingly salient component of many civil conflicts in Africa, this theory has limited applicability to the Kenyan districts we study, where organized ethnic violence is unknown. Thus, the impact of diversity on public goods provision in this setting is likely to work through other channels.

We present below a stylized model that does not rely on divergent public good preferences or differences in the returns to the public good across ethnic groups to generate

theoretical predictions. Instead, the model highlights the importance of social sanctions in overcoming collective action problems in diverse areas and on the difficulties in imposing and enforcing sanctions across ethnic groups. This is consistent with the sociocultural evidence on many African societies and can plausibly explain low funding even for public goods without a clear ethnic dimension, such as water provision in our context.

2.2. *A theory of ethnic diversity, social sanctions, and local public goods*

The model examines the role that social sanctions play in overcoming the free-rider problem in an ethnically diverse area. The key assumption of the model is that there is more social cohesion and social capital within ethnic groups than across groups. In particular, we assume that social sanctions and coordination are possible within groups due to the dense networks of information and mutual reciprocity that exist in groups but are not possible across groups.⁷ This is a compelling assumption for rural African communities in which ethnic and kinship relations regulate access to resources and retain significant control over the use and disposition of land (Shipton, 1985; Berry, 1993; Bates, 1999; Platteau, 2000). Social sanctions in these communities may take the form of partial exclusion from ethnic networks that provide insurance, regulate access to resources, adjudicate disputes, and provide friendship and other social benefits to members. Gugerty (2000) describes the importance of social and kinship ties in the use of sanctioning mechanisms within rotating savings and credit associations in this area. In western Kenya, “members [of a kin group] are bound by exogamy, a general obligation of mutual aid, some degree of collective responsibility for members’ misdeeds, participation in various ceremonies in a member’s life, and inheritance of property” (Were, 1986).

We limit our model to the relatively simple case in which it is costless for individuals to impose social sanctions on their ethnic peers; there is no bargaining or contracting between ethnic groups; and there is no time dimension. These stark assumptions deliver clear predictions that highlight the interplay between diversity and collective action. Moreover, the limitations of our data set—and in particular the lack of information on individual-level public goods contributions—mean we cannot empirically test more intricate theoretical predictions. No single theory can fully explain a social phenomenon as complex as ethnic conflict in all settings, but our theory complements the existing theories by contributing new insights into the role of within-group sanctions in shaping collective action outcomes.

2.3. *Model set-up*

A village funds a local public good from household contributions. There are measure one households in the village, and each household belongs to one of two ethnic groups, A or B (ethnic identity is fixed). The measure of group A (B) households is n_A (n_B), where $n_A + n_B = 1$ and $n_A \geq n_B$, so A is in the majority group. The simplest measure of ethnic

⁷ This stands in sharp contrast to Fearon and Laitin (1996) who examine the case in which repeated interaction creates the possibility of interethnic cooperation either due to the threat of retaliation or within-group policing. Although we recognize the potential importance of such mechanisms, our goal is to emphasize different factors.

diversity in this case is $n_B \in [0, 1/2]$. We abstract from group numerical size for simplicity⁸ and assume that community composition is fixed (though we briefly discuss endogenous sorting in Section 2.5 below).

A household chooses either to contribute to funding the public good at cost $c > 0$ or not to contribute (at cost 0), where the contribution cost can be thought of as the time, effort, and money expended. The variable p_{ij} takes on a value of 1 if household j in group $i \in \{A, B\}$ contributes to the public good and 0 otherwise. The discrete support of p_{ij} simplifies the analysis and is realistic in our setting to the extent that there are fixed costs to contributing to the public good, including the time it takes to, for example, walk to and participate in a community meeting. The measure of households in group A (B) that contribute to the public good is p_A (p_B), and the total proportion of contributing households in the village is $P \equiv p_A + p_B$. All households simultaneously decide whether or not to contribute, and we assume that their contributions are publicly observable.

All households in the community benefit from the aggregate level of public goods contributions regardless of their own contribution. The assumption that free-riders cannot be excluded from the public good is realistic in rural Kenya during the period we study inasmuch as children were essentially never expelled from primary school for nonpayment of fees (although they could be temporarily suspended), and it is also politically difficult to restrict access to public water sources, as we discuss in Section 3. We assume that the public goods benefit function, $b(P)$, is concave increasing in the overall proportion of contributions (Assumption 1) and that $b(0) = 0$.

Assumption 1. $b'' \leq 0$

Households who do not contribute to the public good are punished by community social sanctions.⁹ We assume that social sanctions are strong within ethnic groups and nonexistent across groups. The strength of social sanctions imposed on a household in group $i \in \{A, B\}$ if it does not contribute to the public good is $s_i = s(p_i/n_i)$, which is increasing in the proportion of other households in one's own ethnic group who contribute. The effectiveness of sanctions is increasing in its argument ($s' > 0$) and $s(0) = 0$ and $s(1) = 1$. A natural measure of the total threat of sanctions for free-riders in this village is $S = n_A s_A + n_B s_B$.

The utility function for household j in group $i \in \{A, B\}$ is separable with respect to public good benefits production, the contribution cost, and the sanction punishment, for simplicity. Households choose a contribution $p_{ij} \in \{0, 1\}$ to maximize expected utility

⁸ For simplicity, we focus on proportions of the ethnic groups rather than the total population of the village. As Drazen (2000) notes, the relationship between group size and collective action is sensitive to minor changes in the specification of participation costs and benefits. Chamberlin (1974) and Esteban and Ray (2001) suggest that if anything, larger groups are likely to be more effective at collective action in the provision of pure public goods, which would strengthen the main result in Proposition 1.

⁹ We follow Besley and Coate (1995), Besley et al. (1993), and Akerlof (1980) in positing the existence of social sanctions without formalizing precisely how they are imposed. Fehr and Gächter (2000) provide persuasive evidence that experimental subjects are in fact willing to engage in individually costly punishment strategies to deter free-riding in public goods games.

given common expectations on the contribution levels of both ethnic groups, where expectations are denoted with the superscript e .

$$E(u_{ij}) = b(p_i^e + p_{-i}^e) - p_{ij}c - (1 - p_{ij})s\left(\frac{p_i^e}{n_i}\right) \quad (1)$$

We restrict attention to the interesting case in which fully funding the project is worthwhile in terms of aggregate welfare, but there would be free-riding in the absence of social sanctions (Assumption 2) and in which the maximum level of the sanctions threat is sufficiently powerful to ensure that all households contribute (Assumption 3).

Assumption 2. $0 < c < b(1)$

Assumption 3. $c < 1$

2.4. Solution

All households in an ethnic group face the same contribution decision. The sanction function implies a “tipping” outcome; households contribute to avoid sanctions if they expect a sufficient number of other households to participate, but free-ride if few others are expected to participate.¹⁰ As we discuss in Remark 1, two stable equilibria always exist for each ethnic group: the “low contribution case” where no household in the ethnic group contributes ($p_i=0$) and the “high contribution case” in which all households contribute ($p_i=n_i$). This follows directly from Eq. (1).

Remark 1. All group i households ($i \in \{A, B\}$) contribute to the public good if and only if the expected contribution by other households in their ethnic group is sufficiently high or formally when $\frac{p_i^e}{n_i}s^{-1}(c)$.

There remains the important issue of equilibrium selection in the context of multiple equilibria. We assume that each ethnic group coordinates on the equilibrium that maximizes household utility for members of the ethnic group where this ability to coordinate on the desirable equilibrium is a consequence of the dense information networks within ethnic groups. For a household in group $i \in \{A, B\}$, the utility difference between the high contribution case and the low contribution case is defined as $\Delta_i(n_i, p_{-i}^e) \equiv b(n_i + p_{-i}^e) - c - b(p_{-i}^e)$.

Assumption 4. Given common beliefs about the expected aggregate contribution of the other ethnic group (p_{-i}^e), members of group i ($i \in \{A, B\}$) expect the high contribution case to occur if and only if the expected utility of the high contribution case is greater than the expected utility of the low contribution case or formally when $\Delta_i(n_i, p_{-i}^e) \geq 0$. If this condition does not hold, then members of group i expect the low contribution case to occur.

¹⁰ Refer to Schelling (1978) for related models.

Assumption 4 and Remark 1 together imply that all households contribute in the case of no ethnic diversity ($n_B=0$), and this is the benchmark case against which we compare all subsequent results.

Remark 2. Remark 1 and Assumptions 1–4 imply the following:

- (a) The utility difference between the high contribution case and the low contribution case is weakly decreasing in the expected contribution of the other ethnic group: $\frac{\partial \Delta_i}{\partial p_{-i}^e} \leq 0$.
- (b) The utility difference between the high contribution case and the low contribution case is increasing in the size of group i : $\frac{\partial \Delta_i}{\partial n_i} > 0$.
- (c) The low contribution case is preferable (regardless of the expected contribution of the other group) if the ethnic group is sufficiently small or formally, $n_i \leq b^{-1}(c) \equiv n^* \Delta_i \leq 0$.
- (d) The high contribution case is preferable (regardless of the expected contribution of the other group) if the ethnic group is sufficiently large or formally, $n_i 1 - b^{-1}(b(1) - c) \equiv n^* \Delta_i > 0$, where $0 \leq n^* \leq n^{**}$.

The first result (2a) is a direct consequence of the concavity of b and implies that public good contributions become weakly less valuable as the expected contribution of the other ethnic group increases; this free-riding result is related to the “public goods game” in Drazen (2000). Result (2b) implies that larger groups have more to gain from mobilizing for collective action than small groups. Results (2c) and (2d) imply that ethnic groups collectively choose not to participate for all p_{-i}^e (even 0) when the group is sufficiently small ($n_i \leq n^*$) but always contribute when sufficiently large ($n_i \geq n^{**}$).

2.5. Empirical implications of the model

Before proving the main result in Proposition 1, we restrict model parameters to values in which there is likely to be widespread free-riding. Recall that our ethnic diversity measure is the size of the ethnic minority group (n_B), which takes on values from 0 to 1/2. Below, we discuss the robustness of our results to alternative parameter values.

Assumption 5. $0.5 \leq n^*$

Proposition 1. *Aggregate local public good contributions and the total threat of social sanctions are weakly declining in ethnic diversity.*

At low levels of ethnic diversity (for $0 \leq n_B \leq 1 - n^{**}$), both the total public good contribution and total sanctions threat (S) are positive and strictly decreasing with ethnic diversity, $dP/dn_B < 0$, $dS/dn_B < 0$. For low levels of ethnic diversity, the high contribution case occurs among the majority ethnic group and the low contribution case among the minority group. As ethnic diversity increases in this range, the proportion of the total population made up of the contributing majority group shrinks, and the proportion made up of the free-riding minority group grows, leading to declining total public good funding. For very high levels of ethnic diversity ($1 - n^{**} \leq n_B \leq 0.5$), both the total public good contribution and the total threat of sanctions are 0, as both ethnic groups free-ride. Fig. 1 graphically presents total public good contributions as a function of ethnic diversity.

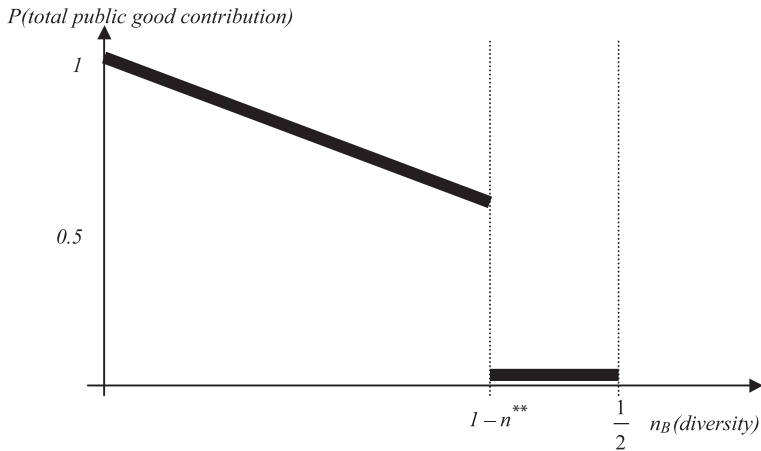


Fig. 1. Ethnic diversity and public good contributions. The thick lines denote equilibrium outcomes. There is an analogous result for threatened sanctions.

One logical extension of the framework would be to incorporate a public goods “taste” difference across ethnic groups, which might provide further incentives for minority groups to contribute to the public good, if in doing so, they gained more say in the type of public good chosen. However, we opt against this approach because as an important goal of this model is to derive collective action predictions without resorting to assumptions about exogenous ethnic taste differences and inasmuch as it is not necessary to derive empirically relevant theoretical predictions. Moreover, evidence presented in Section 5.2 suggests that very few parents mention ethnic or language factors as a factor in choosing a primary school; in models of sharp ethnic taste differences, ethnicity would presumably be a much more salient factor than we find in determining sorting among schools.

Proposition 1 implies that voluntary public goods contributions and social sanctions threats are decreasing in local ethnic diversity, and we test this hypothesis in Section 5. The total sanctions in the model are best understood as threatened sanctions that would be imposed on free-riders although sanctions are actually never imposed in equilibrium in the model. In the empirical section, we assume that threatened sanctions mentioned in the school records correspond to the model’s threat of sanctions rather than to actual sanctions imposed.¹¹

¹¹ The result in Proposition 1 is largely robust to different values of n^* and n^{**} . For example, for $n^* \leq 0.5 \leq n^{**}$ and n_B , sufficiently large members of an ethnic group collectively choose to contribute if the expected contribution of the other ethnic group is expected to be low, but free-ride if the contribution is expected to be high. There are always multiple equilibria in this range, one equilibrium where the numerically larger group (A) has high contribution and the smaller group (B) has low participation and vice versa, and either equilibrium may occur in the absence of coordination across groups. Thus, the theory accommodates cases in which the participation of minority individuals is higher than the participation of the majority group and vice versa. In the range where multiple equilibria are possible, expected aggregate public good contributions are still weakly decreasing in ethnic diversity under the condition that the equilibrium in which the majority A households contribute is at least as likely as the equilibrium in which the B households contribute (results available from the authors upon request). Unfortunately, our data set has incomplete information on ethnic majority and minority group public good contributions, limiting our ability to directly test these theoretical implications in western Kenya.

Another interesting empirical implication of the model relates to incentives for ethnic segregation.¹² As discussed above, alternative theories imply that ethnic groups should seek to completely self-segregate into distinct local schools and wells, barring prohibitive mobility frictions. However, complete self-segregation is not generally an equilibrium in our framework. Imagine an ethnic minority group of measure n_B whose households free-ride off of the public good contributions provided by the majority ethnic group, as in our model. If these minority households had a choice between attending a homogeneous B school or remaining as a minority group in a predominantly A school, it is straightforward to show that they unambiguously prefer to remain in the minority for n_B sufficiently small. In a homogeneous B school, strong social sanctions would compel them to contribute to the public good at cost c , which is a large direct cost for little additional gain in terms of public good provision. A similar thought experiment suggests that others might prefer greater sorting; the majority Group A households that contribute to the public good suffer from free-riding by the minority households and would prefer to attend schools dominated by other A households where free-riding is less severe. These informal arguments suggest that the effect of local sorting on ethnic segregation is ambiguous.¹³ To the extent that observed ethnic sorting among local schools and wells is in fact limited in western Kenya, this would provide further evidence consistent with our theoretical framework but not with the alternative theories discussed above. We present evidence on the extent of sorting among local schools in Section 5.

3. Ethnicity and the identification strategy

We test the theoretical hypothesis that both voluntary public goods contributions and threatened social sanctions are decreasing in local ethnic diversity in Section 5 below. To estimate the relationship between diversity and collective action outcomes, we must rule out the possibility that local unobservable characteristics correlated with ethnic diversity rather than ethnic diversity itself are in fact driving the estimated effects, and we address this issue in several steps. In this section, we first describe the ethnic groups in western Kenya and show that historical ethnic land settlement patterns in this area during the 19th century are largely exogenous and persist to the present day. In Section 4, we present evidence that ethnically diverse areas are similar to homogeneous areas along a range of socioeconomic, agricultural, and school characteristics, ameliorating concerns about unobserved differences across the diverse and homogeneous areas. Finally, in Section 5, we estimate the effect of ethnic diversity including extensive controls for local characteristics.

3.1. Ethnic groups in Western Kenya

The ethnicity data were collected in Busia and Teso districts in western Kenya (Fig. 2). These are primarily rural areas with a local economy based on small-scale farming for

¹² The possibility of endogenous local sorting among schools or wells motivates our use of local residential composition (rather than the endogenously determined diversity of actual school pupils or well users) as the principal empirical measure of local ethnic diversity.

¹³ For a more formal discussion of possible sorting equilibria in a related setting, refer to Miguel (2001).

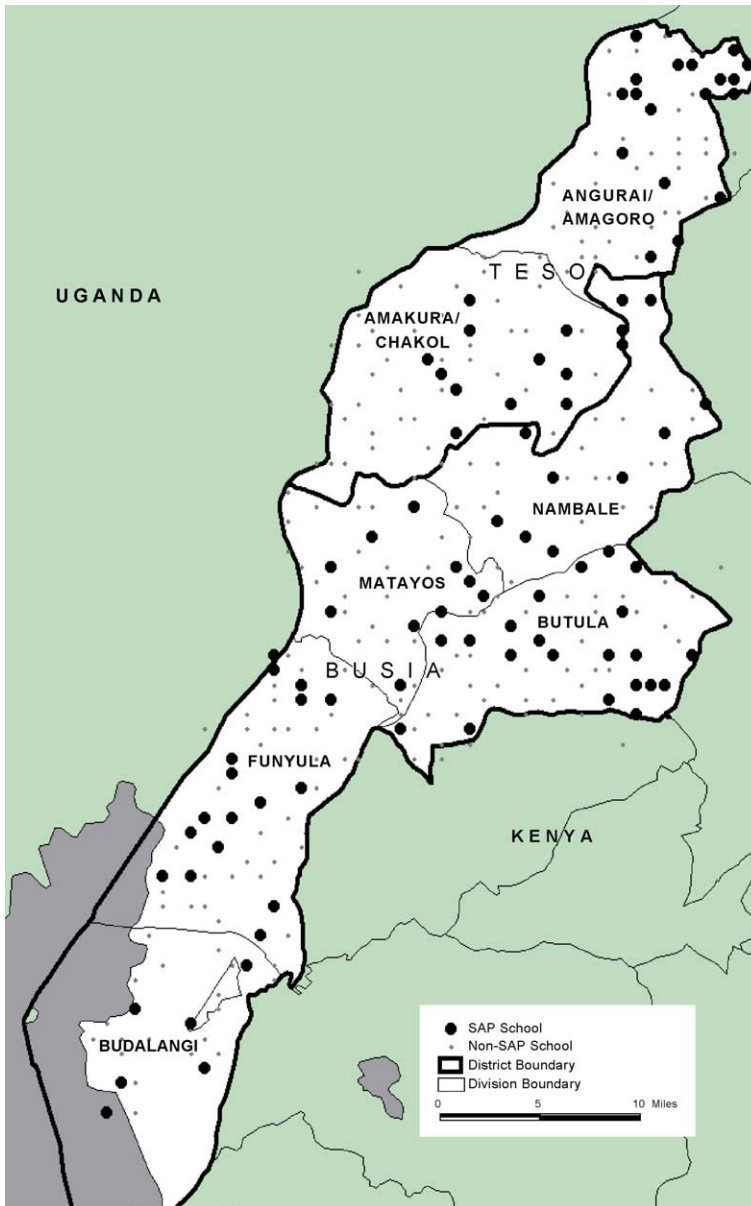


Fig. 2. Map of Busia and Teso Districts, Kenya.

subsistence and local market trade. The combined population of Busia and Teso districts in 1989 was 401,658, and their total area 1652 km² (Government of Kenya, 1994). The districts are relatively poor by Kenyan standards; the daily agricultural wage is approximately 0.85 US dollars.

Busia and Teso districts are moderately ethnically diverse; the largest ethnic groups are the Luhya (67% of the sample), Teso (26%), and Luo (5%).¹⁴ The Luo and Teso are Nilo-Saharan ethnolinguistic groups with pastoralist traditions, and the Luhya are a Bantu (Niger-Kordofanian) group. Luhyas are the majority ethnic group in Busia district, and Tesos are numerically dominant in Teso district, although there are significant minority communities on both sides of the Busia–Teso border.¹⁵

Ethnicity is perhaps the primary cleavage in Kenyan political and social life, and several violent ethnic clashes claiming hundred of lives have occurred in both central and coastal regions of the country during the past decade, most notably in advance of the 1992 and 1997 national presidential polls (Ndegwa, 1997). Although there has not been ethnic violence in Busia or Teso during the postcolonial period, there is evidence that ethnic tension is common in a variety of settings and has an adverse impact on local collective action. Government of Kenya anthropologists write that, “The Teso, as a minority group surrounded by people with whom they were on belligerent terms in the not so distant past, tend to have rather strained relations with their neighbors. . . . It is not uncommon for Teso and Luhya to come to blows in places of work. . . . These attitudes make it difficult to design development projects for the District as a whole which would require inter-ethnic cooperation” (Government of Kenya, 1986).

3.2. Land settlement patterns

The exogeneity of ethnic land settlement patterns in Busia and Teso districts forms a basis for the empirical identification strategy employed in this paper. A variety of evidence suggests that current levels of local ethnic diversity in Busia and Teso districts are largely the product of historical accident rather than recent migration.¹⁶ “The 19th century was a time of considerable unrest throughout the District, with conflict between the Luhya groups, Luo, Teso, and Kalenjin” (Government of Kenya, 1986). Were (1967) writes that “various factors—famine, epidemics, domestic disputes, the spirit of adventure and warfare—made the inhabitants of the region extremely mobile” from the 17th century through the 19th century, when various Nilo-Saharan ethnic groups migrated to western Kenya from present-day Uganda. Successive waves of Teso and Luo migration and the

¹⁴ The data sources are discussed in Section 4 below. School children are generally taught in their vernacular (native) language in western Kenya through Grade 3 although Swahili is usually the medium of instruction in ethnically diverse areas. Starting in Grade 4, classes are conducted in English.

¹⁵ The Luhya are composed of the Khayo, Marachi, Nyala, Samia, Bukusu, Dakho, Kabras, Marama, Sukha, and Wanga subtribes, among others. Other smaller non-Luhya ethnic groups in this area include Kalenjin (Sabot), Kikuyu, Masaai, Somali, Tachoni, and Taita. The Luhya could potentially be considered many distinct ethnic groups inasmuch as certain subtribe dialects are mutually unintelligible, there are historical rivalries between subtribes, and the notion of a single Luhya ethnic group only originated during the colonial period (Government of Kenya, 1986; Were, 1967). However, Miguel (2001) finds that diversity among Luhya subtribes is not generally related to local collective action outcomes in this area.

¹⁶ Religious diversity is not included as an explanatory variable in the analysis inasmuch as local religious affiliation is not plausibly exogenous due to the extensive missionary activity in this area during the past century. A negative correlation between religious fragmentation and school funding cannot be interpreted as causal if evangelical activity is targeted to and is most successful in poor areas, for instance.

Table 1

Ethnic diversity across geographic divisions in Busia and Teso districts in 1962 and 1996

Geographic division	Name in 1962	Proportion of largest residential ethnic group (group in parentheses)	
		1962	1996 (Pupil Questionnaire data)
Budalangi	Bunyala	0.99 (Luhya)	0.94 (Luhya)
Funyula	Samia	0.98 (Luhya)	0.94 (Luhya)
Butula	Marachi	0.92 (Luhya)	0.86 (Luhya)
Amukura/Chakol	South Teso	0.92 (Teso)	0.87 (Teso)
Angurai/Amagoro	North Teso	0.87 (Teso)	0.86 (Teso)
Nambale/Matayos	Bukhayo	0.68 (Luhya)	0.76 (Luhya)

The 1962 data is from the 1962 Kenyan Census (Government of Kenya, 1965). The 1996 data is from the ICS Pupil Questionnaire, which relies on self-described ethnic affiliation.

resulting wars with established Luhya communities largely determined ethnic residential patterns in the area.

The emergence of British colonial authority in western Kenya in 1894 ended wars and cattle raiding, as well as the population movements that accompanied them. Morgan and Shaffer (1966) write that ethnic land claims were “frozen by the Colonial Government by the demarcation of ‘African Land Units.’ This prevented the expansion of tribes into another’s territory and thus eliminated the principal source of major inter-tribal wars. ... Within the African areas the indigenous pattern of ‘water-tight’ units was maintained, but accentuated by the increasing population.” Land demarcation and individual land registration during the postcolonial period “has frozen the previously fluid situation and virtually halted the traditional mobility” (Government of Kenya, 1986). Unlike central Kenya, Busia and Teso were free of European settlement and resulting disruptions of land claims during the colonial period.

We examine the relative stability of residential ethnic composition in Busia and Teso districts during the postcolonial period as a first test of the validity of the identification strategy (Table 1). Comparing residential ethnic composition at the geographic division level in 1996 using pupil survey data (described in Section 4 below) to residential composition in 1962 using Kenyan Census data suggests that ethnic residence patterns have been largely stable. The ordering of residential ethnic diversity across geographic divisions measured by the size of the largest ethnic group is identical in 1962 and 1996 (unfortunately, census data on ethnic composition for smaller administrative units is unavailable). Recent survey evidence also suggests that land sales and residential mobility are extremely rare in Busia and Teso districts. Among 507 local households interviewed for a different study in this area, only three respondents claimed to have bought or sold any land during 1997–1998, and these individuals may have purchased land in the vicinity rather than moving away (results available upon request). Residential mobility is low in rural Kenya for a variety of reasons, including the fact that local land markets are thin as in much of sub-Saharan Africa (Collier and Gunning, 1999) and because approval from relatives is required de facto to sell traditional family or clan land and such approval is often difficult to obtain (Platteau, 2000).¹⁷

¹⁷ By way of contrast, 9.1% of Americans interviewed in the General Social Survey during 1985–1994 had lived in their current community for less than 1 year (DiPasquale and Glaeser, 1999).

Table 2
Pupil descriptive statistics by ethnic group

	Entire sample	Luhyia pupils	Teso pupils	Luo pupils
Number of pupils interviewed	5832	3867	1516	301
Proportion of pupil sample	1	0.66	0.26	0.05
Age in years	14.5	14.4	14.7	14.5
Father, years of education	7.5	7.5	7.4	7.5
Mother, years of education	5.0	5.0	4.9	5.2
Fathers with formal employment	0.24	0.26	0.17	0.29
Mothers with formal employment	0.05	0.04	0.05	0.04
Proportion latrine ownership	0.85	0.84	0.87	0.81
Proportion iron roof ownership	0.26	0.26	0.23	0.30
Proportion livestock ownership	0.78	0.78	0.78	0.80
Proportion cultivates corn (maize)	0.87	0.87	0.86	0.92
Proportion cultivates cash crop	0.39	0.33	0.56	0.29
Attends primary school that is not the closest to home	0.18	0.19	0.17	0.17
Residence and school in different geographic zones	0.12	0.15	0.06	0.06
Lives with a parent if at least one parent is alive	0.85	0.84	0.88	0.84
Average number of full siblings	4.5	4.4	4.6	4.1
Proportion, Catholic	0.56	0.57	0.56	0.64

Data are from the 1996 ICS Pupil Questionnaire administered to pupils in Grades 6 to 8. Other ethnic groups, including Kalenjin, Kikuyu, Masaai, Somali, Tachoni, and Taita, comprise 0.020 of the sample. Formal employment includes government, factory, and corporate jobs but not petty trading (in contrast to the definition of formal employment in Miguel (2001)). Livestock ownership includes cattle, goats, sheep, and pigs (but not chicken). Cash crops include tobacco, sugar cane, and cotton.

Although residential patterns in this area are stable, households can choose which local primary school their children will attend and which water well they will use, creating endogenous school populations and water users. Sorting is an important issue in this context; evidence from a parent survey indicates that nearly a quarter of all households send a child to a primary school that is not the nearest to their home. For example, well-funded and high-quality schools with good headmasters may attract more ethnically diverse pupil populations, if pupils from surrounding areas walk longer distances to good schools to take advantage of their educational quality.¹⁸ This sorting pattern implies an upward bias in ordinary least squares estimates of the impact of ethnic diversity on school funding levels if some aspects of school quality remain unobserved to the econometrician. To avoid biases due to endogenous sorting among schools and wells within walking distance of the household, we employ local residential ethnic diversity as our principal measure of ethnic diversity rather than the ethnic diversity of actual school pupils or water users. The diversity indexes are precisely defined in Section 4.

Another potential identification problem is the possibility that children could move in with relatives to attend a primary school that is not within walking distance of their home, altering effective local ethnic composition even if adult residential patterns remain fixed. Yet, fewer than 15% of pupils are not living with a parent among pupils with at least one surviving parent (Table 2). Inasmuch as some pupils move in with relatives residing in the

¹⁸ If diversity has a negative effect on collective action, these pupil in-flows may somewhat diminish the benefits of having a good headmaster, although in general, they need not completely eliminate the benefits.

same area and relatives often live near each other the proportion of children who move in with relatives in a different village or geographic zone is plausibly considerably less than 15% although data limitations make it impossible to determine the exact proportion. Such rates are unlikely to significantly alter local residential ethnic diversity. The following section describes the primary school and water well data in greater detail.

4. Data description

4.1. Primary school data

Detailed data for 100 of the 337 rural primary schools in Busia and Teso districts were collected from pupil, school, and teacher questionnaires in early 1996 as baseline information for a nongovernmental organization (ICS Africa) School Assistance Project (SAP).¹⁹ Busia and Teso district primary schools are typical for Kenya in terms of educational attainment; the districts combined ranked 26th of 50 districts on 1995 national school exams (Glewwe et al., 1998).²⁰

The nongovernmental organization (NGO) that collaborated with the authors on the project collected a variety of financial and demographic data for these schools in 1996. The pupil questionnaire focused on pupils' schooling background, family educational characteristics and asset ownership, and self-described ethnic affiliation. Over 6000 pupil questionnaires were administered by trained survey enumerators to all Grade 6 through 8 pupils (ages 12–18 years old) present on the day of questionnaire administration in 1996. Younger pupils were not administered questionnaires because of their limited reading and writing skills. In total, 100 school questionnaires and 861 teacher questionnaires were also administered in 1996.²¹

¹⁹ The schools were selected by the Ministry of Education district education office in 1995 using the following criterion (Glewwe et al., 1998). These were needy rural schools that had not been assisted by an earlier World Bank assistance program, which had distributed free textbooks to 100 different schools in this area in 1994. The SAP schools were first stratified by geographic division and then randomly assigned into four groups that were to receive financial assistance in sequence over the following 4 years. Table A1 in Appendix B indicates that the School Assistance Project (SAP) schools had smaller enrollments and performed worse on government exams than other rural schools in the area due to the criteria for selection into the program. However, the interaction term between local ethnic diversity at the level of the geographic zone—a local administrative unit—and the indicator for selection into the SAP is not significantly related to either school enrollment and examination scores, suggesting that sample selection is unlikely to be producing a spurious correlation between local ethnic diversity and school outcomes.

²⁰ In 1996, the original Busia district was split in two: Teso district is the northern part of the original Busia district, and Busia district is the southern part.

²¹ School questionnaires filled by schoolmasters with the assistance of an enumerator contain detailed information on school finances, infrastructure, inputs, and enrollment. Teacher questionnaires focus on qualifications and were completed by teachers themselves. The survey enumerators, Charles Asoka, Robert Namunyu, Polycarp Waswa, and Maureen Wechuli, believe that responses from the school and teacher questionnaires are more reliable than pupil responses. Fortunately, the question on pupil ethnicity is likely to suffer from less response error than other questions.

The Busia District Education Office provided school examination results and examination name lists. NGO enumerators using portable global positioning system (GPS) machines collected school latitude and longitude for all primary schools and wells in the sample. The primary school analysis below includes the 84 of 100 schools with complete pupil, school, teacher, and GPS data and school committee records.²²

4.2. Ethnicity data

Ethnolinguistic fractionalization (ELF) is the principal empirical measure of ethnic diversity.²³ This has the advantage of comparability with the existing literature (Mauro, 1995; Easterly and Levine, 1997 and Alesina et al., 1999). Ethnolinguistic fractionalization is the probability that two people randomly drawn from the population are from distinct groups and is related to a Herfindahl index. Formally,

$$\text{ELF} \equiv 1 - \sum_i (\text{Proportion of group } i)^2 \quad (2)$$

Using the 1996 Pupil Questionnaire data, ethnolinguistic fractionalization among all sample pupils is 0.49. An additional measure of ethnic diversity employed in the analysis is one minus the proportion of the largest ethnic group.

There are two principal sources of local ethnic diversity data employed in this study. The first is diversity of the primary school pupils residing in the geographic zone,²⁴ a measure of diversity independent of local sorting among schools or wells. A second measure of local ethnic diversity is computed among all pupils attending primary schools located within 5 km of either the primary school (when a school outcome is the dependent variable) or the well (when well maintenance is the dependent variable). These data were created from 1996 Government Examination Namelists provided by the Busia and Teso District Education Office. The 5-km radius around each school appears to be a rough upper limit on the distance that people may walk to attend school or fetch water on a daily basis although the empirical results are robust to radiuses of between 3 and 6 km (not shown). Additional information on these measures is provided in Appendix C.

4.3. Primary school organization and funding

Both the central government and local school committees play important roles in Kenyan primary school finance. The national Kenya Ministry of Education pays teacher salaries, while school committees raise funds locally for books, chalk, classrooms, and desks. Although the teacher salaries and benefits paid by the central government account for most primary school spending, approximately 90% of total primary school spending according to the authors' calculations, a reduction in local funding could have an important impact on educational outcomes if local inputs and teachers are complements in educational

²² This data set is available from the authors upon request.

²³ Vigdor (2004) derives the fractionalization index from a simple model of "directed altruism". Esteban and Ray (1994) present a useful methodological discussion of alternative polarization measures.

²⁴ At the time of data collection, Kenya contains 50 districts, each of which is split into divisions, which are in turn composed of zones. There are 22 zones in our data set.

production.²⁵ Locally raised funds are critical to the provision of school supplies, such as chalk and teacher textbooks, as well as paying the salaries of any nonteaching staff such as watchmen. In addition, local funds have an important effect on public perceptions of school quality because they are used largely for construction and other visible capital projects.

Parents raise the bulk of local school funds through two mechanisms: school fees and local fundraising events. Annual school fees are set by the school committee and collected by the school headmaster. Each primary school is managed by its own school committee. The school committee is composed of 12 class representatives directly elected by the parents of schoolchildren, and it typically meets at least three times per year to discuss school finances and plan school projects. School fees in Busia and Teso districts ranged from 200–500 Kenya Shillings (4–10 US dollars) per family in 1995, a nontrivial amount for these families. Participation in primary school activities is one of the most prominent forms of civic participation available to individuals in rural Kenya inasmuch as local government officials are appointed by the central government, and national elections are infrequent and often not free.

The other important source of local primary school funding in western Kenya, accounting for approximately one third of local funding, are village fundraisers called *harambees* in Swahili. At these events, parents and other community members meet and publicly pledge financial support for a planned school investment project, such as the construction of a new classroom. *Harambees* are major local events; planning a *harambee* requires a great deal of effort on the part of the school committee, and all parents and members of the local community are expected to attend. The atmosphere is festival-like, with individuals being called upon to come forward to make payments, while the crowd cheers in support. While contributions at these events are supposedly voluntary, school committees often set *harambee* contribution levels for parents and teachers, and contributions are recorded by the committee.

Harambees are an important source of public finance throughout Kenya, accounting for 40% of total local expenditures on primary schools, water well maintenance, and other local public goods, according to Wilson (1992). Although local political figures, such as Members of Parliament and Chiefs, often attend fundraisers and make contributions (Barkan, 1994), school records indicate that the bulk of *harambee* funds are raised from parent and community contributions, while contributions from politicians and other government officials account for only 19% of *harambee* funds in the area.²⁶

4.4. Social pressure and parental contributions

The authors conducted structured field interviews regarding primary school finance with 12 primary school headmasters in this area on June–July 2000 at six schools in Busia

²⁵ In addition to its school finance role, the National Ministry of Education assigns teachers and headmasters to primary schools. Survey evidence indicates that 80% of teachers in Busia and Teso districts are assigned to teach in their native “home” area. Although it is not unknown for school committees to lobby for the transfer of poorly performing teachers and headmasters, the Ministry of Education generally assigns teachers with minimal local participation. We thank Sylvie Moulin and Robert Namunyu for their observations on these and other issues.

²⁶ These data were gathered by Polycarp Waswa on September–October 2001 for 75 primary schools in Busia district (Funyula and Budalangi divisions), 29 of which are in our sample.

district and six in Teso district, in both ethnically diverse and homogeneous areas (additional information on the interviews is provided in Appendix D), and found that headmasters and school committees employ a variety of methods to encourage *harambee* and school fee payment in western Kenya. These interviews provide further evidence on the importance of social pressure in overcoming free-rider problems. One common method of pressuring parents to contribute is to publicly announce the names of parents who are late with *harambee* or fee payments during parent and school committee meetings. Ten of the 12 headmasters reported that they regularly read out the names of parents late with payments at school meetings, which they believe embarrasses the parents and encourages other community members to pressure them to pay.

For example, the headmaster of Nanderema Primary School in Busia district stated that parents who have already “paid [their fees] become very bitter, and help us collect money from others,” and the headmaster of Buduta Primary School in Teso district stated that parents who have paid “bully those who have not paid.” Additional forms of social pressure to encourage prompt payment of *harambee* contributions or school fees include sending letters to the homes of parents late with fees, asking local church leaders to encourage payment during sermons, and making personal visits to the individual homes of debtors accompanied by the local Chief. *Harambees* themselves also embody a form of social pressure inasmuch as payments are made publicly.

The headmaster interviews also indicate that ethnically diverse communities may be less able to effectively pressure parents to make school payments. Buduta Primary School in Busia district, a predominantly Luhya school with a substantial Luo minority, provides an illustration. The headmaster of Buduta claims that he has had difficulties raising funds from the minority Luo community because they lack “a sense of ownership” for the school and “feel less committed to the school” than the majority Luhyas although he stated that there are “no bad feelings [among Luhyas and Luos] in the school.” The informal mechanisms that sustain Luhya school payments appear to be less effective among members of the minority Luo community, perhaps because they are outside the dense network of reciprocal social relations that sustain high contribution rates among the Luhyas (Woolcock, 1998), as suggested by the theoretical model in Section 2.

While the threat of sanctions and temporary suspension from school are important tools used by headmasters and school committees to enforce payment, the children of noncontributing parents cannot permanently be removed from school. Former President Moi of Kenya repeatedly stated that no child should be refused a primary education because of nonpayment of fees, and this decree limited the discretion of schools to expel students. The content of threatened sanctions therefore includes some potential loss of time in the classroom and less support from social networks but not complete exclusion from the public good.²⁷

4.5. School funding and quality measures

The main school outcome measure is total local school funding collected per pupil in 1995, and this is further subdivided into total local *harambee* donations per pupil and

²⁷ Moi was replaced as President by Mwai Kibaki (the former opposition leader) in December 2002, and primary school fees have since been abolished nationwide. Our description of primary school finance reflects the period prior to this change of regime.

school fees collected per pupil. This measure does not include funds raised from nongovernmental organizations, but most schools received little or no such assistance inasmuch as schools that were receiving larger levels of outside donor assistance were largely excluded from the program. Only six sample schools had received over \$100 in outside funding in 1995, and local fundraising does not appear to be crowded out in these schools (regressions not shown). School facilities and inputs—the number of desks per pupil, latrines per pupil, classrooms per pupil, and school-owned textbooks per pupil in 1996—are also outcome measures and reflect recent local funding levels. Average school performance on 1996 NGO academic examinations (which were based on the format of government exams) for Grades 3 to 8 captures aspects of educational quality.

In addition to the data on school funding, we present information on other forms of community participation including the number of parent meetings at the school and the level of parent and teacher cooperation in the schools as evaluated by NGO field officers. Parent participation is measured by the number of meetings in 1997 held for parents to ratify school committee recommendations. To create rankings of cooperation, three field officers were asked to independently rate schools on a number of characteristics related to participation, motivation, and cooperation. These ratings were used to create an overall rating of the school in terms of parent participation and cooperation and teacher motivation and participation, providing another perspective on local collective action. Primary school teachers in rural western Kenya are typically local community members; 1998 survey evidence indicates that the median teacher lives 3 km from their school, and that nearly 80% of teachers claim to be living on their home compound (i.e., their permanent residence in their native “home” area). This suggests that teacher performance may in part be governed by the same community sanctioning mechanisms that affect parent participation.

4.6. School committee sanctions

The principal empirical measure of social sanctions is the number of times the school committee uses the threat of sanctions to encourage parental contributions or other forms of community involvement in the school. As noted above, school committees manage local school funds and serve as the governing boards for primary schools. The school committee is expected to meet at least three times per school year to set school fees and policies, plan for capital improvements, and monitor fee payments from parents. The primary school headmaster serves as the committee secretary and writes detailed minutes at each school committee meeting, recording agenda items, issues discussed, and decisions made. We analyze the school committee meeting minutes for the 84 schools with complete records and divide all record items into two broad categories.²⁸ The bulk of school committee meeting business (85% of all items, as shown in [Table 3](#)) concerned “administrative” decisions taken on school projects, plans, or routine committee functions

²⁸ The overall sample size of schools falls to 84 from 100 because of missing data, largely data on school committee meetings. [Gugerty and Miguel \(2000\)](#) show that school committee meeting records are somewhat more likely to be missing in ethnically diverse areas. If these ethnically diverse schools have missing records because they have poorly functioning school committees that are also unable to sanction parents late with school contributions, the resulting bias would strengthen our results.

Table 3
Primary school descriptive statistics

	Mean	Standard deviation	Observations
<i>(A) Zone characteristics</i>			
Zonal residential ELF across tribes, 1996 Pupil Questionnaire data	0.23	0.14	84 (Schools)
Proportion of largest ethnic group in zone, 1996 Pupil Questionnaire data	0.86	0.11	84 (Schools)
<i>(B) School and teacher characteristics</i>			
School ELF across tribes, 1996 Pupil Questionnaire data	0.20	0.18	84 (Schools)
School ELF across tribes, 1996 Exam Namelist data	0.21	0.15	84 (Schools)
ELF across tribes for all schools within 5 km of a school (including the school itself), 1996 Exam Namelist data	0.24	0.13	84 (Schools)
Proportion of largest ethnic group in school, 1996 Pupil Questionnaire data	0.79	0.18	84 (Schools)
Total local school funds collected per pupil, 1995 (Kenyan Shillings)	152.6	99.4	84 (Schools)
<i>Harambee</i> donations collected per pupil, 1995 (Kenyan Shillings)	44.8	88.2	84 (Schools)
School fees collected per pupil, 1995 (Kenyan Shillings)	107.8	48.6	84 (Schools)
Desks per pupil, 1995	0.21	0.12	84 (Schools)
Pupil latrines per pupil, 1995	0.016	0.013	84 (Schools)
Classrooms per pupil, 1995	0.030	0.014	84 (Schools)
School-owned texts per pupil, 1995	0.34	0.21	84 (Schools)
Private texts per pupil, 1995	0.07	0.10	84 (Schools)
Pupil enrollment per primary school, 1996	296.3	146.4	84 (Schools)
Average score on 1996 NGO examination, grades 3–8 (in standard deviations)	0.05	0.47	84 (Schools)
School committee record items regarding sanctions or verbal pressure, 1997	3.2	3.0	84 (Schools)
School committee record items regarding administrative activities, 1997	18.9	11.4	84 (Schools)
Parent school meetings, 1997	3.4	1.9	83 (Schools)
Parent cooperation from 0 to 1 (reported by field officers), 1998	0.49	0.33	84 (Schools)
Teacher motivation from 0 to 1 (reported by field officers), 1998	0.54	0.30	84 (Schools)
Pupil–teacher ratio, 1996	29.1	9.8	84 (Schools)
Proportion teachers with high school education, 1996	0.79	0.16	83 (Schools)
Years of teaching experience, 1996	14.0	3.0	83 (Schools)
Proportion of female teachers, 1996	0.26	0.19	83 (Schools)
Latitude (degrees north), GPS data	0.43	0.19	84 (Schools)
Longitude (degrees east), GPS data	34.23	0.13	84 (Schools)
Number of other primary schools within 5 km, GPS data	14.5	3.8	84 (Schools)

Data are from the 1996 ICS School and Pupil Questionnaires, 1996 Government Exam Namelists, and global positioning systems (GPS) readings taken by NGO field workers. Ethnolinguistic fractionalization is defined as $1 - \sum_i p_i^2$ (proportion of ethnolinguistic group i in the population)². School ELF across tribes and the proportion of the largest ethnic group in the school consider Luhya as a single group.

like elections. The remaining 15% of decisions concern pressure placed on parents and school committee members to improve fee payments rates, in-kind contributions, and school discipline, which we refer to as “threatened sanctions”.

The minutes indicate that school committees use a variety of tools and threats to motivate parents to participate in school activities and fundraising, including resolutions in committee meetings, personal appeals, threats to suspend pupils who are behind on their payments, and direct attempts to collect fees. Unfortunately, due to data limitations, we cannot typically identify which parents are being targeted with these sanctions and thus cannot directly test whether ethnic majority or minority group parents are particularly prone to free-riding. We argue, in effect, that these threats are more credible and potent and more often used in ethnically homogenous schools. A more detailed description of these records, including a complete transcription of the meeting minutes for one school and a discussion of the coding system, is presented in Appendix D.²⁹

4.7. *Community water wells*

Water wells are another important local public good in rural western Kenya. The lack of safe drinking water is a major public health problem that contributes to the spread of water-borne diseases, including schistosomiasis, amebiasis, cholera, and other gastrointestinal infections in Kenya and other less developed countries (Government of Kenya, 1986). Well water is generally safer to drink than alternative water sources, such as stream or lake water.

The vast majority of community wells in western Kenya were constructed in 1982–1991 with the assistance of the Finnish government through an organization called the Kenya–Finland Development Cooperation (KEFINCO). KEFINCO identified well sites in cooperation with local communities, dug the original boreholes, and provided the equipment required to operate the wells. Communities were then responsible for forming local well committees in charge of collecting usage and maintenance fees from the community and ensuring that the well remained in good repair. The committees have operated on a voluntary basis with little explicit public authority for revenue collection, so their ability to collect fees depends on their success in exerting social pressure in the local community. A report by a Dutch technical agency finds that many of the water committees attempted to collect fees only when the wells completely broke down, and many committees had particular difficulty collecting funds in wells located in public areas, such as primary schools or village centers (IRC, 2000). The report concludes that the most successful wells were those in which access to water could be strictly regulated and user fees charged at the point of service, but unfortunately, this is rarely the case in practice.

Structured interviews were conducted with a stratified random sample of 33 local water committees in November 2001 to probe the causes of failed collective action regarding well maintenance in Busia and Teso districts. The interviews suggest that most committees in these areas face persistent difficulties in collecting well maintenance contributions from

²⁹ We are grateful to Steve Barham for his assistance with the meeting records. Unfortunately, we are often unable to distinguish between threatened sanctions regarding *harambee* contributions vs. school fees in these records and thus cannot separately examine the impact of these different types of sanctions.

free-riding local residents, and this problem is exacerbated by the fact that many residents believe the well water should be freely available to all local residents. In interviews in ethnically diverse Butula Division, respondents indicated that maintenance fee collection was particularly difficult among ethnic minority individuals from nearby areas who began using the well (additional information on these interviews is provided in Appendix D).

The data on well maintenance comes from a survey of nearly 700 wells conducted in Busia and Teso districts from October 2000 to August 2001. The sample consists of the universe of modern borehole wells constructed in both districts from 1982–1991 by KEFINCO. The current condition of the KEFINCO wells thus reflects the success of local collective action in well maintenance from the 1980s through 2001.³⁰ The survey collected detailed information on the physical condition of the wells, including water flow and missing or broken parts, as well as GPS locations, alternative local water sources, and the perceived performance of the local water committees charged with maintenance. The principal dependent variable for well maintenance is an indicator variable that takes on a value of 1 if water flow in the well was judged to be “normal” by field workers and 0 if either no water flows from the well or if the water flow is “very low”. Only 57% of the wells had “normal” water flow at the time of the survey (Table 9), suggesting widespread collective action failures, and this echoes the findings of an existing Kenyan government report on the state of well maintenance in the region (*Community Water Supply Management Project, 2000*).

A second outcome measure takes on a value of 1 if all well parts are in working condition and 0 if any are missing or broken. The final dependent variable takes on a value of 1 if local residents have access to a functioning alternative well to gauge the possibility of substitution toward other safe water sources. Using GPS data for each well together with the school ethnicity data, we construct our local ethnic diversity measure, the ethnolinguistic fractionalization of children who attend primary schools within 5 km of each well. Unfortunately, we do not have information on social sanctions imposed to sustain water well contributions and thus cannot directly link the ethnic diversity effects we estimate for wells to the theoretical model of social sanctions presented in Section 2.

5. Empirical results

Ethnic diversity is related to sharply lower primary school funding through voluntary fundraising events (harambees) and to lower quality school infrastructure controlling for local socioeconomic characteristics. Moreover, school records indicate that ethnically diverse schools use fewer community social sanctions than more homogenous areas, providing support for the claim that free-riding may be more prevalent in diverse communities because of the inability to create effective community sanctions. However, we cannot entirely rule out all competing hypotheses for the sanctions results. For example, there might be less need to employ sanctions in a community that does not intend to hold large harambees, and weak local support for primary schooling might also drive

³⁰ We were unfortunately unable to obtain data on the precise year of construction for each well and so cannot control for this variable in the analysis below.

both lower school funding and fewer sanctions without necessarily any direct causal link between the two. We discuss these issues in more detail in Section 5.4 below. Nonetheless, the evidence we present below on recorded school committee social sanctions, the interviews with headmasters, the field worker observations on parent cooperation, and the anthropological evidence cited above, taken together, provide considerable evidence that effective social sanctions are in fact a key factor allowing ethnically homogeneous communities to achieve collective action in rural Kenya, as we argue further below.

5.1. Empirical specification

The main empirical specification for all outcomes is presented in Eq. (3). Y is the outcome measure (e.g., school funding, school infrastructure quality, threatened sanctions, or well maintenance). ETHNIC is the measure of local ethnic diversity, and X is a vector of zonal socioeconomic, demographic, and geographic controls (including geographic division indicator variables in some specifications), where i denotes either a school or a well.

$$Y_i = a + \tau \text{ETHNIC}_i + X_i' \beta + \mu_i \quad (3)$$

In the specifications using zonal residential ethnic diversity as the measure of local ethnic diversity, school regression disturbance terms are assumed to be independent across geographic zones but are clustered within zones. In specifications using local ethnic diversity among schools within 5 km of the unit as the measure of diversity, regression disturbance terms are allowed to be correlated across schools as a general function of their physical distance, using the spatial estimation method in Conley (1999).³¹

Observed socioeconomic differences across the major local ethnic groups are minor (Table 2), suggesting that ethnic diversity is unlikely to be proxying for income inequality or average socioeconomic status³², although there are some differences between Tesos and other groups in fathers' formal sector employment and the cultivation of cash crops. To address the possibility that Tesos also have different preferences for educational spending than other groups on average, an indicator variable for the local proportion of Tesos is included as a control variable in certain specifications below. There is considerable variation across schools in diversity, local funding, and other outcome measures (Table 3).

In addition to the OLS specification in Eq. (3), we also present a two-stage instrumental variable estimate of the impact of local ethnic diversity on school funding in Table 5, in which case the specification in Eq. (3) can be thought of as the reduced form. The instrumental variable method imposes the condition that local ethnic diversity only affects funding outcomes through school ethnic diversity and not through other channels.

³¹ Following Conley (1999), spatial standard errors are calculated with a weighting function that is the product of a kernel in each direction (north to south, east to west). The kernels start at one and decrease linearly until they are zero at 8 km from the school. Results are robust to varying this cut-off between 5 and 8 km (results not shown).

³² There is insufficient information on household income, consumption, and land ownership in the data set to directly examine the relationship between public goods and local income inequality.

Beyond any direct impact it may have on local collective action, ethnic diversity may potentially be associated with local school funding through its relationship with other local characteristics. For example, ethnically diverse regions may be poorer than other areas because it is difficult to enforce contracts within heterogeneous communities, leading credit, land, and labor markets to function less efficiently. Diverse areas may also have worse school quality if they are assigned lower quality teachers by the National Ministry of Education. However, socioeconomic and school characteristics are not strongly correlated with local ethnic diversity (Table 4), and hence, unobserved socioeconomic variation correlated with ethnic diversity is unlikely to be driving the collective action outcomes so there is no decisive evidence challenging the validity of the instrumental variable estimation strategy.³³ Twelve of the 16 measures of socioeconomic variation and school quality are not significantly associated with local ethnic diversity in either specification, including mother and father years of education and all of the school teacher characteristics, and no measure is significantly associated with diversity at 90% confidence in both specifications. It is unclear if ethnically diverse areas are generally richer or poorer than homogeneous areas, however, inasmuch as formal sector employment appears to be lower in diverse areas, but the proportion of households growing cash crops, principally tobacco, is significantly higher. The data set does not contain reliable income data, so we are unable to directly estimate income differences. There is no systematic pattern to the remaining coefficient estimates on ethnic diversity. Variables that are significantly related to local ethnic diversity in either specification in Table 4 are included as explanatory variables in the subsequent analysis to control for possible socioeconomic differences across areas.³⁴ Aspects of school quality that are centrally funded by the government, such as the pupil–teacher ratio and the average education of teachers, do not vary with local ethnic diversity (Table 4B), but as shown below, locally funded aspects of school quality are associated with local ethnic diversity.

5.2. Ethnic diversity and school funding

Ethnic diversity is negatively and significantly related to school funding and the quality of school facilities. This result is primarily driven by a reduction in funding through *harambees* (as shown in Table 6 below). Table 5 presents the 1995 school funding results, where the dependent variable is total local school funds collected per pupil in 1995. The first stage regression of school ethnic diversity (dependent variable) on local ethnic

³³ It may be surprising that diverse areas are not in fact significantly poorer than homogeneous areas in light of the large differences in collective action outcomes estimated in Section 5 and the cross-country economic growth results in Easterly and Levine (1997). The most plausible explanation is that educational and other local investments made when today's parents were themselves children (in the 1940s–1960s) were overwhelmingly allocated by the colonial government in Nairobi or by Christian missionary societies and, hence, were not subject to the local collective action problems described in this paper. The decentralization of public goods funding through the *harambee* self-help movement only took root in Kenya during the postindependence period in the late 1960s and 1970s (Barkan, 1994), so the long-term impacts of decentralization on income levels in ethnically diverse areas may not yet be apparent.

³⁴ Although inclusion of these variables may lead to bias if they are themselves directly affected by ethnic diversity (Angrist and Krueger, 1999), this bias is unlikely to be large since these variables are typically not statistically significantly related to local school funding (Table 5).

Table 4
Ethnic diversity and local characteristics

Dependent variable	Coefficient estimate on zonal residential ELF across tribes (OLS)	Coefficient estimate on ELF across tribes among schools within 5 km (spatial OLS)	Number of schools	Mean dependent variable
<i>(A) Pupil characteristics (1996 Pupil Questionnaire)</i>				
Father, years of education	0.5 (1.0)	−0.4 (1.2)	84	7.3
Mother, years of education	1.2 (1.3)	0.2 (1.4)	84	4.9
Fathers with formal employment	−0.09 (0.07)	−0.24*** (0.07)	84	0.23
Mothers with formal employment	−0.01 (0.02)	0.01 (0.02)	84	0.04
Proportion of latrine ownership	0.13 (0.09)	0.06 (0.09)	84	0.84
Proportion of iron roof ownership	0.04 (0.11)	0.02 (0.10)	84	0.25
Proportion of livestock ownership	0.16* (0.08)	0.12 (0.11)	84	0.78
Proportion, cultivates corn (maize)	−0.03 (0.06)	−0.16** (0.08)	84	0.87
Proportion, cultivates cash crop	0.26 (0.31)	0.67*** (0.20)	84	0.40
Average number of full siblings	1.7 (1.5)	2.5 (1.6)	84	7.4
Proportion, Catholic	−0.03 (0.19)	0.07 (0.17)	84	0.57
<i>(B) School and teacher characteristic</i>				
Pupil enrollment per primary school, 1996	72.2 (103.0)	−13.1 (104.4)	84	296.3
Pupil–teacher ratio, 1996	−4.2 (10.0)	−8.8 (6.4)	84	29.1
Proportion, teachers with HS education, 1996	0.10 (0.08)	−0.10 (0.14)	83	0.79
Years of teaching experience, 1996	0.3 (3.0)	2.7 (2.4)	83	14.0
Proportion of female teachers, 1996	−0.12 (0.15)	−0.12 (0.16)	83	0.26

Huber robust standard errors in parentheses. Significantly different than 0 at 90% (*), 95% (**), 99% (***) confidence. Regression disturbance terms are clustered at the zonal level. Ethnolinguistic fractionalization is defined as $ELF = 1 - \sum_i (\text{proportion of ethnolinguistic group } i \text{ in the population})^2$. School ELF considers Luhya as a single group. The coefficient estimate on zonal residential ELF across tribes uses data from the 1996 Pupil Questionnaire. In these specifications, observations are assumed to have independent error terms across geographic zones but not necessarily within zones. The coefficient estimate on ELF across tribes among schools within 5 km uses 1996 Exam Namelist data. In these specifications, regression disturbance terms are allowed to be correlated across schools as a general function of their physical distance, using the estimation strategy developed in Conley (1999).

diversity indicates that local ethnic composition is a strong predictor of school ethnic composition (Table 5, regression 1). Local ethnic diversity alone captures over 40% of the variation in school-level ethnic diversity. Moreover, although the coefficient estimate is less than 1, it is quite large (0.86), indicating that there is in fact limited local ethnic sorting among schools.

Survey data further supports the notion that household sorting decisions among primary schools in this area are not driven mainly by the desire to segregate by ethnic group. Of 2251 parents asked on the main reason why they chose a particular primary school for their children in 1996, 78% claimed to do so because of proximity to their home, 7% because lower school fees, and 7% because of superior academic quality, while less than 1% of the responses can be interpreted as a desire to sort by ethnicity. Although alternative theories of diversity and collective action (including Alesina et al., 1999 and Alesina and

Table 5
Ethnic diversity and local primary school funding

Explanatory variable	Dependent variable								
	School ELF across tribes								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS		OLS	IV-2sls	OLS	OLS	OLS	Spatial OLS	Spatial OLS
	1st stage								
<i>Ethnic diversity measures</i>									
Zonal ELF across tribes	0.86*** (0.07)		-185.7** (77.9)		-145.2*** (49.6)	-143.6* (82.1)			
School ELF across tribes		-32.9 (64.0)		-216.4** (88.4)					
1-(Proportion largest ethnic group in zone)							-162.9** (66.6)		
ELF across tribes for all schools within 5 km								-174.0** (76.3)	-174.0** (80.8)
<i>Zonal controls</i>									
Proportion fathers with formal employment					189.5 (165.1)	-220.6* (120.5)	184.6 (170.9)		142.8 (167.3)
Proportion of pupils with a latrine at home					-431.6*** (139.9)	-286.3 (228.0)	-429.8*** (150.3)		-466.9 (250.2)
Proportion livestock ownership					120.1 (136.9)	186.2 (130.4)	110.6 (148.3)		116.9 (117.7)
Proportion cultivates cash crop					35.7 (61.4)	22.2 (106.9)	27.8 (62.4)		85.2 (78.4)
Proportion Teso pupils						67.9 (181.4)			
Geographic division indicators	No	No	No	No	No	Yes	No	No	No
Root MSE	0.14	99.8	96.7	105.5	95.0	93.0	95.4	97.1	95.0
R ²	0.40	0.00	0.06	-	0.14	0.25	0.12	0.06	0.09
Number of schools	84	84	84	84	84	84	84	84	84
Mean dependent variable	0.20	152.6	152.6	152.6	152.6	152.6	152.6	152.6	152.6

Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. Observations are assumed to have independent error terms across geographic zones but not necessarily within zones for regressions 1 to 7. Ethnolinguistic fractionalization is defined as $1 - \sum_i (\text{proportion of ethnolinguistic group } i \text{ in the population})^2$. School ELF across tribes and the proportion of the largest ethnic group in the school consider Luhya a single group. Regression disturbance terms are allowed to be correlated across schools as a general function of physical distance in regressions 8 and 9 (Conley 1999). Geographic indicators are indicators for six (of the seven) geographic divisions. The instrumental variable in regression 4 is zonal ELF across tribes.

LaFerrara, 2000) imply that there should be substantial pressure for ethnic groups to sort into segregated local schools and water groups, the social sanctions model presented in Section 2 does not generally imply ethnic segregation, and thus, this model appears to be consistent with the data.

There is an insignificant negative relationship between the ethnolinguistic fractionalization of pupils within the school and school funding in the ordinary least squares specification (Table 5, regression 2), and this coefficient may be explained by endogenous local school sorting towards lower cost and higher quality schools. For example, if schools with (unobservably) better quality headmasters and teachers attract pupils from farther away, and these pupils tend to be more ethnically diverse than local pupils on average, this leads to an upward sorting bias on OLS coefficient estimates on ethnic diversity.³⁵

The point estimates on local residential ELF are negative and significantly different than 0 at 95% confidence (Table 5, regression 3), and the instrumental variable specification yields similar results (regression 4); the IV results are largely robust to the inclusion of geographic division indicators, with a coefficient estimate on ELF at -125.7 (standard error 74.9, regression not shown). The coefficient estimate on local residential ELF is also robust to the inclusion of zonal socioeconomic controls (regression 5), geographic division indicator variables, and the proportion of Tesos in the zone (regression 6), suggesting that measured ethnic diversity is not proxying for average socioeconomic status or cultural differences across ethnic groups. The zonal socioeconomic controls include the proportion of fathers with formal sector employment, the proportion of pupils residing with a latrine at home, the proportion of pupils whose households own livestock, and the proportion of pupils whose households cultivate a cash crop.

Fig. 3 graphically presents the negative relationship between average school funding and residential ELF across geographic zones. The diversity effect remains significantly different than 0 at 95% confidence in all cases when schools from one zone at a time are dropped from the sample (results not shown). An interpretation of the coefficient estimate on ELF in regression 5 is that the drop in local school funding associated with a change from complete ethnic homogeneity to average school-level ethnic diversity is 29 Shillings or approximately 20% of average local funding. Inasmuch as an average primary school textbook costs approximately 150 Shillings, and the ratio of textbooks to pupils in these schools is one to three (Table 3), eliminating the “costs” of higher diversity would allow schools to double their textbook stocks in 2 years.

It remains theoretically possible that ELF could be capturing a nonlinearity in the relationship between funding and the size of a particular ethnic group rather than the impact of ethnic diversity per se. However, the functional form of the ethnolinguistic fractionalization index does not appear to be driving the results. A linear measure of ethnic diversity, the proportion of the largest ethnic group in the school, is also

³⁵ Further empirical evidence consistent with this sorting pattern is presented in Miguel (2001), including the finding that within geographic zones, more ethnically diverse schools have statistically significantly higher average test scores and considerably—although not statistically significantly—higher school fundraising per child and total school population (regressions not shown). Inasmuch as schools within a given geographic zone are located near each other, these patterns shed light on parent choices for their children’s schooling and suggest that good quality schools become larger and more ethnically diverse due to local sorting.

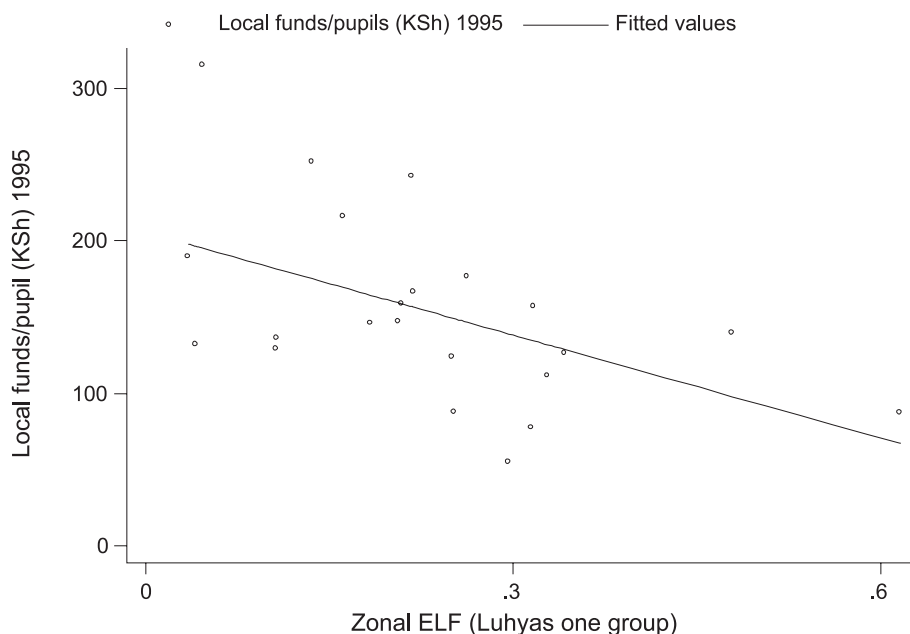


Fig. 3. Total local school funds per pupil (Kenyan Shillings) in 1995 (geographic zone average) versus residential ethnolinguistic fractionalization in the geographic zone (Pupil Questionnaire data).

negatively and significantly related to the level of local school funding per pupil (Table 5, regression 7). Fig. 4 indicates that ethnic diversity is not proxying for the proportion of Tesos in the area. Among predominantly Teso geographic zones, more ethnically diverse zones have lower average funding than homogeneous zones, and the same pattern holds for predominantly Luhya areas, generating a U-shaped relationship.

We next use our alternative diversity measure, ethnic diversity among all schools located within 5 km of the school, in a specification where standard errors are corrected to allow regression disturbance terms to be correlated across schools as a function of their physical distance (Conley, 1999). Local ethnic diversity is strongly associated with lower school funding, and the coefficient estimate on ethnic diversity is significantly different than 0 at 95% confidence (Table 5, regression 8). The point estimate is similar in magnitude to the analogous coefficient using the first diversity measure (regression 3), suggesting that the results are robust to this alternative data source. The estimated relationship between diversity and school funding remains large, negative, and statistically significant when geographic zone socioeconomic controls are included (regression 9). However, when geographic division indicators are included, the point estimate on school ELF remains negative but is no longer statistically significantly different than 0 (results not shown), indicating that the relationship between diversity and funding across geographic divisions is driving much of the overall relationship.

Following Vigdor (2002), we next estimate the relationship between diversity and funding while also including controls for the local population shares of the two main

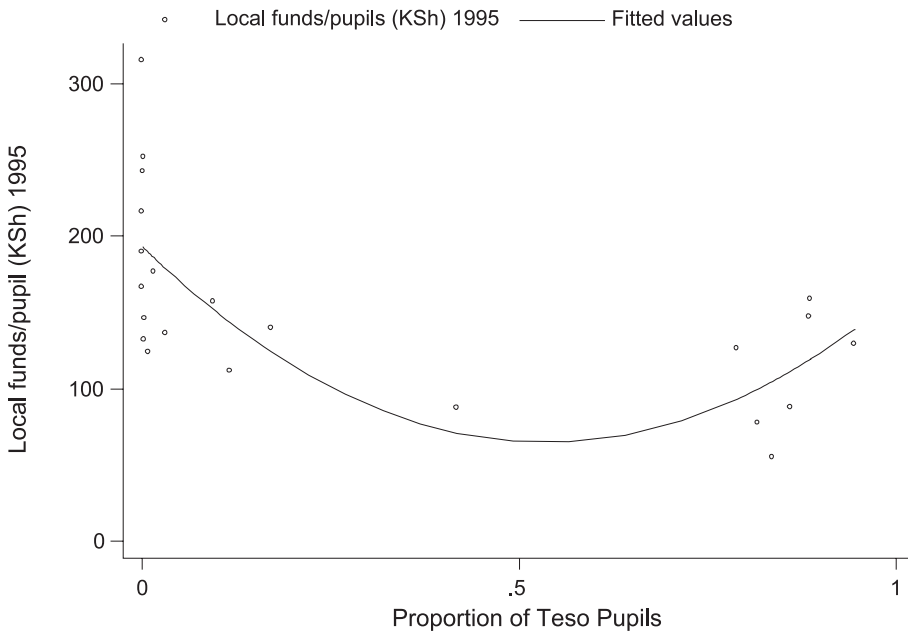


Fig. 4. Total local school funds per pupil (Kenyan Shillings) in 1995 (geographic zone average) versus proportion of Teso pupils residing in the geographic zone (Pupil Questionnaire data). Fig. 4 also contains the quadratic regression fit.

ethnic groups in this area, the Luhya and Teso, who together account for over 93% of the sample population (Luos, who account for the bulk of the remaining population, 5%, together with several smaller groups, comprise the omitted ethnic category). All of the main empirical results are robust to the inclusion of ethnic population shares, and none of the coefficient estimates on the ethnic share terms are statistically significant. The point estimates on zonal ELF are nearly identical without the ethnic population shares (-185.7 , standard error 77.9, Table 6 regression 1) and with the shares as controls (-189.1 , standard error 77.5, regression 2) and similarly for IV specifications without (-216.4 , standard error 88.4, regression 3) and with the shares (-208.4 , standard error 96.1, regression 4). As a further robustness check, we find nearly identical estimates of the impact of a linear diversity measure, one minus the proportion of the largest ethnic group in the geographic zone, on school funding without ethnic population controls (-222.0 , standard error 82.9, regression 5) and with the controls (-187.9 , standard error 81.3, regression 6) and once again similarly for the IV specifications (regressions 7 and 8).

The negative relationship between ethnic diversity and school funding is largely driven by *harambee* funding, but diversity is not significantly associated with school fees collected per pupil (Table 7A). Recall that *harambees* are public events in which primary school parents and other community members and ethnic peers are able to observe individual contributions to the school, so it is plausible that communities may more effectively impose social sanctions on those parents who do not pay *harambee*

Table 6
Ethnic diversity impacts, controlling for ethnic population shares

Explanatory variable	Dependent variable—total local primary school funds collected per pupil in 1995 (Kenyan Shillings)							
	(1) OLS	(2) OLS	(3) IV-2sls	(4) IV-2sls	(5) OLS	(6) OLS	(7) IV-2sls	(8) IV-2sls
Zonal ELF across tribes	−185.7** (77.9)	−189.1** (77.5)						
School ELF across tribes			−216.4** (88.4)	−209.4** (96.1)				
1−(Proportion largest ethnic group in zone)					−222.0** (82.9)	−187.9** (81.3)		
1−(Proportion largest ethnic group in school)							−264.8*** (93.9)	−239.7** (110.4)
Proportion Luhya pupils		−196.5 (393.8)		−136.4 (415.7)		−56.3 (370.5)		−87.0 (410.9)
Proportion Teso pupils		−247.3 (366.6)		−184.8 (386.2)		−120.1 (343.9)		−142.7 (381.4)
Root MSE	96.7	94.7	105.5	104.0	97.2	95.0	104.8	102.4
R ²	0.06	0.12	–	–	0.06	0.12	–	–
Number of schools	84	84	84	84	84	84	84	84
Mean dependent variable	152.6	152.6	152.6	152.6	152.6	152.6	152.6	152.6

Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. Observations are assumed to have independent error terms across geographic zones but not necessarily within zones. Ethnolinguistic fractionalization is defined as $1 - \sum_i (\text{proportion of ethnolinguistic group } i \text{ in the population})^2$. School ELF across tribes and the proportion of the largest ethnic group in the school consider Luhya as a single group. Geographic division indicators and socioeconomic controls are not included in any of the specifications in this table. The instrumental variable in regressions 3 and 4 is zonal ELF across tribes and in regressions 7 and 8 is 1−(proportion largest ethnic group in zone).

Table 7

Other primary school outcomes

Dependent variable	Coefficient estimate on zonal residential ELF across tribes (OLS)	Coefficient estimate on ELF across tribes among schools within 5 km (spatial OLS)	Number of schools	Mean dependent variable
<i>(A) Local school funding</i>				
<i>Harambee</i> donations collected per pupil, 1995 (Kenyan Shillings)	-157.1** (61.6)	-182.1** (68.5)	84	44.8
School fees collected per pupil, 1995 (Kenyan Shillings)	11.9 (35.2)	8.1 (64.6)	84	107.8
<i>(B) School facilities, inputs</i>				
Desks per pupil, 1996	-0.20** (0.08)	-0.31*** (0.08)	84	0.21
Pupil latrines per pupil, 1996	-0.007 (0.009)	-0.007 (0.013)	84	0.016
Classrooms per pupil, 1996	-0.016 (0.016)	-0.023* (0.013)	84	0.030
School-owned textbooks per pupil, 1996	-0.17 (0.13)	-0.27 (0.17)	84	0.34
Private texts (at home) per pupil, 1996	-0.03 (0.07)	-0.10 (0.09)	84	0.07
Number of other primary schools within 5km	-10.2*** (3.5)	-12.2*** (3.7)	84	14.5
<i>(C) Test scores</i>				
Average school score on 1996 NGO exams, Grades 3–8 (in standard deviations)	0.10 (0.52)	0.11 (0.52)	84	0.05
Socioeconomic controls (zonal averages)	Yes	Yes		

Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. Regression disturbance terms are clustered at the zonal level. Ethnolinguistic fractionalization is defined as $ELF = 1 - \sum_i (\text{proportion of ethnolinguistic group } i \text{ in the population})^2$. School ELF considers Luhya as a single group. The coefficient estimate on zonal residential ELF across tribes uses data from the 1996 Pupil Questionnaire. In these specifications, observations are assumed to have independent error terms across geographic zones but not necessarily within zones. The coefficient estimate on ELF across tribes among schools within 5 km uses 1996 Exam Namelist data. In these specifications, regression disturbance terms are allowed to be correlated across schools as a general function of their physical distance, using the estimation strategy developed in Conley (1999). Socioeconomic controls include the proportion of fathers in the geographic zone with formal sector employment, the proportion of pupils residing in the geographic zone with a latrine at home, the proportion of pupils whose households own livestock, and the proportion of pupils whose households cultivate a cash crop. The test score results also an additional explanatory variable, an indicator for having received financial assistance through another NGO program.

contributions than on those parents who do not pay school fees, and this stark difference between *harambee* and school fee results can be interpreted as further evidence in favor of the sanctions theory. An important constraint preventing schools in diverse areas from simply increasing school fees to make up for lower *harambee* contributions is the possibility of sizeable pupil transfers to other nearby schools in response to the higher fees.

5.3. Ethnic diversity and other school outcomes

Ethnic diversity is negatively and significantly related to the quality of school facilities and school textbook ownership (Table 7B). Zonal socioeconomic controls are included as explanatory variables in all specifications. The relationship between ethnic diversity and primary school facilities reflects the cumulative impact of past educational investments, and these results suggest that primary schools in ethnically diverse areas have worse facilities. The coefficient estimates on ethnic diversity are large, negative, and significantly different than 0 for desks per pupil. The drop associated with a change from complete ethnic homogeneity to median school-level ethnic diversity is over 20% of average desks per pupil. For the classrooms and pupil latrine regressions, the coefficient estimates on ethnic diversity are large and negative in all specifications although only sometimes significantly different than 0 at traditional confidence levels. In addition to their impact on learning, infrastructure investments directly enhance pupil utility; for example, classrooms with a sturdy roof shield children from rain, and latrine construction is important for public health—especially given the high rates of intestinal helminth (worm) infections in this area (Miguel and Kremer, 2004).

The stock of school textbooks per pupil is negatively related to ethnic diversity and nearly significantly different than 0 at 90% confidence in one specification (Table 7B). The relationship between ethnic diversity and the number of privately owned textbooks per pupil is also reported to explore the possibility of substitution from publicly provided to privately owned textbooks in diverse areas. The coefficient point estimates on ethnic diversity in this case are near 0 and not statistically significant. This result serves as an important specification check. Unobserved differences in the taste for education or in income across areas should affect both school outcomes and private textbook ownership, so the observed weak relationship between private textbook ownership and diversity strengthens the argument that unobservables are not driving the funding results.

There are significantly fewer primary schools in diverse areas, perhaps due to collective action problems with regard to establishing or maintaining primary schools (Table 7C). Recall that local ethnic diversity is not significantly associated with total pupil enrollment (Table 4), implying that the lower density of schools in diverse areas has not led to school crowding. Taken together, these results suggest that diverse areas are either less densely populated on average or have lower school enrollment rates, perhaps as a result of the lower density of schools, but unfortunately, the data do not permit us to rule out either possibility.

Finally, the estimated relationship between ethnic diversity and average school scores on 1996 NGO examinations for pupils in Grades 3 to 8 is close to 0 in both specifications, which may be surprising in light of the negative relationship between ethnic diversity and local school funding and inputs (Table 7C). However, other recent studies from rural western Kenya have found that average school examination scores respond little to increases in educational inputs, including textbooks, classroom construction, and school health programs (Glewwe et al., 1998; Miguel and Kremer, 2004), and there are similar findings across countries and from the United States (Hanushek and Kimko, 2000).

There are a number of plausible reasons why examination scores may respond little to educational inputs (the following draws on Kremer, 2003). First, the primary school curriculum in Kenya is oriented towards the strongest students most likely to continue on

to secondary school, and many students not at the top of the class may fail to benefit from additional textbooks because they have difficulty understanding them. Moreover, there is little evidence that the provision of additional inputs has any impact on teacher motivation or performance. Finally, many factors important in determining test scores, including children's innate ability and their home learning environment, are not directly affected by increased school spending, dampening impacts.

5.4. Ethnic diversity and threatened social sanctions

Ethnic diversity is associated with significantly fewer threatened sanctions against parents who have not contributed at harambees, paid fees, or participated in school projects or decision making (Table 8A). There is no difference in the total number of recorded administrative (i.e., nonsanction) items between homogenous and diverse schools, suggesting that the sanction result is unlikely to be the result of poor record keeping in

Table 8

School committee records and field officer observations

Dependent variable	Coefficient estimate on zonal residential ELF across tribes (OLS)	Coefficient estimate on ELF across tribes among schools within 5 km (spatial OLS)	Number of schools	Mean dependent variable
<i>(A) School Committee Records</i>				
School committee record items regarding sanctions or verbal pressure, 1997	-3.7** (1.6)	-4.2* (2.3)	84	3.2
School committee record items regarding administrative activities, 1997	5.7 (6.1)	6.2 (10.3)	84	18.9
Parent school meetings, 1997	-1.6 (1.1)	-1.3 (1.6)	84	3.4
<i>(B) Field Officer observations</i>				
Parent cooperation from 0 to 1 (reported by field officers), 1998	-0.77*** (0.26)	-0.84** (0.35)	84	0.49
Teacher motivation from 0 to 1 (reported by field officers), 1998	-0.39** (0.17)	-0.49* (0.29)	84	0.54
Socioeconomic controls (zonal averages)	Yes	Yes		

Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. Regression disturbance terms are clustered at the zonal level. Ethnolinguistic fractionalization is defined as $ELF = 1 - \sum_i (\text{proportion of ethnolinguistic group } i \text{ in the population})^2$. School ELF considers Luhyas a single group. The coefficient estimate on zonal residential ELF across tribes uses data from the 1996 Pupil Questionnaire. In these specifications, observations are assumed to have independent error terms across geographic zones but not necessarily within zones. The coefficient estimate on ELF across tribes among schools within 5 km uses 1996 Exam Namelist data. In these specifications, regression disturbance terms are allowed to be correlated across schools as a general function of their physical distance, using the estimation strategy developed in Conley (1999). Socioeconomic controls include the proportion of fathers in the geographic zone with formal sector employment, the proportion of pupils residing in the geographic zone with a latrine at home, the proportion of pupils whose households own livestock, and the proportion of pupils whose households cultivate a cash crop.

diverse areas. The sharp drop in sanctioning activity in ethnically diverse areas—areas with average levels of ethnic diversity have 31% fewer threatened sanctions than homogeneous areas—is consistent with the theoretical claim that heterogeneous communities are significantly less likely to threaten community sanctions.³⁶

Schools in ethnically diverse areas also exhibit significantly worse subjective ratings of both parent cooperation and teacher motivation in 1997, as assessed by NGO field workers (Table 8B). Schools in diverse areas also hold fewer parent meetings (to discuss school affairs) on average in diverse areas, although these effects are not statistically significant.

As noted above, we cannot entirely rule out all competing hypotheses for these sanction results. For instance, the school committee evidence is based on measures of sanction supply, but we have no explicit measure of sanction demand. Yet, there is suggestive evidence that sanction demand is unlikely to be lower in more diverse schools. The majority of sanctions in the school committee minutes concern the payment of school and *harambee* fees that, although a small percentage of total funding, are critical to the day-to-day functioning of the school, as they pay for chalk, nonteacher staff salaries, and school supplies. In fact, more diverse schools have much lower funding as well as lower levels of observed parent and teacher cooperation, implying that there is likely to be greater demand for sanctions in diverse schools. This should lead to a bias against our empirical finding of strongly negative ethnic diversity effects on recorded sanctions, suggesting that our estimate is, if anything, an underestimate of the true effect of diversity on sanctions.

Another concern is that school committee time spent discussing sanctions may also not be highly correlated with actual sanctions imposed; however, while we are unfortunately unable to directly test this in our data, many of the sanctions included in the school committee records are actually explicit sanctions in and of themselves, such as reading out aloud the names of parents who have not paid fees during PTA meetings. Finally, there might also be less need to employ sanctions in a community that does not intend to hold large *harambees* and similarly in areas where ethnic groups have divergent public good preferences. However, as we argue in Section 5.2, there is little evidence that preferences for education diverge sharply along ethnic lines in this part of Kenya.

5.5. *Ethnic diversity and well maintenance*

Community water well maintenance is negatively and significantly related to local ethnic diversity. Twenty-three wells were dropped from the sample due to missing data, leaving a sample of 667 wells (Table 9). Local ethnic diversity within 5 km of the well is negatively and significantly related to the likelihood that the KEFINCO well has “normal” water flow, using a probit specification (Table 10, regression 1). The point estimate on ethnic diversity implies that areas with average levels of local ethnic diversity are 6 percentage points less likely to have a functioning well than homogeneous areas. The

³⁶ We also estimated a two-stage least squares specification in which local diversity is an instrumental variable for threatened school committee social sanctions and find that each additional recorded sanction is associated with 39.2 additional shillings of funding per pupil (standard error 21.7) in the spatial OLS model. However, it is likely that the exclusion restriction fails to hold in this case, and recorded committee sanctions may only be a small fraction of total community pressure on defaulters, complicating the interpretation of these estimates.

Table 9
Well descriptive statistics

	Mean	Standard deviation	Observations
ELF across tribes for all primary schools within 5 km of the well, 1996 Exam Namelist data	0.23	0.14	667
Indicator variable for “normal” water flow from well, 2000–2001 survey	0.57	0.49	667
Indicator variable for no broken or missing well parts, 2000–2001 survey	0.66	0.48	667
Indicator variable for people in the area get water from another local well (if not normal water flow), 2000–2001 survey	0.32	0.47	196
Year well stopped functioning (if not normal water flow), 2000–2001 survey	1997.5	3.1	196
Latitude (degrees north), GPS data from 2000–2001 survey	0.36	0.17	667
Longitude (degrees east), GPS data from 2000–2001 survey	34.20	0.12	667

Data are from the 1996 ICS School and Pupil Questionnaires, 1996 Government Examination Namelists, and global positioning systems (GPS) readings taken by NGO field workers. Ethnolinguistic fractionalization is defined as $1 - \sum_i (\text{proportion of ethnolinguistic group } i \text{ in the population})^2$. School ELF across tribes and the proportion of the largest ethnic group in the school consider Luhyas a single group. Well parts include the pump handle, the cover and base, and the external and internal pipes and seals.

effect of ethnic diversity is not quite significant at 90% confidence (regression 2) when regression disturbance terms are correlated across schools as a general function of their physical distance, using the spatial estimation method in Conley (1999). The basic results are robust to examining ethnic diversity within a larger radius around each well and varying the kernel used in the spatial standard error correction method although the

Table 10
Ethnic diversity and well maintenance

Explanatory variables	Dependent variable					
	Indicator variable for “normal” water flow from well		Indicator variable for no broken or missing well parts		Indicator variable for people in the area get water from another local well (if the KEFINCO well does not have normal water flow)	
	(1) Probit	(2) Spatial OLS	(3) Probit	(4) Spatial OLS	(5) Probit	(6) Spatial OLS
ELF across tribes among schools within 5 km	−0.26* (0.14)	−0.26 (0.17)	−0.25* (0.13)	−0.25 (0.22)	−0.73** (0.30)	−0.72* (0.36)
Number of wells	667	667	667	667	196	196
Root MSE	–	0.49	–	0.47	–	0.46
Mean dependent variable	0.57	0.57	0.66	0.66	0.32	0.32

Huber robust standard errors in parentheses. Observations are assumed to have independent error terms across geographic zones but not necessarily within zones in regressions 1, 3, and 5. Regression disturbance terms are allowed to be correlated across schools as a general function of their physical distance, using the estimation strategy developed in Conley (1999), in regressions 2, 4, and 6. Significant at 90% (*), 95% (**), 99% (***) confidence. Geographic indicators are included for six (of the seven) geographic divisions.

magnitude of the coefficient estimate on diversity falls when geographical division indicators are included (results not shown).

The inability to maintain wells in diverse areas is likely to be driving this result (rather than the possibility that there is simply less ground water in these areas) inasmuch as ethnically diverse areas also have more wells with missing or broken parts (Table 10, regressions 3 and 4), which the Water Committee interview evidence suggests is often due to theft. There is a significant negative relationship between local ethnic diversity and the use of another local well in areas where the KEFINCO well does not have “normal” water flow, which restricts the sample to 196 wells (regressions 5 and 6).

These results are limited by a lack of data on the specific date of well construction, and it remains possible that poor maintenance is the result of age rather than a failure of collective action (although note that there is no evidence to suggest that the date of well construction varies systematically with ethnic diversity). The empirical results on well maintenance on their own are also not as statistically significant or robust as the school funding results. Even with these caveats in mind, the results suggest that ethnically diverse communities both failed to maintain their KEFINCO wells and also failed to construct additional wells and, thus, that ethnic diversity has implications for collective action beyond the school setting.

6. Conclusion

To summarize, in rural western Kenya, ethnic diversity is associated with sharply lower local school funding through voluntary fundraisers (harambees), worse school facilities, fewer recorded community social sanctions, and there is suggestive evidence of worse well maintenance as well, despite the fact that diverse areas are largely similar to homogeneous areas along a range of socioeconomic and other characteristics. The finding that school committee sanctions were weaker in diverse communities has potentially far-reaching implications inasmuch as a variety of informal collective action, contracting, and credit market outcomes are thought to rely on effective sanctions in less developed countries.

Although the results highlight difficulties raised by decentralized public good provision in less developed countries, it is theoretically uncertain whether centralized local public goods funding at the regional or national government levels would lead to better outcomes. Central governments in many less developed countries are notorious for underproviding recurrent expenses, like textbooks for schools and road maintenance, and the Kenyan government has been singled out as a particularly egregious example of this failure (Easterly, 2001). Further centralization of school funding could also lead to more regional and ethnic favoritism in the allocation of national government funds, which is common in Kenya and other African countries.³⁷ Alternatively, central governments could subsidize the creation of additional primary schools or wells in diverse areas (or mandate segregated

³⁷ Barkan and Chege (1989) study the allocation of national road construction funds in Kenya during the 1970s and 1980s and find that the proportion of road funds allocated to the ethnic homeland of former Kenyan President Jomo Kenyatta fell from 44% in 1979–1980 to 16% in 1987–1988 after Kenyatta’s Kikuyu ethnic group lost its dominant position in the central government, while the ethnic homeland of Kenyan President Daniel Arap Moi, who replaced Kenyatta, saw its share of road funds rise from 32% to 57% during the same period. Refer to Easterly and Levine (1997) for related examples.

schools) to facilitate sorting into ethnically homogeneous groups and thereby avoid the efficiency cost of diversity. However, promoting ethnic separatism is likely to have extremely deleterious long-term social and political implications.

A more attractive approach for addressing the efficiency costs of ethnic diversity lies in designing policies and institutions that promote successful cooperation across ethnic groups. This paper points to the important role that social sanctions may play in sustaining local public goods provision in less developed countries, mechanisms that are most effectively applied within social groups. A long-standing theme among observers of economic development is that the formation of meaningful economic linkages extending beyond the immediate community is a necessary precondition for modern economic growth (Simmel, 1971 [1908]; Greif, 1993; Woolcock, 1998). The design of policies that build social capital across ethnic groups, perhaps including central government nation-building efforts or power-sharing arrangements within organizations, remains a poorly understood yet promising research agenda with critically important implications for economic development in sub-Saharan Africa and elsewhere.³⁸ Several African countries, most notably Tanzania (Barkan, 1994), have engaged in concerted nation-building efforts during the postindependence period, but to our knowledge, few systematic empirical evaluations have been conducted of the impact of these policies on interethnic cooperation.³⁹ We believe that exploring how trust, cooperation, and social capital are constructed and maintained will be a fruitful line of research for economists in the future.

Acknowledgements

George Akerlof, Alberto Alesina, Abhijit Banerjee, Caroline Hoxby, Lawrence Katz, Michael Kremer, Gerard Roland, Paul Schultz, numerous seminar participants, and two anonymous referees have provided valuable comments. We are indebted to the staff of ICS Africa, Sylvie Moulin, Elizabeth Beasley, and especially Michael Kremer for their generosity. Gugerty gratefully acknowledges financial support from the MacArthur Foundation and the World Bank and the Government of Denmark through the Social Capital Initiative. Miguel gratefully acknowledges financial support from the US National Science Foundation Graduate Fellowship and SGER-#0213652, the Harvard Weatherhead Center for International Affairs and the MacArthur Foundation. All errors are our own.

Appendix A. Theory appendix

Remark 1. Proof: This follows directly from Eq. (2).

³⁸ Horowitz (1985) is the seminal discussion of ethnic conflict, and Carroll and Carroll (2000) review the current state of this literature. Varshney (2000) claims that the local density of interethnic organizations determines the extent of communal violence in India and asserts that in the long-run, these are politically constructed.

³⁹ Miguel (2004) is one recent attempt in this direction.

Remark 2. Proof:

- (a) This follows from Assumption 4 and from the concavity of b (Assumption 1).
- (b) Follows from $b' > 0$.
- (c) This follows if we take b^{-1} of both sides of $b(n_i + p_{-i}^e) - c \geq b(p_{-i}^e)$ and set $p_{-i}^e = 0$.
- (d) This follows if we take b^{-1} of both sides of $b(n_i + p_{-i}^e) - c \geq b(p_{-i}^e)$ and set $p_{-i}^e = n_{-i}$.

The proof that $n^* > 0$ follows from the assumptions that $c > 0$, $b(0) = 0$, and $b' > 0$.

The proof that $n^{**} \geq n^*$ follows directly from the concavity of b . b concave implies that b^{-1} convex. The proof is by contradiction. Assert that $n^{**} < n^*$ or equivalently that $1 - b^{-1}(b(1) - c) < b^{-1}(c)$, which is equivalent to $b^{-1}(b(1) - c) + b^{-1}(c) > 1$. We know that $1 = b^{-1}(b(1) - c + c)$, hence, this implies that $b^{-1}(b(1) - c) + b^{-1}(c) > b^{-1}(b(1) - c + c)$. This violates the weak convexity of b^{-1} and is a contradiction. Therefore, $n^{**} \geq n^*$.

Proposition 1. Proof: For $n_B < 1 - n^{**}$, $P = n_A = 1 - n_B$. Therefore, $(dP/dn_B) = -1$, where n_B is the measure of ethnic diversity. The analogous result holds for (dS/dn_B) .

For $n_B > 1 - n^{**}$, Remark 2(c) and (d) together imply the result.

Appendix B. Empirical appendix

Table A1
Selection into NGO assistance program in 1995

Explanatory variable	1995 Pupils enrollment (District Educational Office records)		1995 Average government exam result, Grades 6–8	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Indicator for selection into NGO assistance program	-99.6*** (17.1)	-115.6** (43.2)	-72.0*** (10.7)	-53.2** (22.4)
Zonal residential ELF across tribes in 1996		73.3 (110.2)		116.4 (89.1)
(Indicator for selection into NGO assistance program)* (zonal residential ELF across tribes in 1996)		65.0 (128.5)		-89.5 (73.2)
R^2	0.06	0.07	0.10	0.11
Root MSE	182.3	182.5	102.5	102.0
Number of observations	300	300	300	300
Mean of dependent variable (standard deviation)	379.8 (187.9)	379.8 (187.9)	871.7 (107.7)	871.7 (107.7)

Huber robust standard errors in parentheses. Data are from official District Education Office records. One hundred of the 331 primary schools in Busia and Teso districts were selected for NGO assistance. There are fewer than 331 observations inasmuch as not all schools have Grades 6, 7, 8 classes, and these schools having missing test scores.

Appendix C. Local ethnicity measures

Measure 1 of local ethnic diversity is the ethnolinguistic fractionalization among surveyed students in sample schools residing in the corresponding administrative zone. The information on ethnicity was self-reported by pupils on the 1996 Pupil Questionnaire, which surveyed students in Grades 6 through 8. Although the pupil questionnaire data do not contain ethnic information for pupils below Grade 6, they do indicate that drop-out rates are similar across ethnic groups in Grades 6 to 8, suggesting that differential school participation across groups is unlikely to significantly alter measured diversity.

Measure 2 of local ethnic diversity is ethnolinguistic fractionalization calculated among pupils attending all primary schools located within 5 km of either the primary school (when a school outcome is the dependent variable) or the well (when well maintenance is the dependent variable). The principal advantage of this alternative measure, constructed from Government Exam Namelists, is that it includes information for nearly all primary schools in Busia and Teso districts (326 of 337 schools) rather than just the 100 schools in the NGO assistance program. However, a drawback of these data is that pupil ethnic affiliation was assigned by NGO staff based upon children's names rather than being determined by pupils themselves. We are grateful to Charles Asoka and Maureen Wechuli for assigning pupil ethnicity to these thousands of pupil names.

The assignment of ethnicity by NGO staff is likely to introduce some error into measured school ethnic diversity inasmuch as certain surnames are common across ethnic groups in this area, and names and ethnic affiliation do not always match up. It is particularly difficult to distinguish Luhya and Luo children inasmuch as many Luhyas possess Luo surnames. Approximately 19% of all pupils in the examination name list sample have such "ambiguous" Luo names. Pupils with ambiguous names are assigned with Luhya and Luo ethnicity in proportion to their group's representation within the geographic zone in the 1996 Pupil Questionnaire sample. In other words, pupils with ambiguous names are more likely to be assigned Luo ethnicity in areas in which the pupil survey data indicate that there are more Luos. Despite this possible noise in the data, these two measures of local ethnic diversity are highly correlated (the correlation coefficient is 0.7).

Appendix D. Interview data and school committee records

(1) 2000 Primary school Headmaster interviews

Edward Miguel conducted the 12 Headmaster interviews in June and July 2000. Six schools in Funyula Division, a primarily Luhya area, were interviewed and six schools in Angurai division, a primarily Teso area. The headmasters were asked about school funding levels, mechanisms for collecting school fees (including informal social sanctions), ethnic and clan relations in the area, and pupil transfer patterns across schools.

(2) 2001 Well committee interviews

Gideon Osoma and Franklyn Makokha, ICS Africa field officers, conducted the thirty-three interviews between 5–15 November 2001 among a stratified random sample of

KEFINCO wells in Angurai Division (Teso District), Butula Division (Busia District), Nambale Division (Busia District), and Budalangi Division (Busia District).

(3) 1997 Primary school committee records

The primary school committee consists of 13 individuals. Parents of students in each grade directly elect a representative from among their number, producing nine representatives: eight (one each) for Grades 1 through 8 and one representative from the nursery class. Voting is usually carried out democratically by a show of hands. Serving on the school committee carries social prestige and status within the community, and elected representatives are often individuals of some prominence in the community. The committee then elects two of its elected members to serve as treasurer and chairman, while the headmaster always serves as school committee secretary. Four additional members of the school committee are externally appointed, two by the school's sponsor (often a church) and two by the District Education Office although these individuals are often less involved in school activities.

Table A2

Proportion of all school committee minute items in each coding category

Coding category from above	Proportion
<i>School committees members (example quotes in parentheses)</i>	
(1) Explicit sanctions (members who do not supervise school projects to be disciplined)	0.003
(2) Verbal pressure (school committee must pay <i>harambee</i> contributions)	0.008
(3) Exhortation about behavior (all members must attend all meetings)	0.033
(4) Contribution or fee setting (each member to donate desk to school)	0.010
<i>Parents</i>	
(5) Explicit sanctions (parents not paying fees to be visited by the chief)	0.034
(6) Verbal pressure (all parents to pay fee balances immediately)	0.026
(7) Exhortation about behavior (parents should discipline pupils who are misbehaving)	0.008
(8) Contribution or fee setting, regular (activity fee to be 50 Shillings)	0.066
(9) Contribution or fee setting, special event (each parent to pay 20 Shillings towards <i>harambee</i>)	0.044
<i>Teachers</i>	
(10) Explicit sanctions (teachers who miss class will receive letter from the Headmaster)	0.002
(11) Verbal pressure (teachers to contribute to <i>harambee</i> immediately or face consequences)	0.002
(12) Exhortation about behavior (teachers to uphold school discipline)	0.009
(13) Contribution or fee setting (teachers to participate in construction by carrying stones)	0.002
<i>General</i>	
(14) Exhortations about behavior (the school should unite and community members should stop their gossiping)	0.013
(15) School projects (decisions or discussion about projects, i.e., starting a project, project administration, content; excludes decisions about contributions for projects)	0.133
(16) General school administration (school to hire watchman)	0.462
(17) <i>Harambee</i> organization (general discussion of <i>harambee</i> organization)	0.078
(18) Elections	0.060
(19) Other	0.000
(20) Lack of meeting quorum (meeting disbanded due to lack of quorum)	0.001
(21) Conflict leading to meeting termination (conflict reported between school committee members, parents, and/or teachers)	0.004

Table A3

Sample school committee minutes (for School ID #131)

Meeting number	Agenda item (quotes from school records)	Code
1	School committee and head teachers to work closely to improve academic standards	3
	Parent to avail textbooks to ease the teachers' work	8
	School committee to ensure that the projects are completed if any is begun	3
	School committee members to provide poles for classroom construction	4
	The committee agreed to buy the school the new teaching syllabus	16
	The head teacher to form a committee responsible for running the preprimary section	16
	The head teacher to be given money to go to Eregi School for the PRISM course	16
	Members of the committee are to pay their school funds promptly	3
	Parents with outstanding balances should be followed up for payment	5
	2	Iron sheets for classroom construction to be bought
Syllabus to be bought as a priority		16
Members should cooperate with the head teacher to enable the school to come up academically		3
Every child to pay Ksh 5 towards the district library books		8
Priority areas to be given a fair chance in the development plan		16
Renovation of classrooms to start immediately		15
All parents are to pay Ksh 250/=towards <i>harambee</i>		9
Parents are to assess in a brick making project initiated by the head teacher		9
The committee agreed to help the head teacher with his costs for the PRISM course.		16
3		Parents to be informed about payments for the <i>harambee</i>
	Classroom to be completed by January 1998	15
	For 1998, the school should buy books, desks, build toilets, and build a classroom	15
	The school's debts should be paid when funds become available	16

D.1. Coding categories for 1997 school committee minutes

Complete records of all school committee meetings were reviewed, and the coding scheme given below was used by the authors to code all available meeting minute items. Table A2 presents the proportions of meeting items in each coding category, and Table A3 presents complete school committee meeting records for a representative school. Examples of meeting minutes items that fall within this category are provided in parentheses for each code. Decisions falling into the following coding categories were considered to be threatened sanctions: Codes 1, 2, 3, 5, 6, 7, 10, 11, 12, or 14. Other categories are classified as administrative records.

References

- Akerlof, George, 1980. A theory of social custom, of which unemployment may be one consequence. *Quarterly Journal of Economics* 85, 749–775.
- Alesina, Alberto, LaFerrara, Eliana, 2000. Participation in heterogeneous communities. *Quarterly Journal of Economics* 115 (3), 847–904.
- Alesina, Alberto, Baqir, Reza, Easterly, William, 1999. Public goods and ethnic divisions. *Quarterly Journal of Economics* 114 (4), 1243–1284.
- Angrist, Joshua, Krueger, Alan, 1999. Empirical strategies in labor economics. *Handbook of Labor Economics* vol. 3A. Elsevier Science, North-Holland.
- Barkan, Joel D., 1994. Divergence and convergence in Kenya and Tanzania: pressures for reform. In: Barkan, Joel D. (Ed.), *Beyond Capitalism vs. Socialism in Kenya and Tanzania*. Lynne Rienner Publishers, Boulder.

- Barkan, Joel D., Chege, Michael, 1989. Decentralising the state: district focus and the politics of reallocation in Kenya. *Journal of Modern African Studies* 27 (3), 431–453.
- Bates, Robert H., 1999. Ethnicity, capital formation, and conflict, Harvard Weatherhead Center for international Affairs Working Paper.
- Berry, Sarah, 1993. *No Condition is Permanent: The Social Dynamics of Agrarian Change in Sub-Saharan Africa*. The University of Wisconsin Press, Madison.
- Besley, Timothy, Coate, Stephen, 1995. Group lending, repayment incentives and social collateral. *Journal of Development Economics* 46 (1), 1–18.
- Besley, Timothy, Coate, Stephen, Loury, Glenn, 1993. The economics of rotating savings and credit associations. *American Economic Review* 83 (4), 792–810.
- Carroll, Barbara Wake, Carroll, Terrance, 2000. Accommodating ethnic diversity in a modernizing democratic state: theory and practice in the case of Mauritius. *Ethnic and Racial Studies* 23 (1), 120–142.
- Chamberlin, John R., 1974. Provision of collective goods as a function of group size. *American Political Science Review* 68, 707–716.
- Coleman, James, 1989. *Foundations of Social Theory*. Harvard University Press, Cambridge.
- Collier, Paul, Gunning, Jan Willem, 1999. Explaining African economic performance. *Journal of Economic Literature* 37 (1), 64–111.
- Community Water Supply Management Project (CWSP). (2000). “Report of the Status of Piped Schemes in Western Province”, Kakamega, Kenya, mimeo.
- Conley, Timothy, 1999. GMM estimation with cross sectional dependence. *Journal of Econometrics* 92 (1), 1–45.
- Costa, Dora L., Kahn, Matthew E., 2002. Civic engagement and community heterogeneity: an economist’s perspective. Conference of Social Connectedness and Public Activism, Harvard University, May 2002.
- Des Forges, Alison, 1999. *Leave None to Tell the Story: Genocide in Rwanda*. Human Rights Watch, New York.
- DiPasquale, Denise, Glaeser, Edward L., 1999. Incentives and social capital: are homeowners better citizens? *Journal of Urban Economics* 45 (2), 354–384.
- Drazen, Allan, 2000. *Political Economy in Macroeconomics*. Princeton University Press.
- Easterly, W., 2001. *The Elusive Quest for Growth*. MIT Press, Cambridge.
- Easterly, William, Levine, Ross, 1997. Africa’s growth tragedy: policies and ethnic divisions. *Quarterly Journal of Economics* 112 (4), 1203–1250.
- Esteban, Joan-Maria, Ray, Debraj, 1994. On the measurement of polarization. *Econometrica* 62 (4), 819–851.
- Esteban, Joan-Maria, Ray, Debraj, 2001. Collective action and the group size paradox. *American Political Science Review* 95 (3), 663–672.
- Fearon, James D., Laitin, David D., 1996. Explaining interethnic cooperation. *American Political Science Review* 90 (4), 715–735.
- Fehr, Ernst, Gächter, Simon, 2000. Cooperation and punishment in public goods experiments. *American Economic Review* 90 (4), 980–994.
- Glewwe, Paul, Michael Kremer, Sylvie Moulin. (1998). Textbooks and test scores: evidence from a prospective evaluation in Kenya”, mimeo, MIT.
- Goldin, Claudia, Lawrence F. Katz. (1997). “Why the United States led in education: lessons from secondary school expansion, 1910 to 1940”, NBER Working Paper 6144.
- Government of Kenya, 1965. *Kenya Population Census 1962*. Central Bureau of Stats, Nairobi.
- Government of Kenya, 1986. In: Were, G. (Ed.), *Kenya Socio-Cultural Profiles: Busia District*. Ministry of Planning and National Development, Nairobi.
- Government of Kenya, 1994. *Kenya Population Census 1989*. Central Bureau of Stats, Nairobi.
- Greif, Avner, 1993. Contract enforceability and economic institutions in early trade: the Maghribi traders’ coalition. *American Economic Review* 83 (3), 525–548.
- Gugerty, Mary Kay. (2000). “You can’t save alone: testing theories of rotating savings and credit organizations,” unpublished manuscript, Harvard Program in Political Economy and Government.
- Gugerty, Mary Kay, Edward Miguel (2000). “Community participation and social sanctions in Kenyan schools,” unpublished manuscript, Harvard Department of Economics.
- Hanushek, Eric A., Kimko, Dennis D., 2000. Schooling, labor force quality, and the growth of nations. *American Economic Review* 90 (5), 1184–1208.
- Horowitz, Donald L., 1985. *Ethnic Groups in Conflict*. University of California Press, Berkeley.

- IRC. (2000). "Community Management of Water Supplies in Kenya, Uganda, and Tanzania." Report of the IRC Staff. (www.irc.nl)
- Khwaja, Asim Ijaz. (2000). "Can good projects succeed in bad communities?: collective action in the Himalayas", unpublished manuscript, Harvard University.
- Knack, Stephen, Keefer, Philip, 1997. Does social capital have an economic payoff? A cross-country investigation. *Quarterly Journal of Economics* 112 (3), 1252–1288.
- Kremer, Michael, 2003. Randomized evaluations of educational programs in development countries: some lessons. *American Economic Review, Papers and Proceedings* 93 (2), 102–106.
- Krueger, Alan B., Mikael Lindahl. (2000). Education for growth: why and for whom? NBER Working Paper #7591.
- Mauro, Paulo, 1995. Corruption and growth. *Quarterly Journal of Economics* 110, 681–712.
- Miguel, Edward. 2001. "Ethnic diversity and school funding in Kenya," UC-Berkeley CIDER WP #C01-119.
- Miguel, Edward, 2004. Tribe or nation? Nation-building and public goods in Kenya versus Tanzania. *World Politics* 56 (3, April).
- Miguel, Edward, Kremer, Michael, 2004. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica* 72 (1), 159–217.
- Morgan, W.T.W., Shaffer, N. Manfred, 1966. Population of Kenya, Density and Distribution: A geographical introduction to the Kenya Population Census, 1962. Oxford University Press, Nairobi.
- Ndegwa, S., 1997. Citizenship and ethnicity: an examination of two transition moments in Kenyan politics. *American Political Science Review* 91 (3), 599–616.
- Platteau, J.-P., 2000. Institutions, Social Norms, and Economic Development. Routledge Press.
- Poterba, James, 1997. Demographic structure and the political economy of public education. *Journal of Policy Analysis and Management* 16 (1), 48–66.
- Putnam, R., 1993. Making Democracy Work. Princeton University Press.
- Schelling, Thomas C., 1978. *Micromotives and Macrobehavior*. W.W. Norton and Company.
- Shipton, Parker. (1985). "Land, credit, and crop transactions in Kenya." Ph.D Thesis, Columbia University.
- Simmel, Georg, 1971 [1908]. Group expansion and the development of individuality. In: Levine, D. (Ed.), *Georg Simmel: On Individuality and Social Forms*. University of Chicago Press.
- Varshney, Ashutosh, 2000. *Civic Life and Ethnic Conflict: Hindus and Muslims in India*. Yale University Press.
- Vigdor, Jacob, 2002. Interpreting ethnic fragmentation effects. *Economics Letters* 75 (2), 271–276.
- Vigdor, Jacob, 2004. Community composition and collective action: analyzing initial mail response to the 2000 census. *Review of Economics and Statistics* 86 (1), 303–312.
- Wade, Robert, 1994. *Village Republics: Economic Conditions for Collective Action in South India*. Institute for Contemporary Studies Press, San Francisco.
- Were, Gideon, 1967. *A History of the Abaluyia of Western Kenya*. East African Publishing House, Nairobi.
- Were, Gideon, editor. (1986). *Socio-cultural profile of Busia district*. Government of Kenya: Ministry of Planning and National Development.
- Wilson, L.S., 1992. The harambee movement and efficient public good provision in Kenya. *Journal of Public Economics* 48 (1), 1–19.
- Woolcock, Michael, 1998. Social capital and economic development: toward a theoretical synthesis and policy framework. *Theory and Society* 27 (20), 151–208.
- Wydick, B., 1999. Can social cohesion be harnessed to repair market failures? evidence from group lending in Guatemala. *Economic Journal* 109 (457), 463–475.