ETHNIC DIVERSITY, MOBILITY AND SCHOOL FUNDING: THEORY AND EVIDENCE FROM KENYA

BY

TED MIGUEL*
Department of Economics
Harvard University
Cambridge, MA 02138, USA
miguel@fas.harvard.edu

First draft: March 1998
This draft: June 1999

DEDPS No.14
June 1999

The Development Economics Discussion Paper Series
The Suntory Centre
Suntory and Toyota International Centres for Economics
And related disciplines
London School of Economics
Houghton Street
LONDON WC2A 2AE
Tel: 0171 955 6698

---

* I am grateful to Alberto Alesina, Abhijit Banerjee, Timothy Besley, Miguel Braun, Mary Kay Gugerty, Caroline Minter Hoxby, Lawrence Katz, Michael Kremer, Paul Schultz, and participants in the Harvard Development and Labor/Public Finance Workshops, the 1998 NEUDC Conference at Yale University, the 1999 African Studies Conference at Boston University, University of Namur, the London School of Economics, Bocconi University, and the University of Oxford have provided valuable comments. I am especially indebted to ICS Africa for their generosity. I gratefully acknowledge financial support from the U.S. National Science Foundation and the MacArthur Foundation. All errors are my own.

ETHNIC DIVERSITY, MOBILITY AND SCHOOL FUNDING; THEORY AND EVIDENCE FROM KENYA

BY

TED MIGUEL
Harvard University
Department of Economics
Cambridge, MA 02138, USA

This draft: June 1999
First draft: March 1998

ABSTRACT

This paper explores the relationship between ethnic diversity and local school funding in Kenyan primary schools. The empirical results paint a picture of pervasive local collective action problems in ethnically diverse Kenyan primary schools. Local ethnic diversity is robustly associated with lower local school funding, less parental involvement in school functions, and fewer desks, latrines, and classrooms per pupil in ninety-seven rural Kenyan primary schools. However, local ethnic diversity is not related to average school test score performance in these schools. The theory examines the school choice and school funding process when student mobility between schools is limited by land market imperfections, and some aspect of educational quality – such as headmaster competence – differs markedly across schools. The implications for human capital accumulation, economic growth, and local collective action are discussed, especially for Africa.

Keywords: Education, Ethnicity, Kenya, Mobility, Public Goods

JEL Classification: H41, H71, I22, O15

Ethnic Diversity, Mobility and School Funding: Theory and Evidence From Kenya

Ted Miguel*
Harvard University

First Version: March 1998
This Version: June 1999

Abstract

This paper explores the relationship between ethnic diversity and local school funding in Kenyan primary schools. The empirical results paint a picture of pervasive local collective action problems in ethnically diverse Kenyan primary schools. Local ethnic diversity is robustly associated with lower local school funding, less parental involvement in school functions, and fewer desks, latrines, and classrooms per pupil in ninety-seven rural Kenyan primary schools. However, local ethnic diversity is not related to average school test score performance in these schools. The theory examines the school choice and school funding process when student mobility between schools is limited by land market imperfections, and some aspect of educational quality – such as headmaster competence – differs markedly across schools. The implications for human capital accumulation, economic growth, and local collective action are discussed, especially for Africa.

---

* Department of Economics, Harvard University, Cambridge MA 02138, USA; miguel@fas.harvard.edu. Alberto Alesina, Abhijit Banerjee, Timothy Besley, Miguel Braun, Mary Kay Gugerty, Caroline Minter Hoxby, Lawrence Katz, Michael Kremer, Paul Schultz, and participants in the Harvard Development and Labor/Public Finance Workshops, the 1998 NEUDC Conference at Yale University, the 1999 African Studies Conference at Boston University, University of Namur, the London School of Economics, Bocconi University, and the University of Oxford have provided valuable comments. I am especially indebted to ICS Africa for their generosity. I gratefully acknowledge financial support from the U.S. National Science Foundation and the MacArthur Foundation. All errors are my own.

1 Introduction

Recent economic research suggests that ethnically diverse societies may be prone to slow economic growth.[1] This issue is particularly important for sub-Saharan African countries, which as a group are both the poorest and the most ethnically diverse countries. Easterly and Levine [1997] find that national ethnic diversity may explain much of cross-country differences in public policies, political instability, and other factors associated with economic growth from 1960 to 1990, and conclude that ethnic diversity has been a principal cause of African economic failure in the post-colonial period.[2]

This paper explores the relationship between ethnic diversity and local primary school funding in a poor, rural and ethnically diverse region of western Kenya. Important public primary school investments – including books, desks, and classrooms – are locally funded in Kenya. The empirical results of this paper may contribute to the current debate on the sources of poor African economic performance, since primary school quality is an important determinant of human capital accumulation in poor countries, and human capital accumulation is associated with subsequent economic growth.[3]

A lack of cooperation between ethnic groups may in theory lead them to undervalue common public enterprises like schools, or be unwilling to fund projects that will predominantly benefit other groups. Ethnic heterogeneity may affect many important collective action situations, including involvement in community organizations, and political participation. Exploring the role that ethnic diversity plays in a particular collective action problem – in this case, the funding of primary schools – may contribute to understanding its impact in related settings.

The estimation strategy addresses the possibility of endogenous pupil mobility between schools. The hierarchy of geographic areas in Kenya in terms of size is province (largest area), district, division,

[1] Alesina and Drazen [1991] address the paralysis of divided governments in a "war of attrition" framework. Benhabib and Rustichini [1996] discuss how conflict between social groups may lead to a political growth trap. Alesina and Spolaore [1997] examine the role of cultural heterogeneity in nation formation and disintegration.
[2] Collier and Gunning [1999] and Sachs and Warner [1997] discuss other explanations for African economic failure.
[3] Barro [1997] and Krueger and Lindahl [1998] estimate the relationship between schooling and economic growth.

and zone (smallest area), and the 97 schools in the current study are located within two geographic districts, eight divisions, and 22 geographic zones, in Western province, Kenya. Zonal ethnic diversity is used as an instrumental variable for school-level ethnic diversity, since pupils' mobility between schools (generally by foot) makes school-level ethnic diversity endogenous. Indicator variables for the geographic divisions are included to control for variation in both income and the taste for education across regions, and the relationship between ethnic diversity and school funding is then identified using the variation in school outcomes across zones within a geographic division. Ethnic composition is stable in this area of Kenya, being tied to traditional group lands, and is plausibly exogenous.

Empirically, local ethnic diversity is robustly associated with sharply lower local school funding, less parental participation in school affairs, and fewer desks, latrines, and classrooms per pupil in ninety-seven rural Kenyan primary schools. The drop in local school funding associated with median school ethnic diversity is approximately 30 percent of average local school funding. The effect of ethnic diversity on local school funding and educational investments is robust to the inclusion of a range of socioeconomic, demographic, and school quality controls.

Other researchers have found a similar relationship between ethnic diversity and school funding in the United States. Alesina, Baqir, and Easterly [1997] find that high levels of ethnic diversity are associated with up to 25 percent lower local school funding, and lower funding for other public goods, in U.S. municipalities. Poterba [1996] finds that an increase in the share of ethnic minorities in the school-aged child population from zero to fifty percent – holding other factors constant – is associated with a 27 percent drop in local school spending per child. Goldin and Katz [1997] argue that public secondary schooling expanded slowly in ethnically diverse U.S. school districts from 1910 to 1940.

The model presented in Section 3 is related to the recent theoretical literature examining school choice and funding decisions (Benabou [1993]; Durlauf [1996]; Epple and Romer [1991]; Fernandez and Rogerson [1996]). These existing models assume costless mobility between jurisdictions and active land markets, unrealistic assumptions for many less developed countries; focus on income and

ability differences as driving forces for sorting into separate communities; and feature school segregation as the only stable sorting equilibrium.

The theoretical goal of this paper is to model the primary school funding and school choice process in rural Africa. Three assumptions in particular distinguish the model from previous theoretical work. First, pupils are locally mobile by foot but not by residence due to frictions in rural land markets and the cost of moving across ethnic boundaries. Second, ethnic diversity – rather than income inequality – is the source of agent heterogeneity, and the dimension along which pupils sort. Third, school educational quality is determined both by endogenous pupil composition and by an exogenous quality component representing the important role that headmasters play in the success of rural primary schools.

Since distinct ethnic groups have different educational instruction preferences – such as the language in which classes are conducted – households may face a trade-off between choosing a school for its educational quality, or for its ethnic composition. Ethnic integration or segregation may result depending on the size of these opposing effects. The model's empirical implications for the relationship between ethnic diversity, pupil mobility, and local school funding levels are tested using the data from rural western Kenya.

The paper is structured as follows: the setting in western Kenya is described in Section 2, the theory is presented in Section 3, and the empirical results discussed in Section 4. Section 5 concludes. Appendix A contains all tables, and Appendix B is the mathematical appendix.

2 The setting

2.1 Primary schools in Busia and Teso districts, Western province, Kenya

Kenya's Busia and Teso districts are agricultural regions in Kenya's Western province, bordered by Uganda to the west and Lake Victoria to the south. Busia and Teso districts are mediocre in terms of educational attainment: in 1995, Busia ranked 26th of 50 Kenyan districts on national primary-school

exams. (Glewwe, et al. [1998]). 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. (The 1995 exam results are for the combined district.) The combined population of Busia and Teso districts in 1989 was 420,000, and their total area 1,819 square kilometers.[4]

The material poverty of primary schools in Busia and Teso is striking: few classrooms for younger pupils have desks, so children often sit on the dirt floor; pupil textbooks are rare, and chalk is in short supply; classes are sometimes held outside under a tree for lack of permanent classroom structures.

The national Kenya Ministry of Education regulates the primary school curriculum, administers national examinations, and hires, transfers, and pays teachers. Local school committees composed principally of parents raise funds for books, desks, chalk, and classrooms. Although teacher salaries paid by the central government account for most primary school spending – local school funding makes up only ten percent of total expenditures on teacher salaries on average, according to the author's calculations – a reduction in local inputs could have an important impact on educational output if local inputs and teacher inputs are complements in educational production.

Most local school funds are collected in the form of fees which parents pay to the school headmaster. Annual school fees ranged from 200-500 Kenya shillings (7-17 U.S. dollars) per family in 1995. The collection of school fees entails a complicated process of negotiation between the generally cash-strapped parents and the headmaster, who threatens to suspend the pupils who are late with fees.

The second source of local primary school funding – accounting for approximately one third of local funding in Busia (Table 5) – are village fundraisers called *harambees* in Swahili, at which parents and other community members publicly pledge their financial support for the school. *Harambees* are an important source of public finance in Kenya, accounting for 40 percent of local expenditures on

[4] Refer to *http://www.kenyaweb.com/ourland/western/busia/busia.html* for more information on Busia and Teso districts.

primary schools and other local public goods; nearly 90 percent of Kenyans claim to have participated in at least one *harambee* (Wilson [1992]).

2.2 Data

Detailed financial and demographic data for ninety-seven of the 333 primary schools in Busia and Teso districts were collected from pupil, school, and teacher questionnaires in early 1996, as baseline information for the Dutch non-governmental organization International Christian Support Fund's (ICS) School Assistance Programme. Primary schools in towns, market centers and wealthier rural areas were excluded from the assistance program, leaving a sample of poor rural primary schools.

The pupil questionnaire focuses on pupils' schooling background and family characteristics. 6,626 pupil questionnaires were administered to all grade six through eight pupils (ages twelve to eighteen) present on the day of questionnaire administration. Younger pupils were excluded from the questionnaire because of their limited reading and writing skills. In total, 97 school questionnaires and 861 teacher questionnaires were also administered in early 1996. School questionnaires – filled by schoolmasters with the assistance of a trained enumerator – contain detailed information on school finances and characteristics. Teacher questionnaires focus on teacher qualifications and teaching techniques, and were completed by the teachers themselves. School district examination results were provided by the Busia District Education Office.

The survey enumerators believe that responses from the school questionnaire are the most reliable of the three questionnaires, with teacher responses also generally reliable, but pupil responses often inaccurate. Fortunately, the response to the question on pupil ethnic affiliation is likely to suffer from less response error than other questions. Table 1 in Appendix A describes the data in greater detail.

2.3 Ethnicity

The largest ethnic groups in Busia and Teso districts are Luhya, Teso, and Luo. Table 2 presents the proportion of each group in the pupil sample. The Tesos and Luos are Nilotic groups with pastoralist

traditions, and the Luhyas are Bantu cultivators composed of the Khayo, Marachi, Nyala, and Samia subtribes, among others.[5] It is unclear if Luhya subtribes should be considered a single ethnic group, since certain Luhya dialects are mutually unintelligible (Grimes [1996]), there are historical rivalries between subtribes, and the concept of a Luhya group is recent, originating in the 1950's. In the empirical analysis, I consider ethnic diversity measures in which Luhya subtribes are considered distinct groups, as well as measures in which Luhyas are aggregated.

There is no clear pattern of ethnic socioeconomic stratification in rural Busia and Teso districts. Table 2 indicates that levels of parental educational attainment, and latrine and iron roof ownership are similar for different ethnic groups in this area.

Although relations between ethnic groups in this area are peaceful, the author's observations suggest that ethnic minorities are often treated with suspicion in the rural communities of western Kenya, discouraging household residential mobility across traditional ethnic boundaries. Moreover, there is considerable anecdotal evidence of ethnic tension on school committees in western Kenya, as illustrated by the case of Matumbai Primary School, one of the most ethnically diverse schools in the study. In recent years, many parents in Matumbai have refused to participate in funding the parent-teacher association or to attend school meetings when the elected chairman of the parent teacher association is not a member of their ethnic group. Perhaps as a result of such ethnic tensions, annual per pupil local school funding in Matumbai is only one-third of average funding in the area, and no permanent classrooms have been constructed; most classes take place under a tree.[6]

Following Easterly and Levine [1997], I use ethno-linguistic fractionalization as the principal empirical measure of ethnic diversity. Ethno-linguistic fractionalization is the probability that two people randomly drawn from the population are from distinct groups. Formally,

$$ETHNIC \equiv 1 - \sum_i (\text{Proportion of Ethno-linguistic group}_i \text{ in the population})^2 \qquad (1)$$

[5] Other Luhya subtribes in Busia and Teso districts include the Bukusu, Dakho, Kabras, Marama, and Sukha. Other non-Luhya ethnic groups in this area include the Kikuyu, Masaai, Saboat, Somali, Tachoni, and Taita.

[6] I thank Mary Kay Gugerty, Sylvie Moulin and Robert Namunyu for their observations on this and other issues in Kenya.

Ethno-linguistic fractionalization among sample pupils is 0.82 when Luhya subtribes are considered separate ethnic groups (*ETHNIC1*), and 0.48 if the Luhya subtribes are aggregated (*ETHNIC2*). There is considerable variation in ethno-linguistic fractionalization across schools: while median school-level *ETHNIC1* is 0.27, *ETHNIC1* is greater than 0.70 in five schools and less than 0.05 in five schools. Local ethnic diversity is not significantly related to the year of school founding, the density of primary schools in the area, or the school population (regressions not shown). Most primary schools in Busia and Teso districts were founded in the 1960s and 1970s, often with the assistance of Christian missionary groups.

2.4 Pupil mobility and land markets

Pupils in Busia often transfer between primary schools in search of better quality education and lower school fees. Nearly half of the pupils in the sample have attended more than one primary school, and eighteen percent are currently attending a school that is not the closest school to their home (Table 2). Students not attending the school closest to their home are called transfer students. Table 3 indicates that the proportion of transfer students is large even in geographically isolated schools, suggesting that many pupils walk considerable distances to school.

Table 4 presents the proportion of transfer students (in grades 6 to 8) that are not from the school's dominant ethnic group, and suggests that pupils often travel to schools in which their ethnic group is not in the majority. Although Table 4 overstates the proportion of students who choose to become ethnic minority transfer students – some students do not live within walking distance of a school in which their group is dominant – ethnically mixed pupil flows are clearly the norm in western Kenya.

The land sales market is generally thin in poor countries (Ray [1998]), including Kenya. The thin land market and the inability to sell traditional family and subtribal lands without prior community approval continue to limit residential mobility in rural Kenya (Platteau [1999]). Communal rather than

individual property rights to land were traditional in western Kenya before British colonization, and explicit land sales were rare (Leo [1984]).

3 A theory of ethnic divisions, mobility and school funding

The theory builds on Alesina, Baqir and Easterly's [1997] model of public goods funding in ethnically diverse settings. The model examines school choice and funding decisions when distinct ethnic groups have different educational preferences; when exogenous differences in educational quality across schools are important; and when pupil mobility between schools is costly. These characteristics distinguish the model from existing theories of mobility and public goods funding, and may more realistically portray the school funding process in many less developed countries.

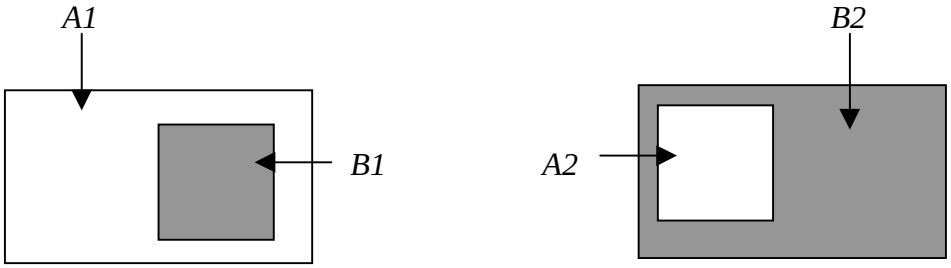
3.1 The set-up

The decision-making unit is the household. There are two ethnic groups, *A* and *B*, and each household belongs to one ethnic group. Households are distributed among two geographically disjoint areas, Area 1 and Area 2, and each area has one school, School 1 and School 2, respectively. The number of schools is fixed.[7]

Each area contains a continuum of households of unit measure, and one child from each household attends school. Households choose the school that their child attends. Households are either *A* households from Area 1 (*A1*), *B* households from Area 1 (*B1*), *A* households from Area 2 (*A2*), or *B* households from Area 2 (*B2*). The proportion of *A* households in Area 1 (Area 2) is $\delta_{1,0}$ ($\delta_{2,0}$). Assume Area 1 is dominated by *A* households ($\delta_{1,0} > ½$), Area 2 is dominated by *B* pupils ($\delta_{2,0} < ½$), and each group has the same majority in its "home" area: $\delta_{1,0} = 1 - \delta_{2,0}$.

[7] Founding a new primary school in Kenya requires obtaining government approval – which may be time consuming – as well as the cost of classroom construction and textbook purchases. I assume that this cost is large enough in the model that ethnic minority groups choose not to establish additional schools in their home area.

Figure 1:
Distribution of households



Area 1 *Area 2*

All households receive identical exogenous income, *y*. This is a one good economy, and income can be used for education spending, *g*, or private consumption, *c*, such that $y = c + g$. Each household in a school receives educational production g^{α} if each pays *g* in school fees.[8] Educational output is concave in school funding, $\alpha \in [0,1]$, since basic educational investments – like learning to read – may have the highest future payoffs (Psacharopoulos [1994]).

Households vote on the funding level and the type of educational instruction in the school they have chosen. The policy space of educational instruction is $S = [0,1]$. In the context of Kenyan primary schools, educational preferences might relate to the language of instruction, or the language in which the headmaster conducts parent teacher association meetings.[9] Household utility is decreasing in the difference between actual school instruction and the household's ideal. All *A* (*B*) households have identical preferences for the ideal type of educational instruction, $s_A \in S$ ($s_B \in S$). The difference in ideal instruction types between ethnic groups is $\sigma \equiv |s_A - s_B|$, a measure of ethnic polarization.

The proportion of *A* pupils in School 1 (School 2) after households make school choice decisions is δ_1 (δ_2), and ethno-linguistic fractionalization in School *m* is $\gamma_m = 2 \cdot \delta_m \cdot (1-\delta_m)$, $m \in \{1, 2\}$. Households' school choice decisions determine school ethnic composition (γ), and ethnic composition may affect both the type and level of school funding, which in turn determine school choice decisions.

[8] Local primary education in Kenya has characteristics of a private good, since children whose parents do not pay school fees are suspended from school.
[9] Younger pupils are taught in their vernacular (tribal) language in Kenya. Starting in grade four, classes are conducted in English. Fluency in Swahili or English is not universal among parents in rural western Kenya.

Congestion effects from large pupil populations are not considered in the model, although they could be introduced without changing the qualitative nature of the theoretical results.

The effect of ethnic diversity on educational production is represented by $f(\gamma)$. $f'(\gamma) < 0$ corresponds to a negative effect of diversity on educational production, for example if conflict between ethnic groups disrupts school meetings or classroom learning. A positive diversity effect, $f'(\gamma) > 0$, could result from the beneficial exchange of ideas among people with different cultural backgrounds.

The quality of the school's learning environment, which may reflect the motivation of the school headmaster and teachers, affects the efficiency of educational provision. Quality is represented by the exogenous random variable $\varepsilon \in [0,1]$. Taking into account school quality (ε), the type of educational instruction (s), and ethnic diversity (γ), effective educational production for household i is:

$$\varepsilon \cdot f(\gamma) \cdot [1 - |s - s(i)|] \cdot g^{\alpha} \qquad (2)$$

Effective educational production can be thought of as the benefit of education in terms of discounted future income. School funds are utilized more efficiently for household i as school quality improves, as ethnic cooperation increases, and as educational instruction approaches the household's ideal. Equation 2 assumes that local school funding, headmaster quality, and ethnic cooperation are complements in educational production. This may be appropriate in settings like rural Kenya where ethnic tension paralyzes school committees, and where the headmaster's failure to stem teacher absenteeism reduces the efficiency of other educational investments.

Each pupil can costlessly attend the school in her home area, but travelling to the other school requires a payment $\Delta \cdot \omega$, where $\Delta \geq 0$ is the cost per unit traveled and ω is a *Uniform[0,1]* random variable representing distance to the school. This travel cost may be thought of as the cost of walking to the distant school, or as transactions costs in the local land sales market. Households have the following utility function, where I is an indicator variable that equals one if the pupil travels to the distant school:

Definition 1: $U^{\omega}_{tn}(m)$ is utility for a type $(tn, \omega) \in \{A1, A2, B1, B2\}$ x $[0,1]$ child in School $m \in \{1,2\}$.

$$U^{\omega}_{tn}(m) = \text{Effective education} + \text{Consumption} - \text{Travel cost} \quad (3)$$
$$= \varepsilon_m \cdot f(\gamma_m) \cdot [1 - |s_m - s_t|] \cdot g_m^{\alpha} + (y - g_m) - \Delta \cdot \omega \cdot I$$

Agents have rational expectations and parameter values are common knowledge. The timing is:

<u>$t = 1$</u>: Households observe school qualities (ε_1, ε_2), simultaneously make school choice decisions, and pay the travel cost, if necessary. The travel cost is sunk, and the school choice is final.[10]

<u>$t = 2$:</u> Majority voting among parents in each school on the level of funding (g) and the type of educational instruction (s), in that order.[11]

<u>$t = 3$:</u> Households receive income, pay school fees, and consume the education and private goods.

The model is solved working backwards. Given the level of school funding, the majority ethnic group in the school sets the educational instruction type to its ideal. The majority group sets the level of school funding to maximize the utility of a representative member, taking into account the type of educational instruction that will be chosen, as well as school quality and the level of ethnic diversity in the school. The maximization problem faced by majority group households is:

$$Max_g \ \ \varepsilon \cdot f(\gamma) \cdot g^{\alpha} - g + y \quad (4)$$

The solution is:

$$g(\gamma)^* = [\alpha \cdot \varepsilon \cdot f(\gamma)]^{(1/(1-\alpha))} \quad (5)$$

The principal insight of this solution is that parents are willing to spend more on school funding when educational investments are more productive, for example when there is a better quality headmaster, or more cooperation across ethnic groups. Although g corresponds most closely to school fees in the context of Kenyan primary schools, Equation 5 may be understood as a reduced form relationship between educational quality and school funding appropriate in a variety of school finance

[10] The assumption of a sunk travel cost is realistic in western Kenya, since neither the annual school fee nor the cost of a uniform particular to a school is refundable if the child transfers to another school during the school year.

[11] This voting order resembles the school funding process in western Kenya: school committees first set the level of school fees, and then decide how to spend the funds. This voting order also allows me to avoid issues of multidimensional voting. I restrict attention to equilibria in which households vote sincerely (alternatively, play weakly dominant strategies), which eliminates a class of equilibria generated by the fact that infinitesimal households' payoffs are invariant to their vote.

settings. For example, if *harambee* contributions are higher when educational investments are more productive, all subsequent theoretical results also hold for community fundraisers.

For simplicity, I consider the case of $y > g(\gamma)^*$, income levels are high enough – alternatively, credit markets function efficiently enough – for households to afford the optimal primary school fee. However, the theoretical insights hold if this assumption is relaxed. For example, at the corner solution $g_1 = g_2 = y$ for low levels of household income, effective educational production remains an increasing function of ε_m, $f(\gamma_m)$, and $1 - |s_m - s(i)|$.

Assumption 1: $\varepsilon_1 > \varepsilon_2$

Assumption 2: $1 - \alpha - \sigma > 0$

Assumption 3: $\delta_1 = ½ \Rightarrow s_1 = s_A$

$\delta_2 = ½ \Rightarrow s_2 = s_B$

Assumption 1 means School 1 has a more favorable learning environment, without loss of generality. Assumption 2 implies that higher school funding always translates into higher utility. Assumption 3 means that in the event of a tie in voting in School 1 (School 2), *A* (*B*) households – the majority in Area 1 (Area 2) – implement their preferred instruction type.

Assumption 4: $U^{\omega}_{A1}(1) > U^{\omega}_{A1}(2), \forall \omega \in [0,1]$

Assumption 5: $\delta_1 \geq ½$

Assumptions 4 and 5 rule out equilibria in which *A* pupils travel to School 2 because they expect *B* pupils to become a majority in School 1. The equilibria ruled out may not be stable, if pupils incur travel costs that would be avoided by coordination on a similar equilibrium with less mobility. Assumptions 3 and 5 together imply that the school instruction type in School 1 (School 2) is s_A (s_B), the preferred educational instruction type in Area 1 (Area 2).

3.2 A simple solution with zero travel costs and $f(\gamma)=1$

The travel cost $\Delta=0$ and $f=1$ to simplify the solution; they are reintroduced in Section 3.3. Equilibrium utility outcomes for type *B2* are presented below, where equilibrium spending in School *m* is denoted $g_m = [\alpha\varepsilon_m]^{1/(1-\alpha)}$. If a type *B2* student travels to School 1, her utility is:

$$U_{B2}(1) = \varepsilon_1 \cdot (1-\sigma) \cdot g_1^{\alpha} - g_1 + y \qquad (6)$$
$$= g_1 \cdot (1-\alpha-\sigma)/\alpha + y$$

If she attends School 2, her utility is:

$$U_{B2}(2) = g_2 \cdot (1-\alpha)/\alpha + y \qquad (7)$$

The school choice decision is determined by differencing Equations 6 and 7:

$$U_{B2}(1) - U_{B2}(2) = (g_1 - g_2)(1/\alpha - 1) - g_1(\sigma/\alpha) \qquad (8)$$

The first term is the benefit of attending School 1 in terms of school funding, which is positive in this case since $\varepsilon_1 > \varepsilon_2$, and $f = 1$. The second term is the cost of School 1's non-ideal educational instruction type ($s_1 = s_A \neq s_B$) for a group *B* pupil. All *B2* students choose School 1 when the gap in educational preferences between groups is small, relative to the difference in quality between schools:

$$(1-\alpha)(1 - (\varepsilon_2/\varepsilon_1)^{1/(1-\alpha)}) > \sigma \qquad (9)$$

This equilibrium is called the Magnet Outcome, because students from both ethnic groups are drawn to the higher quality school. *A2* pupils always travel to School 1 in this simple case, since they prefer both the higher funding and the instruction type in School 1.

If the ethnic educational preference gap (σ) is large relative to the difference in quality between the two schools, the inequality in (9) is reversed and all *B1* pupils travel to School 2, while all *A2* pupils travel to School 1. The resulting equilibrium is called the Segregation Outcome.

3.3 Equilibria

Subgame-perfect Nash equilibrium is the appropriate equilibrium concept for this dynamic game of complete information. I restrict attention to pure strategy equilibria that fulfill the following

monotonicity condition: if a household $tn \in \{A1, A2, B1, B2\}$ and travel cost ω^* does not travel to the distant school, then *tn* households with $\omega > \omega^*$ do not travel to the distant school.

Definition 2: $\omega^{tn} \in [0,1]$ is the proportion of $tn \in \{A1, A2, B1, B2\}$ pupils who attend the distant school.

Given common knowledge parameters $(\delta_{1,0}, \delta_{2,0}, \varepsilon_1, \varepsilon_2, \Delta, \sigma, \alpha)$, a Nash equilibrium is characterized by a set of cut-off travel costs $(\omega^{A2*}, \omega^{B1*}, \omega^{B2*})$ that satisfy incentive compatibility constraints for all household types $(t,\omega) \in \{A1, A2, B1, B2\}$ x $[0,1]$. Assumption 4 implies that $\omega^{A1} = 0$ (no *A1* pupils travel to School 2). Three claims characterize the solution; all proofs are in Appendix B.

Claim 1:
a) $\omega^{B1*} > 0 \Rightarrow \omega^{B2*} = 0$
b) $\omega^{B1*} = 0 \Rightarrow \omega^{B2*} \geq 0$
c) $\omega^{B2*} > 0 \Rightarrow \omega^{B1*} = 0$
d) $\omega^{B2*} = 0 \Rightarrow \omega^{B1*} \geq 0$

Claim 2:
a) $\omega^{B2*} < 1 \Rightarrow \omega^{A2*} > \omega^{B2*}$
b) $\omega^{B2*} = 1 \Rightarrow \omega^{A2*} = 1$

Claim 3: $\omega^{A2*} > 0$

Claim 1 implies that either *B* pupils travel from School 1 to School 2, or from School 2 to School 1, but never both. The intuition is if some *B2* pupils are willing to pay a travel cost to attend School 1, then all *B1* pupils – who can attend School 1 without paying the travel cost – also prefer School 1. Claim 2 implies that at a given travel cost *A2* pupils are more likely to travel to School 1 than *B2* pupils, since *A2* pupils prefer the educational instruction type at School 1 ($s_1 = s_A$) to the instruction type at School 2 ($s_2 = s_B$). Claim 3 implies that *A2* pupils with low travel costs attend School 1. There are two generic pure strategy equilibria:[12]

The Magnet Outcome:
$\omega^{A2*} > \omega^{B2*} > 0$

[12] In addition, there is a knife-edge case in which $\omega^{B1*} = \omega^{B2*} = 0$.

$$\omega^{B1*} = \omega^{A1*} = 0$$

Proposition 1: In the Magnet Outcome
a) School 1 is more ethnically diverse than School 2: $\gamma_1 > \gamma_2$
b) School 1 is better funded than School 2: $g_1 > g_2$
Proof: Refer to Appendix B.

Proposition 1 indicates that school funding and ethnic diversity are positively correlated across the two schools in a Magnet Outcome, even if ethnic diversity has a negative effect on educational efficiency ($f' < 0$), as some Area 2 pupils from both ethnic groups attend School 1 for its superior educational quality. School 2 becomes a poorly funded, low quality and relatively ethnically homogeneous "ghetto".

The Segregation Outcome:

$$\omega^{A2*} > 0, \quad \omega^{B1*} > 0$$
$$\omega^{A1*} = \omega^{B2*} = 0$$

Some *B1* pupils travel to School 2 despite both better quality education in School 1 and the travel cost. Perfect sorting is a special case of the Segregation Outcome (Proposition 2) when mobility is costless, school quality is the same in both schools, and ethnic diversity is not beneficial for the efficiency of educational production, a finding reminiscent of Tiebout's [1956] seminal result.

Proposition 2: If mobility is costless ($\Delta = 0$), quality is the same in both schools ($\varepsilon_1 = \varepsilon_2$), educational instruction preferences differ across ethnic groups ($\sigma > 0$), and diversity is not beneficial for educational production ($f'(\gamma) \leq 0 \ \forall \gamma$), then there is complete ethnic sorting ($\delta_1 = 1, \delta_2 = 0$).
Proof: Refer to Appendix B.

The following example briefly illustrates the workings of the model when $f(\gamma) \neq 1$. For simplicity I consider $\varepsilon_2 = \varepsilon < 1 = \varepsilon_1$, and $f(\gamma) = 1$ for $\gamma < \gamma^*$ and $f(\gamma) = \rho \neq 1$ for $\gamma \geq \gamma^*$, a functional form that accommodates both cases where increased ethnic diversity either reduces educational efficiency ($\rho < 1$) or improves efficiency ($\rho > 1$). I restrict attention to equilibria in which $\gamma_1 \geq \gamma^*$ and $\gamma_2 < \gamma^*$. The Magnet Outcome occurs when $\omega^{B2*} > 0$, or when:

$$(1 - \alpha) \cdot (1 - (\varepsilon / \rho)^{1/(1-\alpha)}) > \sigma \qquad (10)$$

The Magnet Outcome does not occur if ρ is sufficiently small, as integration is costly when ethnic diversity is associated with inefficient educational production, but does occur if the gap in educational quality between the two schools is sufficiently large (ε small). The Segregation Outcome occurs when the inequality in (10) is reversed. Multiple equilibria may be possible if integration is associated with more efficient educational production ($\rho > 1$), as proved in Claim A1 in Appendix B.

3.4 Empirical implications

The following implications of the theory are tested in western Kenya primary schools. γ_m is ethno-linguistic fractionalization in School m, and $\gamma_{m,0}$ is ethno-linguistic fractionalization among households residing in Area m. τ denotes the true effect of ethnic diversity on school funding, and τ_{ols} denotes the ordinary least squares regression estimate of the effect of ethnic diversity on school funding.

(i) $Corr(\gamma_m - \gamma_{m,0}, g_m) > 0$

Proposition 1 implies that the difference between school and local ethnic diversity is *positively* correlated with the level of school funding in a Magnet Outcome, as good schools with high levels of school funding attract ethnically diverse pupil populations from surrounding areas. This positive local relationship between ethnic diversity and school funding among neighboring schools is not inconsistent with a negative relationship among schools not within walking distance of each other, if ethnic diversity is associated with less efficient educational production ($f'(\gamma) < 0$). Table 4 indicates that flows of ethnic minority transfer students are large in western Kenya primary school, suggesting that the Magnet Outcome may be the relevant case there.

(ii) $\tau_{ols} > \tau$

Endogenous pupil mobility in the Magnet Outcome implies that high quality and well-funded schools attract ethnically diverse student populations (Proposition 1). As a result, ordinary least squares regressions of school funding (dependent variable) on ethnic diversity (explanatory variable) that fail

to control for unobserved aspects of school quality suffer from an upward omitted variable bias in the estimated coefficient on ethnic diversity, since both ethnic diversity and school funding are positively related to school quality.

4 Empirical Results

4.1 Identification strategy

The principal empirical insight provided by the model is that school quality may be positively correlated with school-level ethnic diversity due to endogenous pupil mobility (Proposition 1), introducing an omitted variable bias if school quality is imperfectly observed. The ethnic diversity of pupils in the surrounding geographic zone – a measure of diversity largely independent of pupil mobility among neighboring schools – is used as an instrumental variable for school-level diversity to address this potential bias. The ethnic composition of particular zones in western Kenya is stable – being tied to traditional tribal and subtribal lands – and is plausibly exogenous (Government of Great Britain [1929], p. 5; Leo [1984]). Zonal ethnic diversity is computed among all pupils from sample schools in the corresponding zone.

In Kenya, the hierarchy of geographic areas in terms of size is province (largest area), district, division, and zone (smallest area). Since indicator variables for the geographic divisions are included in most empirical specifications to control for regional differences in income and tastes for education, the relationship between ethnic diversity and school outcomes is identified across zones within the same geographic division. In all specifications, I assume that regression disturbance terms are independent across geographic zones, but may be correlated within zones.

Table 5 presents local funding and ethnic diversity for the 22 geographic zones in the sample. Local funds per pupil and ethnic diversity are negatively correlated across zones in Angurai, Budalangi, Butula, Matayos, and Nambale divisions, while there is no clear pattern in Funyula and Amukura divisions. (Amagoro division, which only contains a single sample school, is grouped with

neighboring Angurai division in the empirical analysis.) Donations from local *harambee* fundraisers appear to account for the negative relationship between total funds and ethnic diversity in Table 5.

Socioeconomic status and ethnicity

Although Table 2 suggests that there are no significant socioeconomic differences across ethnic groups in rural areas of Busia and Teso districts, it remains possible that ethnic diversity measures are proxying for the size of particular ethnic groups, whose members differ from other ethnic groups in either income (which is unobserved) or their average taste for education. To control for such demographic and socioeconomic variation across zones, average fathers' education, iron roof ownership, and the proportion of each ethnic group (Khayo, Luo, Marachi, Nyala, Samia, and Teso) in the surrounding geographic zone are included as explanatory variables in many specifications.

Ethnic diversity may affect local school funding indirectly, through its impact on other economic outcomes. Ethnically diverse regions may be poor because contracts are harder to enforce within heterogeneous communities, leading credit, land and labor markets to function less efficiently in such areas (Besley and Coate [1995]; Grief [1993]; LaFerrara [1997]). There may be less incentive to invest in education in areas with poor credit and land markets. Reduced form specifications that exclude socioeconomic controls as explanatory variables capture the total relationship between ethnic diversity and primary school outcomes, including this possible indirect effect through other economic outcomes. When socioeconomic controls are included as explanatory variables, the coefficient estimate on ethnic diversity can be interpreted as the direct effect of ethnic diversity on primary school outcomes.

Pupil mobility across zones

Although pupils may walk across zonal boundaries to attend school, most mobility is likely to occur within zones since pupils cannot plausibly walk more than a few kilometers to school. There is insufficient information on the location of pupil residences to identify which pupils do cross zonal

boundaries, so the zonal ethnic diversity measures are not entirely independent of endogenous pupil mobility.[13] However, if pupils generally travel to zones with good quality schools – a zonal Magnet Outcome effect – there will be an upward bias (toward zero) on zonal ethnic diversity coefficient estimates, strengthening the empirical results.

To address the effect of pupil mobility across zonal boundaries on school ethnic diversity, the ethnic diversity of each school's nearest neighboring geographic zone is considered as a possible instrumental variable for school-level diversity. Regressions 2,4, and 6 in Table 6 suggest that neighboring zone ethnic diversity is a weak predictor of school ethnic diversity, and I do not include it as an instrumental variable in any regressions. Table 6 includes zonal socioeconomic characteristics (fathers' education and iron roof ownership) as additional explanatory variables, and indicates that zonal socioeconomic measures are insignificantly related to school ethnic diversity.

Although family residential mobility is uncommon in western Kenya due to the thin land sale market and reluctance to sell family lands, children could potentially move in with relatives to attend a primary school that is not within walking distance of their home. Table 2 presents evidence that less than 15 percent of pupils are not living with a parent, among pupils with at least one surviving parent. Since some of these pupils moved in with relatives residing in the same geographic zone as their parents – since relatives often live near each other – the proportion of children who move in with relatives to attend a primary school in a different geographic zone is plausibly considerably less than 15 percent, although data limitations make it impossible to determine the exact rate. Such rates of pupil residential mobility are unlikely to dramatically change measured levels of school-level ethnic diversity.

[13] It would be ideal to use Kenyan census data for an independent measure of local ethnic diversity. However, the author's repeated attempts to obtain Kenyan detailed demographic data at the zonal level have not been successful.

4.2 Data

Table 1 contains the data description. The principal empirical measures of ethnic diversity are ethno-linguistic fractionalization disaggregating Luhya subtribes (*ETHNIC1)* and aggregating Luhya subtribes (*ETHNIC2*); ethno-linguistic fractionalization among the Luhya subtribes (*ETHNIC3*); and the size of the largest ethnic group in a school, disaggregating Luhya subtribes (*LARGEGRP).*

Religious diversity is not included as an explanatory variable in the empirical analysis, since 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 a causal relationship if evangelical activity is targeted to and is most successful in poor areas, for example.

The principal educational outcome measure is total local funding per pupil in 1995 (*FUNDS1*) collected from school fees as well as from local fundraising meetings. *FUNDS1* does not include funds raised from outside sources, such as non-governmental organizations. Since the non-governmental organization (ICS) excluded schools that were receiving considerable donor assistance from their program, only six of the 97 sample schools received over $100 in outside funding in 1995, and there is no indication that local fundraising was crowded out in these schools (regressions not shown). Total local funding is composed of funds collected directly from school fees (*FEES1*), and funds collected from donations to the school (*DONAT1*), principally from local community fundraisers (*harambees*). The proportion of parents contributing to the school parent-teacher association fund and the proportion of parents attending the 1995 School Meeting reflect parent participation in the primary school.

School facilities are additional educational outcome measures, since chronically under-funded schools are likely to have fewer educational resources than other schools. The number of desks per pupil, pupil latrines per pupil, and classrooms per pupil in 1996 are measures of schools' physical facilities. School-owned textbooks per pupil is another measure of past local educational investment.

The final educational outcome measure is performance on government primary school examinations for grades 6 to 8.

School characteristics are included as explanatory variables in certain specifications as controls for school quality. These include the number of primary schools within a three kilometer radius, the year of school founding, the pupil population, an indicator variable for a male headmaster, and an indicator variable if the school headmaster believes the school's highest priority is the purchase of textbooks, which may contain information on headmaster quality. Teacher characteristics include the frequency of assigned homework, the proportion of teachers in the school with a high school-equivalent Form IV education, and the proportion of male teachers in the school. It is theoretically unclear how gender affects teacher quality. Ethno-linguistic diversity among teachers, and an indicator variable if the headmaster is not a member of the school's largest ethnic group are additional school characteristics.

4.3 Results

Tables 7 and 8 present results using ethno-linguistic fractionalization disaggregating Luhya subtribes (*ETHNIC1*), and the proportion of the school's largest ethnic group (*LARGEGRP*), respectively, as the measure of ethnic diversity. The instrumental variable for both measures of school ethnic diversity is zonal ethnic diversity disaggregating Luhya subtribes (*ZETHNIC1*). Coefficient point estimates on both measures of ethnic diversity are negative, and the instrumental variable estimates are significantly different than zero at 90 or 95 percent confidence. The coefficient on ethnic diversity remains negative and significantly different than zero at 95 percent confidence when the proportion of each ethnic group in the zone and zonal socioeconomic controls are included, suggesting that the measures of ethnic diversity are not proxying for ethnic socioeconomic stratification. The first stage F-statistics are significantly different than zero at high levels of confidence (Table 6). The Magnet Outcome bias

discussed in Section 3.4 is a potential explanation for both the insignificant negative relationship between ethnic diversity and school funding, and the poor fit of the ordinary least squares regressions.

An interpretation of the instrumental variable coefficient estimate on ethnic diversity in Table 7, regression 5 is that the drop in local school funding associated with a change from ethnic homogeneity to median school-level ethnic diversity is roughly 30 percent of average local funding.

Figure 2 plots local school funding versus ethno-linguistic fractionalization for the 97 schools in the sample, and reveals that no ethnically diverse school (with *ETHNIC1* ≥ 0.5) is well funded. A non-parametric t-test that the proportion of schools with funding greater than 250 shillings per pupil is the same for schools with *ETHNIC1* > 0.5 and for those with *ETHNIC1* ≤ 0.5 is rejected at 99 percent confidence. The unexplained variation in the data may be due to unobserved components of school quality, or to weather and commodity price shocks that affect parents' ability to pay school fees.

Figure 2:
Ethnic diversity versus total school funding per pupil in 1995

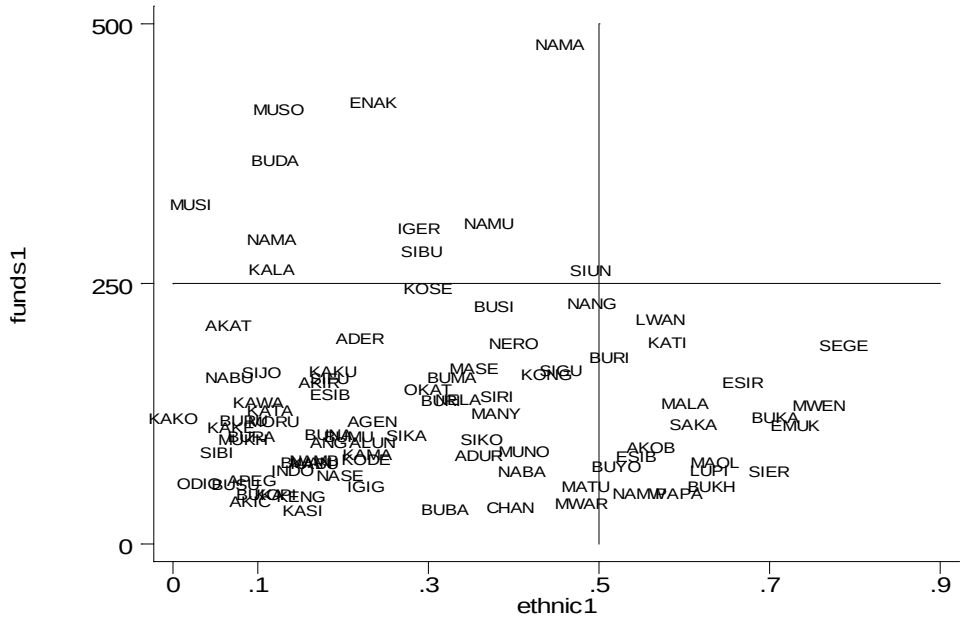


Table 9 suggests that ethnic diversity is associated with a variety of local collective action problems. Ethnic diversity is strongly negatively associated with the collection of donations from community fundraisers (*harambees*), but not significantly related to the level of school fees collected

per pupil. Organizing a successful *harambee* requires considerable community participation, which may be lacking in ethnically diverse areas. Table 9 also suggests that parent participation in the parent-teacher association and in school meetings is significantly lower in ethnically diverse areas than in homogeneous areas.

Table 10 includes a range of headmaster, school, and teacher characteristics as potential controls for school quality. The Magnet Outcome presented in Section 3 suggests that the coefficient point estimate on *ETHNIC1* should become more negative when school quality is included as an explanatory variable, since endogenous pupil mobility implies that quality and funding are positively correlated across neighboring schools. *ETHNIC1* remains significantly different than zero at 95 percent confidence across all specifications when school and teacher characteristics are included as explanatory variables, and coefficient point estimates become more negative in several specifications. However, a change in the school sample due to missing data complicates the interpretation of this change in the coefficient estimate. The headmaster, school and teacher characteristics are insignificant predictors of school funding, with the exception of having a male headmaster, which is significantly negatively associated with the level of school funding.

Table 11 includes both ethnic diversity aggregating Luhya subtribes and ethnic diversity among the Luhya subtribes as explanatory variables, and suggests that ethnic diversity across tribes (*ETHNIC2*) accounts for the observed negative relationship between ethnic diversity and school funding. The effect of ethnic diversity among Luhya subtribes (*ETHNIC3*) on school funding is insignificantly different from zero in both regressions.

The estimated ordinary least squares coefficients on ethnic diversity are uniformly greater than the instrumental variable estimates in Tables 7 and 8. This is consistent with both the theoretical Magnet Outcome presented in Section 3, and with attenuation bias. Measurement error in the ethnic diversity indexes is likely since ethnic affiliation is available for only a subsample of pupils in each school. Table 12 attempts to determine which of the two explanations is driving the pattern between

OLS and IV coefficient estimates. If the difference between school and zonal ethnic diversity (*DIFETHN*) is primarily due to measurement error from pupil absences on the day of questionnaire administration, data entry errors, or random noise, then *DIFETHN* and school funding will be unrelated. However, if zonal ethnic diversity is a reasonable proxy for local ethnic diversity within walking distance of a school, the theory predicts that Magnet schools will be both more ethnically diverse and better funded than surrounding schools (Proposition 1), implying a positive relationship between *DIFETHN* and *FUNDS1*. In Table 12 the coefficient estimates on *DIFETHN1* and *DIFETHN2* are both positive, and the coefficient on *DIFETHN2* is nearly significantly different than zero at 90 percent confidence ($t > 1.4$). I interpret this as suggestive empirical support for the Magnet Outcome effect.

Table 13 presents the relationship between ethnic diversity and primary school facilities in 1996. Current school facilities reflect the cumulative impact of past local educational investment decisions. Levels of desks per pupil, pupil latrines per pupil, and classrooms per pupil are all found to be negatively associated with ethnic diversity. The coefficient estimates on ethnic diversity are large, negative and significantly different than zero at 95 percent confidence in both instrumental variable regressions for desks per pupil and latrines per pupil, and significantly different than zero for one of the two instrumental variable specifications for classrooms per pupil. Taken together, these results suggest that primary schools in ethnically diverse areas have significantly worse physical facilities than schools in more homogeneous areas. An interpretation of the instrumental variable coefficient estimates on ethnic diversity is that the drop in desks per pupil associated with the change from ethnic homogeneity to median school-level ethnic diversity is 35-59 percent of average desks per pupil.

Table 13 also includes the stock of school textbooks per pupil as a dependent variable. Although the coefficient estimates on ethnic diversity are negative in all three specifications, they are not significantly different than zero at traditional confidence levels. The scope for parental conflict over particular educational investments may explain the different results for textbook provision and

infrastructure investments. The purchase of standard government textbooks is likely to be less controversial across ethnic groups than the decision of which school structures to build, and specifically who gets the contract to build them (Gugerty et al. [1999]).

I also report the relationship between ethnic diversity and privately-owned textbooks per pupil, to explore the possibility of substitution from publicly-provided to privately-owned textbooks in ethnically diverse areas, and find that coefficient point estimates on ethnic diversity are insignificantly different than zero in all specifications with private textbooks per pupil as the dependent variable.

Table 14 examines the relationship between ethnic diversity and school average scores on government examinations for pupils in standards 6 to 8. The results suggest that there is no clear relationship between test score performance and ethnic diversity in these schools. This may be surprising given the lower levels of local school funding, and worse physical facilities and parental involvement in ethnically diverse areas. A possible explanation is provided by Regression 5 in Table 14, which indicates that the stock of school-owned textbooks per pupil is positively associated with test scores, and the stock of desks and classrooms per pupil are insignificantly related to test performance. Table 13 suggested that additional local school funds in ethnically homogeneous areas are principally invested in physical infrastructure projects rather than in textbooks.

Several other explanations may account for the weak relationship between ethnic diversity and test scores in these primary schools. First, it is difficult to interpret average school test score results since some schools exclude their worst students from taking exams (Glewwe, et al [1998]). Unfortunately, there is insufficient information on which students sat for the government exams to address this participation bias. Second, local educational investments may affect school performance in ways not captured by Kenyan government examinations, which tend to emphasize rote learning (Somerset [1987]). Finally, the empirical relationship between school funding and academic outcomes in less developed countries remains disputed. Hanushek's [1995] survey concludes that there is little evidence of a systematic causal relationship between educational inputs and test scores in developing

countries. However, Case and Deaton [1998] find that school funding has a strong positive effect on numeracy test scores and pupil enrollment among South African primary school students.

Table 14 (regression 5) suggests that schools with a high proportion of male teachers perform poorly on government examinations, and that schools with large pupil populations outperform smaller schools, once controls are included for school quality. However, the endogeneity of pupil population complicates the interpretation of this coefficient estimate.

5 Discussion

The empirical results paint a picture of pervasive local collective action problems in ethnically diverse primary schools in rural western Kenya. Ethnic diversity is associated with sharply lower total local school funding, less parental participation in parent-teacher associations and school meetings, and with fewer educational investments such as desks, classrooms, and latrines, although not with textbook provision. Ethnic diversity across tribes (Luhya, Luo and Teso) rather than among Luhya subtribes appears to be driving the negative relationship between ethnic diversity and school outcomes, suggesting that collective action problems may be more severe in the presence of larger cultural and linguistic differences.

Lower donations from local fundraisers (*harambees*) – events that require considerable organizational effort and community participation – in ethnically diverse areas account for the drop in local school funding. This pattern of contributions may arise if community sanctions against free-riders are more effective in homogeneous communities, or if local politicians and business people – important *harambee* contributors – gain more by publicly contributing to schools in which their ethnic group is dominant. The negative estimated relationship between ethnic diversity and *harambee* contributions suggests that diversity may be negatively associated with the provision of other local public goods in Kenya, given the central role that *harambees* play in Kenyan local public finance (Wilson [1992]). However, it is unclear if the negative relationship between ethnic diversity and

school funding identified in this paper is robust to settings in which voluntary community fundraisers are not a principal source of funding.

Centralizing funding decisions at the regional or national level is a possible response to the negative impact of local ethnic diversity on school funding. This has occurred to a large extent in Kenya as teacher salaries – primary schools' largest expense – are paid by the national Ministry of Education. However, centralized provision introduces the problem of ethnic and regional conflict at the national level. 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 percent in 1979-1980 to 16 percent 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 percent during the same period. If the central government allocated funds for desks, chalk, and textbooks similarly, politically marginalized areas might be largely excluded from school funds. Besley and Coate [1999] argue formally that the decentralized provision of local public goods may be optimal in countries with such 'winner-takes-all' national politics, even in settings where there are efficiency benefits to centralized provision.

A robust negative relationship between local ethnic diversity and human capital accumulation would have bleak welfare implications for many countries, including most of sub-Saharan Africa. Before drawing broad conclusions about the impact of ethnic diversity on economic growth, however, further research should examine the relationship between ethnic diversity, school finance, and human capital accumulation in a variety of other countries and settings.

References

Alesina, Alberto, Reza Baqir, and William Easterly. (1997). "Public goods and ethnic divisions", NBER Working Paper 6009.

Alesina, Alberto and Allen Drazen. (1991). "Why are stabilizations delayed?" *American Economic Review*, **81**, 1170-1188.

Alesina, Alberto, and Enrico Spolaore. (1997). "On the number and size of nations", *Quarterly Journal of Economics*, **112**, 1027 – 1055.

Barkan, Joel D., and Michael Chege. (1989). "Decentralising the state: district focus and the politics of reallocation in Kenya", *Journal of Modern African Studies*, **27(3)**, 431-453.

Barro, Robert. (1997). *Determinants of Economic Growth: A cross-country empirical study.* MIT Press, Cambridge.

Benabou, Roland. (1993). "Workings of a city: location, education, and production", *Quarterly Journal of Economics*, **108**, August, 619-652.

Benhabib, Jess, and Aldo Rustichini. (1996). "Social conflict and growth", *Journal of Economic Growth*, **1(1)**, 125-142.

Besley, Timothy, and Stephen Coate. (1995). "Group lending, repayment incentives and social collateral", *Journal of Development Economics*, **46(1)**, 1-18.

Besley, Timothy, and Stephen Coate. (1999). "Centralized versus decentralized provision of local public goods: A political economy analysis," NBER Working Paper 7084.

Case, Anne, and Angus Deaton. (1998). "School inputs and educational outcomes in South Africa," mimeo Princeton University.

Collier, Paul, and Jan Willem Gunning. (1999). "Explaining African economic performance," *Journal of Economic Literature*, **37(March)**, 64-111.

Durlauf, Steven. (1996). "A theory of persistent income inequality," *Journal of Economic Growth*, **1(1) March**, 75-93.

Easterly, William, and Ross Levine. (1997). "Africa's growth tragedy: policies and ethnic divisions", *Quarterly Journal of Economics*, **112**, November, 1203-1250.

Epple, Dennis and Thomas Romer (1991). "Mobility and redistribution", *Journal of Political Economy*, **99(4)**, 828-858.

Fernandez, Raquel and Richard Rogerson. (1996). "Income distribution, communities, and the quality of public education", *Quarterly Journal of Economics*, **111(1)**, February, 135-164.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. (1998). "Textbooks and test scores: Evidence from a prospective evaluation in Kenya", mimeo, MIT.

Goldin, Claudia and 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 Great Britain. (1929). *Kenya Native Affairs Department Annual Report, 1928*. London: Millbank, Westminster.

Greif, Avner. (1993). "Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition", *American Economic Review*, **83(3)**, 525-548.

Grimes, Barbara F. (1996). Kenya, *Ethnologue 13th Edition*. Summer Institute of Linguistics. (*www.sil.org/ethnologue/countries/Keny.html*)

Gugerty, Mary Kay, and Michael Kremer. (1999). "The Effects of Development Funding on Social Capital: the Kenya Local Community Action Project," mimeo., Harvard University.

Hanushek, Eric A. (1995). "Interpreting Recent Research in Schooling in Developing Countries", *World Bank Research Observer*, **10(August)**, 227-246.

Krueger, Alan B., and Mikael Lindahl. (1998). "Education for Growth: Why and For Whom?", mimeo, Princeton University and Stockholm University.

LaFerrara, Eliana. (1997). "Ethnicity and reciprocity: A model of credit transactions in Ghana," mimeo, Harvard University.

Leo, Christopher. (1984). *Land and Class in Kenya*. University of Toronto Press, Toronto.

Platteau, Jean-Philippe. (1999). *Institutions, Social Norms, and Economic Development*, Forthcoming, Harwood Publishing.

Poterba, James M. (1996). "Demographic structure and the political economy of public education", NBER Working Paper 5677.

Psacharapolous, G. (1994). "Returns to investment in education: A global update", *World Development*, **22(9)**, September, 1325-43.

Ray, Debraj. (1998). *Development Economics*. Princeton University Press.

Sachs, Jeffrey D., and Andrew M. Warner. (1997). "Sources of Slow Growth in African Economies", *Journal of African Economies*, **6(3)**, 335-376.

Somerset, H. (1987). "Examination reform in Kenya," EDT 64, World Bank: Washington, DC.

Tiebout, Charles. (1956). "A Pure Theory of Local Public Expenditures", *Journal of Political Economy*, **64**, 416-424.

Wilson, L.S. (1992). "The *Harambee* movement and efficient public good provision in Kenya," *Journal of Public Economics*, **48**, 1-19.

Appendix A:

Table 1: Data Description

Variable name	Definition and source
Ethnic diversity	
ETHNIC1	Pupil ethno-linguistic fractionalization index which counts Luhya subtribes as separate groups. (SOURCE: 1996 ICS Pupil Questionnaire)
ETHNIC2	Pupil ethno-linguistic fractionalization index which counts Luhya subtribes as one group. (SOURCE: 1996 ICS Pupil Questionnaire)
ETHNIC3	ETHNIC1 – ETHNIC2. Ethno-linguistic fractionalization among Luhyas. (SOURCE: 1996 ICS Pupil Questionnaire)
LARGEGRP	Proportion of student population accounted for by the largest group, disaggregating Luhya subtribes. (SOURCE: 1996 ICS Pupil Questionnaire)
ZETHNIC1	Zonal pupil ethno-linguistic fractionalization index in the surrounding geographic zone, which counts Luhya subtribes as separate groups, computed among pupils who filled the Pupil Questionnaire in early 1996. (SOURCE: 1996 ICS Pupil Questionnaire)
ZETHNIC2	Zonal pupil ethno-linguistic fractionalization index in the surrounding geographic zone, which counts Luhyas as one group, computed among pupils who filled the Pupil Questionnaire in early 1996. (SOURCE: 1996 ICS Pupil Questionnaire)
ZETHNIC3	ZETHNIC1 – ZETHNIC2. Zonal ethno-linguistic fractionalization among Luhyas in the surrounding geographic zone. (SOURCE: 1996 ICS Pupil Questionnaire)
NETHNIC1	Zonal pupil ethno-linguistic fractionalization index in the nearest neighboring geographic zone to the school, counting Luhya subtribes as separate groups. (SOURCE: 1996 ICS Pupil Questionnaire)
NETHNIC2	Pupil ethno-linguistic fractionalization index in the nearest neighboring geographic zone to the school, counting Luhyas as a single group. (SOURCE: 1996 ICS Pupil Questionnaire)
NETHNIC3	Pupil ethno-linguistic fractionalization index in the nearest neighboring geographic zone to the school, among Luhyas. (SOURCE: 1996 ICS Pupil Questionnaire)
DIFETHN1	ETHNIC1 – ZETHNIC1
DIFETHN2	ETHNIC2 – ZETHNIC2

Educational outcomes	
FUNDS1	Total school funding per pupil raised from school fees and donations, minus "activity funds" which are sent to the District Education Officer, 1995. (SOURCE: 1996 ICS School Questionnaire)
FEES1	School funding per pupil raised from school fees, minus "activity funds" which are sent to the District Education Officer, 1995. (SOURCE: 1996 ICS School Questionnaire)
DONAT1	School funding per pupil raised from donations, principally harambees, 1995. (SOURCE: 1996 ICS School Questionnaire)
PTA2PUP	Parents who participated in funding the parent-teacher association per pupil, 1995 (SOURCE: 1996 ICS School Questionnaire)
MEET2PUP	Parents who attended the school meeting per pupil, 1995 (SOURCE: 1996 ICS School Questionnaire)
DESK2PUP	Desks per pupil in the school, 1996 (SOURCE: 1996 ICS School Questionnaire)
PLAT2PUP	Pupil latrines per pupil in the school, 1996 (SOURCE: 1996 ICS School Questionnaire)
ROOM2PUP	Classrooms per pupil in the school, 1996 (SOURCE: 1996 ICS School Questionnaire)
BKS2PUP	School textbooks per pupil in the school in 1996 (SOURCE: 1996 ICS School Questionnaire)
TEXT2PUP	Personal textbooks per pupil in 1996. (SOURCE: 1996 ICS School Questionnaire)
TEST95	Average school score in standards 6 through 8 on the 1995 KCPE, national government primary school examination (SOURCE: District-All Data, ICS office in Busia)

Table 2:
Average pupil characteristics, by ethnic group[*]

	Entire pupil sample	Luo pupils	Teso pupils	Luhya pupils	Khayo (Luhya) pupils	Marachi (Luhya) pupils	Nyala (Luhya) pupils	Samia (Luhya) pupils
Number of pupils interviewed	6626	328	1686	4414	1194	1154	535	1417
Proportion of pupil sample	1.000	0.050	0.256	0.670	0.181	0.175	0.081	0.215
Age in years	14.5	14.5	14.7	14.5	14.6	14.4	14.7	14.4
Latrine ownership	0.848	0.798	0.880	0.838	0.877	0.888	0.626	0.841
Iron roof ownership	0.256	0.309	0.226	0.258	0.230	0.256	0.267	0.275
Fathers with post-primary education	0.380	0.332	0.381	0.378	0.354	0.306	0.412	0.448
Mothers with post-primary education	0.187	0.214	0.185	0.183	0.177	0.150	0.175	0.214
Attends primary school that is not closest to home	0.180	0.167	0.179	0.180	0.141	0.199	0.195	0.192
Lives with a parent, given that at least one parent is alive	0.851	0.831	0.878	0.844	0.846	0.862	0.835	0.832

[*] Other Luhya subtribes comprise 0.018 of the pupil sample, and other non-Luhya ethnic groups comprise 0.020 of the sample. Other Luhya subtribes in Busia and Teso districts include the Bukusu, Dakho, Kabras, Marama, and Sukha. Other non-Luhya ethnic groups in this area include the Kikuyu, Masaai, Saboat, Somali, Tachoni, and Taita.

Table 3:
The proportion of pupils attending a school not closest to home (1996)
(Source: 1996 ICS School Questionnaire and Pupil Questionnaire)

Number of schools within a three kilometer radius of the school	% of pupils attending a school not closest to home
Zero (19 Schools)	17.4
One (23 Schools)	19.1
Two (15 Schools)	15.2
Three (18 Schools)	20.4
Four (22 Schools)	16.0

Table 4:
The proportion of transfer students not from the school's largest ethnic group (1996)
(Source: 1996 ICS School Questionnaire and Pupil Questionnaire)

Largest ethnic group in the school	Average % of transfer students from other ethnic groups
Luhya – all subtribes (60 schools)	11.1
Khayo (18 schools)	45.6
Marachi (15 schools)	17.1
Nyala (7 schools)	28.7
Samia (20 schools)	20.0
Luo (3 schools)	41.7
Teso (31 schools)	21.6

Table 5: 1995 average district, division, and zone characteristics[*]
(schools with complete demographic and financial information)

Geographic area	Number of schools	Number of pupils	Total local funds per pupil	Local donations per pupil	Ethnic1	Ethnic2	Ethnic3
Entire sample	92	26936	135.5	44.5	0.328	0.204	0.124
Busia district	63	19826	151.6	56.9	0.353	0.189	0.164
Budalangi division	7	2160	243.9	188.3	0.287	0.122	0.165
Bunyala North zone	4	1509	208.3	147.2	0.352	0.144	0.208
Bunyala South zone	3	651	291.2	243.2	0.115	0.058	0.056
Butula division	17	5676	162.6	58.5	0.351	0.161	0.190
Burinda zone	4	1198	124.3	12.1	0.717	0.327	0.390
Butunyi zone	3	906	114.0	16.6	0.404	0.169	0.235
Marachi North zone	4	1728	160.9	58.9	0.311	0.092	0.219
Tingolo zone	6	1844	213.7	110.0	0.328	0.262	0.066
Funyula division	20	5687	136.1	48.8	0.220	0.097	0.124
Agenga/Nanguba zone	5	1839	151.5	48.1	0.167	0.037	0.131
Bwiri zone	4	1234	187.3	107.1	0.370	0.143	0.226
Funyula zone	5	1377	111.8	20.5	0.251	0.217	0.035
Nambuku zone	6	1237	109.2	34.1	0.169	0.041	0.128
Matayos division	8	2959	125.7	30.6	0.355	0.233	0.121
Bukhayo South zone	3	1185	142.2	66.3	0.210	0.081	0.129
Mundika zone	5	1774	115.8	9.1	0.574	0.452	0.122
Nambale division	11	3344	122.9	4.6	0.623	0.398	0.224
Bukhayo Central zone	6	1792	138.7	7.6	0.645	0.328	0.317
Bukhayo East zone	2	574	99.2	0	0.738	0.326	0.412
Bukhayo North zone	3	978	107.0	2.8	0.679	0.602	0.076
Teso district	29	7110	100.6	17.6	0.256	0.248	0.009
Amagoro division	2	567	61.1	0	0.271	0.267	0.004
Kocholia zone	2	567	61.1	0	0.271	0.267	0.004
Amukura division	12	2831	124.1	23.5	0.207	0.197	0.010
Aremit zone	4	711	127.3	24.6	0.070	0.069	0.001
Kaujakito zone	4	1088	145.5	18.3	0.242	0.234	0.008
Kotur zone	4	1032	99.5	27.6	0.334	0.315	0.019
Angurai division	15	3712	87.1	15.3	0.284	0.275	0.008
Angurai zone	5	1012	122.9	47.5	0.228	0.226	0.002
Chamasiri zone	7	1758	78.2	0	0.371	0.359	0.012
Katakwa zone	3	942	48.0	20.1	0.297	0.294	0.003

[*] Three schools were the only schools in their zone. These schools were assigned to the nearest zone: Apegei (schid 105) to Aremit zone; Mukhobola (schid 158) to Bunyala South zone; Sifuyo (schid 193) to Nambuku zone. The ethnic diversity figures are for all schools in the sample, not only those with complete financial and demographic information.

Table 6: The instrumental variables, *ZETHNIC1, ZETHNIC2, ZETHNIC3* [*]

	Dependent variable: ETHNIC1		Dependent variable: ETHNIC2		Dependent variable: ETHNIC3	
	(1)	(2)	(3)	(4)	(5)	(6)
Explanatory variable	ols	ols	ols	ols	ols	ols
Pupil ethno-linguistic fractionalization in surrounding geographic zone, Luhyas disaggregated (ZETHNIC1)	0.869***	0.826***				
	(0.065)	(0.092)				
Pupil ethno-linguistic fractionalization in nearest neighboring geographic zone, Luhyas disaggregated (NETHNIC1)		0.102				
		(0.089)				
Pupil ethno-linguistic fractionalization in surrounding geographic zone, Luhyas aggregated (ZETHNIC2)			0.849***	0.833***		
			(0.087)	(0.105)		
Pupil ethno-linguistic fractionalization in nearest neighboring geographic zone, Luhyas aggregated (NETHNIC2)				0.125		
				(0.113)		
Pupil ethno-linguistic fractionalization in surrounding geographic zone among Luhyas (ZETHNIC3)					0.815***	0.740***
					(0.121)	(0.098)
Pupil ethno-linguistic fractionalization in nearest neighboring geographic zone among Luhyas (NETHNIC3)						0.089
						(0.087)
Proportion of fathers in the geographic zone with some post-primary education (ZFEDUC)		-0.013		-0.087		0.006
		(0.201)		(0.189)		(0.138)
Proportion of pupils in the geographic zone with iron roofing (ZIRONROO)		0.089		0.022		0.178
		(0.277)		(0.245)		(0.193)
R^2	0.56	0.56	0.43	0.44	0.55	0.56
Root MSE	0.142	0.144	0.136	0.137	0.090	0.090
Number of observations	97	97	97	97	97	97
Mean of dependent variable	0.314		0.202		0.112	

[*] Notes: Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence.

Table 7: Dependent variable, Total local school funds raised in 1995 (*FUNDS1*)[*]

Explanatory variable	(1) ols	(2) ols	(3) iv-2sls	(4) iv-2sls	(5) iv-2sls
Pupil ethno-linguistic fractionalization disaggregating Luhya subtribes (ETHNIC1)	-10.0 (39.3)	-10.3 (49.1)	-142.7* (69.2)	-110.6** (49.5)	-155.4** (62.9)
Proportion of Khayo pupils in the geographic zone (ZKHAYO)				-4.1 (54.0)	41.4 (64.0)
Proportion of Luo pupils in the geographic zone (ZLUO)				89.8 (172.5)	182.6 (210.4)
Proportion of Nyala pupils in the geographic zone (ZNYALA)				233.7* (131.2)	238.3 (144.7)
Proportion of Samia pupils in the geographic zone (ZSAMIA)				-412.5** (163.0)	-458.8*** (158.5)
Proportion of Teso pupils in the geographic zone (ZTESO)				-38.9 (55.2)	-4.3 (78.1)
Proportion of fathers in the school with some post-primary education (FEDUC)					113.3 (93.2)
Proportion of fathers in the geographic zone with some post-primary education (ZFEDUC)					30.6 (126.0)
Proportion of pupils in the geographic zone with iron roofing (ZIRONROO)					-90.7 (105.8)
Geographical division indicators	No	Yes	Yes	Yes	Yes
R^2	0.00	0.19	-	-	-
Root MSE	92.6	86.4	89.2	88.6	91.1
Number of schools	92	92	92	92	92
Mean of dependent variable	135.5				

[*] Notes: Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. The instrumental variable for *ETHNIC1* is zonal ethno-linguistic fractionalization disaggregating Luhyas (*ZETHNIC1*). Observations are assumed to have independent error terms across geographic zones, but not necessarily within zones.

Table 8: Dependent variable, Total local school funds raised in 1995 (*FUNDS1*)[*]

Explanatory variable	(1) ols	(2) ols	(3) iv-2sls	(4) iv-2sls	(5) iv-2sls
Proportion of pupils in the school's largest ethnic group, Luhyas disaggregated (LARGEGRP)	23.8 (42.9)	28.3 (53.8)	185.4** (86.7)	156.9** (67.3)	214.1** (82.4)
Proportion of Khayo pupils in the geographic zone (ZKHAYO)				-12.7 (53.9)	23.1 (64.0)
Proportion of Luo pupils in the geographic zone (ZLUO)				109.1 (169.2)	203.6 (203.4)
Proportion of Nyala pupils in the geographic zone (ZNYALA)				277.0** (123.1)	296.6** (131.3)
Proportion of Samia pupils in the geographic zone (ZSAMIA)				-357.2** (160.4)	-376.5** (157.4)
Proportion of Teso pupils in the geographic zone (ZTESO)				-43.5 (53.9)	-17.9 (74.8)
Proportion of fathers in the school with some post-primary education (FEDUC)					104.2 (96.3)
Proportion of fathers in the geographic zone with some post-primary education (ZFEDUC)					26.6 (128.6)
Proportion of pupils in the geographic zone with iron roofing (ZIRONROO)					-90.7 (101.9)
Geographical division indicators	No	Yes	Yes	Yes	Yes
R^2	0.00	0.19	-	-	-
Root MSE	92.5	86.4	89.0	88.8	91.6
Number of schools	92	92	92	92	92
Mean of dependent variable	135.5				

[*] Notes: Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. The instrumental variable for *ETHNIC1* is zonal ethno-linguistic fractionalization disaggregating Luhyas (*ZETHNIC1*). Observations are assumed to have independent error terms across geographic zones, but not necessarily within zones.

Table 9:
Sources of funding, and parent participation in school activities[*]

Dependent variable	Coefficient estimate on ETHNIC1 (1) ols	(2) iv-2sls	(3) iv-2sls	Number of schools	Mean of dependent variable
Sources of local funding					
Total school fees collected per pupil, 1995 (*FEES1)*	2.1 (29.2)	16.9 (30.1)	-5.1 (38.3)	92	91.0
Total donations collected per pupil, 1995 (*DONAT1*)	-12.4 (58.7)	-159.5* (77.2)	-150.4** (60.0)	92	44.5
Parent participation in the school					
Parents who participated in funding the parent-teacher association per pupil, 1995 (*PTA2PUP*)	-0.039 (0.073)	-0.173** (0.074)	-0.356*** (0.092)	92	0.217
Parents who attended the school meeting per pupil, 1995 (*MEET2PUP*)	-0.125*** (0.026)	-0.150*** (0.048)	-0.110 (0.094)	95	0.195
Geographical division indicators	Yes	Yes	Yes		
Demographic and socioeconomic controls	No	No	Yes		

[*] Notes: Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. The instrumental variable for *ETHNIC1* is zonal ethno-linguistic fractionalization disaggregating Luhyas (*ZETHNIC1)*. Observations are assumed to have independent error terms across geographic zones, but not necessarily within zones. The demographic and socioeconomic controls include *FEDUC, ZFEDUC, ZIRONROO, ZKHAYO, ZLUO, ZNYALA, ZSAMIA, ZTESO.*

Table 10: Controlling for headmaster, school, and teacher quality
Dependent variable, Total local school funds raised in 1995 (*FUNDS1*)[*]

Explanatory variable	(1) iv-2sls	(2) iv-2sls	(3) iv-2sls	(4) iv-2sls
Pupil ethno-linguistic fractionalization disaggregating Luhya subtribes (ETHNIC1)	-401.7** (155.0)	-150.9** (61.4)	-138.8** (75.5)	-350.7** (135.2)
Headmaster is male (HMSEX)	-225.8* (114.7)			-249.5** (104.7)
Headmaster does not belong to school's largest pupil ethnic group (HMDIFFER)	46.4 (57.8)			13.9 (78.2)
School headmaster thinks that textbook purchase is school's highest priority (TEXTBEST)	-7.8 (22.2)			3.3 (21.6)
Number of other primary schools within 3 kilometers of the school (DIST_3)		6.1 (7.4)		1.4 (12.7)
Total pupil population in the school (TOTSTUD)		-0.058 (0.123)		-0.035 (0.109)
Year of school founding (YEARBEGI)		-0.16 (1.67)		-1.35 (1.42)
Frequency of assigned homework, reported by teachers (FREQWORK)			22.8 (14.3)	15.7 (24.5)
Teacher ethno-linguistic fractionalization aggregating Luhya subtribes (TETHNIC)			130.4** (61.7)	168.5 (125.6)
Proportion of teachers in the school with high-school equivalent education, Form IV (TCHIV)			-12.8 (60.5)	79.9 (99.4)
Proportion of male teachers in the school (TCHSEX)			34.4 (64.4)	59.2 (106.3)
Geographical division indicators	Yes	Yes	Yes	Yes
Demographic and socioeconomic controls	Yes	Yes	Yes	Yes
R^2	-	-	-	-
Root MSE	99.1	93.0	90.1	98.8
Number of schools	74	90	91	72
Mean of dependent variable	135.5			

[*] Notes: Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. The instrumental variable for *ETHNIC1* is *ZETHNIC1*. Observations are assumed to be independent across geographic zones, but not necessarily within zones. The demographic and socioeconomic controls include *FEDUC, ZFEDUC, ZIRONROO, ZKHAYO, ZLUO, ZNYALA, ZSAMIA, ZTESO*. The reduced sample size is due to omitted responses to *TEXTBEST*.

Table 11: Alternate measures of ethnic diversity
Dependent variable, Local school funds raised in 1995 (*FUNDS1*)*

Explanatory variable	(1) iv-2sls	(2) iv-2sls
Pupil ethno-linguistic fractionalization across tribes (ETHNIC2)	-135.3*	-287.0***
	(63.2)	(85.5)
Pupil ethno-linguistic fractionalization among Luhya subtribes (ETHNIC3)	-159.1	34.2
	(128.6)	(90.7)
Geographical division indicators	Yes	Yes
Demographic and socioeconomic controls	No	Yes
R^2	-	-
Root MSE	89.4	102.3
Number of schools	92	92
Mean of dependent variable	135.5	

Table 12: Testing the Magnet Outcome
Dependent variable, Total local school funds raised in 1995 (*FUNDS1*)

Explanatory variable	(1) ols	(2) ols
Difference between school and zonal pupil ethno-linguistic fractionalization, disaggregating Luhya subtribes (DIFETHN1)	58.2	
	(56.1)	
Difference between school and zonal pupil ethno-linguistic fractionalization, aggregating Luhya subtribes (DIFETHN2)		90.3
		(63.6)
Geographical division indicators	Yes	Yes
R^2	0.19	0.20
Root MSE	86.1	85.5
Number of schools	92	92
Mean of dependent variable	135.5	

* Notes: Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. The instrumental variable for *ETHNIC2* is *ZETHNIC2*, and for *ETHNIC3* is *ZETHNIC3*. Observations are assumed to have independent error terms across geographic zones, but not necessarily within zones. The demographic and socioeconomic controls include *FEDUC, ZFEDUC, ZIRONROO, ZKHAYO, ZLUO, ZNYALA, ZSAMIA, ZTESO*.

Table 13: Primary school facilities and textbooks*

Dependent variable	Coefficient estimate on ETHNIC1 (1) ols	(2) iv-2sls	(3) iv-2sls	Number of schools	Mean of dependent variable
Physical infrastructure					
Desks per pupil in 1996, reported by headmasters (DESK2PUP)	-0.15*** (0.05)	-0.27** (0.10)	-0.45*** (0.13)	96	0.204
Pupil latrines per pupil in 1996, reported by headmasters (PLAT2PUP)	-0.008 (0.008)	-0.017** (0.008)	-0.036** (0.015)	96	0.016
Classrooms per pupil in 1996, reported by headmasters (ROOM2PUP)	-0.014* (0.007)	-0.016 (0.018)	-0.046*** (0.014)	96	0.031
Textbooks					
School-owned textbooks per pupil in 1996, reported by headmasters (BKS2PUP)	-0.12 (0.12)	-0.10 (0.17)	-0.19 (0.19)	96	0.346
Privately-owned textbooks per pupil 1996, reported by headmasters (TEXT2PUP)	0.003 (0.056)	-0.002 (0.078)	0.060 (0.076)	96	0.078
Geographical division indicators	Yes	Yes	Yes		
Demographic and socioeconomic controls	No	No	Yes		

* Notes: Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. The instrumental variable for *ETHNIC1* is zonal ethno-linguistic fractionalization disaggregating Luhyas (*ZETHNIC1)*. Observations are assumed to have independent error terms across geographic zones, but not necessarily within zones. The demographic and socioeconomic controls include *FEDUC, ZFEDUC, ZIRONROO, ZKHAYO, ZLUO, ZNYALA, ZSAMIA, ZTESO.*

Table 14: School test scores[*]

Dependent variable, Average school score on 1995 government examinations, standards 6-8 (*TEST95*)

Explanatory variable	(1) ols	(2) iv-2sls	(3) iv-2sls	(4) iv-2sls	(5) iv-2sls
Pupil ethno-linguistic fractionalization disaggregating Luhya subtribes (ETHNIC1)	22.2 (46.5)	52.8 (84.1)	108.4 (78.4)	13.1 (94.9)	-14.8 (74.9)
School-owned textbooks per pupil in 1996 (BKS2PUP)				82.8 (49.2)	146.0*** (51.2)
Privately-owned textbooks per pupil in 1996 (TEXT2PUP)				181.3 (125.7)	112.0 (94.6)
Desks per pupil in 1996 (DESK2PUP)				16.4 (104.0)	8.4 (110.8)
Classrooms per pupil in 1996 (ROOM2PUP)				-1828.7* (886.3)	-1006.5 (1373.6)
Headmaster is male (HMSEX)					-38.3 (33.6)
Headmaster does not belong to school's largest pupil ethnic group (HMDIFFER)					21.7 (25.2)
Teacher ethno-linguistic fractionalization aggregating Luhya subtribes (TETHNIC)					17.9 (74.2)
Proportion of male teachers in the school (TCHSEX)					-127.2** (46.2)
Total pupil population in the school (TOTSTUD)					0.211** (0.087)
Geographical division indicators	Yes	Yes	Yes	Yes	Yes
Demographic and socioeconomic controls	No	No	Yes	Yes	Yes
R^2	0.14	-	-	-	-
Root MSE	81.9	82.1	82.1	76.4	69.1
Number of schools	94	94	94	93	88
Mean of dependent variable	822.8				

[*] Notes: Huber robust standard errors in parentheses. Significantly different than zero at 90% (*), 95% (**), 99% (***) confidence. The instrumental variable for *ETHNIC1* is zonal ethno-linguistic fractionalization disaggregating Luhyas (*ZETHNIC1)*. Observations are assumed to have independent error terms across geographic zones, but not necessarily within zones. The demographic and socioeconomic controls include *FEDUC, ZFEDUC, ZIRONROO, ZKHAYO, ZLUO, ZNYALA, ZSAMIA, ZTESO.*

Appendix B:
Mathematical Appendix

- **Ethnic diversity**

For School 1, ethnic diversity after transfers is:

$$\delta_1 = [\delta_{1,0} + \omega^{A2}\delta_{2,0}]/[\delta_{1,0} + \omega^{A2}\delta_{2,0} + (1 - \omega^{B1})(1 - \delta_{1,0}) + \omega^{B2}(1 - \delta_{2,0})] \quad (A\text{-}1)$$

When there are two ethnic groups, $\gamma_j = 2\cdot\delta_j\cdot(1 - \delta_j)$ for $j \in \{1,2\}$:

$$\gamma_1 = 2[\delta_{1,0} + \omega^{A2}\delta_{2,0}][(1 - \omega^{B1})(1 - \delta_{1,0}) + \omega^{B2}(1 - \delta_{2,0})]/[\delta_{1,0} + \omega^{A2}\delta_{2,0} + (1 - \omega^{B1})(1 - \delta_{1,0}) + \omega^{B2}(1 - \delta_{2,0})]^2 \quad (A\text{-}2)$$

δ_2 and γ_2 are defined analogously.

- **Utility outcomes**

(A-3) $U^{\omega}_{B1}(1) = g_1\cdot(1 - \sigma - \alpha)/\alpha + y$

(A-4) $U^{\omega}_{B1}(2) = g_2\cdot(1 - \alpha)/\alpha - \Delta\cdot\omega + y$

(A-5) $U^{\omega}_{A2}(2) = g_2\cdot(1 - \sigma - \alpha)/\alpha + y$

(A-6) $U^{\omega}_{A2}(1) = g_1\cdot(1 - \alpha)/\alpha - \Delta\cdot\omega + y$

(A-7) $U^{\omega}_{A1}(1) = g_1\cdot(1 - \alpha)/\alpha + y$

(A-8) $U^{\omega}_{A1}(2) = g_2\cdot(1 - \sigma - \alpha)/\alpha - \Delta\cdot\omega + y$

- **Proofs**

Claim 1: a) $\omega^{B1*} > 0 \Rightarrow \omega^{B2*} = 0$
b) $\omega^{B1*} = 0 \Rightarrow \omega^{B2*} \geq 0$
c) $\omega^{B2*} > 0 \Rightarrow \omega^{B1*} = 0$
d) $\omega^{B2*} = 0 \Rightarrow \omega^{B1*} \geq 0$

Proof:

a) (*) $\omega^{B1*} > 0 \Rightarrow U^{\omega}_{B1}(2) > U_{B1}(1), \forall\, \omega < \omega^{B1*}$

(**) $U_{B2}(2) = U^{\omega=0}_{B1}(2)$

(***) $U_{B1}(1) \geq U^{\omega}_{B2}(1), \forall\, \omega$

(*), (**), (***) $\Rightarrow U_{B2}(2) = U^{\omega=0}_{B1}(2) > U_{B1}(1) \geq U^{\omega'}_{B2}(1), \forall\, \omega'$
$\Rightarrow \omega^{B2*} = 0$

b) (*) $\omega^{B2*} = 0 \Rightarrow U_{B2}(2) > U^{\omega}_{B2}(1) \;\forall\, \omega > 0$
$U_{B2}(2) \geq U^{\omega=0}_{B2}(1)$

(**) $U^{\omega=0}_{B2}(1) = U_{B1}(1)$

(***) $U^{\omega=0}_{B1}(2) = U_{B2}(2)$

(*), (**), (***) $\Rightarrow U^{\omega=0}_{B1}(2) = U_{B2}(2) \geq U^{\omega=0}_{B2}(1) = U_{B1}(1)$
$\Rightarrow \omega^{B1*} \geq 0$

c) and d) proven analogously.

Claim 2: a) $\omega^{B2*} < 1 \Rightarrow \omega^{A2*} > \omega^{B2*}$
b) $\omega^{B2*} = 1 \Rightarrow \omega^{A2*} = 1$

Proof:

a) (*) $\sigma \in (0, 1-\alpha) \Rightarrow U^{\omega}_{A2}(1) > U^{\omega}_{B2}(1)\ \forall\, \omega$, $U_{B2}(2) > U_{A2}(2)$

(**) $\omega^{B2*} < 1 \Rightarrow U^{\omega}_{B2}(1) > U_{B2}(2)\ \forall\, \omega < \omega^{B2*}$, $U^{\omega=\omega B2*}_{B2}(1) = U_{B2}(2)$

(***) $U^{\omega=\omega A2*}_{A2}(1) = U_{A2}(2)$

(*), (**), (***) $\Rightarrow U^{\omega=\omega B2*}_{A2}(1) > U^{\omega=\omega B2*}_{B2}(1) = U_{B2}(2) > U_{A2}(2) = U^{\omega=\omega A2*}_{A2}(1)$

$\Rightarrow U^{\omega=\omega B2*}_{A2}(1) > U^{\omega=\omega A2*}_{A2}(1)$

$\Rightarrow g_1(1/\alpha - 1) - \Delta\cdot\omega^{B2*} > g_1(1/\alpha - 1) - \Delta\cdot\omega^{A2*}$

$\Rightarrow \omega^{A2*} > \omega^{B2*}$

b) Proved analogously, with ω^{A2*} *bounded from above at one.*

Claim 3: $\omega^{A2*} > 0$

Proof:

Assert that $\omega^{A2*} = 0$ *and work toward a contradiction.*

(*) $\omega^{A2*} = 0 \Rightarrow U_{A2}(2) \geq U^{\omega=0}_{A2}(1)$

(**) $U^{\omega=0}_{A2}(1) = U_{A1}(1)$

(***) $U^{\omega=0}_{A1}(2) = U_{A2}(2)$

(*), (**), (***) $\Rightarrow U^{\omega=0}_{A1}(2) \geq U_{A1}(1)$, *contradicting Assumption 4.*

Proposition 1: In the Magnet Outcome

a) School 1 is more ethnically diverse than School 2: $\gamma_1 > \gamma_2$

b) School 1 is better funded than School 2: $g_1 > g_2$

Proof:

a) Define $\omega^{B2} \equiv \omega \in (0,1)$, *and* $\omega^{A2} = \omega + \mu > \omega^{B2}$ *by Claim 2. After pupils travel,*

(A-9)

$$\frac{\gamma_1}{\gamma_2} = \left(1 + \frac{\omega + (1-\delta_{1,0})\mu}{(1-\omega)(1-\omega-\mu)\delta_{1,0}(1-\delta_{1,0})}\right)\frac{(1-(\omega+(1-\delta_{1,0})\mu))^2}{(1+\omega+(1-\delta_{1,0})\mu)^2}$$

When $\mu = 0$, *this becomes:*

$\gamma_1/\gamma_2 = [(1-\omega)^2 + \omega/(\delta_{1,0} - \delta_{1,0}^2)]/(1+\omega)^2 > 1$, *since* $\delta_{1,0} > \frac{1}{2}$.

The derivative of A-9 with respect to μ *is strictly positive for* $\mu > 0$, *proving the claim.*

c) Magnet Outcome $\Leftrightarrow U^{\omega}_{B2}(1) - U^{\omega}_{B2}(2) = (g_1 - g_2)(1-\alpha)/\alpha - g_1(\sigma/\alpha) - \Delta\cdot\omega > 0$, *for some* ω

$\Rightarrow g_1 - g_2 > (\alpha/(1-\alpha))\cdot(g_1(\sigma/\alpha) + \Delta\omega) > 0$

$\Rightarrow g_1 > g_2$

Proposition 2: If mobility is costless ($\Delta = 0$), school quality is the same across schools ($\varepsilon_1 = \varepsilon_2$), there are different educational preferences across ethnic groups ($\sigma > 0$), and ethnic diversity is not beneficial for educational production ($f'(\gamma) \leq 0\ \forall\, \gamma$), then there is complete ethnic sorting ($\delta_1 = 1$, $\delta_2 = 0$).

Proof:
Prove for $f(\gamma) \equiv 1$. The result holds a fortiori for $f'(\gamma) \leq 0 \ \forall \gamma$.

(*) $\varepsilon_1 = \varepsilon_2$ *and* $f(\gamma) \equiv 1 \ \forall \gamma \Rightarrow g_1 = g_2$

(**) $\Delta = 0 \Rightarrow \Delta \cdot \omega = 0 \ \forall \omega$

(*), (**) $\Rightarrow U^{\omega}_{B2}(1) - U^{\omega}_{B2}(2) = -g_1(\sigma/\alpha) < 0 \ \forall \omega$, *in Equation 8.*
Analogously, $U^{\omega}_{A1}(2) - U^{\omega}_{A1}(1) = -g_2(\sigma/\alpha) < 0 \ \forall \omega$.

- **Multiple Equilibria**

A relevant case for multiple equilibria, in the example described in section 3.3, is when $\gamma_1 \geq \gamma^*$ is expected in a Magnet Outcome, $\gamma_1 < \gamma^*$ is expected in a Segregation Outcome, and $\gamma_2 < \gamma^*$ is expected in both cases.

Claim A1: A sufficient condition for multiple equilibria in the example in section 3.3 is:
$(\varepsilon/\rho)^{1/(1-\alpha)} < 1 - \sigma/(1-\alpha) < \varepsilon^{1/(1-\alpha)}$

Proof*: The Segregation Outcome occurs when*
$(1-\alpha)(1-\varepsilon^{1/(1-\alpha)}) < \sigma$ *(A-13)*

Solving (10) and (A13) simultaneously determines the range of multiple equilibria. The condition implies that $\rho > 1$.

The range of multiple equilibria in Figure A1 is ME. When f′ > 0, inflows of B2 pupils into School 1 make education production more efficient, making School 1 more attractive for other pupils. The strategic complementarity generates multiple equilibria.

Figure A1:
The range of multiple equilibria