

**Sanctioning, Clientelism and Politicization:
The Impact of Ethnicity on Primary and Secondary Education in Africa**

Anke Weber^a

September 2009

Abstract: Ethnic diversity is widely assumed to negatively affect a country's education outcomes. Yet, comprehensive explanations for this link are still missing. This paper combines the existing theory of the sanctioning ability of villages with general theories on clientelism and politicized ethnicity. The theories are tested using a novel dataset of district level school enrollment in 31 African countries in combination with newly coded indicators for district ethnic diversity, district shares of presidents' co-ethnics and existing ethnic parties. Estimating a hierarchical model supports the sanctioning and clientelism theory and provides first evidence for a *positive* effect of ethnic parties on enrollment. Low-income countries with weak institutions benefit significantly from the existence of ethnic parties by an increase in enrolment of over 30 percent. Moreover, nation building policies targeted at uniting diverse populations are not found to provide a cure to the ethnic diversity problem.

JEL classification: H41, I22, 010

Keywords: Public goods, politicized ethnicity, sanctioning, Africa

^a Center for Comparative and International Studies (CIS), Institute of Political Science, University of Zurich, Hirschengraben 56, CH-8001 Zurich, Switzerland, anke.weber@pw.uzh.ch. I am grateful to Nils Weidmann and Kathleen Cunningham, as well as Kimuli Kasara for generous provision of information and data. Also, I thank Ilia Rainer for useful comments and David Schönholzer and Maximilian Stern for great research assistance.

1. Introduction

The role of ethnicity in the development process is still poorly understood. Ever since the work by Horowitz (1985), researchers tried to disentangle the various dimensions of ethnicity and its impact on numerous aspects of a country's development. Generally, ethnically diverse countries seem to bear a particularly high burden in the development process. Ethnically diverse countries exhibit lower macroeconomic stability (Alesina and Drazen, 1991), lower growth rates (Easterly and Levine, 1997), and increased corruption (Mauro, 1995).

Additional evidence suggests that ethnicity might also hamper the provision of public goods, such as education. Evidence derived from the impact of ethnicity on education provides an interesting example for the impacts of ethnicity on the provision of public goods (or near public goods) in general.¹ Education is widely assumed to be of particular importance for the development process. Improving educational outcomes seems to have a positive impact on productivity and earnings, and hence on growth rates (Mankiw, Romer, and Weil, 1992; Michaelowa, 2000; and Hanushek and Kimko, 2000). Furthermore, increasing mothers' education is assumed to improve children's health and reduce fertility rates. This, in turn, is viewed to increase growth rates and thereby to drive development (Cochrane, 1979; and Glewwe, 1998). Moreover, studies suggest that education plays an important role in fighting the spread of HIV/Aids (Kelly, 2000). The development effects of education are particularly strong for primary and lower secondary education. This is also understood by international organizations and the donor community, who target their policies specifically on primary education in developing countries.²

Evidence on the impact of ethnicity on education derived from cross-country studies is, so far, rather ambiguous. Recent research started examining the underlying mechanisms of the effect of ethnicity. However, this research focused entirely on local communities and emphasized the sanctioning ability of homogenous villages found in a two country comparison (Miguel, 2004; and Miguel and Gugerty, 2005). This *sanctioning theory* posits that ethnically diverse villages are unable to sanction parents that do not contribute to village funding of local schools. This leads to an under-provision of education in heterogeneous villages.

¹ Note that education fails to fulfill the requirement of non-excludability for a public good, since children can be excluded from the classroom. However, education is to a large extent non-rival and generates multiple externalities.

² For example, the Millennium Development Goal 2 targets universal primary education.

Other strands of the literature, which focus on clientelistic distribution of government funds, have been, so far, neglected when explaining the impact of ethnicity on education. Although the clientelism argument has not yet been applied to the effect of ethnicity in the education sector, it provides valid explanations for government distribution patterns. More precisely, the *clientelism theory* posits that ethnic parties distribute funds primarily to their co-ethnics, which leads to an improvement of education outcomes in the region of co-ethnics. Clear empirical evidence, however, still lacks (e.g. Rainer and Franck, 2009; Kasara, 2007; Miguel and Zaidi, 2003).

Related to both, the sanctioning and the clientelism theory, is the theory of politicization of ethnicity. This theory has, however, been formerly neglected in the discussion. Whether ethnicity influences education through village diversity (*sanctioning theory*) or distribution of government funds (*clientelism theory*) might also depend on whether ethnicity is perceived as a politically relevant factor in the first place. Insights into the mechanism of ethnic parties are provided by Posner (2005) and Chandra (2004). The *theory of politicization*, hence, posits that the impact of ethnicity on education (via village sanctioning and clientelistic distribution of state funds) depends critically on the political relevance of ethnicity.

This paper contributes to the literature by combining the hitherto distinct theories of sanctioning and clientelism and attempting to formulate a theory of politicization of ethnicity. The theories are tested with a novel dataset and newly coded ethnic indicators. Data on primary and secondary education on district level for 31 African countries is combined with specific indicators of ethnic groups designed to test the distinct theories of ethnicity. In addition, due care is paid to possible interactions between the ethnic indicators, as well as to influences of institutional and economic factors on the impact of ethnicity. The dataset is analyzed estimating a two-level hierarchical model.

Following this introduction, section 2 reviews the theories explaining the impact of ethnicity on education and derives relevant hypotheses. The data and variables used for the econometric analysis are briefly discussed in section 3 with detailed coding rules provided in Appendix G and H. Section 4 presents the econometric results for primary enrollment (4.1) and secondary enrollment (4.2). Section 5 discusses the results and section 6 concludes.

2. How Does Ethnicity Affect Education?

The question arises, how and to which extent do the existing ethnic groups³ influence a country's educational outcomes. Empirical evidence has so far produced only a vague picture of the impact of ethnicity on education. Studies conducted on the impact of national ethnic diversity on primary and secondary education in the U.S. and Africa, found both negative effects of ethnicity (Easterly and Levine, 1997; Alesina, Baqir, and Easterly, 1999; Goldin and Katz, 1999; and Habyarimana, Humphreys, Posner, and Weinstein, 2007), insignificant effects (Keefer, 2005), as well as effects depending on the quality of a country's institutions (Easterly, 2001).

This empirical ambiguity might be due to the following problems. First, the aforementioned studies focus on the national level ethnic diversity measured by the so-called ethno-linguistic fractionalization (ELF) (see Appendix G). In particular, by using the ELF, the studies are not able to account for the different channels through which ethnicity may influence education and which operate on different levels (country and district). The second drawback is the neglect of the studies to include an indicator of the political relevance of ethnicity, i.e. the politicization, in their regressions.

The following paragraphs, therefore, discuss the different mechanisms through which ethnicity influences education and derives hypotheses. From the extensive literature on ethnicity, one can identify three major theories explaining why ethnic groups might influence education. The first theory is placed at the community level, where village funding for schooling depends on the ability of communities to sanction non-contributors. The second theory is based on the clientelism argument, which posits that politicians distribute funds to their co-ethnics. Third, effects of ethnicity on education might be influenced by the role of ethnicity in the political process. The following paragraphs outline the three effects and present the hypotheses.⁴

The Sanctioning Effect

³ In this paper, all social cleavages, such as race, tribe, language, and religion are subsumed in the term "ethnic group" and ethnic identity is assumed to be socially and politically constructed. For a discussion of the construction of ethnic identity, see Fearon and Laitin (2000).

⁴ Note that several authors raise the issue of different preferences of ethnic groups (Easterly and Levine, 1997; and Alesina, Baqir, and Easterly, 1999). However, more recent research rejects a significant impact of preferences on education (Habyarimana, Humphreys, Posner, and Weinstein, 2007).

The first theory explaining the impact of ethnicity on education posits that ethnically diverse villages exhibit lower school funding because they are unable to sanction non-contributors. A number of researchers provide evidence from the U.S. on a significant negative impact of ethnic diversity on participation in groups (Alesina and La Ferrara, 2002), on local public goods (Vigdor, 2004), and on education spending (Cutler, Elmendorf, and Zeckhauser, 1993; and Poterba, 1997). The most compelling argument for the impact of ethnic diversity on the local level in developing countries is provided by Miguel and Gugerty (2005). They examine the interethnic cooperation in the education sector of Kenyan villages. In these villages, primary schooling is financed through contributions made by the parents. If parents do not contribute, they can be sanctioned. Common sanctions are the exclusion from the village, and therefore, exclusion from networks that provide social insurance. Social insurance provided by the village is especially important in countries with otherwise weak infrastructure and poor public insurance systems as found in Africa. Miguel and Gugerty (2005) argue that villages, being composed of only a few different ethnic groups, are able to impose sanctions on parents that do not contribute to the school. However, parents not being members of the predominant ethnic group in the village are less likely to be affected by such sanctions. Therefore, villages that are composed of a large variety of ethnic groups are unable to impose credible sanctions.⁵ On the contrary, more homogenous villages exhibit higher trust and lower transactions costs, which helps them to impose sanctions (Fearon and Laitin, 1996; and Alesina and La Ferrara, 2002). According to Miguel and Gugerty (2005), lower sanctions in diverse villages, then, translate into lower contributions for primary education. Consequently, the authors expect villages with more homogenous ethnic composition to have higher funding for primary schools. Therefore, the following hypothesis is posited:

H₁: Higher ethnic diversity is associated with lower education outcomes.

The Clientelism Effect

The second theory explaining the impact of ethnicity on education is based on the clientelism argument. Clientelism is generally associated with an under-provision of goods to all citizens and an over-provision of goods targeted to specific groups.⁶ In the education sector, politicians are expected to distribute state resources for education primarily to their specific

⁵ A similar argument is made by Kimenyi (2006).

⁶ For an overview over different types of clientelism and definitions, see Clapham (1982) and Lemarchand (1972).

clientele, which are often their co-ethnics. Since state resource distribution is primarily determined by the incumbent president, her co-ethnics are the likely beneficiaries. Therefore, one could posit the following hypothesis:

H₂: President's co-ethnics are associated with higher education outcomes.

The majority of research supports the clientelism theory (Rainer and Franck, 2009; Miguel and Zaidi, 2003; Collier and Garg, 1999; and Barkan and Chege, 1989). However, Kasara's (2007) study provides an interesting example of president's co-ethnics being negatively affected by substantially higher taxation than members of other ethnic groups. The empirical ambiguity might be explained by the neglect of the aforementioned studies to account for the role of ethnicity in politics, i.e. the politicization.

The Effect of Politicization and Institutions

Although some clientelistic behavior might be found in all developing countries, Posner (2005) and Chandra (2004) argue that clientelistic behavior might be particularly pronounced in countries with ethnically diverse populations and parties based on ethnicity. They argue that the struggle for state resources encourages politicians to emphasize ethnic affiliations to attract voters. Emphasizing the ethnic identity seems necessary, since voters have only limited information on how politicians distribute state funds. Voters expect that politicians distribute funds primarily to their own ethnic group. As a consequence, citizens vote for the politician belonging to their ethnic group.⁷ Once a party is elected, it is, then, expected to distribute national resources to their ethnic members. If political parties in a country are, however, not based on ethnic identity, then politicians are viewed to distribute state resources more evenly. Hence, the impact of ethnicity on education via clientelistic distribution might depend critically on the relevance of ethnicity in politics:

H₃: The clientelism effect is more pronounced in countries with politicized ethnicity.

In addition, the relevance of ethnicity in politics might also influence the sanctioning ability of villages. Whether ethnicity is perceived as a relevant political factor and a driving force in dividing voters, might also affect how members of different ethnic groups interact in a local community. If voters are strongly divided by ethnic identity, then this division might also hamper inter-ethnic cooperation on the village level. On the contrary, if citizens do not vote according to their ethnic membership, then ethnic diversity in the village will not be perceived

⁷ A similar argument was made by Wantchekon (2003).

as a factor hindering inter-ethnic cooperation. Hence, in environments with non-politicized ethnicity, even very diverse villages might not suffer from the inability of sanctioning non-contributing parents (as predicted by H_1).

Empirical evidence supports the relationship between the sanctioning theory and politicized ethnicity. A comparison of the sanctioning mechanism in Kenya and Tanzania reveals significant variation of the effect of ethnic diversity on village funding (Miguel, 2004). While the effect was found to be strong and significant in Kenya, it was insignificant in Tanzania. Differences in the effect of ethnicity were attributed to different nation building policies pursued by the two countries. Miguel (2004) argues that while Tanzanian politicians emphasized unity, Kenyan politics was strongly divided by ethnic identities. One could therefore argue that the negative sanctioning effect in Kenya and the insignificant effect in Tanzania are caused by the politicization and the non-politicization of ethnicity in the two countries. From this one could posit the following hypothesis:

H₄: The sanctioning effect is more pronounced in countries with politicized ethnicity.

Closely linked to the role of the political relevance of ethnicity is the role of institutional quality. Easterly (2001) showed that the negative effect of national level ethnic diversity might be mitigated by good institutional quality. Although this interaction was established using a national level ethnic indicator, hence neglecting the underlying mechanisms, it provides some evidence for the relevance of institutional quality. The assumed negative impact of ethnic parties on a country's education outcomes might be negated by sound institutions. Moreover, stable institutions might provide a favorable environment for village funding, independent of a village's diversity and for an even distribution of government funds. Hence, institutional quality might also directly affect the impact of village sanctioning and clientelism in the education sector:

H_{5a,b}: High institutional quality leads to a diminished sanctioning (H_{5a}) and clientelism effect (H_{5b}).

3. Data and Variable Selection

Assessing the impact of ethnicity on education is particularly important for a country exhibiting very diverse populations. Almost all countries with high numbers of ethnic groups, are found in Africa (Easterly and Levine (1997), p. 1219). In addition, this region is still the least developed part of the world and in dire need of improving educational outcomes. This

paper will, therefore, focus on the impact of ethnicity on education in Africa and test the posited theories by using data for African countries.

Dependent Variable

Disaggregated data on district level educational outcomes for 31 African countries can be drawn from the Demographic and Health Survey (DHS) (MEASURE DHS, 2008). Data on school enrollment for household members between 6-10 years and 11-15 years is used from the part “Household characteristics” of the DHS survey. According to Gardner (1998), the official school age for primary education in the DHS surveys range from 6 to at least 10 years and hence enrollment for children aged 6-10 is used to capture the primary enrollment rate (**primary enrollment**). While pupils aged 10-15 years might also be still attending primary school, the majority is expected to have transited to secondary schools. Therefore, enrollment of pupils aged 10-15 is used to capture effects on **secondary enrollment**.⁸

School enrollment data is available at the district level with an average of 7 districts per country.⁹ DHS have been carried out in different time periods (the earliest surveys dating from 1991, Cameroon, and the most recent from 2006, Niger), and for several countries more than one DHS has been carried out. To use all available information, all surveys of African countries are included in this analysis.¹⁸

Explanatory Variables

To test the diversity, clientelism and politicization effect, three distinct measures of ethnicity are coded. A detailed description of the rules for coding can be found in Appendix G and H.

Ethnic Diversity

The **diversity** variable measures ethnic diversity of districts and is used to test the sanctioning theory.¹⁰ Information on existing ethnic groups and their location on the district level are

⁸ Note that the theoretical arguments in section 2 are based on education spending (by villages and the state), rather than education outcomes. However, due to lack of comparable and disaggregated data on education expenditure education outcomes are used to proxy education spending. Using enrollment rates to proxy education expenditure might only be problematic if enrollment is close to 100 percent, since then expenditure is likely to increase the quality and not the quantity of education (UNESCO, 2004; p.193). However, in the dataset less than 7 percent of the observations exhibit enrollment rates greater than 90 percent and hence enrollment rates seem a reasonable proxy for education expenditure.

⁹ According to Rabe-Hesketh and Skrondal (2008, p. 62), the number of clusters (61 countries) and size (average 7 districts) of the underlying dataset are sufficient to estimate country- and district level variations including random slopes (as presented in section 5).

¹⁰ Note that ideally, ethnic diversity should combine information on the number and size of ethnic groups in a village. Unfortunately, information on ethnic diversity is not available for villages and, hence, will be proxied by the district level.

drawn from Cunningham and Weidmann (2008), and Cederman, Rød, and Weidmann (2007). Ethnic diversity of districts is calculated using the ethno-linguistic fractionalization (ELF) measure and ranges from 0 (complete homogeneity) to 0.99 (maximum heterogeneity). Districts from Cunningham and Weidmann (2008) are carefully matched with the differing DHS districts (see Appendix G).

Presidents' Co-ethnics

The **co-ethnics** variable measures the district population shares of presidents' co-ethnics and is used to test the clientelism theory. Politics and especially state resource distribution in Africa are viewed to be primarily shaped by incumbent presidents (e.g. Kasara, 2007). Therefore, the ethnic group of the incumbent president is the one assumed to benefit primarily from state resource reallocation.¹¹ Ethnic identity of the incumbent president is coded as the identity that is politically relevant. Information on presidents' ethnic identity is drawn from Fearon, Kasara, and Laitin (2007) and district population shares of the presidents' co-ethnics are obtained from Cunningham and Weidmann (2008). The co-ethnics indicator is a continuous variable and shares of co-ethnics range from zero to 100 percent of district population shares (see Appendix H).

Politicization of Ethnicity

The **politicization** variable measures the political relevance of ethnicity and is used to test the theory of politicization. Referring to Fearon (2006), the political relevance of ethnicity can be measured by the degree to which political parties are formed along ethnic identity in contrast to parties being based on ideologies. Hence, the politicization indicator measures whether the majority of political parties (including the ruling parties) are based on ethnic identity rather than ideology or programs.

Information on the formation base on parties is drawn from the section "Political Participation" in the Polity IV Country Reports by Marshall and Jaggers (2008) for the respective countries. The information from Polity IV was used to code the politicization variable with outcome values zero, one and two. The value zero denotes that parties are entirely based on ideologies or programs. The value one denotes that voting for certain parties and a party's campaigning might evolve around ethnic identity. The value two denotes that

¹¹ In few cases the incumbent president cannot be considered as the politician governing state affairs. Following Kasara (2007), in those cases information on the effective leader was drawn from Goemans, Gleditsch, and Chiozza (2008).

there exists an ethnic party which is supported by distinct ethnic groups and which includes ethnic identity in its campaigning.

As an example, for Kenya in 1998, the Kenya African National Union (KANU) is coded as an ethnic party due to its “extremely narrow constituency bases”¹² as well as its “ethnically based patronage rule”¹². In comparison, Benin in 1996 is coded as exhibiting moderately ethnic parties due to existence of a clear “regional foundation”¹³ of political parties but lack of specific ethnic support bases. Almost 50 per cent of the countries studied do not exhibit characteristics of ethnic parties and are therefore coded as exhibiting programmatic parties (politicization=zero).

Control Variables

To ensure that the observed significant relationships between the various ethnic indicators and education are not caused by other countries’ characteristics, variables on countries’ institutional and economic background, as well as countries’ education system characteristics, presidents’ incumbency, and time dummies are included in the regression.

A country’s institutional quality is proxied using data from the Freedom House index, which measures political rights and civil liberties (Freedom House, 2008). This index is denoted as **institutions** and was re-coded into the range of -2.25 to 2.75 with positive values denoting strong institutions and negative values weak institutional quality.

Countries’ income (**income**) is particularly relevant for educational outcomes, with richer countries having more resources to distribute in the education sector. To capture variation on the district level within countries, income is approximated using the percentage of households in the district that use “piped water” as their primary source of drinking water (MEASURE DHS, 2008). This approach also takes into account that general GDP levels might be reversely caused by education as education is related to increased productivity and wages. In contrast, access to water is mainly determined by state-and district infrastructure spending and hence will be considered as exogenous. The aggregated piped water indicator on country level is highly correlated (0.73) with GDP (drawn from the World Development Indicator database, World Bank, 2008).¹⁴ Hence, access to piped water can be seen as an instrument for GDP and will be estimated in a reduced form equation.

¹² Marshall and Jagers, 2008, Polity IV Country Report 2003, Kenya

¹³ Marshall and Jagers, 2008, Polity IV Country Report 2003, Benin

¹⁴ The corresponding regression coefficient is highly significant at the 1 percent level and the R-square is 0.53 (not shown).

In addition, a count variable of the years the incumbent president held power (**president's incumbency**) is included in the estimation since the effect of presidents' co-ethnics might depend on the lengths of her incumbency.

Following Michaelowa and Weber (2007), general characteristics of countries' education systems, such as national education expenditure¹⁵, institutional quality, the share of the school-aged population, pupil-teacher ratio, are considered to influence education outcomes. Education expenditure, measured as current education expenditure in percent of GNI (**education expenditure**), is included in the estimation to control for the overall level of funds allocated to education. In addition, the share of school-age children (**children**) is used to account for the demand for schooling and includes children aged 0-14 years as a percentage of overall population. Furthermore, the pupil-teacher ratio in primary schools (**pupil-teacher ratio**) is included in the estimation of primary enrollment rates. Information on education expenditure; number of children and pupil-teacher ratios are available from the World Development Indicator (WDI) database (World Bank, 2006).¹⁶

Furthermore education outcomes might increase slowly over time. Therefore, dummies for three different five-year periods are included allowing for a general time trend over these periods. Dummy variables for the period 1990-1994 (**1990**), 1995-1999 (**1995**), and 2000-2006 (**2000**) are created and 1995 and 2000 are included in the estimation.¹⁷

4. Econometric Results

The dataset includes pooled observations of countries and multiple time points (for example Eritrea in 1995 and Eritrea in 2002)¹⁸ and variables on district and country level. Due to the two-level structure of the data, estimating a hierarchical model is required. Standard OLS assumes that observations are independent. If this is not the case standard errors are too small and effects might become spuriously significant. Because of the hierarchical structure of the dataset, the independency assumption is not met. More precisely, estimating the similarity of

¹⁵ Note that from the theoretical part, ethnicity is expected to affect education spending on the national level not by changing the absolute amount of money spent on education but by altering the distribution of funds over country' districts. Hence, including a measure for national education expenditure will not bias estimations of the clientelism effect.

¹⁶ The few missing values in pupil-teacher ratios and income are approximated by values for the respective variable of proximate years for the same country.

¹⁷ All variables except dummy variables are grand mean centered to allow meaningful values of the regression slopes and facilitate interpretation of interaction terms in the hierarchical estimation. Descriptive statistics and correlation matrix are presented in Appendix A and E.

¹⁸ Re-estimation of the final model including only the newest survey data per country (here: Eritrea 2000) yields smaller coefficients for diversity, but unchanged coefficients of all other variables (not show).

two districts in one country reveals that districts are highly correlated 0.78¹⁹, and most of the variation of enrolment rates in the district can be explained by country level grouping structure.

A second argument for the use of a hierarchical model (and against FE with clustered standard errors) is the proposed variation of the district level variables coefficient of diversity and clientelism over countries as posited in hypotheses H_3 - H_5 . Hierarchical models allow for the inclusion of these effects by estimating to which extent country and district level variables contribute to the variance of the dependent variable (Rabe-Hesketh and Skrondal, 2008). In a first step, variations over countries of variables are identified and random slopes for these variables containing a variable specific error term are included. In a second step, predictors of this randomness (besides the variable specific error term) are included. Since the district level variables, clientelism and sanctioning, are posited to depend on national level politicization and institutions, the interaction terms between those variables are included in the regression.

The final model for both primary and secondary enrollment was obtained estimating five succeeding models, which are presented in Appendix B and C and model specifications in Appendix D. Starting off with an intercept-only model (M1), a model including district level variables (M2) and country-level variables (M3) are estimated, before testing for random slopes (M4) and including predictors of these random slopes (M5). The final model presented below contains all significant effects found for district- and country-level variables as well as significant random slopes. The assumption of normality of the error terms was tested and outliers were excluded for both primary and secondary enrolment.²⁰ Estimation results reported in *Table 1* were obtained estimating the model without the outliers. Coefficients do not change substantially when outliers are excluded (see Appendix B and C, column 7) and differences in estimations are reported in footnotes under *Table 1*.

The sanctioning and clientelism hypothesis (H_1 and H_2) are tested by including **diversity** and **co-ethnics** as explanatory variables in the regression (equation 1). Testing the relation

¹⁹ Intra-class correlation for primary enrolment: $\rho = \sigma^2_{u0j(\text{Country-level})} / (\sigma^2_{u0j(\text{Country-level})} + \sigma^2_{eij(\text{District-level})})$
 $= 482.78 / (482.78 + 133.36)$, refer to Appendix B, M1.

Intra-class correlation for secondary enrolment: $\rho = \sigma^2_{u0j(\text{Country-level})} / (\sigma^2_{u0j(\text{Country-level})} + \sigma^2_{eij(\text{District-level})})$
 $= 443.85 / (443.85 + 119.10)$, refer to Appendix C, M1.

²⁰ Outlying countries were Nigeria (1990, 1999, and 2003) and Rwanda (2000). Rwanda might be an exception due to its small size and to the unusual high foreign aid inflows after the 1994 genocide (OECD/DAC, 2008). This increased level of aid might have boosted enrollment in the succeeding years. In the case of Nigeria, the problem might be the coding of ethnicity as adopted from Kasara (2007). Kasara coded the leaders in Nigeria as being members of "Middle Belt", which is the association of the administrative districts in the middle part of the country. This part is inhabited by various different ethnic groups, with one being the Hausa ethnic group. In contrast to Kasara, other researchers (Miles, 1987) report presidents in Nigeria to specifically belong to the Hausa group.

between **politicization**, **institution**, **diversity** and **co-ethnics** (H_3 , H_4 , and H_5) requires two steps. Since politicization and institutions are variables on the country-level, a test of their influence on diversity and co-ethnics requires estimating the variation of the diversity and co-ethnics variables over country. This means β_{1j} and β_{2j} (equation 1) are estimated as random slopes by including the error terms u_{1j} and u_{2j} (equation 3 and 4). As a second step, **politicization** and **institutions** are included as predictors for the randomness of the diversity and co-ethnics coefficient (equations 3 and 4).

The final model for primary and secondary enrollment is estimated as follows (insignificant coefficients are in brackets):

District level model: Enrollment at the district level (Enrollment_{ij}) depends on the ethnic diversity, the percentage of presidents' co-ethnics and the level of income plus district level error for district i in country j ²¹:

$$\text{Enrollment}_{ij} = \beta_{0j} + \beta_{1j}\text{diversity}_{ij} + \beta_{2j}\text{co-ethnics}_{ij} + \beta_{3j}\text{income}_{ij} + e_{ij} \quad (1)$$

Country level model: In addition, enrollment depends on country level variables, namely the politicization, institutional quality, interaction between institutions and politicization, national education expenditure and the share of the school-aged population. In addition, the error term u_{0j} captures unobserved country level heterogeneity:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{politicization}_j + \gamma_{02}\text{institutions}_j + \gamma_{03}\text{institutions}_j \times \text{politicization}_j + \gamma_{04}\text{education expenditure}_j + \gamma_{05}\text{children}_j + \gamma_{06}2000 + u_{0j} \quad (2)$$

Moreover, the impact of the district level variables diversity, co-ethnics and income on enrollment varies over countries, with parts of the variation being explained by politicization and institutions:

$$\beta_{1j} = \gamma_{10} + [\gamma_{11}\text{politicization}_j] + [\gamma_{12}\text{institutions}_j] + [u_{1j}] \quad (3)$$

$$\beta_{2j} = \gamma_{20} + [\gamma_{21}\text{politicization}_j] + [\gamma_{22}\text{institutions}_j] + u_{2j} \quad (4)$$

$$\beta_{3j} = \gamma_{30} + \gamma_{31}\text{politicization}_j + [\gamma_{32}\text{institutions}_j] + u_{3j} \quad (5)$$

²¹ Estimating the model using fixed effects with clustered standard errors on the country level does not change the district level coefficients substantially and hence supports the validity of the multilevel estimates (refer to Appendix B and C, last column).

4.1 Results for Primary Enrolment

Table 1, column 2, reports the results for primary enrollment.²¹ As theoretically expected, ethnic diversity is negatively related to enrollment and co-ethnics is positively related to enrollment. Both coefficients are strongly significant, with diversity on the 1 percent level and co-ethnics on the 5 percent level. This provides evidence to accept hypotheses 1 and 2. Changing from a complete homogeneous district (diversity=0) to a complete heterogeneous district (diversity = 0.99) is associated with a decrease in primary enrollment of 8.9 points²² on the scale of primary enrolment from 8.5 to 97.2. This is equivalent to a 10 percent decrease in primary enrolment.

In comparison, the maximum effect of co-ethnics is only half the size of the diversity effect. Changing from a district with zero co-ethnics (co-ethnics=0) to a district with 100 percent co-ethnics (co-ethnics=1) is associated with an increase in primary enrollment by 4.51 points²³ which is equivalent to an increase of 5 percent in primary enrollment rates.

Most interestingly, the coefficient of diversity does not appear to vary significantly over countries ($\sigma^2_{u1j(\text{Diversity})}$ is insignificant). This indicates that the effect of diversity is independent from country level variables, such as **politicization** and **institutions** and provides counterevidence to the theory. More precisely, these econometric results provide evidence to reject the hypothesis 4 and 5a, which postulated that the diversity effect depends on the level of politicization and institutional quality.

The random slope of the co-ethnic variable, in contrast, turns significant ($\sigma^2_{u2j(\text{Co-ethnics})}$). However, politicization and institutions are not able to explain variations of co-ethnics. Both interaction terms, Co-ethnicsXpoliticization and Co-ethnicsXinstitutions, appear insignificant in the estimation. This provides evidence to reject the hypothesis 3 and 5b, which posited a significant relationship between politicization, institutions and co-ethnics.

Besides the insignificant relation between politicization with diversity and co-ethnics, the interaction term of politicization with institutions and with income, turns out significant. Following Brambor, Clark, and Golder (2006) the interaction term between the three variables politicization, income and institutions is re-estimated and significance levels for all combination of the three variables are obtained. *Graph 1* depicts the marginal effect of politicization on primary enrollment as institutional quality and income varies.

²² Maximum effect of diversity=(max_diversity - min_diversity)* β_1 = (0.99-0)*-8.997.

²³ Maximum effect of co-ethnics=(max_co-ethnics - min_co-ethnics)* β_2 = (1-0)*4.522.

Table 1: Results for Primary and Secondary Enrollment

	Dependent variable: Primary enrollment	Dependent variable: Secondary enrollment
Fixed Part		
Diversity	-8.997*** (<0.01) ^{a)} (H_1)	-4.665** (0.033) ^{a)c)} (H_1)
Co-ethnics	4.522** (0.043) ^{a)} (H_2)	4.292** (0.049) ^{a)} (H_2)
Politicization	3.903 (0.184) ^{b)}	1.122 (0.719)
Co-ethnicsXpoliticization	insignificant ^{d)} (H_3)	insignificant ^{d)} (H_3)
DiversityXpoliticization	insignificant ^{d)} (H_4)	insignificant ^{d)} (H_4)
Co-ethnicsXinstitutions	insignificant ^{d)} (H_5)	insignificant ^{d)} (H_5)
DiversityXinstitutions	insignificant ^{d)} (H_5)	insignificant ^{d)} (H_5)
Institutions	4.417* (0.052) ^{a)}	2.346 (0.208) ^{a)}
InstitutionsXpoliticization	-4.955** (0.019) (H_5)	-3.835* (0.086) (H_5)
IncomeXpoliticization	-0.073* (0.079)	-0.040 (0.352)
Income	0.299*** (<0.01) ^{a)}	0.218*** (<0.01) ^{a)}
Education expenditure	4.335*** (<0.01) ^{a)}	4.236*** (<0.01) ^{a)}
Children	-1.681** (0.018) ^{a)}	-0.896 (0.132) ^{a)}
2000	7.637 (1.09)	9.098* (0.076)
Random Part		
σ^2_{eij} (District level)	84.41	87.22
σ^2_{u0j} (Country level)	294.30*** (<0.01)	329.37*** (<0.01)
σ^2_{u1j} (Diversity)	insignificant	Insignificant
σ^2_{u2j} (Co-ethnics)	116.56*** (<0.01)	102.63* (0.086)
σ^2_{u3j} (Income)	0.020 (0.114)	0.023* (0.073)
Wald	chi2(10)=116.58 (<0.01)	chi2(10)=62.10 (<0.01)
Log restricted-likelihood	-1621.19	-1630.44
N	418	418
Countries	61	61

P values in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

a) p-values for directed hypotheses.

b) When outliers are included politicization turns significant at the 5 percent level (Appendix B, column 7).

c) When outliers are included diversity is no more fully significant at the 10 percent level (p-value=0.103) (Appendix C, column 7).

d) The random slope for diversity did not turn significant (σ^2_{u1j} (Diversity)), which provides evidence that national level politicization and institutions do not change the coefficient of diversity. Co-ethnics is found to significantly vary over countries (σ^2_{u2j} (Co-ethnics)) and interaction term Co-ethnicsXpoliticization and Co-ethnicsXinstitutions were included as predictors of this randomness (see Appendix B and C, model M5).

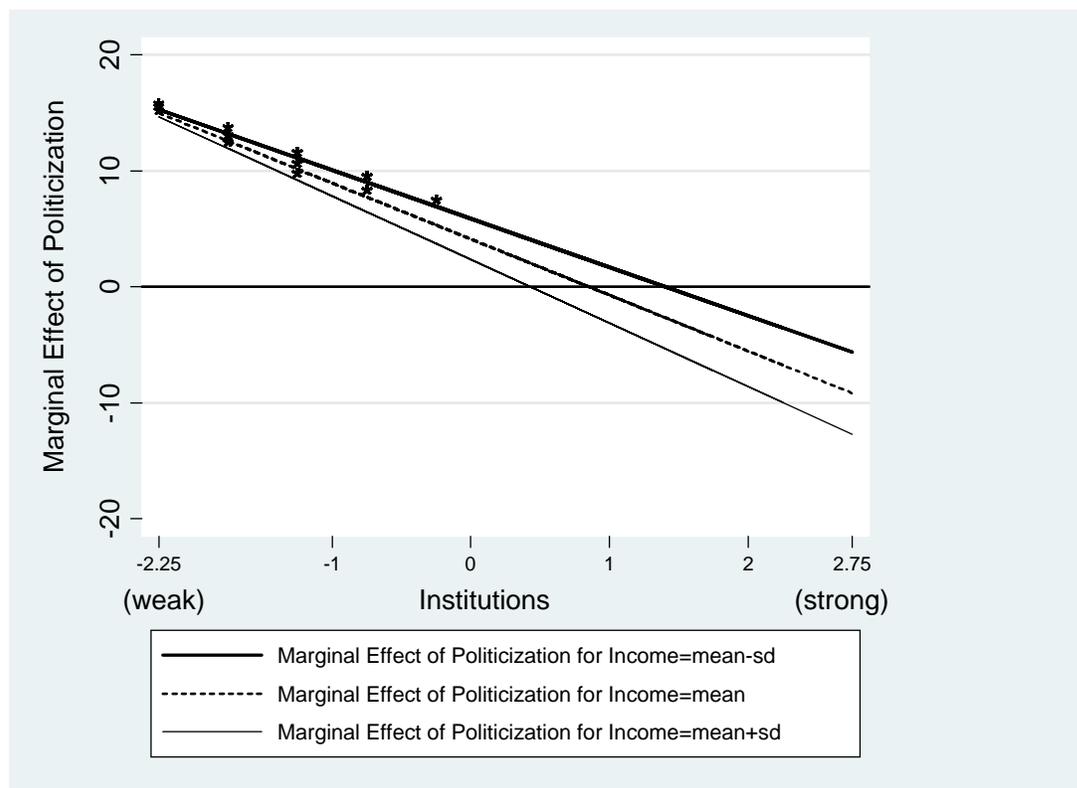
Constant term is not presented here.

Results obtained using a restricted maximum likely method and independent covariance structures.

For an overview of variable definitions and sources, see Appendix A.

The marginal effect of politicization is depicted on the y-axis. The x-axis depicts institutional quality, which varies from weak (-2.25) to strong institutions (2.75). In addition, the marginal effect of politicization is estimated for three typical income levels (mean income, and one standard deviation above and below the mean), which is depicted by the three different lines in *Graph 1*. The stars denote significance at the 95 percent level.

Graph 1: Marginal Effect of Politicization on Primary Enrollment as Institutions and Income Change



* denotes significance at the 95% level.

sd denotes standard deviation.

From *Graph 1*, one can infer that countries with average and good institutional quality (0-2.75) do not exhibit significant marginal effects of politicization in none of the three income levels.²⁴ However, once institutional quality declines (institutions<0), politicization seems to have a significantly *positive* influence on enrolment. Indeed, the change from programmatic to ethnic parties in countries with weak institutions is associated with an increase in enrollment rates by more than a 30 percent. For countries with low income, the marginal

²⁴ Estimating the marginal effect for countries with highest institutional quality (South Africa) reveals a significant negative effect of politicization (see Appendix F, category IV).

effect of politicization turns significant if institutional quality is below average (institutions<0). For wealthier countries, the marginal effect of politicization only turns significant for very low institutional quality (institutions<-1.3). It seems, therefore, that politicized ethnicity compensates for negative effects of weak institutions on enrolment rates.²⁵ Note that correlation between politicization and institutions is negligible (0.022) which provides support for the independence of the measures of politicization and institutions.

Control variables of a country's education system (education expenditure, children) enter significantly in the regression. Furthermore, the dummy for the period 2000-2006 remains insignificant. The Wald test indicates a good general fit of the final model and log-restricted likelihood increases from the model including only the intercept (M1), to the final model including district and country-level variables and random slopes (see Appendix B).

4.2 Results for Secondary Enrollment

Results for secondary enrollment are presented in *Table 1*, column 3.²¹ As for primary enrollment, one finds a negative effect of diversity and a positive effect of co-ethnics on secondary enrollment, both significant at the 5 percent level. This provides evidence to accept hypothesis 1 and 2. The maximum effect of ethnic diversity equals 4.61 points²⁶, which corresponds to an increase in secondary enrollment of 5.3 percent. With respect to the impact of diversity on primary enrollment this effect is comparatively low.

While the effect of diversity differs from primary to secondary education, the effect of co-ethnics seems to almost be the same for the different enrollment levels. For secondary enrollment, one finds that the maximum effect of co-ethnics is 4.3 points²⁷. This is equivalent to an increase in secondary enrollment by 5 percent.

Testing the hypotheses 4 and 5 by including random slopes for the diversity and co-ethnics variable reveals insignificant variation of diversity over countries ($\sigma^2_{u1j(\text{Diversity})}$ is insignificant). This indicates that none of the country level variables, such as politicization and institutions, significantly changes the diversity coefficient. Therefore, hypotheses 4 and 5a can be rejected.

The co-ethnics coefficient, however, exhibits significant variation over countries ($\sigma^2_{u2j(\text{Co-ethnics})}$ is significant). Trying to explain some of the variation of co-ethnics, by

²⁵ Appendix F depicts the country-by-country plots of the marginal effects of all countries where the interaction terms turned significant in *Graph 1*.

²⁶ Maximum effect of diversity=(max_diversity - min_diversity)* β_1 = (0.99-0)*-4.665.

²⁷ Maximum effect of co-ethnics=(max_co-ethnics - min_co-ethnics)* β_2 = (1-0)*4.292.

including **politicization** and **institutions** as predictors (by including the interaction terms), fail. This provides evidence to reject the hypothesis 3 and 5a as politicization and institutions cannot explain the variation of co-ethnics over countries.

In addition, the interaction between politicization and institutional quality turns significant. However, re-estimating the marginal effects of politicization on secondary enrollment for varying degrees of institutions and income (similar to *Graph 1*) reveals insignificant effects (not shown).

Countries' national education expenditure and income have significant positive effects on secondary enrolment and income does vary significantly over countries ($\sigma^2_{u3j(\text{Income})}$). However, while the share of the school-aged population is not significantly related to secondary education, there is a significant time effect. In comparison to the omitted time periods, a country's secondary enrollment in the period 2000-2006 is increased. The restricted-log likelihood and the Wald test indicate a good general fit of the final model.

5. Discussion

Summing up, the econometric results support hypothesis 1 and 2, reject hypothesis 3, 4 and 5, and provide interesting insights into the role of politicization on enrollment rates.

Hypothesis 1 was supported by showing a significant negative coefficient of the **diversity** variable in *Table 1*. Communities being composed of several ethnic groups exhibit significantly lower enrollment rates. This lends strong evidence for the validity of the sanctioning theory by Miguel and Gugerty (2005). Heterogeneous communities seem, indeed, to suffer from a collective action problem, namely the inability to sanction non-contributing parents. This inability leads to lower school finances which translate into significantly lower educational outcomes.

As seen in *Table 1*, the coefficient of diversity is much smaller for secondary than for primary enrollment. Since secondary enrollment comprises children going to primary and secondary school, the difference in coefficients might be driven by the portion of children attending secondary schools. Data from developing countries reveals that there are substantial differences in the financing resources for primary and secondary education.²⁸ Village funding might be particularly important for primary and less important for secondary education. Primary schools are mainly located in one village, while the fewer secondary schools belong to various communities. While village funding is particularly relevant for primary education,

²⁸ Households' contribution to primary education as percent of total expenditure ranges from about 20 percent of total education expenditure (Malawi) to nearly 50 percent (Zambia), see UNESCO (2008), Figure 4.5, p.151.

secondary schooling might be mostly funded by the government. Hence, the ethnic composition of a village plays a minor role for secondary school funding and this might have resulted in smaller coefficients of the diversity variable for secondary enrollment.

The second hypothesis (H_2) was supported by a significant positive coefficient of the **co-ethnics** variable in *Table 1*. Districts with high shares of president's co-ethnics exhibit comparably higher enrollment rates. This provides strong evidence for the clientelism theory. Incumbent politicians seem to distribute state education resources primarily to their ethnic clientele, which then translates into higher enrollment rates of president's co-ethnics.

Hypothesis 3 and 5b concerning the relation between clientelism with politicization and with institutions was not supported by the econometric results. The coefficient of clientelism varies substantially over countries, which was shown by the significant slope variance ($\sigma^2_{u2j(\text{Co-ethnics})}$ is significant).²⁹ However, this variation could neither be explained by politicization nor by institutions, demonstrated by the insignificant interaction terms between clientelism with politicization and institutions. This provides evidence for an effect of clientelism that is independent from the degree of politicization and institutional quality. Whether ethnicity is politically relevant does not substantially influence clientelistic distributions of state education resources. This lends support that policies targeted at reductions of the political relevance of ethnicity, for example through nation building policies (Miguel, 2004), will not be able to mitigate the negative effect of the clientelistic distribution.

The role of politicization and institutions was also tested for the sanctioning effect. Again, the hypotheses (H_3 and H_{5a}) are not supported by the econometric results as shown by the insignificant slope variance of the diversity variable ($\sigma^2_{u1j(\text{Diversity})}$ is insignificant). This strongly rejects Miguel's (2004) notion that the effect of ethnic diversity differs over countries. Miguel argues that countries with strong nation building in the past (as in Tanzania) do not suffer from ethnic diversity. This is strongly rejected by the econometric results, which do not find a significant impact of politicization on the diversity coefficient. Ethnic diversity seems to negatively affect education outcomes independent of the political relevance of ethnicity. This challenges Miguel's (2004) notion that nation building policies, designed to decrease the relevance of ethnicity in the political process, might mitigate the negative diversity effect.

Besides the insignificant influence of politicization on sanctioning and clientelism, the econometric results provided interesting insights into the direct influence of politicization on enrollment rates. As demonstrated by the significant marginal effects of politicization in

²⁹ Similar results are found by Franck and Rainer (2009).

Graph 1, politicization is found to significantly influence primary education in countries with weak institutions. Countries with average and below average institutional quality benefit from higher degrees of politicization. For countries with very weak institutions and low income, higher politicization is associated with increases in enrollment of over 30 percent. Politicization is measured by the existence of ethnic parties. In contrast to programmatic parties, ethnic parties seem to perform better in worse environments. They seem to maintain a minimum level of education spending resulting in increased enrollment. This lends evidence to the idea that ethnic parties depend more critically on rewarding their members for their support than programmatic parties. While programmatic parties might have other means to maintain a positive relationship with their voters, ethnic parties seem to depend mainly on distribution of state resources. This explains why even in worst environments, ethnic parties are associated with increased enrollment rates.

6. Conclusion

Countries with a high number of ethnic groups are seen to bear a particular high burden in providing public goods. Ethnic diversity is assumed to downsize the amount of local public goods provided. However, a clear understanding of why this is happening has been, so far, not discovered. Detecting the channels through which ethnicity influences public good provision is, therefore, crucial to define policies helping to overcome the burden of ethnic diversity.

This paper contributes to these efforts by providing a test of the underlying mechanisms driving the effect of ethnicity on education provision. In particular, the study identified the sanctioning, clientelism and politicization theory on education provision. Econometric results strongly support the relevance of these theories for primary and secondary education provision in Africa. In particular, heterogeneous communities seem to suffer from increased ethnic diversity by their inability to raise sufficient funds for schooling. Lower school funding then translates into substantially lower enrollment rates. Hence, policies targeted at supporting village fundraising activities through strengthening the local community's organizational structure would substantially improve enrollment rates.

In addition, the econometric results point to an unequal distribution of state resources from presidents primarily to their ethnic clientele. Indeed, members of the president's ethnicity exhibit significantly higher enrollment rates than members of other ethnic groups. This effect, however, varies substantially over countries and lends credit to the idea that clientelistic distribution of state funds might be influenced by other factors still uncovered. In the econometric analysis, neither politicization nor institutional quality or income was found

to explain the variation of clientelistic distribution. It is left to future research to establish more detailed understanding of the mechanisms at work.

Furthermore, the econometric results provide interesting insights into the role of politicization of ethnicity. The relevance of ethnicity in the political process is neither found to influence village fund raising nor state distribution of education resources. This clearly contrasts earlier findings, which proposed investing in nation-building policies as a cure to the ethnicity problem (Miguel, 2004). Generally, nation building policies denote efforts to unite the different ethnic groups in a country and create a united identity (for example by introducing a lingua franca, such as Swahili in Tanzania). This might then lessen the political relevance of ethnicity. Unfortunately, such policies seem neither to effectively support village fundraising nor to lead to a more equal distribution of state resources.

A second intuition derived from the econometric estimation regarding politicization points to a positive influence of politicization on enrollment rates in countries with weak institutions. In contrast to programmatic parties, ethnic parties seem to maintain a minimum level of education spending even in adverse environments. Further research should investigate the relationship between institutional quality and ethnic parties and derive a clearer understanding of the interactions between politicians and voters which seem to differ in programmatic and ethnic parties.

References

- Africa Almanac. (2008). Africa's Historical Timeline. Retrieved August 8, 2008, from <http://www.africaalmanac.com/history.html>
- Alesina, A., and Drazen, A. (1991). Why Are Stabilizations Delayed? *American Economic Review* , 81 (5), pp. 1170-1188.
- Alesina, A., and La Ferrara, E. (2002). Participation in Heterogeneous Communities. *Quarterly Journal of Economics* , 85 (2), pp. 207-234.
- Alesina, A., Baqir, R., and Easterly, W. (1999). Public Goods and Ethnic Divisions. *World Bank Policy Research Working Paper* (2108).
- Antoun, N., and Campling, L. (2008). *John Agyekum Kufuor Biography - The Development of a Political Mind, In and Out of National Politics*. Retrieved August 9, 2008, from <http://biography.jrank.org/pages/2864/Kufuor-John-Agyekum.html>
- Barkan, J. D., and Chege, M. (1989). Decentralising the State: District Focus and the Politics of Reallocation in Kenya. *Journal of Modern African Studies* , 27 (3), pp. 431-453.
- Bruk, S. I. (1964). *Atlas Narodov Mira*. Moscow: Academy of Science USSR.
- Campling, L. (2008). *Paul Kagame Biography - A Refugee's Childhood, Uganda and the National Resisitance Movement*. Retrieved August 8, 2008, from <http://biography.jrank.org/pages/2859/Kagame-Paul.html>
- Cederman, L. E., Rød, J. K., and Weidmann, N. (2007). *Geo-Referencing of Ethnic Groups: Creating a New Dataset*. Chicago, IL: paper presented at the Annual Convention of the International Studies Association.
- Chandra, K. (2004). *Why Ethnic Parties Succeed - Patronage and Ethnic Head Counts in India*. Cambridge: Cambridge University Press.
- CIA. (2008). *CIA - World Factbook*. Retrieved August 7, 2008, from <http://www.cia.gov/library/publications/the-world-factbook/>
- Clapham, C. (1982). Clientelism and the State. In C. Clapham, *Private Patronage and Public Power: Political Clientelism and the Modern State*. London: Frances Printer.
- Clark, W. R., Doces, J. A., and Woodberry, R. D. (2006). *Aid, Protestant Missionaries, and Growth*. paper presented at the Leitner Political Economy Workshop, Yale University.
- Cochrane, S. H. (1979). *Fertility and Education: What Do We Really Know?* Baltimore: John Hopkins University Press.
- Collier, P., and Garg, A. (1999). On Kin Groups and Wages in the Ghanaian Labour Market. *Oxford Bulletin of Economics and Statistics* , 61 (2), pp. 133-151.

- Cunningham, K., and Weidmann, N. B. (2008). *Shared Space: Ethnic Groups, State Accommodation and Local Conflict*. San Francisco, CA: paper presented at the Annual Meeting of the International Studies Association.
- Cutler, D. M., Elmendorf, D. W., and Zeckhauser, R. J. (1993). *Demographic Characteristics and the Public Bundle*. NBER Working Paper No. 4283.
- Easterly, W. (2001). Can Institutions Resolve Ethnic Conflict? *Economic Development and Cultural Change* , 49 (4), pp. 687-706.
- Easterly, W., and Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics* , 112 (4), pp. 1203-1250.
- Fearon, J. D. (2006). Ethnic Mobilization and Ethnic Violence. In B. R. Weingast, and D. Wittman, *Oxford Handbook of Political Economy*. Oxford: Oxford University Press.
- Fearon, J. D., and Laitin, D. D. (1996). Explaining Interethnic Cooperation. *American Political Science Review* , 90 (4), pp. 715-735.
- Fearon, J. D., and Laitin, D. D. (2000). Violence and Social Construction of Ethnic Identity. *International Organization* , 54 (4), pp. 845-877.
- Fearon, J. D., Kasara, K., and Laitin, D. D. (2007). Ethnic Minority Rule and Civil War Onset. *American Political Science Review* , 101 (1), pp. 187-193.
- Freedom House . (2008). *Freedom House Country Ratings*. Retrieved July 11, 2008, from <http://www.freedomhouse.org/ratings/index.htm>
- Gardner, R. (1998). *Education, DHS Comparative Studies* (Vol. 29). Calverton, Maryland: Macro International Inc.
- Glewwe, P. W. (1998). Why Does Mother's Schooling Raise Child Health in Developing Countries? Evidence from Morocco. *Journal of Human Resources* , 1 (34), pp. 124-159.
- Goemans, H. E., Gleditsch, K. S., and Chiozza, G. (2008). *ARCHIGOS - A Data Set on Leaders 1875-2004, Version 2.9*. Retrieved August 8, 2008, from <http://mail.rochester.edu/~hgoemans/Archigos.pdf>
- Goldin, C., and Katz, L. F. (1999). Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910-1940. *Journal of Interdisciplinary History* , 29 (4), pp. 683-723.
- Gordon, R. G. (2008). *Ethnologue: Languages of the Word, Fifteenth Edition*. Retrieved August 8, 2008, from <http://www.ethnologue.com/>

- Habyarimana, J. P., Humphreys, M., Posner, D. N., and Weinstein, J. M. (2007). Why Does Ethnic Diversity Undermine Public Goods Provision? *American Political Science Review* , 101 (4), pp. 709-725.
- Hanushek, E. A., and Kimko, D. D. (2000). Schooling, Labor-Force Quality, and the Growth of Nations. *American Economic Review* , 90 (5), pp. 1184-1208.
- Haruna, I. (2004). *The Middle Belt: History and Politics*. Retrieved August 9, 2008, from Nasaraw State Weekly Newspaper:
<http://www.nasarawastate.org/newsday/news/culture/11129114540.html>
- Horowitz, D. L. (1985). *Ethnic Groups in Conflict*. Berkley, University of California Press.
- Kasara, K. (2007). Tax Me If You Can: Ethnic Geography, Democracy, and the Taxation of Agriculture in Africa. *American Political Science Review* , 101 (1), pp. 159-172.
- Keefer, P. (2005). *Clientelism, Credibility and the Politics of Young Democracies*. paper presented at the Annual Meeting of the American Political Science Association, Washington D.C.
- Kelly, M. J. (2000). *The Encounter between HIV/AIDS and Education*. Harare, Zimbabwe: UNESCO, Sub-Regional Office for Southern Africa.
- Kimenyi, M. (2006). Ethnicity, Governance and the Provision of Public Goods. *Journal of African Economies* , 15 (AERC Supplement 1), pp. 62-99.
- Lemarchand, R. (1972). Political Clientelism and Ethnicity in Tropical Africa: Competing Solidarities in Nation-Building. *American Political Science Review* , 66 (1), pp. 68-90.
- Mankiw, N. G., Romer, D., and Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* , 107 (2), pp. 407-437.
- Marcus, R. R. (2004). *Political Change in Madagascar: Populist Democracy or Neopatrimonialism by Another Name?* South Africa: Institute for Security Studies (ISS) Paper No. 89.
- Marshall, M. G., and Jaggers, K. (2008). *Polity IV Country Reports*. Retrieved August 8, 2008, from <http://www.systemicpeace.org/polity/polity06.htm>
- Mauro, P. (1995). Corruption and Growth. *Quarterly Journal of Economics* , 110 (3), pp. 681-712.
- MEASURE DHS. (2008). Demographic and Health Surveys. Retrieved March 8, 2008, from <http://www.statcompiler.com/>
- Michaelowa, K. (2000). *Returns to Education in Low Income Countries: Evidence for Africa*. paper presented at the Annual meeting of the Committee on Developing Countries of the German Economic Association, Hamburg.

- Michaelowa, K., and Weber, A. (2007). Aid Effectiveness in the Education Sector: A Dynamic Panel Analysis. In S. Lahiri, *Theory and Practice of Foreign Aid* (pp. 357-385). Amsterdam: Elsevier.
- Miguel, E. (2004). Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania. *World Politics* , 56 (3), pp. 327-362.
- Miguel, E., and Gugerty, M. K. (2005). Ethnic Diversity, Social Sanctions, and Public Goods in Kenya. *Journal of Public Economics* , 89 (11-12), pp. 2325-2368.
- Miguel, E., and Zaidi, F. (2003). *Do Politicians Reward Their Supporters? Regression Discontinuity Evidence from Ghana*. Retrieved June 8, 2008, from http://www.sscnet.ucla.edu/polisci/wgape/papers/2_MiguelZaidi.pdf
- Miles, W. F. (1987). Partitioned Royalty: The Evolution of Hausa Chiefs in Nigeria and Niger. *Journal of Modern African Studies* , 25 (2), pp. 233-258.
- Minorities at Risk Project. (2005). *Assessment for Black Moors in Mauritania*. Retrieved April 7, 2008, from <http://www.cidcm.umd.edu/mar/assessment.asp?groupId=43502>
- OECD/DAC. (2008). *International Development Statistics (IDS) Databank*. Retrieved June 8, 2008, from <http://www.oecd.org/dataoecd/50/17/5037721.htm>
- Posner, D. N. (2005). *Institutions and Ethnic Politics in Africa*. Cambridge: Cambridge University Press.
- Poterba, J. M. (1997). Demographic Structure and the Political Economy of Public Education. *Journal of Policy Analysis and Management* , 16 (1), pp. 48-66.
- Rabe-Hesketh, S., and Skrondal, A. (2008). *Multilevel and Longitudinal Modeling Using Stata*. Texas: StataCorp LP.
- Rainer, I., and Franck, R. (2009). *Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education and Health in Sub-Saharan Africa*. paper presented at the Annual Meeting of the Public Choice Society March 2009. Las Vegas.
- STATOIDS. (2008). *Administrative Divisions of Countries Website*. Retrieved June 26, 2008, from <http://www.statoids.com/statoids.html>
- UNESCO. (2008). *EFA Global Monitoring Report, Education for All by 2015: Will We Make It?* Retrieved June 9, 2008, from <http://unesdoc.unesco.org/images/0015/001547/154743e.pdf>
- UNESCO. (2004). *EFA Global Monitoring Report, From Targets to Reform: National Strategies in Action*. Retrieved June 10, 2008, from www.unesco.org/education/efa_report/2003_pdf/chapter5.pdf

Vigdor, J. L. (2004). Community Composition and Collective Action: Analyzing Initial Mail Response to the 2000 Census. *Review of Economics and Statistics* , 86 (1), pp. 303-312.

Wantchekon, L. (2003). Clientelism and Voting Behavior - Evidence from a Field Experiment in Benin. *World Politics* , 55 (3), pp. 399-422.

World Bank. (2006). *World Development Indicators CD-ROM* . Washington D.C.

World Bank. (2008). *World Development Indicators CD-ROM* . Washington D.C.

Appendix A: Variables and Descriptive Statistics

Variable name	Definition	Sources	Level	Mean	Standard Deviation	Min	Max	Obs.
Primary enrollment	Average district school enrolment for household members between 6-10 years	MEASURE DHS (2008)	District	55.62	24.30	8.5	97.2	418
Secondary enrollment	Average district school enrolment for household members between 11-15 years	MEASURE DHS (2008)	District	66.6	21.80	11.9	98.3	418
Diversity	Probability that two randomly drawn individuals in the same district are members of different ethnic groups (range:0-1)	Cunningham and Weidmann (2008)	District	.0019	.2521	-.409	.427	418
Co-ethnics	Population-share of leader's ethnic co-members in the district (range: 0-1)	Cunningham and Weidmann (2008), Kasara (2007), Goemans <i>et al.</i> (2008)	District	.0029	.379	-.288	.711	418
Politicization	Equals 0 if parties are not based on ethnicity; equals 1 if parties are partially based on ethnicity; equals 2 if parties are strongly based on ethnicity	Marshall and Jagers (2008)	Country	.660	.660	0	2	418
Institutions	Freedom House index of political rights and civil liberties (-2.25 to 2.75, whereby 2.75 denotes the highest degree of freedom)	Freedom House (2008)	Country	0	1.29	-2.25	2.75	418
Income	Average district percentage of households that have access to piped water	MEASURE DHS (2008)	District	.00238	24.60	-31.87	67.22	418
Education expenditure	National education expenditure (percent of GNI)	WDI (World Bank, 2006)	Country	.0609	1.66	-1.67	6.28	418
Children	Population aged 0-14 years (percent of total population)	WDI (World Bank, 2006)	Country	-.1140	3.43	-12.29	6.09	418
Pupil-teacher ratio	Pupil-teacher ratio in primary education	WDI (World Bank, 2006)	Country	3902	12.79	-21.76	38.05	418
President's incumbency	Incumbency of political leader (in years)	Various sources	Country	.384	8.17	-10.68	23.81	418

Note: Descriptive statistics are reported for grand mean centered variables and excluding outliers.

Appendix B: Results for Primary Enrollment

Dependent variable: Primary enrollment	M1: Intercept- only	M2: district level variables	M3: country level variables	M4: M3+random slopes	M5: M4+cross-level interactions	Final model including outliers	FE (clustered s.e.)
Fixed Part							
Diversity		-8.870*** (<0.01) ^{a)}	-8.906*** (<0.01) ^{a)}	-8.967*** (<0.01) ^{a)}	-8.607*** (<0.01) ^{a)}	-8.248*** (<0.01) ^{a)}	-8.346*** (0.008) ^{a)}
Co-ethnics		5.723*** (<0.01) ^{a)}	5.423*** (<0.01) ^{a)}	4.898** (0.033) ^{a)}	4.273 (0.112) ^{a)}	4.654*** (0.039) ^{a)}	6.601** (0.018) ^{a)}
Income		0.240*** (<0.01) ^{a)}	0.232*** (<0.01) ^{a)}	0.248*** (<0.01) ^{a)}	0.235*** (0.001) ^{a)}	0.291*** (<0.01) ^{a)}	0.231*** (<0.01) ^{a)}
DiversityXco-ethnics		-0.413 (0.956)					
DiversityXincome		-0.049 (0.650)					
Co-ethnicsXincome		0.035 (0.649)					
President's incumbency			0.392 (0.157)		0.366 (0.189)		
Politicization			2.442 (0.414)	4.290 (0.139)	3.045 (0.295)	5.238** (0.047)	
Children			-1.763** (0.011) ^{a)}	-1.712** (0.011) ^{a)}	-1.243* (0.053) ^{a)}	-1.788*** (<0.01) ^{a)}	
Education expenditure			4.618*** (<0.01) ^{a)}	4.377*** (<0.01) ^{a)}	4.771*** (<0.01) ^{a)}	3.515*** (<0.01)	
Pupil-teacher ratio			0.285 (0.185)				
Institutions			3.674* (0.081) ^{a)}	4.333* (0.054) ^{a)}	4.274* (0.059) ^{a)}	3.965* (0.065)	
InstitutionsXpoliticization			-4.029** (0.049)	-5.079** (0.015)	-4.413** (0.036)	-4.253** (0.035)	
InstitutionsXincome					-0.033 (0.246)		
PoliticizationXincome					-0.075* (0.089)	-0.077* (0.062)	
Co-ethnicsXpresident's incumbency					-0.354 (0.366)		
IncomeX1995					0.131 (0.181)		
IncomeX2000					0.064 (0.507)		
Co-ethnicsXpoliticization					0.800 (0.807)		
Co-ethnicsXinstitutions					-1.805 (0.408)		
1995			4.259 (0.485)		4.553 (0.463)		
2000			7.643 (0.197)		9.956 (0.093)	5.922 (0.189)	
Constant	57.15*** (<0.01)	56.89*** (<0.01)	49.51*** (<0.01)	54.02*** (<0.01)	48.65*** (<0.01)	50.78*** (<0.01)	55.61*** (<0.01)
Random Part							
$\sigma^2_{\epsilon ij}$ (district level)	133.36	102.85	102.18	84.07	83.92	123.01	102.14
σ^2_{u0j} (country level)	482.78*** (<0.01)	398.50*** (<0.01)	276.55*** (<0.01)	287.98*** (<0.01)	276.91*** (<0.01)	273.57*** (<0.01)	422.32
σ^2_{u1j} (Diversity)				insignificant	insignificant	insignificant	
σ^2_{u2j} (co-ethnics)				126.06*** (<0.01)	126.99*** (<0.01)	77.47 (0.771)	
σ^2_{u3j} (income)				0.0237*** (<0.01)	0.0244*** (<0.01)	0.0098 (0.423)	
Wald (F-test for FE)		chi2(6)=124.8 (<0.01)	chi2(12)=163.9 (<0.01)	chi2(8)=110.0 (<0.01)	chi2(18)=122.7 (<0.01)	chi2(10)=114.6 (<0.01)	F(3,60)=19.8 (<0.01)
Log restricted-likelihood	-1708.65	-1652.79	-1627.27	-1624.24	-1617.77	-1785.92	
R ² (within)							0.239
R ² (between)							0.205
N	418	418	418	418	418	445	418
Countries	61	61	61	61	61	65	61

P values in parentheses;* p<0.1. ** p<0.05. *** p<0.01; a) p-values for directed hypotheses.

Results obtained using a restricted maximum likely method and independent covariance structures.

For an overview of variable definitions and sources, see Appendix A.

Appendix C: Results for Secondary Enrollment

Dependent variable: Secondary enrollment	M1: Intercept- only	M2: district level variables	M3: country level variables	M4: M3+random slopes	M5: M4+cross-level interactions	Final model including outliers	FE (clustered s.e.)
Fixed Part							
Diversity		-5.036** (0.033) ^{a)}	-4.490** (0.045) ^{a)}	-4.696** (0.032) ^{a)}	-4.524** (0.039) ^{a)}	-3.706 (0.103) ^{a)}	-3.840* (0.079) ^{a)}
Co-ethnics		4.674** (0.024) ^{a)}	5.082*** (0.006) ^{a)}	4.603** (0.035) ^{a)}	5.102* (0.068) ^{a)}	4.598** (0.040) ^{a)}	6.725** (0.012) ^{a)}
Income		0.170*** (<0.01) ^{a)}	0.161*** (<0.01) ^{a)}	0.193*** (<0.01) ^{a)}	0.193*** (<0.01) ^{a)}	0.209*** (<0.01) ^{a)}	0.158*** (<0.01) ^{a)}
DiversityXco-ethnics		-5.229 (0.486)					
DiversityXincome		0.026 (0.809)					
Co-ethnicsXincome		0.015 (0.851)					
President's incumbency			0.137 (0.652)		0.248 (0.409)		
Politicization			1.258 (0.691)		0.318 (0.918)	2.782 (0.331)	
Children			-0.882 (0.136) ^{a)}			-1.257 (0.054) ^{a)}	
Education expenditure			4.137*** (<0.01) ^{a)}	4.076*** (<0.01) ^{a)}	4.212*** (<0.01) ^{a)}	3.203** (0.016) ^{a)}	
Institutions			2.217 (0.225) ^{a)}		-0.939 (0.328) ^{a)}	2.793 (0.326)	
InstitutionsXpoliticization			-3.513 (0.122)			-3.416 (0.119)	
InstitutionsXincome					-0.010 (0.722)		
PoliticizationXincome					-0.045 (0.327)	-0.047 (0.286)	
Co-ethnicsXpresident's incumbency					-0.133 (0.732)		
IncomeX1995					0.083 (0.412)		
IncomeX2000					0.010 (0.923)		
Co-ethnicsXpoliticization					-0.950 (0.768)		
Co-ethnicsXinstitutions					0.128 (0.953)		
1995			4.458 (0.509)		6.305 (0.352)		
2000			11.460* (0.075)	10.467** (0.035)	13.822** (0.030)	6.998 (0.153)	
Constant	64.76*** (<0.01)	64.44*** (<0.01)	56.81*** (<0.01)	59.84*** (<0.01)	55.75*** (<0.01)	59.61*** (<0.01)	66.65*** (<0.01)
Random Part							
$\sigma_{\epsilon ij}^2$ (District level)	119.10	104.82	104.00	87.21	87.57	125.47	103.78
σ_{u0j}^2 (Country level)	443.85*** (<0.01)	404.98*** (0.01)	347.25*** (<0.01)	335.81*** (<0.01)	348.74*** (<0.01)	326.99*** (<0.01)	439.05
σ_{u1j}^2 (Diversity)				insignificant	insignificant	insignificant	
σ_{u2j}^2 (Co-ethnics)				97.13*** (<0.01)	108.46*** (<0.01)	64.36 (0.298)	
σ_{u3j}^2 (Income)				0.0245*** (<0.01)	0.0271*** (<0.01)	0.0152 (0.247)	
Wald (F-test for FE)	.	chi2(6)=60.5 (<0.01)	chi2(11)=79.0 (<0.01)	chi2(5)=55.7 (<0.01)	chi2(16)=57.8 (<0.01)	chi2(10)=55.7 (<0.01)	F(3,60)=7.3 (<0.01)
Log restricted-likelihood	-1685.90	-1656.60	-1635.89	-1637.11	-1633.13	-1795.05	
R ² (within)							0.135
R ² (between)							0.069
N	418	418	418	418	418	445	418
Countries	61	61	61	61	61	65	61

P values in parentheses; * p<0.1, ** p<0.05, *** p<0.01.

a) p-values for directed hypotheses.

Results obtained using a restricted maximum likely method and independent covariance structures.

For an overview of variable definitions and sources, see Appendix A.

Appendix D: Model Specifications

Model specifications for Appendix B: Primary enrolment

$$M1: \text{Primary}_{ij} = \gamma_{00} + e_{ij} + u_{0j}$$

$$M2: \text{Primary}_{ij} = \gamma_{00} + \beta_{1j}\text{diversity}_{ij} + \beta_{2j}\text{co-ethnics}_{ij} + \beta_{3j}\text{income}_{ij} + \beta_{4j}\text{diversity}_{ij}\text{Xco-ethnics}_{ij} + \beta_{5j}\text{diversity}_{ij}\text{Xincome}_{ij} + \beta_{6j}\text{co-ethnics}_{ij}\text{Xincome}_{ij} + e_{ij} + u_{0j}$$

$$M3: \text{Primary}_{ij} = \gamma_{00} + \beta_{1j}\text{diversity}_{ij} + \beta_{2j}\text{co-ethnics}_{ij} + \beta_{3j}\text{income}_{ij} + \gamma_{01}\text{politicization}_j + \gamma_{02}\text{institutions}_j + \gamma_{03}\text{institutions}_j\text{Xpoliticization}_j + \gamma_{04}\text{education expenditure}_j + \gamma_{05}\text{children}_j + \gamma_{06}\text{pupil-teacher ratio}_j + \gamma_{07}1995 + \gamma_{08}2000 + \gamma_{09}\text{president's incumbency}_j + e_{ij} + u_{0j}$$

$$M4: \text{Primary}_{ij} = \gamma_{00} + \beta_{1j}\text{diversity}_{ij} + (\gamma_{20} + u_{2j})\text{co-ethnics}_{ij} + (\gamma_{30} + u_{3j})\text{income}_{ij} + \gamma_{01}\text{politicization}_j + \gamma_{02}\text{institutions}_j + \gamma_{03}\text{institutions}_j\text{Xpoliticization}_j + \gamma_{04}\text{education expenditure}_j + \gamma_{05}\text{children}_j + e_{ij} + u_{0j}$$

$$M5: \text{Primary}_{ij} = \gamma_{00} + \beta_{1j}\text{diversity}_{ij} + \beta_{2j}\text{co-ethnics}_{ij} + \beta_{3j}\text{income}_{ij} + \gamma_{01}\text{politicization}_j + \gamma_{02}\text{institutions}_j + \gamma_{03}\text{institutions}_j\text{Xpoliticization}_j + \gamma_{04}\text{education expenditure}_j + \gamma_{05}\text{children}_j + \gamma_{06}1995 + \gamma_{07}2000 + \gamma_{08}\text{president's incumbency}_j + e_{ij} + u_{0j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}\text{politicization}_j + \gamma_{22}\text{institutions}_j + \gamma_{23}\text{president's incumbency}_j + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + \gamma_{31}\text{politicization}_j + \gamma_{32}\text{institutions}_j + \gamma_{33}1995 + \gamma_{34}2000 + u_{3j}$$

Model specifications for Appendix C: Secondary enrolment

$$M1: \text{Secondary}_{ij} = \gamma_{00} + e_{ij} + u_{0j}$$

$$M2: \text{Secondary}_{ij} = \gamma_{00} + \beta_{1j}\text{diversity}_{ij} + \beta_{2j}\text{co-ethnics}_{ij} + \beta_{3j}\text{income}_{ij} + \beta_{4j}\text{diversity}_{ij}\text{Xco-ethnics}_{ij} + \beta_{5j}\text{diversity}_{ij}\text{Xincome}_{ij} + \beta_{6j}\text{co-ethnics}_{ij}\text{Xincome}_{ij} + e_{ij} + u_{0j}$$

$$M3: \text{Secondary}_{ij} = \gamma_{00} + \beta_{1j}\text{diversity}_{ij} + \beta_{2j}\text{co-ethnics}_{ij} + \beta_{3j}\text{income}_{ij} + \gamma_{01}\text{politicization}_j + \gamma_{02}\text{institutions}_j + \gamma_{03}\text{institutions}_j\text{Xpoliticization}_j + \gamma_{04}\text{education expenditure}_j + \gamma_{05}\text{children}_j + \gamma_{06}\text{pupil-teacher ratio}_j + \gamma_{07}1995 + \gamma_{08}2000 + \gamma_{09}\text{president's incumbency}_j + e_{ij} + u_{0j}$$

$$M4: \text{Secondary}_{ij} = \gamma_{00} + \beta_{1j}\text{diversity}_{ij} + (\gamma_{20} + u_{2j})\text{co-ethnics}_{ij} + (\gamma_{30} + u_{3j})\text{income}_{ij} + \gamma_{01}\text{politicization}_j + \gamma_{02}\text{education expenditure}_j + \gamma_{03}2000 + e_{ij} + u_{0j}$$

$$M5: \text{Secondary}_{ij} = \gamma_{00} + \beta_{1j}\text{diversity}_{ij} + \beta_{2j}\text{co-ethnics}_{ij} + \beta_{3j}\text{income}_{ij} + \gamma_{01}\text{politicization}_j + \gamma_{02}\text{institutions}_j + \gamma_{03}\text{institutions}_j\text{Xpoliticization}_j + \gamma_{04}\text{education expenditure}_j + \gamma_{05}\text{children}_j + \gamma_{06}1995 + \gamma_{07}2000 + \gamma_{08}\text{president's incumbency}_j + e_{ij} + u_{0j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}\text{politicization}_j + \gamma_{22}\text{institutions}_j + \gamma_{23}\text{president's incumbency}_j + u_{2j}$$

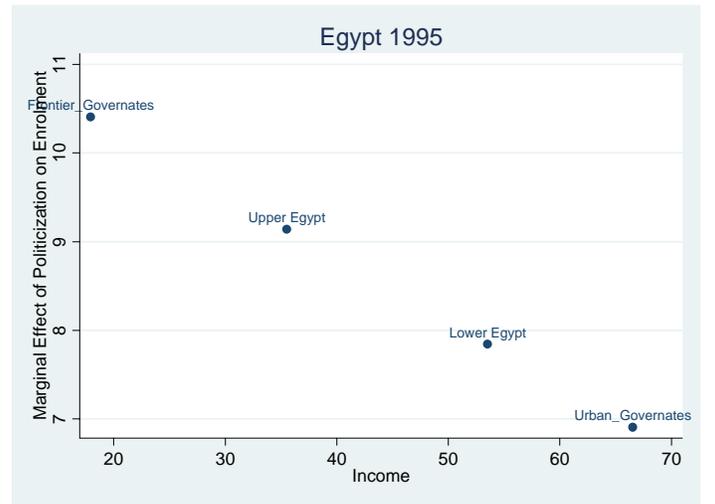
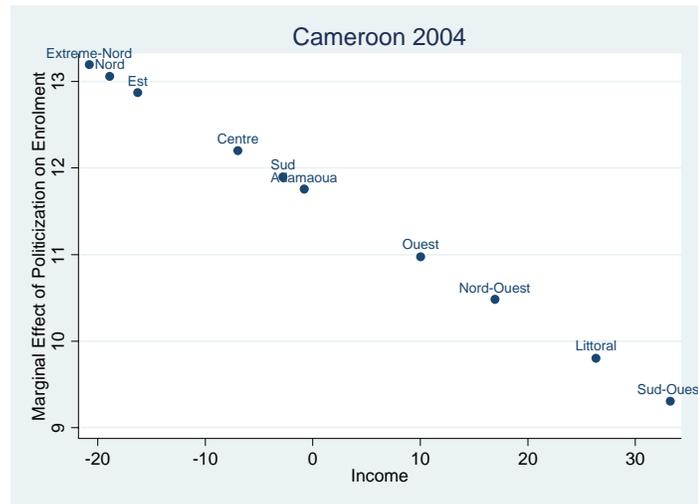
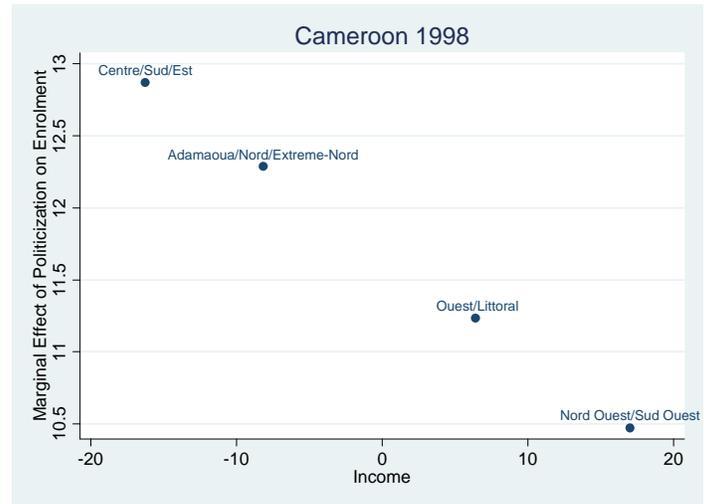
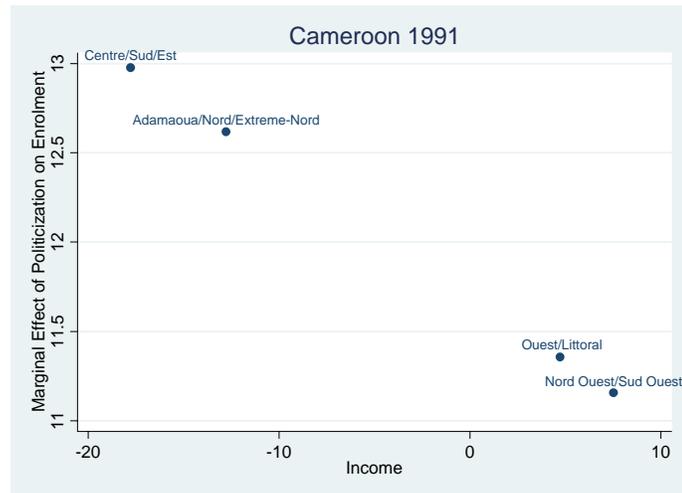
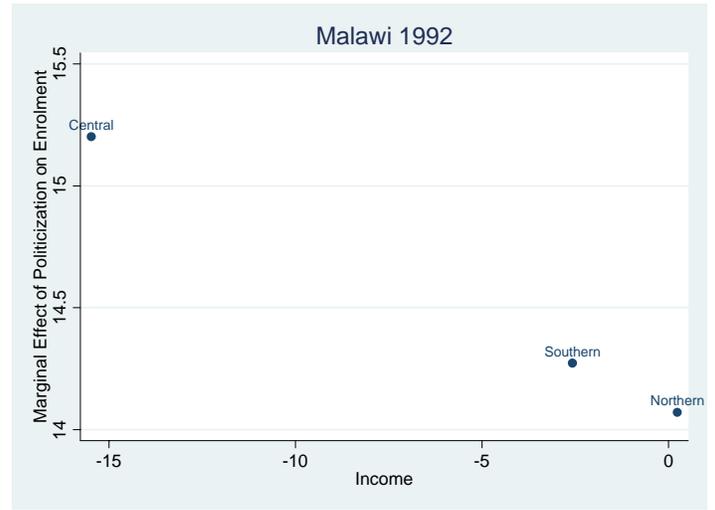
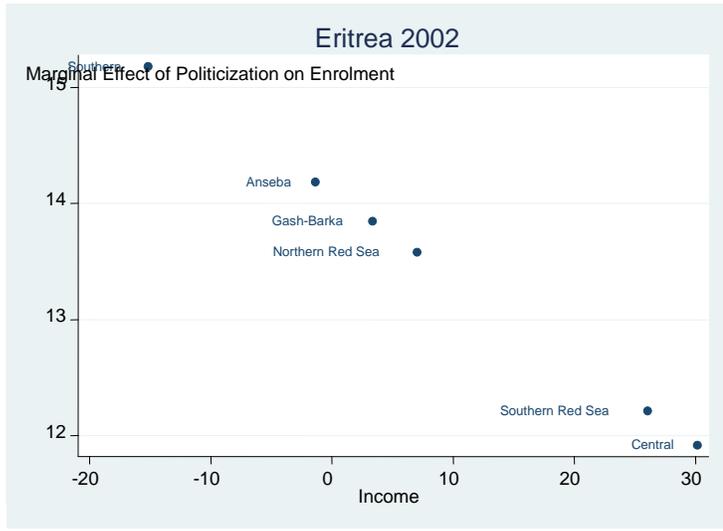
$$\beta_{3j} = \gamma_{30} + \gamma_{31}\text{politicization}_j + \gamma_{32}\text{institutions}_j + \gamma_{33}1995 + \gamma_{34}2000 + u_{3j}$$

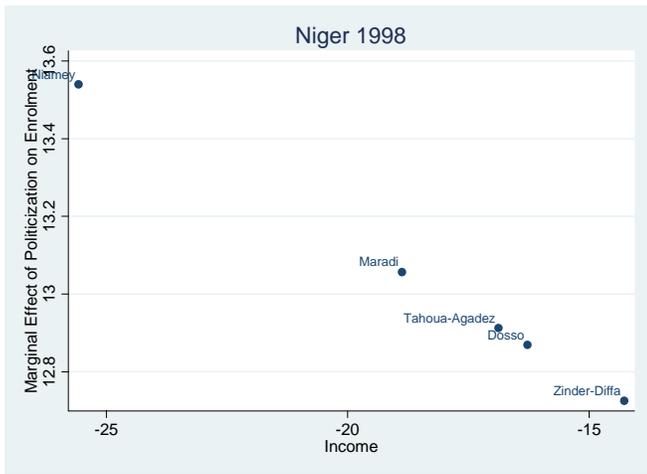
Appendix E: Correlation Matrix

	Diversity	Co-ethnics	President's incumbency	Politicization	Income	Children	Education expenditure	Pupil-teacher ratio	Institutions	1995	2000
Diversity	1.00										
Co-ethnics	-0.456	1.00									
President's incumbency	-0.076	0.078	1.00								
Politicization	0.225	-0.129	0.165	1.00							
Income	-0.254	0.239	0.046	-0.036	1.00						
Children	0.268	-0.226	-0.020	0.110	-0.532	1.00					
Education expenditure	-0.163	0.310	0.236	0.219	0.358	-0.405	1.00				
Pupil-teacher ratio	0.1115	-0.169	-0.069	0.272	-0.387	0.306	-0.334	1.00			
Institutions	-0.060	-0.028	-0.284	0.022	0.062	-0.157	0.033	0.090	1.00		
1995	0.0360	-0.034	-0.193	-0.043	0.006	0.037	-0.070	-0.213	0.002	1.00	
2000	-0.063	0.057	0.026	0.066	0.052	-0.201	0.003	0.302	-0.182	-0.577	1.0

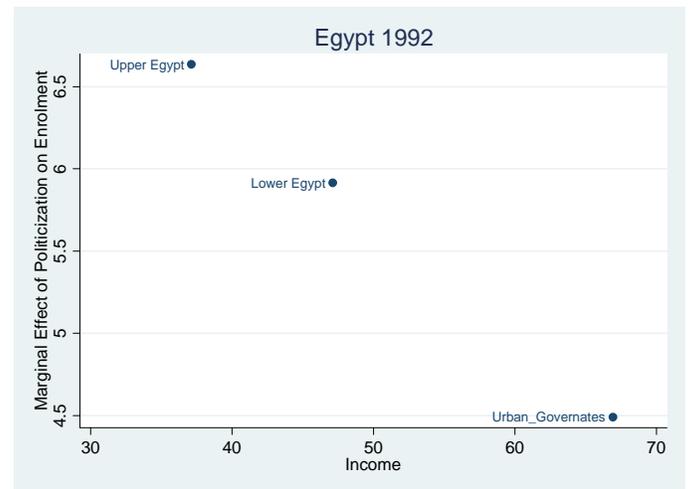
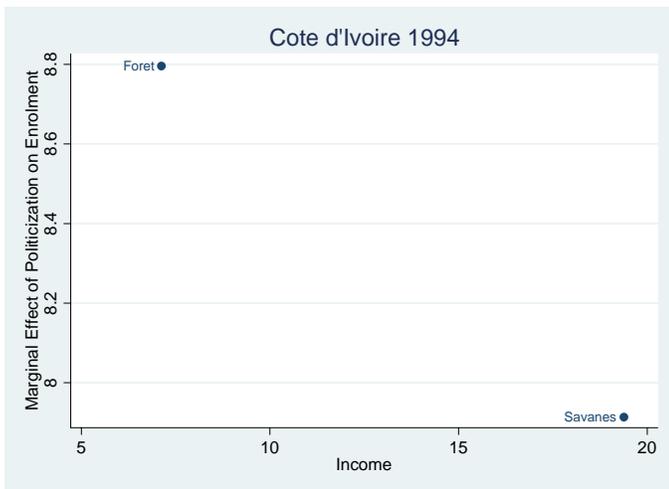
Appendix F: Country-by-Country Marginal Effects of Politicization

I. Countries with lowest institutional quality

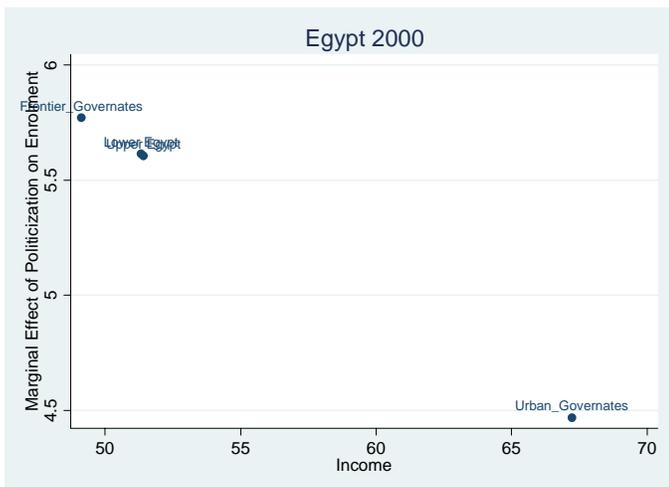




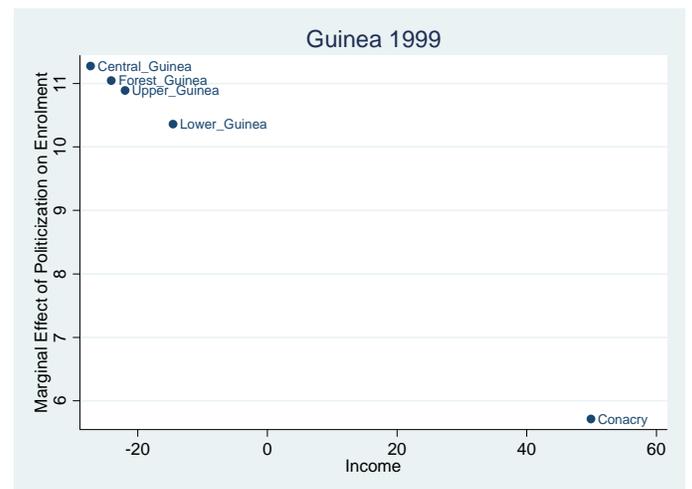
II. Countries with low institutional quality



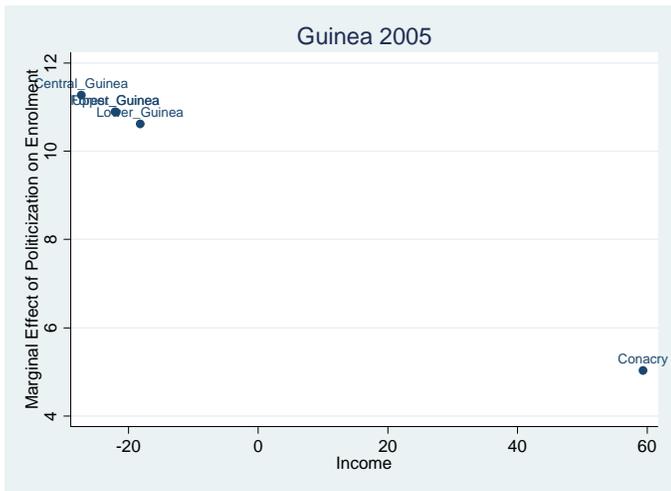
Note: The marginal effect of politicization for Urban_governates is not significant.



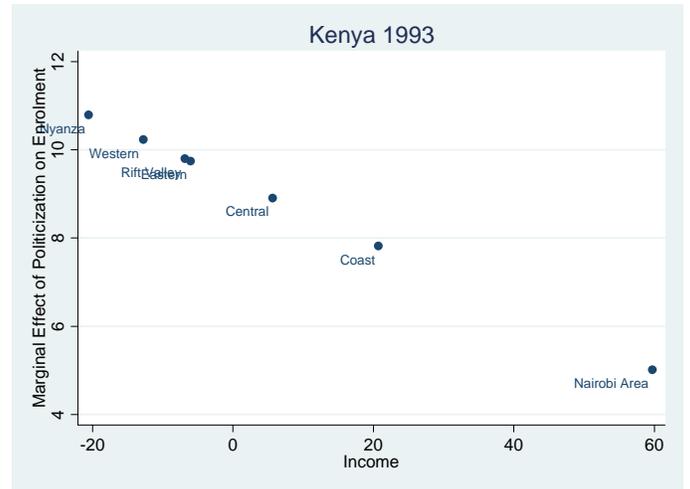
Note: The marginal effect of politicization for Urban_governates is not significant.



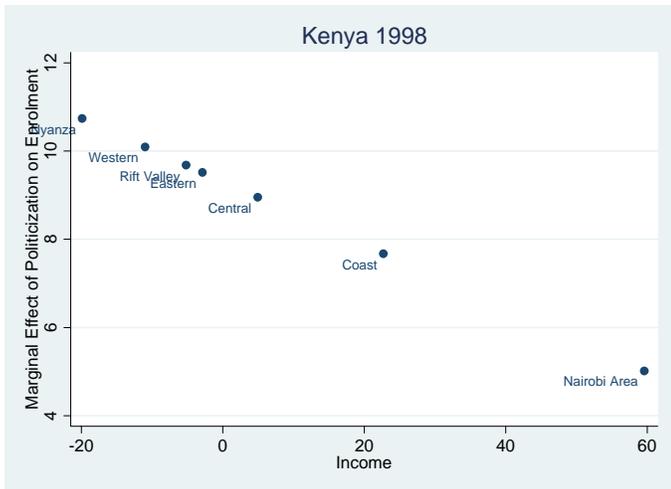
Note: The marginal effect of politicization for Conacy is not significant.



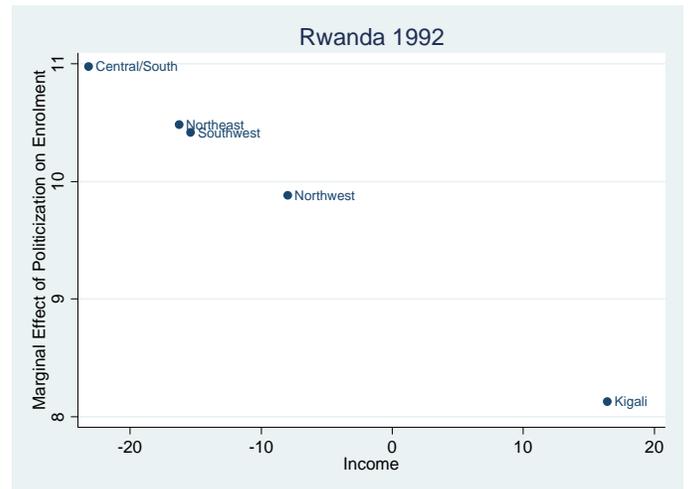
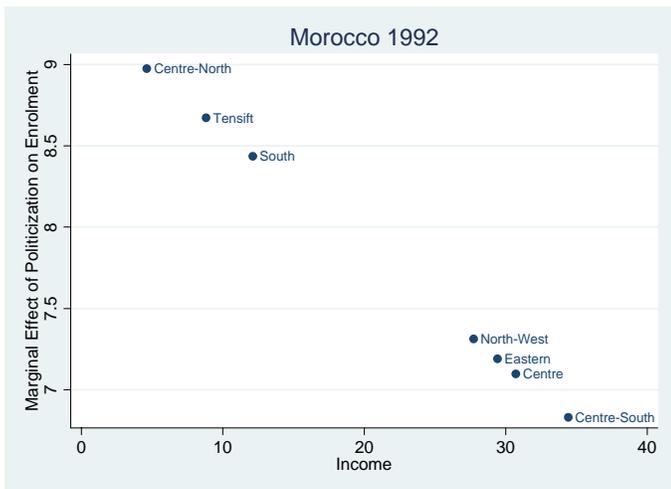
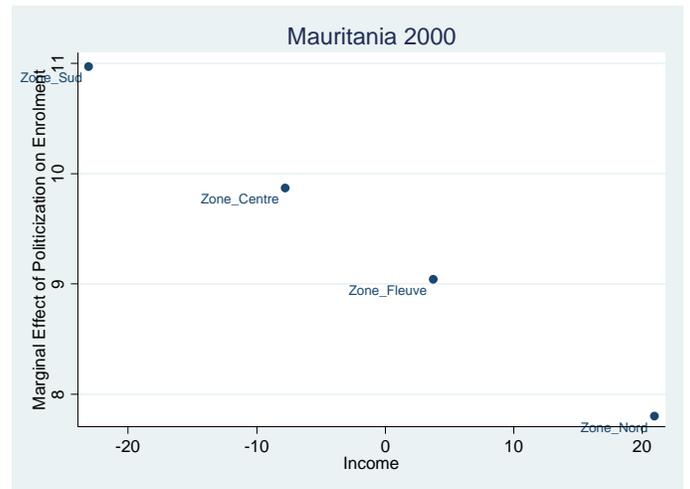
Note: The marginal effect of politicization for Conacyr is not significant.

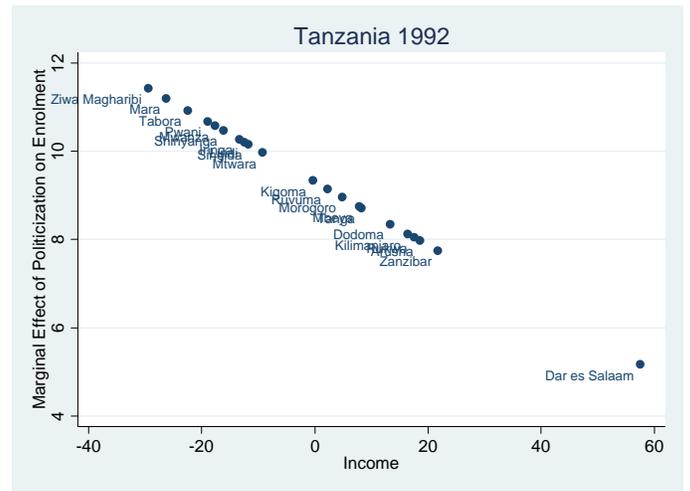
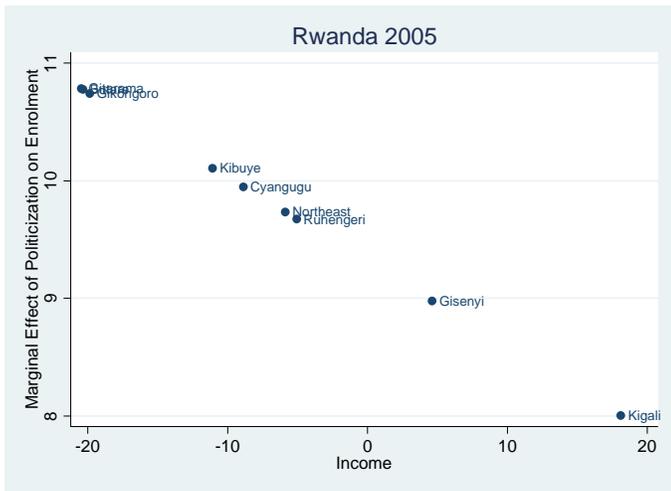


Note: The marginal effect of politicization for Nairobi Area is not significant.

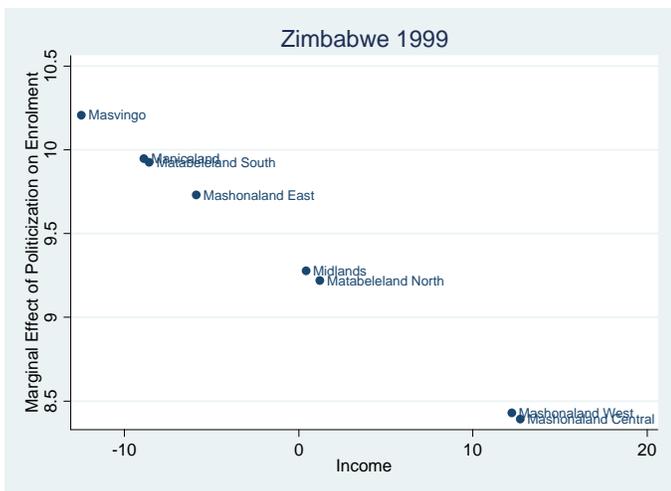
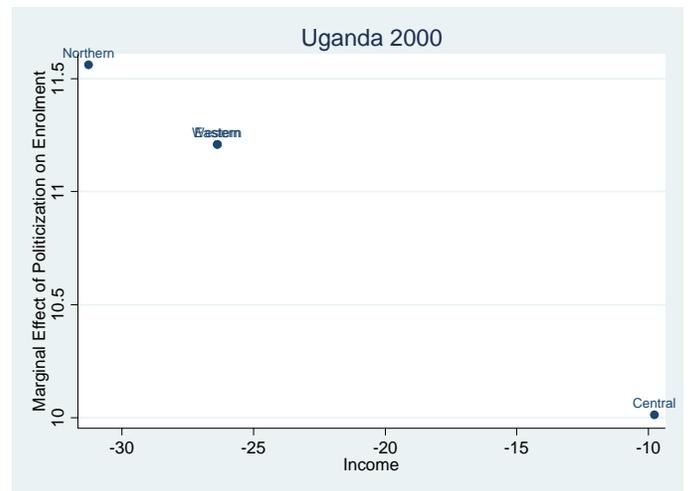
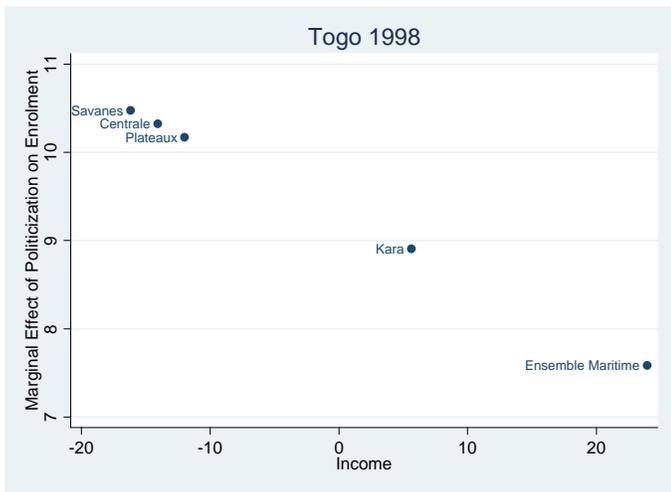


Note: The marginal effect of politicization for Nairobi Area is not significant.

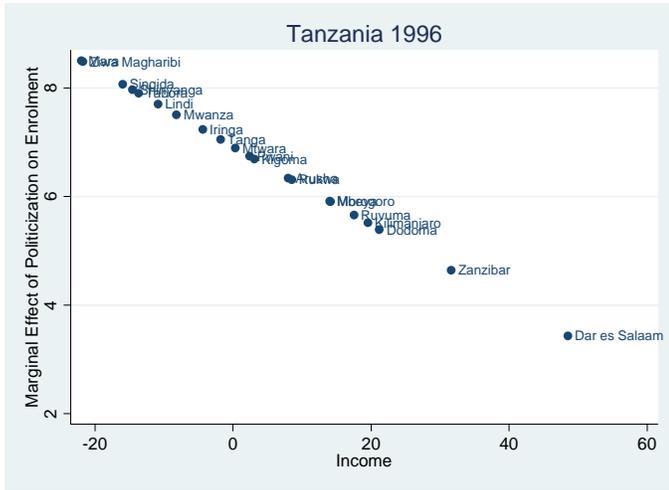




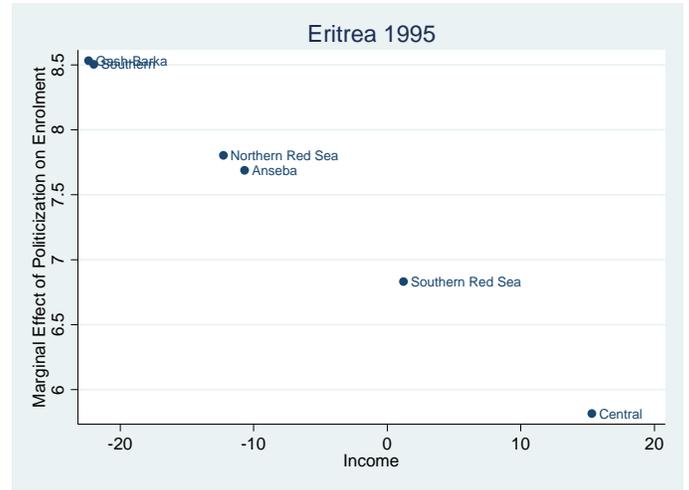
Note: The marginal effect of politicization for Dar es Salaam is not significant.



III. Countries with average institutional quality

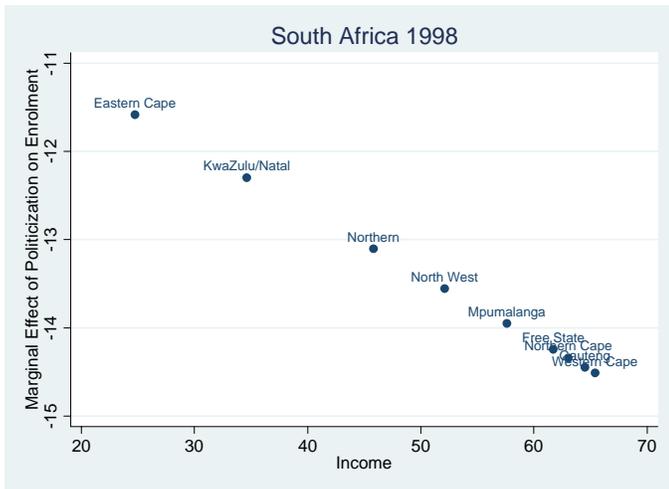


Note: The marginal effect of politicization is only significant for districts close to mean income (=0.0023).



Note: The marginal effect of politicization for Central is not significant.

IV. Countries with sound institutions



Specification of Appendix F

The graphs depict the marginal effect of politicization on primary enrolment for those countries that exhibit a significant interaction term as identified in *Graph 1* (29 out of 61). Calculations are performed using the regression coefficients reported in *Table 1*, column 2 for primary enrollment and the following formula:

$$d(\text{Primary enrollment})/d(\text{politicization})=3.90 -4.955*\text{institutions} - 0.073*\text{income}$$

Appendix G: Coding Rules for Diversity³⁰

The diversity variable is designed to capture the heterogeneity of the population at the district level. This variable has per se *no political dimension* but is a pure summation of all ethnic groups in the district. To code the ethnic diversity on the district level information is needed on (i) the districts used and (ii) the calculation of diversity. The following section will describe the two steps (i,ii) in greater detail and concludes with a presentation of some limitations of the data (iii).

(i) Which districts?

The districts from Cunningham and Weidmann (2008) are matched with the DHS districts. Information on alternate names of districts and development of district organization that permits matching of the two databases was found at the STATOIDS (2008).

(ii) Calculation of diversity

To calculate the diversity on district level, the widely used formula for the ethno-linguistic fractionalization measure as used by Alesina *et al.* (2003) is used:

$$\text{diversity}_j=1-\sum_{i=1}^N s_{ij}^2$$

Where s_{ij} = share of ethnic group i in country j ³¹.

When several smaller districts from Cunningham and Weidmann (2008) are included in a larger district in DHS, then the calculation is as follows. First, the population of one ethnic group in all three smaller districts is added and divided by the sum of the population of the three smaller districts. This group-share is then used to calculate the diversity.

(iii) Limitations of diversity data

Several limitations have to be kept in mind when working with the data from Cunningham and Weidmann (2008). First and most important, the list of ethnic groups and their population shares do *not* vary over time but is constant to the original coding of the Atlas Narodov Mira in 1964 (Bruk (1964)). Note that consequently the diversity measure does not vary within one country at different time points. Second,

³⁰ Nils Weidmann, pers. comm..

³¹ Alesina *et al.* (2003), p. 158-159.

Cunningham and Weidmann's model is not able to account for nomadic ethnic groups (without at specific settlement area) and for overlapping of several settlement areas. The complete list of matched districts is available upon request to anke.weber@pw.uzh.ch.

Appendix H: Coding Rules for co-ethnics

To code the share of population belonging to the ethnic group of a country's leader on district level, one needs information on (i) which leader, (ii) which ethnic identity, and (iii) the geographic location of leader's ethnic group denoted in district level population shares. The following section will describe the three steps (i-iii) in greater detail. The complete list reporting shares of leaders' ethnic co-members is available upon request to anke.weber@pw.uzh.ch.

(i) Which leader?

Since the object of the co-ethnics variable is to capture possible resource distribution by the leader to his/her ethnic group, it is necessary to measure the leader that effectively has the power over state resources. In most cases, this is the incumbent president at the time of the DHS survey. However, in rare cases power does not lie with the president but with other politicians. Following Kasara's (2007) approach, this paper uses the list of "effective leaders" by Goemans, Gleditsch and Chiozza (2008).

(ii) Which ethnic identity?

Since ethnic identity varies and depends on the situation that individuals face, ethnic identity of the leader is coded as the identity *known to be politically relevant*. Fortunately, Fearon, Kasara and Laitin (2007) already coded the ethnicity of effective leader. This data on ethnic identity of the leader is used for all surveys dating 2000 and back. For newer DHS surveys, own coding is used, since Fearon, Kasara and Laitin's dataset is limited to the year 2000.

(iii) Geographic location of leader's ethnic group

The information on leader's ethnic identity by Kasara (2007) was matched with district level population shares of ethnic groups provided by Cunningham and Weidmann (2008). Particular care was attributed to this procedure because of different names for ethnic groups and different grouping criteria (language, tribe, ethnicity, and race), which had to be matched. Information from the Ethnologue Country Index by Gordon (2005) was used to identify specific groups and their geographic location. In rare cases, no information on ethnic identity was available from Fearon, Kasara and Laitin (2007) because the dataset does not code ethnicity of leaders after the year 2000. In addition, for some countries that clearly exhibited ethnic conflicts (such as Rwanda, the 1994 genocide) only one ethnic group (Banyuranda) was coded in the country in Cunningham and Weidmann (2008). In both of the two cases, where information on ethnicity was missing, leaders' ethnic identity was coded using other sources (as described in Table H).

Table H summarizes the list of countries (and leaders) in cases where (1) Fearon, Kasara and Laitin's (2007) ethnic identity matches Cunningham and Weidmann (2008) or several smaller ethnic groups were summarized into one larger group, (2) alternative names or a subgroup description for ethnic groups have been used, and (3) in which the author had to code ethnic group-shares. Although, one might argue that if leader's real ethnic identity is a subgroup of the one used in this analysis, this approximation can still be regarded as a valid approximation of the real ethnic identity.

Table H: Coding of Leaders' Ethnic Identity

(1) Perfect match	(2) Alternative names or subgroup	(3) Coding Weber (2008)
<ul style="list-style-type: none"> ▪ Burkina Faso (1992, 1998, 2003) ▪ Ghana (1993,1998) ▪ Guinea (1999, 2005) ▪ Niger (1998, 2006) ▪ Mozambique (1997, 2003) ▪ Namibia (1992, 2000) ▪ Zambia (1996, 1992) ▪ South Africa (1998) ▪ Egypt (1992, 1995, 2000) ▪ Morocco (1992, 2003) ▪ Mali: FKL=Mande==Mandingo+Soninke+Susu in CW [Explanation: Mande comprises the three ethnic groups Mandingo, Soninke and Susu] 	<ul style="list-style-type: none"> ▪ Benin: FKL=Fon = Tem in CW [Explanation: in CW all Gbe-language groups count as being a member of Ewe (Gbe is the upper-level grouping and Fon and Ewe are subgroups)] ▪ Cameroon: FKL=Beti==Fang in CW [Explanation: Beti is upper-level grouping, and Fang is a subgroup] ▪ Central African Republic: FKL=Sara==Bagirmi in CW [Explanation: Sara is a subgroup of Bagirmi] ▪ Republic of Congo: FKL=Kouyou==Bakele in CW [Explanation: Scanning of geographic location of all ethnic groups in Congo in CW showed that Kouyou are located in the same region as Bakele, hence Bakele are used as an approximation of Kouyou] ▪ Cote d'Ivoire: FKL=Baule==Akan in CW [Explanation: Baule are a sub-group of the Akan people] ▪ Eritrea: FKL=Tigrinya==Tigrai in CW [Explanation: Tigrai is an alternative name of Tigrinya] ▪ Ethiopia: FKL=Tigre==Tigrai in CW [The Tigre are a related ethnic group of the Tigrai in Eritrea (and located at the border to Eritrea) and can be used as an approximation of the Tigre group (see language map of Ethiopia, Gordon (2005))] ▪ Gabon: FKL=Teke==Bateke in CW [Explanation: Bateke is alternative name for Teke] ▪ Kenya: FKL=Kalenjin==Nandi in CW [Explanation: Nandi is a dialect of Kalenjin] ▪ Lesotho: FKL=Sotho==Basuto in CW [Explanation: Basuto is a sub-group of Sotho] ▪ Niger: FKL= Djerma (Zarma)== Songai in CW [Explanation: Zarma is one of the Songai-languages] ▪ Tanzania, President Benjamin Mkapa: FKL=Ngoni==Angoni in CW ▪ Togo: FKL=Kabre(Cabrai, B==Tem in CW [Explanation: same dialect] ▪ Uganda: FKL=Ankole==Banyoro in CW [The tribe Banyankore is located in the Ankole region and Bayoro matches geographic location of Ankole] ▪ Zimbabwe: FKL=Shona==Mashona in CW [Explanation: Mashona is alternative name for Shona] 	<ul style="list-style-type: none"> ▪ Ghana, President John Agyekum Kufuor: (Antoun and Campling (2008)) =Asante=Akan in CW [Explanation: Asante is a dialect of Akan] ▪ Madagacsar: The only ethnic group in CW is "Malagasy". However, Malagasy comprises a number of political relevant ethnic groups, such as: Betsimisaraha, Plateau, Sakalava and Bara (Marcus (2004), and Gordon (2005)). Geographic location and groupshare in districts were estimated using the language map of Madagascar from Gordon (2005). President Didier Ratsiraka: FKL=Betsimisaraha== Betsimisaraha; President Ratsirahonana: FKL=Merina==Plateau [Explanation: Plateau Malagasy comprises more than just the Merina people, but the other people are very close to Merina and the primary ethnic divide is between Merina and cotier. President Ravalomanana: (Africa Almanac (2008))= Merina and Merina is approximated with Plateau ethnic group in CW. ▪ Malawi: CW does not include Chewa (=ethnic identity of president Hastings Kamuzu Banda). Geographic location and groupshare in districts were estimated using the language map of Malawi from Gordon (2005). President Banda: FKL=Chewa==Chewa; President Muluzi=Yao==Yao. ▪ Mauritania: FKL codes ethnic identity of President as "Bidan (White) M", whereas CW has only one category summarizing white and black Moors, namely "West Saharan Arabs". Recoding is necessary since ethnic conflict runs along the black Moor versus white Moor divide (Minorities at Risk Project (2005)). Since white and black moors are evenly dispersed throughout the country with no clear geographic location of either of the ethnic groups (Minority at Risk Project (2005)), the overall share of group members drawn from the Polity IV country report 2003: Mauritania (Marshall and Jagers (2008) are used as group shares. The shares are as follows: White Moors and Black Moors represent 70 percent of the total Mauritanian population. 30 percent of the population is white Moors and 40 percent is black Moors. Hence $0.3/0.7=0.43$ percent of West Saharan Arabs are white_Moors and $0.4/0.7=0.57$ percent of West Saharan Arabs are black_Moors. ▪ Nigeria: FKL=middle Belt, but CW has no such coding. Hence, recode all inhabitants of the Middle Belt region (= districts Adamawa, Taraba, Benue, Kogi, Kwara, Niger, Abuja, Plateau, Nassaraw (Haruna Izah (2004)) as being member of "middle Belt" and keep the ethnic groupshares of all other inhabitants not living in the Middle Belt area. ▪ Rwanda: CW does not list Hutu and Tutsi as distinct groups but only "Banyuranda". Recode district shares by using the following population shares: Tutsi=0.15 percent, Hutu=0.85 percent (CIA (2008)), The Twa group (approx. 1 percent is omitted). President Kagame: (Campling (2008) = Tutsi ▪ Tansania, Ali Hassan Mwinyi: FKL=Zanzibar, missing in CW, code district Zanzibar as 1 and all other districts as 0 ▪ Senegal, Abdoulaye Wade: (Africa Almanac (2008)): Wolof= Wolof in CW

FKL= Fearon, Kasara and Laitin (2007)

CW= Cunningham and Weidmann (2008)