Sub-Saharan Growth Surprises: Geography, Institutions and History in an All African Data Panel

Matthias Cinyabuguma and Louis Putterman*
University of Maryland and Brown University

November 12, 2006

Abstract

We use two sets of cross-country growth regression models to investigate the determinants of the growth rates in African countries between 1960 and 2000. Both sets of models contradict conclusions based on global samples: within Africa, we find greater coastal population and distance from the equator unhelpful and greater ethnic heterogeneity helpful to growth. Our results confirm the negative effects of corruption and civil wars and the positive effects of trade openness, political rights and political openness, and suggest that these institutional and policy variables are endogenous to geographic and historical factors including the colonizing power and the religious and ethnic make-up of the country.

JEL Classification: O40, O11, O55, O17

Keywords: Sub-Saharan Africa, Institutions, Geography, Social and Historical Factors, GMM, IV

*Assistant Professor and Professor of Economics, respectively, at University of Maryland and Brown University. This paper draws upon Cinyabuguma's doctoral dissertation completed at Brown University in 2005. He wishes to thank Herakles Polemarchakis and Oded Galor for their guidance and encouragement. We thank Sean Campbell, David Weil, Frank Kleibergen, Ki Young Park, Jonathan Temple, Roland Pongou, Paul Collier, Anke Hoeffler and Marcel Fafchamps for valuable suggestions and comments. Please address questions and comments to Louis_Putterman@Brown.Edu

1 Introduction

In an early study of the relationship between economic growth and policies, Robert Barro (1991) pointed out the unsatisfactory nature of dummy variables for Latin America and sub-Saharan Africa (hereafter, "Africa") that continued to be statistically significant even when additional controls were incorporated. This struck Barro and others as problematic because until one knows which economically relevant features of being located in a region account for its different performance, one lacks a full explanation of what determines growth rates. If we can identify the relevant features and control for them in the regression, the regional dummies should hold no further explanatory power.

The mystery of the Africa dummy attracted attention in part because of the distressingly poor economic outcomes recorded in so many African countries in the 1980s and 1990s.[1] Easterly and Levine (1997) suggested that much of the difference between African and other countries was due to ethnic heterogeneity, which they found to negatively impact growth rates in a global sample with a quarter or so of the observations being African.[2] While accounting for ethnic heterogeneity alone failed to fully eliminate the Africa dummy from their regressions, they found that the Africa dummy become insignificant when they added the average growth rate of neighboring countries. Sachs and Warner (1997) argued that poor growth outcomes in Africa could be substantially attributed to lack of coastal access, the disease ramifications of tropical climates, and protectionist trade policies. They incorporated these variables in some of their specifications, found significant positive effects of coastal populations and trade openness, and significant negative effects of the fraction of each country's area located within the tropics, in their global sample, and found that the Africa dummy became insignificant.

[1] The World Bank calculated that the annual growth rate of per capita income between 1985 and 1995 was a negative 1.1% in sub-Saharan Africa, compared to (positive) 0.3% in Latin America and the Caribbean, 2.9% in South Asia, and 7.2% in East Asia and the Pacific.

[2] Subsequent studies have found that the relationship between ethnic heterogeneity and economic growth is not monotonic: having a few major groups contending for power seems to be more harmful than having many small groups, none of which can hope to dominate the others (Reynal-Querol (2002)). Sachs and Warner (1997) found ethnic heterogeneity had no explanatory power in their growth regressions after geographic variables and trade openness are controlled for. Bockstette et al. (2002) suggest that ethnic heterogeneity is a symptom of late development of intensive agriculture, high population density, and the state, and find that the ethnic heterogeneity variable has little or no explanatory power after controlling for population density and early state formation.

Like these papers, the present paper is concerned with what accounts for the poor economic performance of many countries in Africa. Rather than searching for explanations of African differences in the context of global cross section regressions, however, we undertake a cross section growth study taking advantage only of variations within Africa itself. Like other economists, we suspected that Africa's economic problems might be attributable to distinctive geographic or climatic factors and to distinctive social, institutional and historical factors. But we wanted to focus on variations within Africa itself as a test of whether factors that appear likely to be important based on the analysis of global samples play the same roles in a purely intra-African context. In our study, we find strong support for the usual positive relationship between investment rates and growth, and for the standard negative relationship between initial income and growth, which provides assurance that Africa is not all that different from other continents and that data for the continent's countries is not all that poor. But we were surprised to find that some of the main factors that Easterly and Levine (1997) and Sachs and Warner (1997) proposed as explanations for the African growth difference had precisely the reverse impacts on differences in performance within Africa itself. Ethnic heterogeneity is positively, not negatively, correlated with our growth residuals and with institutions and policies favoring growth. And countries in our sample were doing better if they were closer to, not further from the equator, and if their populations lived inland, not near a coast.

We follow a number of papers that employed an instrumental variables procedure to deal with the endogeneity of institutions. Mauro (1995) introduced corruption in the empirical growth literature and used ethno-linguistic fractionalization as an instrument. Hall and Jones (1999) employed measures of language, a trade share indicator due to Frankel, Romer and Cyrus (1996), and a distance from the equator indicator as instruments for social capital. Acemoglu et al. (2001, 2002) argue that both settler mortality and indigenous population density in 1500 can be used as instruments for modern day political institutions. Easterly and Levine (2003) found that endowments of tropical climate, germs, and crops affect economic development only through institutions and policies. Our study is broader in the sense that it utilizes instruments drawn from geographic, social and historical characteristics, it involves institutional and policy variables,

and uses suitable regression procedures that account for simultaneity problems.

We know of a few studies that have worked with country samples exclusively from Africa. Bertocchi and Canova (2003) used African observations to investigate whether different colonizing powers affected the consequent economic performance of the countries they colonized in different ways. They found that colonial legacy impacts the differing growth performances in Africa. To understand the ramifications of colonization they used factors such as the legal origin and the degree of economic penetration of the colonial powers. Savvides (1995) uses a fixed effect framework covering 28 countries (Maghreb included) with four seven year periods over 1960-87. His study focuses on trade policy and he finds economic growth to be correlated with growth in the trade-GDP ratio, investment, initial income, schooling, and growth of the government. Odjo and Oshikoya (1995) apply OLS and GLS techniques to study the economic performance of 17 African countries. On average, their findings contend that investment, external debt, population growth, human capital and proxies for macroeconomic environment significantly determine long-run growth in Africa.[3] Gyimah-Brempong et al. (1999) explore the relationship between instability, investment and economic growth in sub-Saharan Africa. Through a dynamic panel approach they confirm the inverse relationship between political instability and economic growth identified by earlier studies. Garner (2005) examines whether Acemoglu et al.'s "reversal of fortunes" applies to Africa. He finds, consistent with them, that population density in 1500 is a negative predictor of average country income today. Gennaioli and Rainer (2005) test the theory that countries whose ethnic groups were more centrally organized before colonization have better public goods provision, and finds the evidence to be supportive. While these analyses are interesting, their focuses are still relatively narrow, and none check for robustness by using the alternative methodologies which we discuss presently.

We succeeded in assembling a sufficiently large data set to generate statistically significant findings with respect to many variables that interest us. Some of the variables studied are available in cross section only, but others have both cross sectional and time

[3]Hadjimichael et al., (1995) have focused on the evolution of savings, investment, and net financial balances of the government and private sectors. They also evaluate the relative contributions of policy and exogenous factors – such as terms of trade losses – to the growth, savings, and investment performance of sub-Saharan African countries, as well as evaluating the impact of foreign assistance.

series variation. One strategy we adopted for dealing with this problem was to conduct a panel data analysis with a GMM model using variables that change over time and according to country, and then to conduct an OLS regression analysis of the relationship between the time-invariant variables and the country fixed effects from the panel regressions.[4]

A different methodological issue leads us to adopt a second method as an alternative to the GMM-fixed effects approach. We wanted to include in our analysis historical and geographical variables that seem likely to be exogenous relative to important institutional and policy variables. For example, the occurrence of civil wars is a contemporary political phenomenon likely to have a direct and significant influence on growth outcomes, while ethnic heterogeneity might also affect growth, in part by affecting the likelihood of civil wars. More generally, factors determined well before a country's establishment, for instance latitude, dominant religion, colonizing power, and ethnic heterogeneity, might be treated as exogenous, while such phenomena as political stability, political rights, civil wars and corruption are potentially influenced by them. To deal with the likely differences between the two types of factors' impacts on growth, we estimated a number of two-stage least squares models and tested for endogeneity and the role of instrumental variables, also carrying out over-identifying restrictions tests. To the extent possible, we tried to include the same variables in both our GMM-fixed effects analyses and in our 2SLS-IV analyses, so that findings could be compared and contrasted, increasing our confidence in some of the qualitative findings.

Like other studies we find that African economies have grown more slowly when they suffered from more corruption, more civil wars, less political rights, and less economic openness. However, each of the latter proximate determinants of growth is confirmed to be better treated as an endogenous product of historical and geographic factors than as an exogenous variable. Other findings include the discovery that countries in which state-level polities existed in the centuries prior to colonization have performed better than countries that did not have state-level polities, or in which such polities had been less enduring. Countries with larger Muslim populations exhibit slower growth. Former

[4]Details on the two-stage GMM dynamic estimation are provided in section 3.

British and French colonies have outperformed other countries, with former Belgian, Italian or Portuguese colonies left far behind. Countries with a high predicted trade share according to Frankel and Romer have grown faster than their counterparts with low predicted trade shares. Having more natural capital has reduced growth rates.

The rest of the paper is organized as follows: in Section 2, we discuss possible determinants of economic growth among African economies and introduce the variables used in our growth models. In Section 3 we introduce the procedure for GMM estimation and OLS analysis of how time-invariant factors determine country fixed effects, and we also explore our 2SLS-IV models and their results. Section 4 summarizes and concludes the paper.

2 Determinants of Performance

The countries of sub-Saharan Africa occupy a distinctive place in the world economy that emerged hand in hand with the modern nation-state, industrial technology, and the organizational forms of modern commerce and finance. Sub-Saharan Africa was the birthplace of humanity and still exhibits more genetic, cultural, and linguistic diversity than any other region of similar size. However its burden of tropical climates and disease, and its isolation from the other continents by oceans to the West and East, and the desert to the north, meant that until recent times, foraging and nomadic peoples coexisted in it with agriculturalists who, in the main, worked small plots of land without the benefit of plows, draught animals, and fertilizer. Their societies were often knit together mainly by small or decentralized tribes, interspersed among which were various small kingdoms and in a few cases (e.g. Mali and Ghana), larger empires. While part of the "Old World," sub-Saharan Africa had such limited contact with the civilizations of the Mediterranean and Asia that the source of the Nile in East Africa was still unknown to Europeans more than three centuries after the voyages of Columbus and Magellan. It was only after Europeans had learned of the malaria-inhibiting powers of quinine that colonial administrations, the production of tropical export crops, modern mining, and the nation state in its modern form came to most of sub-Saharan Africa. European colonial presence, though in most countries extending less than seventy years, dramatically affected African political and

economic structures. Decolonization left fragile modern infrastructures, oriented mainly toward serving primary export industries, with states initially manned by thin cadres of educated citizens, funded disproportionately by export taxes and foreign aid. These infrastructures became highly susceptible to clientalist corruption, ethnic conflict, and change of government by military coup.[5]

The picture is not entirely bleak, as there have been enormous advances in education and life expectancy. For example the primary school enrollment rate went from 50 percent in 1970 to 72 percent in 1992, while the secondary school enrollment was 7 percent in 1970 and 24 percent in 1992. The literacy rate of adults aged 15 and over increased by 121.4 percent between 1970 and 2000. Life expectancy increased by 17.5 percent during the period 1960-2000. Expatriate staffs have been replaced by well-educated local personnel in government departments, public and private sectors, professional posts, and universities. 42 of 47 African countries have returned to democratically elected governments between 1990 and 2000 (Barkan, 2002), and some of the conflicts of the past, as in Mozambique, Angola and Uganda, have been resolved or greatly diminished in scope. Still, exports of cocoa, palm oil and coffee have yet to regain their heights of the 1960s and '70s in countries like Ghana, Nigeria and Kenya, Ivory Coast, and Democratic Republic of Congo. Devastating civil wars have flared up in the past decade in Liberia, the Democratic Republic of Congo, Sierra Leone, Rwanda and Sudan, among other countries, and AIDS is claiming some 2.2 million lives a year in populations in which few if any can afford anti-retroviral drugs.[6] Analyzing cross-national and sub-national data, Lorentzen, McMillan and Wacziarg (2005) argue that high adult mortality, exacerbated to some extent by the ongoing AIDS epidemic, explains almost all of Africa's growth tragedy.[7]

We begin with these historical remarks because we believe that the fact that the 50

[5]Data from the World Bank and other sources indicate that most sub-Saharan African economies attained positive growth in the 1960's and 1970's and reached their peak per capita GDP in the latter decade. 41% of the countries in our sample had a per capita income in 1995 that was less than its 1960's level, 35% a 1995 income below their 1970 level, 56% a 1995 income below their 1980 level, and 62% percent a 1995 income below their 1990 level. Only six percent of sub-Saharan Africa's population live in nations with higher per capita income in 1995 than they had ever achieved. Interested readers can refer to Table 1 in Rodrik, D. (1999) for details on some of these numbers.

[6]For details see UNAIDS 2004 Report on the global AIDS epidemic

[7]A good survey of proposed explanations of African economic performance is provided by Collier and Gunning (1999).

poorest countries in the world include 33 African states cannot be understood independently of Africa's history and the geographic and political factors that have shaped it. If this is correct, then differences in performance among African countries are likely also to be linked to the same factors. Our empirical analysis of African growth will include a number of these factors. Many of these will be treated as ultimate and hence exogenous causes, including equatorial location, landlocked geography, an ecology conducive to the spread of malaria, and early historical realities shaped by these factors (i.e. the late appearance of states, their ethno-linguistic diversity, and the ability of European powers to carve up the continent in the 1890s). Others will be treated as endogenous, that is as heavily influenced by the exogenous variables, including the level of corruption in government and the occurrence of civil wars.[8]

A first geographic factor included in our study is proximity to the equator. Since many have suggested that tropical climate is among Africa's disadvantages due to its impacts on disease and agriculture, it seems important to check whether African countries with the most tropical climates have been more disadvantaged than the others[9]. The second included variable relates to costs and ease of access to international trade. We tried both a dummy variable for being landlocked and the measure by Gallup, Sachs with Mellinger (1999) of the proportion of the population living within 100 kilometers of an ice-free coast or ocean-navigable river; we present results with the latter (called Coast Pop. Shr.), which were more frequently significant. A third geographic variable is included to test the hypothesis that having more natural resources constitutes, however surprisingly, a "curse" (perhaps because it leads to rent-seeking rather than investment promoting policies, or because it contributes to civil wars and other forms of political instability). Sachs and Warner (1997) included for this purpose the share of primary product exports in GDP, but this seems inadequate, since a high primary export share is as likely to be a consequence as a cause of underdevelopment. We use instead the measure of natural capital developed by the World Bank (1997). A fourth variable relates more directly to disease and in particular malaria. Using malaria incidence measures is inadequate because economic

[8] A list of all included variables, their definitions, and the data sources, appears in Appendix 1.

[9] We also experimented with the use of more direct measures of climate, like the proportion of the country's land area within the geographical tropics, but found them to give less significant results.

development exerts reverse causal influence upon the incidence of malaria, so we instead use the malaria ecology measure from Kiszewski et al. (2004).[10] Finally, we included a measure of the degree to which geography, as opposed to policy, favors trade: the trade share predicted by a gravity model that uses a country's population and geographical features (Pred. Trd. Shr.) only, from Frankel, Romer & Cyrus (1996)—sometimes called "natural openness."[11]

Turning to potentially important social and historical variables that can be treated as exogenous when looking at post-1960 growth rates, we considered first several alternative measures of ethnic heterogeneity, fragmentation, or polarization. In the end, we found most consistently significant and thus adopted as our representation of these factors the principal measure from Easterly and Levine (1997), which is the average of five different indices of ethno-linguistic heterogeneity (Eth. Heter.).[12] Religious conflicts may also be important in some countries, for example Nigeria and the Sudan. We include two alternative religious measures–religious fractionalization and religious polarization—used by Reynal-Querol (2002). In addition we use shares of the population classified as Muslim and as Catholic, respectively. Our next measure is the colonizing power, grouped as Eng. Col. (English colony dummy), Fr. Col. (French colony dummy), or other Col. (Other colony dummy).[13] We take from Hall and Jones (1999) a measure of the fraction of

[10] Malaria ecology provides an instrument for malaria risk that controls for the fact that causation may run not only from malaria to income but also from income to malaria. According to Kiszewski et al. (2004), the basic formula for ME includes temperature, species abundance, and vector type. "Because ME is built upon climatological and vector conditions on a country-basis, it is exogenous to public health interventions and economic conditions".

[11] Frankel and Romer (1999) and Frankel et al. (1996) point out that openness, as a policy variable, matters for economic development. They use the country's natural propensity to trade, based on the gravity model, as instrument for openness. In the gravity model, predicted trade between two countries goes up with the area and population size of the trading partner and down with the distance between two countries. Trade also increases with greater population size of the home country and decreases if the country is landlocked.

[12] Quite a few studies (e.g., Alesina, Easterly, and Baqir, 1999; Easterly and Levine, 1997) have indicated that ethno-linguistic diversity may affect economic development both directly, and indirectly by shaping the underlying institutions and policies that, in turn, determine economic growth. We decided against the simultaneous inclusion of alternative indicators of the ethnic structure, such as including both an ethnic heterogeneity and an ethnic polarization measure, because of their relatively high correlations with one another.

[13] Following Graziella Bertocchi and Fabio Canova, (2002), metropolitan countries are assigned to the colonial power that ruled longest. The former German colonies – Burundi, Rwanda, Tanzania, Cameroon and Togo – were divided among the countries that took them over after WWI. Burundi and Rwanda are

the population speaking one of the five primary Western European languages (including English) as mother language (Eur. Lng. Shr.), and the fraction of the population speaking English (Eng. Lng. Shr.) as a first language.[14] Finally, we include a measure of the early development of societies proposed by Bockstette, Chanda and Putterman (2002), an index of the depth of experience with state-level polities since 1 C.E. (Statehist01).[15]

We consider five main indicators of institutional quality, political stability, and policy. We anticipate that these factors may significantly affect economic performance, but they may also be endogenous to some or all of the geographic and social-historical factors mentioned above. First, among the various measures of the state quality,[16] the political openness (Polit. Open.) variable from the Polity IV dataset[17] produced consistently significant effects, so it is incorporated in our analysis. Second, we explored measures of political rights (Polit. Rights), of which we adopted the Freedom House index.[18] Because the economies of a number of African polities are well known to have suffered severely

assigned to Belgium, while Tanzania goes to Britain. Though Cameroon and Togo were jointly mandated by France and Britain, we list Togo and Cameroon under France alone since the British part of Togo was annexed to Ghana, and Cameroon is currently in the CFA-franc zone; while Somalia is placed under Italy (ignoring the few British and French influences there). In our specification, for reason of parsimony, we use only three categories of colonizing powers: British, French and Others. The variable "Other Col." is a dummy variable equal to one if the country listed was under Belgian, Italian or Portuguese rule, and zero otherwise. One country in our sample, Liberia, is listed as never colonized and accordingly none of the three colonial power dummies is applied to it.

[14] We take care not to use both variables in the same specification.

[15] Some African countries, for example Zambia, were apperently without kingdoms, states or empires until the 19^{th} century, while other, for example Nigeria, contained kingdoms as early as the 8^{th} century. Statehist01 sums values for half centuries, ranging from 0 for no state to 50 for an indigenous polity covering most of the present territory, and uses a discount rate of 1% per half century before the most recent included period, 1901-1950. Bockstette et al. (2002) interpret statehist as an indicator of pre-modern development and find in a global sample that countries with higher statehist values grew more rapidly from 1960 to 1995. We use a revised and expanded version of their data set (see Putterman, 2004).

[16] We also tried the ICRG80 and Kaufman et al. (1999a)'s variables as measures of political institutions but due to poor coverage of our sample, the results were not reliable and are not shown in this paper.

[17] Polit. Open. measures directly the limits of executive power. It was provided by the Polity IV data set and used by Azam and Hoeffler (2002). The score ranges from 0 to 10, where 10 denotes a highly open regime.

[18] The index of political freedom around the world has been published by Freedom House since 1972. The survey consists of a series of questions grouped under political rights and civil liberties, and each country or territory is given a numerical score for each category. The average scores are used to assign each country the status of "free", "partially free", or "not free". A score of 7 corresponds to countries enjoying the greatest freedom and a score of 1 to those enjoying the least freedom.

from civil wars, we also use the civil wars indicator from Collier and Hoeffler (2002) and test its endogeneity to geographic and historical factors. In line with Mauro (1995) we incorporate an index for corruption to deal with the abuse of office for private gain. It measures the degree to which corruption is perceived to exist among a country's public officials and politicians. Finally, openness to international trade is likely to have been important for economic outcomes. We treat it as a quasi-institutional or policy variable and use our estimates to investigate the influence that geography and history have exerted upon it. We use the trade openness (Yrs. Open) measure from Sachs and Warner (1995), also used by Hall and Jones (1999) among others.[19]

We study the determinants of economic growth in the framework of a standard growth model in which the growth rate of per capita GDP is the dependent variable, and the set of explanatory variables includes initial GDP per capita, average investment ratio (I/GDP), and a measure of human capital, the secondary school enrollment ratio (Sec. Enrol.)[20]

3 GMM and IV Models

3.1 GMM and Residuals Analysis

We have data for all or most of the variables used in this study for up to 33 out of 43 sub-Saharan African countries for all or part of the period from 1960 to 2000.[21] Because this number of countries is not that much larger than the number of determinants of growth that we want to include in our models, it is desirable to utilize additional variations

[19] The Sachs-Warner index measures the fraction of years during the period 1950 to 1994 that the economy has been open. A country was open if (i) non-tariff barriers covered less than 40 percent of trade, (ii) average tariff rates were less than 40 percent, (iii) the black market premium was less than 20 percent during the 1970s and 1980s, (iv) the country was not classified as socialist, and (v) the government did not monopolize major exports. Sala-i-Martin (1997) used the Openness variable to check for robustness, and found that the Sachs-Warner measure of openness is among the variables which are robust and correlated with growth.

[20] The three variables are treated as core explanatory variables in the sensitivity analysis of Levine and Renelt (1992).

[21] We study sub-Saharan Africa, thus excluding countries bordering the Mediterranean due to their different histories and cultures. Because of its domination by its white minority until the mid-1990s and the resulting differences between its economy and those of other sub-Saharan countries, we also excluded South Africa from our study.

available to us in the form of differences in growth rates and in values of the explanatory variables during different sub-periods. Accordingly, we organize our data on growth rates and on the other variables for which measures are available into eight half-decade observation periods, for 1960-64, 1965-69, etc. With some country observations missing for some sub-periods, this gives us up to a maximum of 200 observations in total, with between 5 and 8 observations for most countries.[22] Rather than simply pooling these observations, we have made precise estimates of the explanatory variables' effects by controlling for country and period specific effects. But some variables, including most of the geographic and historical factors, do not vary with time. We deal with this by means of a two step procedure: first, we use the generalized method of moments to estimate variants of the basic production function, including some additional time-variant factors, and then we estimate OLS regressions in which the extracted fixed effects for each included country are dependent variables and various sets of time-invariant factors are explanatory variables. Our core results (as in Caselli et al. (1996) and Hoeffler (2002)) will be based on the specification

$$g_{i,t} = \alpha + \beta y_{i,t-1} + \gamma x_{i,t} + \delta w_i + \eta_i + \upsilon_{i,t}, \tag{3.1.1}$$

and the fixed effects (which may contain both time-and country [observed and unobserved]-specific components) will be extracted as[23]

$$\hat{\upsilon}_{it} \equiv (g_{i,t} - \hat{\alpha} - \hat{\beta} y_{i,t-1} - \hat{\gamma} x_{i,t}) = (\delta w_i + \eta_i) + \upsilon_{i,t}, \tag{3.1.2}$$

where $g_{i,t}$ reflects the average growth rate over a series of five years periods, $y_{i,t-1}$ denotes the log of income per capita at the beginning of each of these periods, and $x_{i,t}$ is a subset of exogenous explanatory variables.[24] η_i accounts for the unobserved country specific effects, $\upsilon_{i,t}$ is the error term and w_i represents measured time invariant country characteristics.[25] Certain country specific variables, like religious fractionalization, are time variant.

[22] Countries included in our sample, their metropolitan powers and their independence dates can be found in Bertocchi et al. (2002).

[23] The two step procedure was also applied in Hoeffler (2002), Blanchlower, Oswald and Sanfey (1996) and Battese and Coelli (1995).

[24] A further ingredient accounted for in all our regressions is the time dummies. We have one dummy for each five-year period. The estimated coefficients from those dummies will not be shown in our tables to save space.

[25] Numerous studies have used cross-section frameworks to study the empirics of long run growth (Mankiw, Romer and Weil (1992), Barro (1997), Levine and Renelt (1992), Sala-i-Martin (1997) and

Among others, Knight et al. (1993), Islam (1995), and Casselli et al. (1996) estimated a neoclassical growth model using panel data methodology and concluded that single equation models, used to study the empirics of long-run growth, were misleading as they produced biased and inconsistent estimates. They argued that the endogeneity of some of the regressors and the unobserved country specific effects problems were not addressed when single equations and cross-country data were used. Fortunately the GMM approach addresses both types of problems. Following Arellano and Bond (1991), we run the first-differenced GMM analysis, with some restrictions on exogenous variables.[26]

In panel estimation, consistent estimation of the structural coefficients depends crucially on the stochastic properties of the error terms, i.e., whether they are serially correlated or not. In order to deal with different sets of restrictions, a large number of techniques including GLS, the within group estimator (FE), and the GMM estimator (Chamberlain, 1983) have been proposed. With endogenous regressors and country characteristics, the Arellano and Bond (1991) dynamic panel estimator performs better. In this paper we follow Casselli et al. (1996) and Gyimah-Brempong et al. (1999) and we apply the Arellano-Bond panel data estimation procedure.[27] In equation (3.1.1) we regress the growth rate of output on a broad set of explanatory variables, including the natural log of the initial level of GDP per capita (Log iGDPpc) and the natural log of the investment share of GDP (Log(I/GDP)). Since the measured country characteristics w_i may be correlated with the unobserved country-specific effects η_i and/or the error

King and Levine (1993)). However, reducing time series to a single (average) observation means that not all available information is being used (Hoeffler (2002); Ojo and Oshikoya (1995)). In addition, it is well established that the cross-country estimator is only consistent under strong restrictions that the individual fixed-effects are uncorrelated with the other right-hand-side variables (Casseli et al.(1996)). This is not the case for our specifications, where at least one of the regressors is conceptually endogenous. For example, it is reasonable to suppose that institutional and policy variables are determined simultaneously with the rate of growth. This suggests that the use of panel data with GMM would be a suitable method.

[26] We assume that the errors are independent across countries and serially uncorrelated for the coefficients on country characteristics to be consistent. We further assume that $y_{i,t-1}$ and x_{it} are predetermined for v_{it} in the sense that $E(v_{it}\,|y_{i,t-1}, x_{it}) = 0,\ \forall t$.

[27] The Arellano and Bond's estimation requires that the variables be measured as deviations from their period means and that the equations be estimated in differenced form. Because of the endogeneity and the correlation with the error terms, an instrumental variable estimator is called for. The Arellano-Bond first-differenced GMM is an IV estimator that uses past values of the explanatory variables as well as all strictly exogenous variables as instruments.

term $v_{i,t}$, the model given by equation (3.1.1) reduces to:

$$g_{i,t} = \alpha + \beta y_{i,t-1} + \gamma x_{i,t} + \eta_i^* + v_{i,t}, \tag{3.1.3}$$

where $\eta_i^* = \eta_i + \delta w_i$, and $v_{i,t}$ is i.i.d. with mean zero and a constant variance and, in particular, $E(v_{it}) = E(v_{it}v_{is}) = 0$ for $s \neq t$. Equation (3.1.3), because of endogenous regressors and its dynamic nature, implies correlation of error terms with regressors, thus violating the orthogonality condition. Using the differenced GMM procedure proposed by Arellano and Bond (1991), we obtain a consistent estimator $(\hat{\beta}, \hat{\gamma})$ for β and γ. These consistent estimates are plugged into equation (3.1.2) to extract the residuals of the growth regression. These residuals are then regressed on the measured country characteristics, w_i, as:

$$g_{i,t} - \hat{\alpha} - \hat{\beta} y_{i,t-1} - \hat{\gamma} x_{i,t} = \delta + \delta_1 w_i + \xi_{i,t}. \tag{3.1.4}$$

3.1.1 First Stage Regressions for GMM Models

Table 1 presents the results of the first-stage dynamic GMM estimator. The growth rate of per capita income is the dependent variable and independent variables are conventional time-variant economic variables (iGDPpc, I/GDP, Sec. Enrol.), institutional variables (including an indicator of civil wars and institutional quality index), and time dummies. We present four alternate specifications three of which drop one or another of the last three explanatory variables to check the robustness of results and to allow for the inclusion of country observations otherwise excluded due to missing data. The Arellano-Bond model (as any dynamic panel-data model) allows past realizations of the dependent variable (Lagged Gr(GDPpc)) to affect its current level[28]. Table 1 also presents , a first – and a second – order serial correlation test, as well as the Sargan test of over-identifying restrictions. It additionally presents Wald test statistics for the significance of some subsets of regressors.[29] These tests determine the correctness of the dynamic GMM

[28] Perhaps due to the modest sample size and the small number of observations for some countries, the maximum number of lags for the dependent variable which gives good results in this model is one.

[29] The Wald Test is a way of testing the significance of particular explanatory variables in a statistical model. If for a particular explanatory variable, or some subsets of regressors, the Wald Test is significant, then we conclude that the parameters associated with these variables are not zero, so that the variables should be included in the model. If otherwise, these explanatory variables can be omitted from the model. We use one-step results for inference on coefficients, as recommended by Arellano and Bond.

estimator used here. A key condition exploited in this analysis is the absence of serial correlation among the error terms. After examining our results, we found negative first order correlation in the residuals. The M1 statistic rejects the null hypothesis of lack of first order-serial correlation between error terms, validating our conjecture that the error term contains country specific effects.[30] However, the M2 statistic in Table 1 fails to reject the null of lack of second-order autocorrelation in the first-difference residuals in all specifications.[31]

[30] The fact that the error term contains country specific characteristics is a strong argument for the use of the second-stage GMM procedure. The Sargan tests of over-identifying restrictions are rejected in all four specifications that use the dynamic GMM estimator without robust standard errors. We only show results with robust standard errors in this paper.

[31] Test statistics and p-values for the first and second order autocorrelation in residuals are given by M1 and M2 respectively and are distributed as standard normal. The M2 statistic tests for lack of second-order serial correlation in the first-difference residuals. This is the case if the errors in the model in levels are not serially correlated, but also if the errors in levels follow a random-walk process (Arellano et al. 1991). The M1 statistic however tests for lack of first-order serial correlation in the differenced residuals. Since $(v_{it} - v_{it-1})$ is the first difference of serially uncorellated errors, M1 need not be statistically zero; but the consistency of the GMM estimators hinges heavily upon the assumption that M2=0.

Table 1: First Stage Regressions for GMM Models

Dependent variable: Growth rate of GDP per capita (Gr(GDPpc))

Indep. Var.	**1**	**2**	**3**	**4**
Lagged Gr(GDPpc)	-0.111*	-0.137**	-0.116*	-0.120*
	(0.062)	(0.069)	(0.067)	(0.072)
Log iGDP	-3.813*	-3.436*	-3.911**	-3.469**
	(2.106)	(2.081)	(2.006)	(1.415)
Log. (I/GDP)	5.160***	5.400***	5.432***	4.213***
	(1.148)	(1.204)	(1.079)	(0.977)
Civil Wars	-1.887***		-1.260***	-1.514***
	(0.496)		(0.389)	(0.348)
Polit. Open.	0.376**	0.261		0.126
	(0.162)	(0.184)		(0.137)
Sec. Enrol.	-0.529	-0.527	-0.763	
	(0.689)	(0.656)	(0.627)	
Constant	0.017	-0.051	0.146	-0.212*
	(0.324)	(0.298)	(0.261)	(0.124)
Wald χ^2	80.04	86.49	84.84	78.71
Sargan χ^2	25.79	28.03	27.83	19.03
	(0.17)	(0.103)	(0.11)	0.52
# Obs.	136	136	141	190
# Countries	24	24	24	33
M1	-2.40**	-2.92***	-3.12***	-3.16***
P-value	(0.016)	(0.004)	(0.001)	(0.002)
M2	-0.73	-1.45	-0.35	-1.01
P-value	(0.464)	(0.148)	(0.724)	(0.312)

Note: Table 1 represents a GMM estimation of model coefficients via the Arellano-Bond (1991) dynamic panel data model. Entries for variables in this table, and the following tables, are estimated coefficients followed by robust standard errors, in parentheses, and *, **and *** indicate significance at the 10, 5 and 1% levels respectively.

Turning to the estimates, in columns 1 and 2, the coefficients on the initial GDP per capita are negative and significant at the 10 percent level, while significant at the 5 percent level in columns 3 and 4. The estimated speed of convergence is about 3.6 percent

per annum. The investment share of GDP is highly significant ($p < 0.001$) and positive in all specifications, with an average coefficient of 5.04. Civil wars is very significant with the anticipated sign in all regressions in which it is included. However, the variable "political openness," while of the expected sign, is significant ($p < 0.05$) only in the regression of column 1.[32] The coefficient on civil wars implies that an increase of one standard deviation in civil wars directly decreases economic growth by 0.24 standard deviations, a 0.60% change in growth rate when political openness and schooling are controlled for in our GMM specification, and by 0.16 standard deviations, a 0.40% change in growth rate, when only schooling is controlled for in our GMM specification. This is a relatively large effect compared to the average growth for the sample period $(1960-2000)$ of about 0.73%.[33] These results provide support for the view that institutional problems, especially civil wars, affected growth negatively in Sub-Saharan African countries.[34]

3.1.2 Second Stage Regressions for GMM Residuals

In stage two GMM, we use the coefficients from the fourth column of Table 1. By dropping secondary schooling, which is not significant in the other regressions, this specification permits a significant gain in sample size, without major differences in the other coefficients.

In the second stage, we run an OLS regression with the residuals of the column 4 GMM regression on the time invariant characteristics. The results shown in Table 2 can tell us which measured country characteristics are correlated with variations in the country specific effects.[35]

[32] While Barro (1991) and Collier and Hoeffler (2002) looked at actual democracy, we are looking at political openness (which is defined as democracy by Polity IV data set), and at corruption. Note that almost all of the coefficients in our regressions are larger when we use robust standard errors, as we do in the estimates shown here.

[33] Estimates of the growth rate vary among the columns of Table 1 due to changes in country coverage, but all are within the narrow range from 0.71% to 0.76%. The average growth rate estimated by Tahari et al. (2004) for the sample period (1960-2000) is 0.73%.

[34] A growing literature on civil wars and economic performance, including a recent book by Clement (2005), confirms the negative effects of civil wars on growth in sub-Saharan Africa.

[35] We need to stress that the coefficients on the fixed effects regression are consistent if and only if all the country characteristics are uncorrelated with the unobserved county specific effects. This is, however, a strong restriction, which would hold only in very rare cases. We shall thus regard these results with caution, as in Hoeffler 2002.

Table 2: Second Stage Regressions for GMM Residuals

Dependent Variable: GMM Residuals

Indep. Var.	**1**	**2**	**3**	**4**
Latitude	-0.353**	-0.341**	-0.336**	-0.654***
	(0.166)	(0.163)	(0.175)	(0.162)
Latitude_sq	0.017**	0.019***	0.019***	0.039***
	(0.008)	(0.008)	(0.008)	(0.009)
Coast Pop. Shr.	-1.679	-3.747***	-3.711***	-4.027***
	(1.135)	(1.206)	(1.125)	(1.167)
Natural Cap.				-0.0001
				(0.0001)
Malaria Ecol.			-0.008	
			(0.035)	
Pred. Trd. Shr.		2.279***	2.286***	3.110***
		(0.765)	(0.766)	(0.705)
Yrs. Open	5.883**	0.301	0.367	
	(3.039)	(3.632)	(3.702)	
Eth. Heter.	2.349**	4.129***	4.237***	4.595***
	(1.207)	(1.421)	(1.460)	(1.117)
Statehist01	4.745**	9.383***	9.421***	8.677***
	(2.101)	(2.914)	(2.926)	(3.088)
% Muslim	-0.022**	-0.037***	-0.036***	-0.033***
	(0.010)	(0.012)	(0.012)	(0.011)
Constant	-1.702	-9.7676***	-9.691***	-11.487***
	(1.126)	(2.802)	(2.805)	(2.495)
R^2	0.29	0.32	0.32	0.43
# Obs	169	169	169	140
# Countries	29	29	29	25

Since the GMM regressions of Table 2 include multiple observations for the same country in different periods, the country-specific error terms extracted from the column 4 regression are in fact country-and-period terms, which we analyze in two alternatives ways. First, columns 1 – 4 of Table 2 report regressions in which a given country can be represented by several observations because the dependent variable—an error term from the GMM estimate—differs across periods, although the explanatory variables do

not. Second, in regressions not reported in Table 2 we run a cross section version of our fixed effects OLS regression, in which the single observation per country uses as dependent variable the average of the error terms for that country from the period-specific observations in Table 1.[36] The results are similar, although the standard errors on the cross-section regressions were somewhat larger. In both the Table 2 regressions and the cross section variant, we find that coefficients on the time invariant characteristics (except malaria ecology, natural capital and trade openness) are all significant at the 5% level or better.

The estimates for three of the geographic measures deserve particular attention. After experimenting with both a linear and a quadratic specification, we find that the latter performs best for latitude, and that the significant coefficients and their signs are incompatible with the widespread belief, also supported in many regressions using world samples, that countries further from the equator grow faster.[37] The surprising negative and significant coefficients ($p < 0.05$) on the Coast Pop. Shr. variable are contrary to Sachs and Warner (1997)'s finding (see also Gallup, Sachs and Mellinger, 1999) that countries in the vicinity of an ocean or an ocean-navigable river tend to grow faster than landlocked countries. Our result suggests that, among African countries, those close to the ocean did not grow any faster than those further from it, and indeed, performed worse. The size of the coefficient implies that Coast Pop. Shr. would have a large impact on economic development. The coefficient on log of predicted trade share (Pred. Trd. Shr.) is positive and significant in the country effect estimates that include it.[38]

Turning to the social and historical variables, we find positive and significant coefficients on statehist01, indicating that countries with long histories of polities above the

[36] See Hoeffler, 2002 for a detailed discussion on these issues.

[37] See, for example, Sala-i-Martin 1997. Some West African countries such as Nigeria and Ghana could have hotter and more humid climates than countries like Kenya, Rwanda and Burundi, which lie closer to the equator but have much of their populations living at high and thus relatively temperate altitudes. In Africa, too, distance from the equator, especially northward, can mean prevalence of deserts, which may have economic disadvantages equal to those of tropical climate. Note again that more direct measures of climate were not significant in our regressions, and thus are left out of the specifications shown.

[38] As for other geographic measures, natural cap. is never significant but always with the familiar sign. This captures an unexpected facet of economic life, but one that has by now been confirmed for world samples— resource-poor economies often outperform resource-rich economies in economic growth. Malaria Ecol. (Malaria ecology) has the expected sign, but is not significant.

tribal level tended to have faster growth than those without such histories, all else being equal. The coefficients on ethno-linguistic fractionalization (Eth. Heter.) are all highly significant, and their sign contrary to that found by Easterly and Levine (1997) for the related ethnic fractionalization measure. The included religious measure variable (% Muslim) indicates that countries with a large proportion of the population affiliated with muslim faith achieved significantly lower levels of growth than did other African countries. [39]

3.2 2SLS and IV Regression

One of the drawbacks of the method used in the previous section is that it fails to distinguish between variables that are clearly exogenous to the current policy environment and the current performance of institutions and the economy, and variables that may be influenced by them. While a country's latitude, its ethnic makeup, and the power that colonized it in the nineteenth century cannot possibly have been influenced by its adoption of trade openness or experience of civil wars in recent times, the reverse could easily be the case.

In this section we reanalyze our data, taking as the pivot of our methodology a distinction between endogenous and exogenous variables, rather than between time-variant and time-invariant variables. We estimate a set of two-stage least squares regression models in which geographic and historical variables are used as instruments to predict civil wars, trade openness, political rights, political openness, and corruption, presenting only those results that pass tests for both the endogeneity of the latter variables and for over-identifying restrictions. Assuming that the unobserved country characteristics, η_i, are not significant and using measured geographical and socio-historical factors as instruments, the 2SLS-IV regression follows this system of equations:

$$\begin{aligned} g_{i,t} &= a_o + a_1 Y_{i,t} + a_2 X_{i,t} + v_{i,t}, \\ Y_{i,t} &= b_o + b_1 Z_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{3.2.1}$$

[39] This result contrasts with Sala-i-Martin (1997)'s finding that the proportion of Muslims is positively correlated with growth of GDP per capita in regressions using world samples. % Catholic was insignificant and thus dropped from these regressions; it appears in one IV model.

where X is a set of exogenous controls, which act directly on the outcome, and Y is a vector of endogenous institutional and policy variables, while Z is a vector of time-variant and time-invariant exogenous instruments, and $\varepsilon_{i,t}$ is an error term. In this specification, the Z vector is being used to instrument for the set of endogenous variables contained in the Y vector.

Our identification methodology imposes restrictions that the instruments should be significantly correlated with the endogenous variables, uncorrelated with the error term, and should not influence the independent variable by themselves; that is $E(Y_{i,t}Z_{i,t}) \neq 0$, $E(Z_{i,t}v_{i,t}) = 0$, and $E(Z_{i,t}g_{i,t}) = 0$. Under these conditions, we can use instrumental variable estimation, provided that there is at least one instrument for each endogenous variable. We use versions of the Sargan and Basmann tests of over-identifying restrictions for a regression estimated via instrumental variables in which the number of instruments exceeds the number of regressors. To test for endogeneity, we use the "Durbin-Wu-Hausman" (DWH) test, which is numerically equivalent to the standard "Hausman test."[40] The null hypothesis states that an OLS estimator of the same equation would yield consistent estimates.

3.2.1 First Stage Regressions for IV Models

In Table 3, we show estimates of one first-stage regression for each of five institutional and policy variables. For each 1^{st}-stage regression, different sets of instruments were experimented with until one was found that (a) had good predictive power, (b) passed over-identification restrictions tests with respect to the 2^{nd}-stage regressions, and (c) passed the endogeneity test for each regressor specified as endogenous.[41] We note that while some of the 1^{st}-stage regressions have more individually significant coefficients than do others, each regression as a whole explains enough of the variance in its dependent variable to satisfy these requirements.[42]

[40] Davidson and MacKinnon (2004, pp338-340)

[41] The tests for endogeneity and over-identifying restrictions for the validity of results presented in Table 4 have been conducted and are presented in Table 5.

[42] Note that if column 1 were a stand-alone regression to explain the frequency of civil wars, it would be preferable to use a tobit regression, since the dependent variable has minimum and maximum limits of 0 and 1, respectively that are often observed in the data. According to Wooldridge, page 84 (2002), however, first stage IV should be performed by OLS even when the variable to be instrumented is binary,

Table 3: First Stage Regressions for IV Models

	Dependent Variable.				
Indep. Var.	**Civil Wars**	**Yrs. Open**	**Polit. Rights**	**Polit. Open.**	**Corruption**
Pred. Trd. Shr.		0.093***		2.162***	-1.072***
		(0.01)		(0.435)	(0.078)
Latitude		-0.029***		-0.535***	0.195***
		(0.003)		(0.094)	(0.023)
Latitude.sq		0.002***		0.039***	-0.013***
		(0.0002)		(0.004)	(0.001)
Eur. Lng. Shr.		-0.392**			
		(0.200)			
Relig. Frac.	0.005				
	(0.003)				
Relig. Polar.			0.834		
			(0.637)		
% Catholic				0.034**	
				(0.011)	
% Muslim	0.002				0.01***
	(0.002)				(0.002)
Eth. Heter.	-0.099	0.206***	-0.269	3.492***	-0.584**
	(0.151)	(0.018)	(0.396)	(0.894)	(0.284)
Statehist01	-0.228	0.343***	1.675		-4.358***
	(0.291)	(0.065)	(1.280)		(0.589)
Eng. Lang. Shar.			3.042		
			(4.328)		
Other. Col.	0.313***		-1.076***	-0.584	
	(0.113)		(0.280)	(0.673)	
Constant	-0.055	-0.319***	2.819***	-6.234***	7.12***
	(0.210)	(0.035)	(0.674)	(1.166)	(0.324)
R^2	0.14	0.49	0.19	0.54	0.77
# Obs.	191	191	138	200	136
# Countries	25	25	24	26	18

discrete, or limited in range. To check the robustness of the results shown in the Table 3, we also estimated each regression using tobit/probit, and found that the results were not dramatically different.

To understand the implications of individual results in Table 3, we note that Table 4 will show Civil Wars and Corruption to be strong negative predictors, and Yrs. Open, Polit. Rights and Polit. Open. strong positive predictors, of economic growth. By implication, then, a factor that has a significant, say, positive effect on, say, a negative determinant of growth, in Table 3, can be seen to ultimately exert a negative effect on growth, albeit indirectly. Keeping this in mind, we can more conveniently discuss results for the table as a whole. For example, Other Col., signifying Belgian, Italian or Portuguese colonization, is a significant positive predictor of civil wars, a significant negative predictor of political rights, and is insignificantly negatively related to political openness, all of which point to negative indirect effects of Other Col. on economic growth. Predicted trade share has significant positive effects on Yrs. Open and Polit. Open., and a significant negative effect on Corruption, all pointing to positive indirect effects on growth.

Statehist01 appears in four of the five 1^{st}-stage regressions, in all cases with signs consistent with its positive correlation with growth in the GMM residuals analysis, and significant at the 1% levels in the equations for Yrs. Open and Corruption. Ethnic heterogeneity and latitude also figure in the regressions with significant results consistent with those in the GMM analysis but contrary to conventional expectations. More ethnically heterogeneous countries have significantly less corruption, more political openness and more trade openness. Although both level and square terms for latitude are significant in the same equations, the larger magnitudes of the coefficient on the level term make clear that countries further from the equator were less open to trade, less politically open, and suffered from more corruption.

Most of the remaining variables appear in fewer regressions, with only the Catholic and Muslim population shares ever attaining independent significance. The Catholic share appears to aid growth by way of its effect on political openness, while the Muslim share appears to harm growth by increasing corruption.

3.2.2 Second Stage Regressions for IV Models

The second stage regression results are shown in Table 4. As in Hall and Jones (1999), only one of the five endogenous variables predicted in Table 3 is used in each column to

help explain the rate of growth of per capita GDP.[43] A glance at Table 4 shows that, entered individually, each instrumented variable is statistically significant, four of them at the 1% level. The models also perform nicely in that the initial GDP and investment share variables attain significant coefficients of the predicted sign in all five regressions.

Table 4: Second Stage Regressions for IV Models

Dependent variable: Gr(GDPpc).

Indep. Var.	**1**	**2**	**3**	**4**	**5**
Civil wars	-4.651**				
	(2.013)				
Yrs. Open		14.900***			
		(2.865)			
Polit. Rights			1.755***		
			(0.681)		
Pol. Open.				0.529***	
				(0.116)	
Corruption					-1.451***
					(0.249)
Log iGDPpc	-1.080*	-1.084**	-1.906**	-1.500***	-1.600***
	(0.646)	(0.547)	(0.897)	(0.482)	(0.655)
Log(I/GDP)	1.851***	2.246***	2.902***	2.928***	2.861***
	(0.640)	(0.464)	(0.611)	(0.478)	(0.577)
Coast Pop. Shr.	-1.689	-1.501*	-2.433**		-0.718
	(1.156)	(0.935)	(1.260)		(1.227)
Natural Cap.	-0.0002*	-0.0001	-0.0001	-0.00012	-0.0001
	(0.0001)	(0.00007)	(0.00009)	(0.00008)	(0.00008)
Malaria Ecol.	-0.043*	-0.045**	-0.039	-0.018	-0.039
	(0.027)	(0.023)	(0.030)	0.022	(0.030)
Constant	5.125	1.909	0.733	1.780	8.858**
	(4.804)	(3.439)	(4.574)	3.038	(4.132)
Uncenterd R^2	0.11	0.32	0.23	0.33	0.44
# Obs.	191	191	138	200	136
# Countries	25	25	24	26	18

[43] Including several instrumented variables simultaneously in regressions paralleling those in Table 4 failed to give good results.

The regression in column 1 treats civil wars as endogenous, using the predicted value from the corresponding column of Table 3. It confirms the well known result (Easterly and Levine (1997), Alesina and Perotti (1996a)) that civil wars have been very detrimental for African economies. The negative and significant ($p < 0.05$) coefficient on this variable suggests that a one standard deviation increase in the probability of civil wars will reduce the rate of economic growth by (0.55) standard deviations. Results for the other variables in the five regressions are discussed following remarks on the other endogenous variables.

The results in column 2 treats the Sachs-Warner policy variable, Yrs. Open, as endogenous (predicted by the corresponding regression of Table 3) and finds it to be a significant predictor ($p < 0.01$) of economic growth in Africa. In columns 3 and 4, the Freedom House measure of political rights and civil liberties, and the political openness variable from the Polity IV data set are, respectively, the instrumented variables (predicted by the associated Table 3 regressions). Each variable obtains a highly significant ($p < 0.01$) positive coefficient. Finally, column 5's regression has the corruption perception index as instrumented explanatory variable predicted by column 5 of Table 3. The negative significant ($p < 0.01$) coefficient confirms that corruption has been bad for growth in Africa, consistent with Mauro's (1995) findings for a world sample. This coefficient suggests that a one standard deviation increase in the degree of corruption reduces the rate of growth by (0.43) standard deviations.

Turning to the other variables in Table 4, the estimates for Coast Pop. Shr. reaffirm the surprise found in the GMM model: the coefficients are negative in all four cases where coastal population is included, and significant at the 5% and 10% levels respectively in two of them.[44] This provides additional evidence paralleling results in the second stage of the GMM estimates that greater coastal population is unhelpful for growth.

The now familiar negative effect of natural capital is supported by all five equations, but significant at the 10% level in only the first equation. Finally, the coefficients on Malaria Ecol. are all negative with two of them significant at 10% and 5% levels, respectively.

[44] We dropped Coast Pop. Shr from column 3 to improve the overal fit of that model.

Table 5: **Endogeneity and Over-identifying Restrictions Tests.**

	Tests	Civil Wars	Yrs. Open	Polit. Rights
Overid.	Sargan (χ^2)	5.034	8.45	3.33
		(0.28)	(0.13)	(0.50
	Basmann (χ^2)	4.68	7.96	3.02
		(0.32)	(0.15)	(0.55)
Endog.	WH (F)	5.351^b	9.22^a	2.88^c
		(0.021)	(0.002)	(0.09)
	DWH (χ^2)	5.636^b	9.50^a	3.11^c
		(0.017)	(0.002)	(0.07)
	Instruments	Statehist01	Statehist01	Statehist01
		% Muslim	Eur. Lang. Shr.	Eng. Lang. Shr.
		Relig. Frac.	Pred. Trd. Shr.	Relig. Polar.
		Other Col.	Latitude	Other Col.
		Eth. Heter.	Latitude. sq	Eth. Heter.
		na	Eth. Heter.	na

Table 5 continued.

	Tests	Polit. Open.	Corruption
Overid.	Sargan (χ^2)	9.12	7.84
		(0.10)	(0.16)
	Basmann (χ^2)	8.69	7.15
		(0.12)	(0.20)
Endog.	WH (F)	6.97^a	13.46^a
		(0.008)	(0.0003)
	DWH (χ^2)	7.23^a	13.62^a
		(0.007)	(0.0002)
	Instruments	Pred. Trd. Shr.	Pred. Trd. Shr.
		% Catholic	Statehist01
		Other. Col.	% Muslim
		Eth. Heter.	Eth. Heter.
		Latitude	Latitude
		Latitude. sq.	Latitude. sq.

Overid. stands for over-identifying restrictions and Endog. stands for endogeneity tests. DWH refers to Durbin-Wu-Hausman chi-sq test and WH refers to Wu-Hausman F test. Entries in parentheses are p-values, and a, b and c indicate significance levels at 1, 5 and 10%.

Table 5 presents alternative tests of endogeneity (Wu-Hausman F test and Durbin-Wu-Hausman Chi-sq test) and of overidentifying restrictions (Sargan N*R-sq test and

Basmann test). In all five cases where an endogenous variable is instrumented, these tests a) confirm the endogeneity of each policy or institutional variable at 10% level or better, and b) validate our instruments by failing to reject the null hypothesis, $E(Z_{i,t}V_{i,t}) = 0$ and $E(Z_{i,t}g_{i,t}) = 0$, with $p - value > 10\%$.

3.3 Consistency of GMM and IV Models, and Robustness

By virtue of the negative and significant coefficients on initial GDP per capita, our dynamic GMM and IV technique estimates both confirm the usual conditional convergence result, i.e. income is growing faster in poorer countries, all else being equal. The GMM estimates also indicate that countries' growth rates were slowing during the period covered, with the previous period's growth rate a significant negative predictor of that in the current period. The effect of the investment/GDP ratio is robustly positive and significant in both sets of estimates, with a 1% change in (I/GDP) leading to a 4.2 to 5.2% change in the GDP growth rate according to the GMM estimates and a 1.8 to 2.9% change in GDP growth rate according to the IV estimates. Findings with regard to geographic, historical, institutional and policy variables also show considerable consistency across methods of analysis. We summarize those findings in Section 4.

Turning to robustness concerns, in all regressions we corrected for heteroskedasticity or implemented regressions with robust variance estimates[45]. We checked for their robustness to the exclusion of Botswana, a major outlier, and to different ways of dealing with the correlation between predicted trade share, coastal population share and trade openness. These investigations are summarized in Appendix 2.

4 Summary and Conclusions

This paper has explored the role of political, social, and geographic factors in explaining poor economic performance of countries in sub-Saharan Africa in the past forty years. Our findings indicate that initial GDPpc and the investment share of GDP influence economic

[45] Unlike other tables where we display results with robust standard errors, results shown in Table 4 do not correct for heteroskedasticity because the endogeneity and over-identifying restrictions tests, which are fundamental to our results, cannot be obtained if we correct for heteroskedasticity there.

growth in the conventional manner in sub-Saharan African countries, providing assurance that the data are adequate and that the laws of economics hold south of the Sahara as elsewhere. But the available measure of human capital (average schooling) is never significant and always has a negative sign. We also show that several types of variables, in addition to the conventional economic ones, explain poor performance in sub-Saharan Africa. These include institutional and policy variables (civil wars, economic openness, political rights, political openness and corruption), geographic variables (latitude, coastal population share, natural capital, malaria ecology, and the Frankel-Romer predicted trade share), and social and historical variables (ethnic heterogeneity, the statehist measure of pre-modern development, colonizing power, religion, and proportions speaking Western European languages in general and English in particular). Our results contrast with Sala-i-Martin's finding (for global samples) that a larger Muslim share of population is associated with higher rates of growth, and they accord with Bertocchi and Canova's finding that having been colonized by Belgium, Italy or Portugal conferred a worse fate than colonization by England or France.

Our results contained there major surprises. First, coastal population share has consistently negative effects on the rate of growth in our African country panel according to both the GMM and the IV models. Thus, we find no support for Sachs and Warner's contention that one reason Africa's countries have grown slowly is that so many of them are landlocked. Perhaps those African countries that have coastlines have yet to reap their potential trading advantages–due in part to the dearth of good ports and good port management, in part to various obstacles to the growth of export-oriented manufacturing.

Second, contrary to the oft-mentioned idea that proximity to the equator is a disadvantage, in our African panel countries closer to the equator have been growing faster, all else equal. Our GMM residual estimates give highly significant coefficients implying that the growth rate is decreasing, although at a decreasing rate, with latitude. In the first-stage IV regression models in which it proved helpful to include latitude (columns 2, 4 and 5), the coefficients are also highly significant and imply that trade openness and political openness are decreasing at a decreasing rate, and corruption is increasing at a decreasing rate, with latitude.

Third, contrary to Easterly and Levine's hypothesis and their evidence from global

regressions to the effect that ethnic heterogeneity may account for "Africa's growth tragedy," the ethnic heterogeneity variable, identical to the main measure of that concept used in their paper, consistently attracts significant coefficients of the "wrong" sign in our Africa-only panel. In all four GMM residual regressions shown, ethnic heterogeneity is positively related to growth residual, significant at the 1% level in three of them and significant at 5% in one regression. In the 1st-stage IV regressions, ethnic heterogeneity has signs consistent with an ultimately positive effect on growth in four of five specifications. It is a significant positive predictor of Yrs. Open ($p < 1\%$) and Polit. Open. ($p < 1\%$), and a significant negative predictor of Corruption ($p < 5\%$). Ethnic heterogeneity is also negatively correlated with civil wars, but not significantly so. Its sign in the equation for political rights is the only one that runs contrary to the trend of the other regressions, but that estimate is also insignificant.[46] The positive effect of ethnic heterogeneity may be due to the fact that most African societies are at least somewhat heterogeneous and that moving from few to many groups actually reduces, rather than increasing, the negative impacts of heterogeneity on social, political, and thus economic outcomes.

We introduced a variable not previously focused on in the African context, state history, and we found its inclusion to be statistically warranted and often important in both our GMM and our IV models. In the GMM residual estimates, a longer state history always predicts faster growth, significant at the 1% level in three equations (columns 2-4) and significant at the 5% level in one equation (column 1). In the IV models, statehist01 was found useful in predicting four endogenous policy and institutional variables, though it is only independently significant in the second and fifth columns ($p < 1\%$).[47] Its positive association with the growth rate and its negative association with corruption are

[46] Because other measures of ethnic heterogeneity and of similar concepts have been suggested, and in view of findings that ethnic heterogeneity's effects may be non-linear, we also tried using measures including ethnic fragmentation, social fractionalization, the share of the population not speaking official language and ethnolinguistic dominance, and we estimated versions of our GMM and IV models that included ethnic heterogeneity in both level and square terms. No other measure consistently performed as well as the ethnic heterogeneity variable used by us (and by Easterly and Levine (1997)). Quadratic specifications returned positive level and negative square term coefficients, often significant, supporting the idea of a concave relationship with an interior maximum. Since the coefficient on ethnic heterogeneity is significant more often in the simpler version and since R^2's for the regressions are about the same, we show the linear estimates in our tables.

[47] Statehist is not included in the predicting of political openness since its inclusion in the set of instruments violates the endogeneity and over-idendifying restrictions tests.

consistent with the previous findings of Bockstette *et al.* for world samples.[48]

Most of these results were confirmed by both our GMM and residuals approach and our two-stage least squares analysis, but the latter approach suggests a more specific interrelation among the causal factors, namely that policies and institutions like the existence of greater corruption or more frequent civil wars, while often quite decisive as proximate causes of which countries make progress and which fail to do so, are to a significant degree caused by longer term processes and more exogenous geographical and demographic facts.

How might we view these results if concerned about the near-term prospects of Africa's economies? On the positive side we count the fact that geographic and demographic factors like equatorial and inland location and ethnic heterogeneity do not seem to be bars to growth in their own rights, and that changing policies on trade and creating greater political openness and lessening corruption appear capable of facilitating economic growth. On the negative side, the social, cultural and historical legacies at work in many countries of the region may make the creation of high quality institutions and policies a challenging, though one can hope not impossible, task. Our confirmation of the importance of those institutions suggests to us that studying what does and doesn't work in the area of capacity building in public, private, and civil spheres in Africa is one of the most important directions for future research.

[48] See also Chanda and Putterman (2005). Gennaioli and Rainer (2005) find a positive association between the degree of centralization of African tribes before colonization and measures of the success of provision of certain public goods including paved roads and lower infant mortality during the early post-colonial period. Although their result appears broadly consonant with ours, their concept of centralization applies to many tribes that would not be classified as having states by the definition used in Bockstette *et al* (2002) (and our revised source, Putterman 2004), and their centralization series is not positively correlated with statehist01 for our sample.

References

[1] Acemoglu, Daron, Simon Johnson and James Robinson, 2001, "The Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91 (5): 1369-1401.

[2] Acemoglu, Daron, Simon Johnson and James Robinson, 2002, "The Reversal of Fortune: Geography and Development in Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 117(4) 1231-1294

[3] Alesina, A., Baqir, R., Easterly, W., 1999, "Public Good and ethnic divisions," *Quantitative Journal of Economics* 114, 1243-1284.

[4] Alesina, A., Perotti R., 1996a, "Income Distribution, Political Instability, and Investment." *European Economic Review* 40, 1203-1228

[5] Azam Jean-Paul and A., Hoeffler, 2002, "Violence against Civilians in Civil Wars: Looting or Terror?," *Journal of Peace Research* 39: 461-485

[6] Barkan, Joel, 2002, "The Many Faces of Africa: Democracy Across a Varied Continent", *Harvard International Review* 24 (2): 72-77.

[7] Barro, Robert, 1991, "Economic Growth in a Cross-Section of Countries," *Quarterly Journal of Economics* 106: 401-44.

[8] Barro Robert, 1997, "*Determinants of Economic Growth: A Cross-Country Empirical Study*," MIT Press.

[9] Barro, and Lee, 2000, "International Data on Educational Attainment: Updates and Implications," *CID* Working Paper no. 42

[10] Battese, G. E. and Coelli, T. J. 1995, "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data." *Empirical Economics*, Vol. 20, pp. 325–32.

[11] Bertocchi, Graziella and Fabio Canova, 2002, "Did Colonization Matter for Growth? An Empirical Exploration into the Historical Causes of Africa's Underdevelopment," *European Economic Review*, Vol. 48, pp 1851-1871.

[12] Blanchflower, D. G., Oswald, A. J. and Sanfey P., 1996, "Wages, Profits, and Rent-Sharing", *The Quarterly Journal of Economics*, Vol 111, pp. 227–51

[13] Bockstette, Valerie, Areendam Chanda and Louis Putterman, 2002, "States: The Advantage of an Early Start," *Journal of Economic Growth* 7: 347-69

[14] Casselli, F., Gerardo Esquivel and Fernando Lefort, 1996, "Reopening the convergence debate: a new look at cross-country growth empirics," *Journal of Economic Growth*, vol. 1, issue 3, pp363-89

[15] Chamberlain, G, 1983, "Panel Data", in Z. Griliches and M.D. Intrilligator (eds), *Handbook of Econometrics*, Vol. 2, Amsterdam: North Holland.

[16] Chanda, Areendam and Louis Putterman, 2005, "State Effectiveness, Economic Growth, and the Age of States," in Matthew Lange and Dietrich Rueschemeyer, eds., *States and Development: Historical Antecedents of Stagnation and Advance,* Basingstoke, England, Palgrave-MacMillan.

[17] Clement, Jean A.P., Ed, 2005, "Post conflict Economics in sub-Saharan Africa: lessons of the Democratic Republic of the Congo," *Washington, International Monetary Fund.*

[18] Collier, Paul and Anke Hoeffler, 2002, "On the incidence of Civil War in Africa," *Journal of Conflict Resolution* 46: 13-28.

[19] Collier, Paul and Gunning, Jan Willem, 1999, "Explaining African Economic Performance," *Journal of Economic Literature* 37: 64-111.

[20] Davidson and MacKinnon, 2004, "*Econometric Theory and Methods*," Oxford University Press (New York)

[21] Easterly, William and Ross Levine 2003, "Tropics, Germs, and Crops: How Endowments Influence Economic Development" *Journal of Monetary Economics*, 50:1, 3-40

[22] Easterly, William and Ross Levine, 1997, "Afica's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics* 112: 1203-1246

[23] Frankel, J, Romer, D, Cyrus, T, 1996, "Trade and growth in East Asian countries: cause and effect?" *National Bureau of Economic Research Working Paper* No.5732.

[24] Frankel, J, Romer, D, 1999, "Does trade causes growth?" *American Economic Review*, Vol. 87, pp 379-99.

[25] Gallup, John Luke and Jeffrey Sachs with Andrew Mellinger, 1999, "Geography and Economic Growth," pp 127-178 in Boris Pleskovic and Joseph Stiglitz, eds., *Annual World Bank Conference on Development Economics 1998*. Washington: The World Bank.

[26] Garner, Phillip, 2005, "Reversal in Africa," unpublished paper, Brigham Young University Dept. of Economics.

[27] Gennaioli, Nicola and Rainer, Ilia, 2005 "The Modern Impact of Pre-colonial Centralization in Africa," unpublished paper, Stockholm University and George Mason University.

[28] Gyimah-Brempong, K, Thomas L. Traynor, 1999, "Political Instability, Investment and Economic Growth in Sub-Saharan Africa", *Journal of African Economies*, Vol.8, No.1.

[29] Hadjimichael, M.T., D. Ghura, M Muhleisen, R. Nord, and M. Ucer, 1995, "Sub-Saharan Africa: Growth, Savings, and Investment, 1986-93," *Occasional paper 118*, International Monetary Fund, Washington DC.

[30] Hall, Robert E. and Jones, Charles I., 1999, "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114:83-116.

[31] Hoeffler Anke, 2002, "The Augmented Solow Model and the African Growth Debate," *Oxford Bulletin of Economics & Statistics* Vol.64, No.2, 135-158

[32] Islam, N., 1995, "Growth Empirics: A Panel Data Approach", *Quarterly Journal of Economics,* 110

[33] Kaufman, D., Kraay, A., Zoido-Lobaton, P., 1999a, "Aggregate Governance Indicators," *World Bank Research Working Paper* No. 2195.

[34] King, R., and R. Levine, 1993, "Finance and Growth: Schumpeter Might Be Right", *Quarterly Journal of Economics*, Vol. 153, 717-738.

[35] Kiszewski, Mellinger, Spielman, Malaney, Ehrlich Sachs, Jeffery Sachs, 2004, "A Global index representing the stability of Malaria Transmission," *American Journal of Tropical Medicine and Hygiene*, 70(5): 486-498.

[36] Knight, M., N. Loayza and D. Villanueva, 1993, "Testing the Neoclassical Growth Model", *IMF Staff Papers*, Vol. 40, 512-541

[37] Laporta, Lopez-de-Silanes, Shleifer, Vishny, 1999, "the quality of government", *Journal of Law, Economics, and Organization* 15, 222-279.

[38] Levine, Ross and David Renelt, 1992, "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review* 82: 942-63.

[39] Lorentzen, McMillan and Wacziarg, 2005, "Death and Development", *NBER Working Paper* No. W11620.

[40] Mankiw, Gregory N., David Romer and David N. Weil, 1992, "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107(2): 407-437.

[41] Mauro, P., 1995, "Corruption and growth," *Quarterly Journal of Economics* 110, 681-712.

[42] Ojo, O., and Oshikoya T., 1995, "Determinants of Long term growth: Some African Results", *Journal of African Economies*, 4, 163-191.

[43] Putterman, Louis, 2004, "State Antiquity Index ("Statehist") Version 3," available at http://www.econ.brown.edu/fac/Louis%5FPutterman/

[44] Reynal-Querol, 2002, "Ethnicity, Politial Systems, and Civil Wars," *Journal of Conflict Resolution*, Vol. 46, No.1, 29-54.

[45] Rodrik D., 1999, "Trade Policy and Economic Performance in Sub-Saharan Africa," *NBER Working Paper* No. W6562.

[46] Sachs, Jeffrey and Andrew Warner, 1995, "Natural Resource Abundance and Economic Growth," *NBER Working Paper 5398.*

[47] Sachs, Jeffrey and Andrew Warner, 1997, "Sources of Slow Growth in African Economies," *Journal of African Economies* 6(3): 335-76.

[48] Sala-i-Martin, X., 1997, "I just Run Two million regressions," *American Economic Review*, Vol. 87, pp 178-183.

[49] Savvides, Andreas, 1995, "Economic Growth in Africa," *World Development"* 23(3), 449-458.

[50] Tahari A, Ghura D, Akitoby B and Brou-Aka E, 2004, "Sources of Growth in Sub-Saharan Africa", *IMF working Paper* No. 171.

[51] World Bank, 1997, "Expanding the Measure of Wealth: Indicators of Environmentally Sustainable Development," Environmentally Sustainable Development Studies and Monographs Series, #17.

[52] Wooldridge, Jeffrey M., 2002, *Econometric analysis of cross section and panel data.* Massachusetts Institute of Technology

Appendix 1: Variable definitions and sources

Civil wars: A dummy variable indicating whether the country is at war. It only considers internal wars which resulted in at least 1000 battle related deaths (civilian and military) per year. Data are from Azam and Hoeffler (2002) and were used in Collier and Hoeffler (2002).

Coast Pop. Shr. (Coastal population share): Measures the ratio of population within 100 km of ice-free coast to total population. Calculated using 100 kilometers from ice-free coast buffer. Data come from Gallup, Sachs with Mellinger (1999).

Corruption: Based on the Corruption Perceptions Index (CPI) from Transparency International (2003) (see www.Transparency.org). To obtain a measure that is non-negative and is increasing rather than decreasing in the amount of perceived corruption, we define Corruption as 6 – CPI.

Eng. Lng. Shr. (English language share) : Measures the fraction of population speaking English as first language. Data are from Hall and Jones (1999).

Eth. Heter. (Ethnic heterogeneity): The average of five different indices of ethnolinguistic fractionalization. Gives the probability of two random people in a country not speaking a same language.

Eth. Heter. sq. (Ethnic heterogeneity squared): See Ethnic heterogeneity.

Eur. Lng. Shr. (European language share): Measures the fraction of the population speaking one of five major Western European languages, including English, as mother tongue. Data are from Hall and Jones (1999).

Gr(GDPpc) (Growth rate of GDP per capita): Measures the annual percentage growth rate of GDP per capita based on constant local currency (source: World development indicators).

Lagged Gr(GDPpc)(Lagged growth rate of GDP per capita): The value of Gr (GDPpc) for the previous five year period. See the definition for Gr(GDPpc)).

Latitude: Measures the absolute value of the latitude of the country, where zero is the equator. La Porta et al (1999).

Latitude. sq (Latitude squared): See Latitude.

Log(I/GDP) (Investment share of GDP): Measures the natural log of investment-to-GDP ratio. World development indicators.

Log iGDPpc (Log of initial GDP per capita): Measures the natural log of real gross

domestic product per capita in the first year of the five year period under observation. (source: Own computations based on data from World development indicators).

Malaria Ecol. (Malaria Ecology): A measure of ecological factors associated with the risk of contracting malaria, used in place of actual malaria incidence in order to eliminate possible influences of country income on it. The basic formula for ME includes temperature, species abundance, and vector type. Source: Kiszewski et al. (2004).

Natural Cap. (Natural Capital): Measures the value of a country's agricultural lands, pasture lands, forests and subsoil resources including metals, minerals, coal, oil and natural gas. (source: World Bank, Environmentally Sustainable Development Studies and Monographs Series, #17, June 1997).

Other Col. (Other Colony dummy): Dummy variable indicating that country was a Belgian, Italian, Portuguese or German colony. Germany is not represented in our data due to missing observations and the coding of former German colonies under whichever power controlled them after World War I. Data are from La Porta et al., (1999).

Other Lang. (Other Language): Measures the percent of population not speaking the official language (source: Easterly and Levine, 1997).

Percent Catholic : – Measures the percentage of the population that belonged to the Roman catholic church in 1980 (or 1990-95 for countries formed more recently). Data are from La Porta et al., (1999).

Percent Muslim: Measures the percentage of the population that belonged to the Muslim faith in 1980 (or 1990-95 for countries formed more recently). Data are from La Porta et al., (1999).

Polit. Open. (Political Openness): Measures the general openness of political institutions. This index ranges from 0-10, where 10 denote a highly open regime. It is the variable that is called "democracy" in the Polity IV data set and is also referred to as "democracy" by Collier and Hoeffler (2002) and is used by Azam and Hoeffler (2002).

Polit. Rights (Political Rights): An average of an index of political rights and an index of civil liberties from Freedom House (2003), www.freedomhouse.org. Political rights include the right to vote, the right to compete for a public office; while civil liberties include the freedoms of expression and belief, associational and organizational rights, rule of law and personal autonomy without interference from the state. Data are from an annual survey and are available since

1978 for 198 countries.

Pred. Trd. Shr. (Log of Predicted Trade Share): Measures the natural log of the Frankel-Romer predicted trade share. The predicted trade share is computed from a gravity model based only on population and geography (source: Frankel, Romer and Cyrus (1996)).

Relig. Frac. (Religious fractionalization index): Ranges from 0 to 100. A value of zero indicates that the society is completely homogenous whereas a value of 100 would characterize a completely heterogeneous society (source: Collier and Hoeffler (2002)).

Relig. Polar. (Religious polarizion): Indices taking the form $IRC1 = 1 - \sum_{i}^{N} (0.5 - \pi_i)^2 \pi_i / 0.25$, where π_i is the proportion of the religion or ethnic group and N is the number of religions or ethnic groups. They take their largest values when the maximum share is $\frac{1}{2}$ and their smallest value when the maximum share is 1/6. From Reynal-Querol (2002).

Sec. Enrol (Secondary School Enrollment Ratio.): Average years of schooling for those 25 years old and over (Source: Barro, and Lee(2000)).

Social frac. (Social fractionalization index): The product of the ethno-linguistic fractionalization and the religious fractionalization plus the ethno-linguistic fractionalization or the religious fractionalization index, which ever is the greater. Definition and data are from Collier and Hoeffler (2002).

Statehist01 (State history index): Measures the proportion of years from 1 to 1950 C.E. in which there were one or more state level polities, whether they were indigenous or imposed, and the proportion of the territory they encompassed, discounting backward at a 1% rate for each half century. An updated version of the index used in Bockestette et al., 2002. From Putterman, 2004.

Yrs. Open (Years Open): Measures the fraction of years during the period 1950 to 1994 that the economy has been open according to criteria proposed by Sachs and Warner, which include (1) average tariff rate below 40%, (2) less than 40% of imports covered by quotas and licensing, (3) black market premium below 20%, (4) absence of extreme controls on exports, and (5) absence of a socialist economy. It is measured on a scale from 0–1 (source: Sachs and Warner, 1997).

Appendix 2: **Robustness Checks**

A sensitivity analysis was conducted using Botswana as outlier. We thought that the presence of Botswana — Africa's star economic performer and a landlocked country with a nearly homogeneous population— might drive the strong negative coefficient on coastal population share. But we found no important changes in the GMM results when Botswana was dropped from the sample, and although there is a drop in the significance level in the two IV specifications where the coefficients on coastal population share were significant (11% to 57% and 5.3% to 33%), the coefficients on Coast Pop. Shr. show no sign changes. Regarding the effects of dropping Botswana on Ethnic heterogeneity, on state history, and on latitude, we found no important changes in the GMM results, but an increase in the statistical significance of the latitude variable in the IV regressions. Turning to the robustness of the policy and institutional variables when Botswana is dropped, we found that the significance of these variables was slightly affected, except that corruption and political rights became insignificant. Performing new endogeneity tests without Botswana in the sample, we found that four out of five endogenous variables pass the tests at 10% or better, but the equation for political rights fails to pass this test.

We were also worried about a possible correlation between predicted trade share, coastal population share and the policy variable "Yrs. Open". It happens that the correlation between the trade share and Coast Pop. Shr. is very large. Predicted trade share and trade openness could also be related to each other in two ways if predicted trade share indicates the natural level of trade. First, a high natural level of trade may make it more likely for the country to adopt more open policies, and second if a country has a high trade potential (predicted trade share), then openness will benefit its growth more than if the country has low trade potential. We included, therefore, (in estimates not shown here) an interaction term, the product of predicted trade share with the openness variable. When using the predicted trade share and the openness variables together with the interaction term, only the predicted trade share is significant while both the openness variable and the interaction term attain insignificant coefficients. When the predicted trade share is used with the interaction term, only the predicted trade share is significant with the correct sign. When the interaction term is used with Yrs. Open., neither one is significant. Finally, when the interaction term is used alone, it gives a large positive coefficient which is significant at five percent level. We found, therefore, no evidence of multicollinearity driving our results.

Appendix 3: Country Inclusion by Regression

Tables/ Columns/ Countries

	Table 1				Table 2				Table 3					Table4				
	1	2	3	4	1	2	3	4	1	2	3	4	5	1	2	3	4	5
#	24	24	24	33	29	29	29	25	25	25	24	26	18	25	25	24	26	18
BEN	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
BWA	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
BFA				υ	υ	υ	υ	υ	υ	υ	υ	υ		υ	υ	υ	υ	
BDI				υ	υ	υ	υ	υ	υ	υ	υ	υ		υ	υ	υ	υ	
CMR	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
CAF	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ		υ	υ	υ	υ	
TCD				υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
COG	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
DRC	υ	υ	υ	υ														
GAB				υ	υ	υ	υ											
GMB	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
GHA	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
GNB	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ		υ	υ	υ	υ	
MDG				υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
CIV				υ					υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
KEN	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
LBR	υ	υ	υ	υ														
MWI	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
MLI	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
MRT				υ	υ	υ	υ	υ	υ	υ	υ	υ		υ	υ	υ	υ	
MUS	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
NER	υ	υ	υ	υ								υ					υ	
NGA				υ	υ	υ	υ											
RWA	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ		υ	υ	υ	υ	
SEN	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
SLE	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ		υ	υ	υ	υ		υ	υ
SOM				υ	υ	υ	υ											
SDN	υ	υ	υ	υ	υ	υ	υ											
SWZ	υ	υ	υ	υ	υ	υ	υ											
TGO	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ		υ	υ	υ	υ	
UGA	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
ZMB	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ
ZWE	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ	υ

refers to the number of countries in a column of a particular table. Countries are excluded from regression if data are not available for some of the variables included.