# Artificial States*

Alberto Alesina, William Easterly and Janina Matuszeski
Harvard University, New York University and Harvard University

February 2006
Revised: November 2008

**Abstract**

We define "artificial states" as those in which political borders do not coincide with a division of nationalities desired by the people on the ground. We propose and compute for most countries in the world two measures of the degree to which borders may be artificial. One measures how borders split ethnic groups into two separate adjacent countries. The other measures the straightness of land borders, under the assumption the straight land borders are more likely to be artificial. We then show that these two measures are correlated with several measures of political and economic success.

## 1 Introduction

Artificial states are those in which political borders do not coincide with a division of nationalities desired by the people on the ground. Former colonizers or post-war agreements among major powers regarding borders have often created monstrosities in which ethnic, religious or linguistic groups were thrown together or separated without any respect for those groups' aspirations. Eighty percent of African borders follow latitudinal and longitudinal lines, and many scholars believe that such artificial (unnatural) borders, which create ethnically fragmented countries or, conversely, separate the same people into bordering countries, are at the root of Africa's economic tragedy.[1] Not only in Africa, but

*We thank Jean Marie Baland, Alberto Bravo-Biosca, Ernesto dal Bo, Ashley Lester and participants at conferences at Brown, the NBER and seminars at Harvard for useful comments. The editor, Jordi Geli, and three anonymous referees provided very valuable feedback on the first version of this paper. For much needed help with maps we thank Patrick Florance, Kimberly Karish and Michael Oltmans. Alesina gratefully acknowledges financial support from the NSF with a grant through the NBER.

[1] See Easterly and Levine (1997) for early econometric work on this point. Herbs (2000) and especially Englebert, Tarango and Carter (2002) focus on the arbitrariness of African borders as an explanation of political and economic failures in this region. At the time of decolonization, new rulers in Africa made the decision to keep the borders drawn by former colonizers to avoid disruptive conflicts among themselves.

around the globe, including Iraq and the Middle East, failed states, conflict and economic misery are often very visible near borders left over by former colonizers, borders which bore little resemblance to the natural division of peoples.

There are four ways in which those who drew borders created problems. First, they gave territories to one group, ignoring the fact that another group had already claimed the same land. Second, they drew boundary lines that split ethnic (or religious or linguistic) groups into different countries, frustrating the national ambitions of various groups and creating unrest in the countries formed. Third, they combined into a single country groups that wanted independence. Fourth, even if there were no major ethnic divisions in the new states, they were still a random collection of families, clans, and villages that would not have a strong collective national identity.

The results have sometimes been disastrous. Artificial borders increase the motivation to safeguard or advance nationalist agendas at the expense of economic and political development. As George Bernard Shaw put it, "A healthy nation is as unconscious of its nationality as a healthy man is unconscious of his health. But if you break a nation's nationality it will think of nothing else but getting it set again." When states represent people put together by outsiders, these peoples may find it more difficult to reach consensus on public goods delivery and the creation of institutions that facilitate economic development, compared to states that emerged in a homegrown way.[2] Peoples may have weaker allegiance to various collective agendas in artificial states than in non-artificial ones. In every society there are social norms that sanctions those who do not contribute to the public groups, and these norms may be stronger in less fragmented societies where most people share common goals and culture. The crucial question is what is the relevant group? In a natural state, there is more likely to be a strong nationalist allegiance to that state, making the whole nation the group. In an artificial state, people are likely to have a smaller radius of group identification, such as the kin network or the local village or region.

While the nature of borders has been discussed in the political science (especially) and economic literature, we are not aware of systematic work relating the nature of country borders to the economic success of countries. Our goal is to provide measures that proxy for the degree to which borders are natural or artificial and relate these measures to economic and political development. Just to be clear, by "artificial," we mean a political border drawn by individuals not living in the areas divided by these borders, normally former colonizers. All other borders can be considered "natural," as they were drawn by people on the ground. Needless to say, often borders may start as artificial and then be modified by people on the ground. These adjustments on the ground may or may not reflect the desire of a majority of the people living there, especially if dictatorial regimes make the adjustments.

We provide two measures never before used in econometric analysis of comparative development. One is relatively simple and captures whether or not an

[2] See Alesina, Baqir and Easterly (1999) for a model and evidence on how ethnic fractionalization decreases the quality of public policies.

ethnic group is "cut" by a political border line. That is, we measure situations in which the same ethnic group is present in two bordering countries. The assumption is that a homogenous group would prefer to live in one country, rather than being split by a border. This measure accounts fairly precisely for one of the ways in which borders may be "wrong," that is, when they cut through groups and leave them in separate countries. But it does not capture other ways in which borders may be undesirable; for instance, situations in which two ethnic groups are forced into the same country or a state that is simply a random collection of regions that do not have a national identity.

We then provide a second measure, based upon the assumption that, if a land border is close to a straight line, it is more likely to be drawn artificially, for example by former colonizers. However, if it is relatively squiggly, it is more likely to represent geographic features (rivers, mountains, etc.) and/or divisions carved out in time to separate different people. This second measure probably comes closer to capturing instances in which lines drawn at former colonizers' tables have remained in place. Needless to say, the straight-border measure is not perfect, but much of our paper concerns precisely discussing this measure and its alternatives. In addition, even squiggly borders could be artificial, and the process that led to the formation of some straight borders may actually have been fairly "natural." We are fully aware of these possibilities, but our goal in this paper is twofold: first, to actually provide the measure (a non-trivial task), and second, to begin to investigate some basic correlations of squiggliness.

In summary, we define artificial states as those that have straight borders and/or a large fraction of their population belonging to a group (or groups) split with a neighboring country.[3] After constructing our measures, we explore how they are correlated with GDP per capita, a standard measure of economic development.

Because borders can be changed, as Alesina and Spolaore (1997) emphasized, citizens can rearrange the borders of artificial states. Indeed this happens, for example during the breakdown of the Soviet Union. In fact, it is quite possible that, as time goes by, many currently straight borders will become squiggly as they are rearranged. Relatively newly independent countries have had "less time" than countries that were never colonized in which to re-carve their borders based on an equilibrium reflecting how different peoples want to organize themselves. Englebert, Tarango and Carter (2002) document several instances of border instability in Africa due to the artificial original borders. Even among never-colonized countries, tensions due to "artificial" borders sometimes remain, for example, with the Basque independence movement in Spain.

We are not aware of other papers that have attempted to consider formally (as opposed to narratively) the relationship of the shape of countries to economic development. However, our paper is related to three strands of the literature. One strand is the recent work on the size of countries and its relationship to economic growth, as in Alesina and Spolaore (2003), Alesina,

[3] While a border's straightness is an indication of artificiality in the context of countries, it may be just the opposite in other contexts. For instance, in the context of electoral districts, a squiggly border may be the result of gerrymandering.

Spolaore and Wacziarg (2000), and Alcala and Ciccone (2004), among others. Second, our work builds on the literature concerning the relationship between ethno-linguistic fractionalization and economic growth, as in Easterly and Levine (1997), Alesina et al. (2003), and the survey in Alesina and La Ferrera (2005). Our paper discusses one historical phenomenon that may have led to excess ethnic fractionalization. Third, the role of former colonizers has also been widely studied (see La Porta et al. (1999), Acemoglu, Johnson and Robinson (2001), Glaeser et al. (2004)) but not specifically with regard to the importance of borders. Our paper specifies a new mechanism by which colonizers affected the subsequent development of the areas they colonized. In many ways, we bridge these three strands because we focus on how colonizers created fragmented societies by drawing artificial borders.

The paper is organized as follows: In Section 2, we provide historical examples of the artificial border-drawing. Section 3 describes our basic hypothesis, presents our measures of artificial borders, and discusses the properties of these measures. Section 4 investigates whether artificial states indeed perform less well than other states. The last section concludes the paper.

# 2 Examples of problematic borders

Examples of problematic borders abound. MacMillan (2002), in her analysis of the post-World War I meeting at Versailles, describes how the redrawing of borders around the world was decided based on compromises between the winning powers, often with little regard for preserving nationalities. American President Woodrow Wilson spoke often and eloquently in favor of a nationality principle, namely that political borders had to respect ethnic boundaries and nationality, but that principle was often ignored, even by Wilson himself. The book by MacMillan clearly documents, sometimes even in hilarious ways, how borders were drawn on maps with strikes of a pencil by the leaders of England, France and the US, ignoring the leg work of their experts and without even knowing the names of the nationalities and ethnicities involved. Historians agree that the Treaty of Versailles created many problematic borders that set the seeds for a very large number of future conflicts.

The past and current trouble in the Middle East at least in part originated from this kind of agreement between Western powers. Under the Sykes-Picot agreement between Britain and France during World War I, Northern Palestine would go to the French, Southern Palestine to the British, and Central Palestine, including Jerusalem, would be an allied Condominium shared by the two. After the war, the French agreed to give up any claims to Palestine in return for control over Syria. The British abandoned their protegee (Faisal) in Syria and offered him Iraq, cobbling together three different Ottoman provinces containing Kurds, Shiites and Sunnis. This set the stage for instability and the military coups that led to Saddam Hussein. In Lebanon, the French added Tripoli, Beirut and Sidon to the traditional Maronite area around Mount Lebanon, giving their Maronite Christian allies control of what were originally Muslim areas. A comparison country in the Middle East that is much less "artificial" is Tunisia, which does not have significant numbers of any partitioned group and has had happier development and political outcomes than the states partitioned under Sykes-Picot.

The partition of India and Pakistan is another famous example of artificial borders. The burning issue in the partition of 1947 was whether and how to award separate rights of national self-determination to Hindus and Muslims (the British ignored the national aspirations of smaller groups such as the Sikhs, which would bring its own bitter consequences). The Congress Party of Gandhi and Nehru campaigned for independence for one unitary Indian state, including Hindus, Muslims, and Sikhs from Peshawar to Dhaka. Mohammed Ali Jinnah founded the Muslim League, which called for a separate state for Muslims: Pakistan. But since Hindus and Muslims were mixed together all over the subcontinent, how could you come up with a plan to carve a Muslim nation out of India?

This intermixing was the result of a complex history that included the Muslim Mughal dynasty that the British Raj replaced. Until the last days of the Raj, there were Muslim princes ruling over majority Hindu princedoms and Hindu princes ruling over majority Muslim princedoms. The only areas with a Muslim majority were in the extreme northwest and the extreme northeast, separated by a thousand miles, and still containing large minority Sikh and Hindu communities.

In the Muslim Northwest Frontier Province (NWFP), ethnic Pathans were separated from their fellow Pathans in Afghanistan by the Durand Line, an arbitrary boundary between Afghanistan and British India laid down by a previous British bureaucrat. Peshawar, the capital of NWFP, was the traditional winter home of the Afghan kings. The Pathans preferred either an independent Pukhtoonwa uniting all Pathans or a Pathan-led Greater Afghanistan. At the time of partition, NWFP had a Congress-allied government led by a charismatic advocate of nonviolence, Khan Abdul Ghaffar Khan (the “Frontier Gandhi").

Back in British India, two other provinces of the future Pakistan were Sindh and Balochistan. Sindhi feudal landowners initially opposed the Pakistan idea and only later gave their grudging support under the naïve hope that Sindh would be largely autonomous. Balochi tribesmen (also divided from ethnic compatriots by a colonial boundary with Iran) preferred an independent Balochistan, which would lead to a secessionist attempt in the 1970’s, met with murderous repression by the Pakistani state. Punjab and Bengal were on the borders and had slight Muslim majorities. Congress leaders would not consent to hand them over to the Muslims. This meant that the British partitioned the mosaic of Hindus and Moslems in each state. Prior to partition, the Unionist government in Punjab wanted a united Punjab and backed neither the Muslim League nor Congress.

The unhappiest heir of the partition of 1947 is Pakistan. Jinnah complained that he got a “moth-eaten” Pakistan, with missing halves of Bengal and Punjab, little of Kashmir, some frontier territory, and two disjointed areas of West and East Pakistan. As late as 1981, only 7 percent of the Pakistani population were primary speakers of the supposed national language, Urdu. So to sum up, Pakistan wound up as a collection of Balochistan, NWFP, Sindh (all of whom entertained secession at various times), East Bengal (which successfully seceded in 1971 to become Bangladesh, although only after a genocidal repression by West Pakistani troops), Mohajir migrants from India (many of whom regretted the whole thing), and West Punjab (which had its own micro-secessionist movement by the Seraiki linguistic minority).[4] A much less artificial comparator state to Pakistan is Thailand (partly because it escaped colonial control), with Thailand having much better development outcomes.

Many other artificial borders were drawn during the colonial period, and few borders changed after de-colonization. Africa is the region most notorious for arbitrary borders. Historian Roel Van Der Veen (2004) points out that, prior to the era of de-colonization, states had to prove their control of a territory before being recognized by the international system. Virtually all new African states would have failed this test at the time of de-colonization. However, with de-colonization in Africa (and to some extent in other regions), the leading international powers changed this rule to recognize nations that existed principally on paper as the heir to a former colonial demarcation. As Van Der Veen put its, "letterbox sovereignty" was conferred upon whatever capital and whichever ruler the letters from the UN, the IMF, and the World Bank were addressed to. This left the new rulers more accountable to international organizations and leading industrial powers than to their purported citizens.[5] States consisted of little more than a few former independence agitators, the indigenous remnant of the colonial army, and a foreign aid budget. The new rulers of African states had no incentive to change a system of which they were the main beneficiaries, and hence the Organization of African Unity adopted a convention in the 1960's to treat colonial boundaries as sacrosanct (only rarely violated since). Englebert, Tarango and Carter (2002) have many more examples of problematic borders in Africa that led to disputes, political instability and economic failures.

One example of an African nation that is less "artificial" is Botswana, which consists mainly of the unpartitioned Tswana ethnic group and has relatively squiggly borders on three out of four sides. One that is more artificial than average is Equatorial Guinea, whose Fang ethnic group is partitioned with neighbors Cameroon and Gabon and whose mainland borders (not including the island part of the nation) are mainly straight lines. Although both nations have natural resources (diamonds and oil, respectively), Botswana has been a politically stable democracy with mostly good development outcomes, while Equatorial Guinea is an unstable dictatorship with a miserable ranking on most development indicators.

A case study of one African ethnic group shows some of the problems that can arise from divided loyalties. A border the British drew splits a group known as the Kakwa in two. The British drew the border originally (without consulting the Kakwa) as a quarantine line to halt the spread of sleeping sickness, but it

[4] These examples are from Easterly (2006).
[5] Van De Veen (2004), p29

remained as the border between independent Sudan and Uganda long after the original rationale was forgotten. The Kakwa were one of the many African groups in Southern Sudan that felt aggrieved by Northern dominance. The Kakwa in Sudan joined the civil war against the North. In the 1960's, the Kakwa in Uganda supported their fellow Kakwa in Sudan, providing cross-border refuges and arms. The Ugandan military actively supported this effort, because history's most famous Kakwa - Idi Amin - was chief of the Ugandan army. Amin arranged for weapons and food to go from Uganda to the Southern Sudanese, even including supplies purchased for Uganda's army. Amin made a number of trips to southern Sudan to meet with resistance leaders, and they met him in Uganda.

When Ugandan president Milton Obote and Idi Amin later got into a power struggle, the Sudanese Kakwa repaid the favor. They provided support to Idi Amin - at least 500 Southern Sudanese guerillas participated in his military coup against Obote. After his takeover, Amin reorganized the army, putting Kakwa, as well as some Southern Sudanese guerillas, in key positions. By late 1973, Kakwa or Southern Sudanese accounted for twenty-one of the top twenty-four positions in Amin's military. When the Tanzanian army overthrew Amin after his catastrophic rule, he escaped through his friends in Southern Sudan (Ade (1985)). Thus, the long-ago partition of one small ethnic group had terrible consequences for two separate "artificial" states - Uganda and Sudan.

Latin America is a lesser known (and much earlier) example of artificial borders drawn by a colonial power: in this case, Spain. The Spanish created administrative units (viceroyalties, captaincies, audiencias, etc.) in the Americas that had virtually nothing to do with indigenous groups on the ground. For example, the various Mayan groups in southern Mexico, Guatemala, and what became other Central American states were split between units. The province of Upper Peru, which later became Bolivia, split the Quechuas between Bolivia and Peru and combined the Quechuas with the Aymaras in Bolivia. When independence arrived in the early 19th century, the new states were controlled by the European elites, who formed states based on these colonial demarcations. In the words of one historian, "The new 'sovereign' states were often little more than a loose collection of courts, custom houses, and military units." (Winn 1992, p. 83). Although there were some wars that altered a few borders, today's Latin American states still correspond closely to Spanish colonial divisions. A state that suffered relatively little from colonial demarcations and partitioned ethnic groups is Chile (because of the lower pre-colonial population density and the natural border conveyed by the Andes). Chile is obviously much more of a development success story than Bolivia.

## 3 Artificial states: hypotheses and measures

Our main hypothesis is that artificial states perform less well than non-artificial ones. We are fully aware of a critical methodological trade-off here: often, objective measures clearly induce classification mistakes. But using an excessive amount of judgment calls leads to tautological results; the knowledge of *ex post* success and failure of states is very likely to influence those judgment calls. In the present paper, we choose to use objective measures almost exclusively, as we discuss in detail below.

We construct two measures. The first is completely new, and its construction is, we hope, a significant contribution in itself; this is the fractal measure.The second measures the degree to which ethnic groups were split by borders and is based upon a calculation for each pair of adjacent nations using detailed data of ethnic groups within nations assembled in Alesina et al. (2003). The first measure captures a more general sense of artificiality that we hypothesize has consequences even when there are no ethnic issues (see the discussion above on nationalism as a device to control free-riding on public goods).

## 3.1 The fractal measure

The basic idea is to compare the borders of a country to a geometric figure. If a country looks like a perfect square with borders drawn with straight lines, the chances are these borders were drawn artificially. On the contrary, borders that are squiggly lines (perhaps meant to capture geographic features and/or ethnicities) are less likely to be artificial. Squiggly geographic lines (such as mountains) are likely to separate ethnic groups, for reasons of patterns of communication and migration.

But how can we measure squiggliness? We first present the measure, and then we discuss its properties and alternatives.

Fractal dimension is analogous to the typical concept of the dimension of an object, although, unlike the simple definition of dimension, the fractal dimension can be a fractional number. A point has a fractal dimension of zero, a straight line a fractal dimension of one, and a plane a fractal dimension of two. However, unlike the traditional definition of dimension, as a line stops being perfectly straight and begins to meanders more and more, i.e. to become more and more squiggly, the fractal dimension increases. In the limit that a curve meanders so much that it essentially fills a whole page, then the fractal dimension becomes much closer to 2 than to 1. This is because the "line" is behaving more like a "plane."

Our measure is meant to capture how close a border is to a straight line that would have a fractal dimension of 1 versus a line so squiggly that it fills a plane and has a fractal dimension of 2. In practice, the fractal measure of actual borders is much closer to 1 than to 2, but there is variation. Figure 1 shows two countries, Sudan and France. Visually, they are quite different, as the borders of Sudan are very straight, and those of France are quite squiggly. It will turn out that the fractal dimension for France is 1.0429 and that of Sudan is 1.0245, reflecting the fact that Sudan's borders are much closer to being straight lines (dimension 1.0000) than France's borders.

The fractal dimension can be calculated in several ways. We use the box-count method, which is the most straightforward (Peitgen, Jurgens and Saupe (1992), p 218-219). For this method, a grid of a certain size/scale is projected onto the border, and the number of boxes that the border crosses is tallied. The scale of this grid is also recorded, as measured by the length of a side of a box in the grid. This gives a pair of numbers: box-count and box-size. The process is then repeated using grids with different box-sizes, each time recording both the box-size and the number of boxes that the border crosses. Given the pairs of data, box-size and box-count, the log-log plot of this data gives the fractal dimension as follows, where the negative of the slope ($b$) is the fractal dimension of the line:

$$In(boxcount) = a - b * In(boxsize)$$

Some intuition for this method can be gained by considering two extreme cases, a perfectly straight line and a line so wiggly that it covers a whole page (Figure 2a-2d). Figures2a and 2b show two different grids projected onto a perfectly straight line. The length of the side of a box (or the "box size") in Figure 2a is twice that of Figure 2b, and we can normalize the box sizes to 2 and 1, respectively. Counting the number of squares that the line crosses in each case, we get a box count of 24 for Figure 2a when the box size is 2, and a box count of 48 for Figure 2b when the box size is 1. Thus, for the straight line, the box count doubles (or increases by a factor of $2^1$) when the box size is halved (or "increases" by a factor of $2^{-1}$). Plotting $In(boxcount)$ versus $In(boxsize)$ yields a downward-sloping line with a slope of $-1$ (Figure 1g and Table 1). Thus, the fractal dimension for the straight line depicted in Figures 2a and 2b is determined to be 1. This makes sense because the normal notion of dimension for a perfectly straight line is 1.

Next consider Figures 2c and 2d, which show a line so squiggly that it covers the whole page. Here the box count is 176 when the box size is 2 (Figure 2c), and the box count is 704 when the box size is 1 (Figure 2d). Thus, the box count quadruples (increases by a factor of $2^2$) when the box size is halved ("increases" by a factor of $2^{-1}$). In this case, the plot of $In(boxcount)$ versus $In(boxsize)$ yields a downward-sloping line with a slope of $2/-1 = -2$ (Figure 2g and Table 1). Consequently, for this line, which is so squiggly that it fills the whole page, the fractal dimension is 2. This is also in agreement with the standard notion of dimension in which a plane or a page has two dimensions.

The borders of countries will be in between these two extremes of a perfectly straight line with fractal dimension 1 and a very squiggly line that fills a whole page and has a fractal dimension of 2. Consider the somewhat less squiggly line in Figures 2e and 2f. Here, when we calculate the fractal dimension using the box counting method, we find that the box count increases from 54 (Figure 2e) to 130 (Figure 2f) when the box size is reduced from 2 to 1, respectively. Thus, the box count is more than doubling when the box size is halved. Yet the box count is not quadrupling, as was the case with the very squiggly line (Figures 2c and 2d). We would thus expect that a plot of $In(boxcount)$ versus $In(boxsize)$

would have a slope that is steeper than $-1$ but not quite as steep as $-2$. In fact, when we do the calculation for this example, the slope is $-1.267$ (Figure 2g and Table 1). Based on this result, we would a sign a fractal number of 1.267 to this squiggly line. In practice, the fractal dimension of most country borders is between 1.000 and 1.100. Squiggly borders have fractal dimensions closer to 1.100, while straighter borders have fractal dimensions closer to 1.000.

These examples use only two data points to determine the fractal dimension of a line form. In practice, when calculating the fractal dimension of country borders, we use twelve different box sizes. The smallest box size is the smallest possible, given the digital nature of our data. This smallest box size corresponds to about 0.001 of a degree latitude or longitude. In addition to this box size, which we normalize to 1, we also use grids with box sizes of 2, 3, 4, 6, 8, 16, 31, 64, 128, 256, and 512. As in the examples above, for each box size, we project a grid with that box size onto our country border. We then count the number of boxes that the border crosses, resulting in a data point of box count and box size. Using all twelve box sizes gives us twelve data points with which to regress $In(boxcount)$ on $In(boxsize)$. Recall that the general formula for the fractal dimension is given by

$$In(boxcount) = (\text{constant intercept}) - (\text{fractal dimension}) * In(boxsize)$$

Thus, we take the negative of the slope of the regression of $In(boxcount)$ on $In(boxsize)$ as the fractal dimension for the country.

It is useful to present an example, using the case of Colombia. Figure 3 shows our method for determining the fractal dimension for Colombia. The graph plots $In(boxcount)$ versus $In(boxsize)$ and has twelve points, corresponding to the twelve different box sizes. For each box size, we have projected a grid of that size onto the border for Colombia and counted the number of boxes that the border crosses. Taking logs of this data, we arrive at our twelve data points, representing the pairs of data, $In(boxsize)$ and $In(boxcount)$. Regressing $In(boxcount)$ on $In(boxsize)$ using these twelve data points gives the straight line pictured on the graph. This line has a slope of $-1.0354$. Using the equation above, we take the negative of the slope of the regression line as the fractal dimension. Thus, the fractal dimension for Colombia is 1.0354. Finally, for the purposes of our analysis, we calculate a fractal index for each country, which is the log of the fractal dimension. Returning to our example, since the fractal dimension of Colombia is 1.0354.

For the purposes of our analysis, we calculate a fractal index for each country, which is the log of the fractal dimension. Returning to our example, since the fractal dimension of Colombia is 1.0354, the fractal index for Colombia is $In(1.0354) = 0.0348$.

Finally, because coastlines along oceans are extremely squiggly compared to non-coastline borders, we choose to consider the fractal dimension of only the non-coastline portion of a country's border. Since islands have no non-coastline political boundaries, they cannot have a fractal variable according to our definition. Thus, we calculate the fractal variable for 144 non-island countries.

### 3.2 Properties

A measure of the straightness or squiggliness of country borders ideally exhibits several properties. One desirable property is scale-invariance, meaning the ideal measure should not differ systematically for large or small countries. Scale-invariance also means we should be able to apply our measure to a particular country and get consistent results, regardless of the scale of the analysis for that country. Our measure is indeed scale-invariant.[6] [7]

A second desirable property of a "squiggliness" measure is the degree to which it measures larger-scale irregularities, as opposed to smaller-scale ones. Small-scale deviations from a smooth curve or line may well be the result of ethnic considerations or other local politics determining whether a particular parcel of land should be on one side of a border or another. Since we are interested in comparing borders where local and ethnic considerations were taken into account, with more "artificial" borders, we prefer our measure to focus on these small-scale irregularities, rather than measuring the overall shape of a country. Unlike measures such as this circumscribed/inscribed circle ratio, the fractal measure emphasizes the small-scale variation that we are interested in

[6] To be precise, our measure is not 100 percent scale-invariant, but it is close to scale invariance. Analyzing a country when at differing degrees of being "zoomed in" or "zoomed out" may yield slightly different values for the fractal dimension. However, these numbers do not vary greatly for each country, and the relative rankings of countries are maintained. More importantly, our measure allows us to consistently compare large and small countries. By using the same set of 12 box-sizes (as measured in degrees latitude and longitude) for each country, our analysis for each country is on the same "human" scale as for the other countries. By contrast, other measures of compactness, such as the ratio of the area of a circumscribed and an inscribed circle for the country border, may differ systematically for large and small countries.

[7] We chose to use unprojected map data. Projecting country borders by continent creates distortions among countries within the same continent, while projecting each country individually causes imbalances between large and small countries in the degree of distortion. When using unprojected map data, at very high latitudes the boxes for the box count are rectangular and smaller in area than at lower latitudes. The rectangular shape is not a problem for straight borders and for most squiggly borders, although it could be an issue for a border running generally north-south, with a lot of smaller-scale squiggliness in the east-west direction. The smaller average area means that the set of box sizes used for the high latitude countries is proportionally smaller. Since we use box sizes that increase geometrically, we are only concerned if this area is smaller by a factor of 2 which occurs for latitudes greater than 60 degrees. Only two sets of borders fall into this category. The Alaska border between the U.S. and Canada is only a fraction of each country's border. Also, it is mostly straight and so unlikely to be affected by either rectangular shape or smaller area boxes. The remaining borders are all in Scandinavia, among Norway, Sweden, Finland and Russia (only a small fraction of its border). A test using the 11 smallest versus the 11 largest box sizes revealed that the distortions from using unprojected data for these countries leads to lower fractal dimensions (more artificial) for these borders. This bias runs counter to our findings since these countries are highly successful.

measuring.

Finally, and most importantly, we would like a measure that allows us to consider only part of the border at a time. In particular, we disregard coastlines in this analysis, since they are determined by nature and not by politics and are very squiggly. The fractal measure can be applied to selected portions of the border, such as just the political boundaries. Most other measures of compactness must use the entire boundary, including coastlines. For instance, other common compactness measures include: the ratio of the longest axis to the maximum perpendicular length; the ratio of the minimum shape diameter to the maximum diameter; various ratios among the area of the shape, the area of an inscribing circle and the area of a circumscribing circle; the moment of inertia of the shape; and the ratio of the area of the shape to the area of a circle with the same perimeter.[8] All of these measures require a closed shape in order to be calculated.

## 3.3 The partitioned measures

Our second new measure focuses on the specific case of borders cutting across an ethnic group and dividing it into two adjacent countries. The variable is defined as the percent of a country's population that belongs to a partitioned group. The latter is a group that is present in two bordering countries.

The intuition for this variable is that, when the same ethnic group is split between two adjacent nations, the border is very likely to be "artificial," i.e., not reflective of the desires of people on the ground. The presumption is that, in most cases, ethnic groups would like to be in the same country. Obviously, if an ethnic group is distributed among countries very far apart, it is out of the question that such a group could form a country including all its members (except by migrating, of course). But this does not seem to be the case when the same group is split between two adjacent countries. When this situation occurs, it may signal an artificial border.

Our point is that this kind of artificiality also hinders good development outcomes. An important mechanism for inducing sacrifice for public goods is group loyalty. Both ethnic groups and nations command group loyalty from their members. If ethnic group = nation, the group loyalty on the basis of ethnicity and nation is the same, and we think it is easier to achieve cooperation and consensus for public goods, as well as for overall pro-development policies. In cases of split ethnic groups, this leads to divided loyalties, as individuals of the partitioned group are conflicted whether to ally themselves with their co-ethnics in another nation. If they care more about their co-ethnics than they do about their co-nationals, then they are less willing to invest in public goods in their home nations. Note that if groups wanted to be in the same country they could migrate. Not only it is costly but countries may prevent migrations (both in and out) and in same cases they may invest resources in repressions of migratory movements.

[8] For more on this, see Niemi, Grofman, Carlucci, and Hofeller (1990) and Flaherty and Crumplin (1992).

Of course, the devil is in the details. How do you decide what constitutes a "partitioned" group? It is important that the coding be as objective and mechanical as possible, because if we did subjective coding, we might be influenced by a knowledge of the outcomes in each country. A purely mechanical procedure may lead to some obvious mistakes, but it is arguably better to have some mistakes than to open the door to subjective coding. Only where the mechanical procedure leads to a systematic egregious error should a correction be made, and the correction should follow some clearly defined principles.

Our mechanical procedure was to define partitioned ethnic groups as identifiable groups listed in the ethnic fractionalization exercise of Alesina et al. (2003) that appear in two or more adjacent countries. That paper, which produced a widely used data set on ethnicity, defined an ethnic group as any group defined on the basis of physical characteristics (e.g., skin color) and/or language that was politically salient enough in each country to be listed separately in the "ethnic breakdown" given in descriptions of the nation in general sources such as encyclopedias or world yearbooks.[9] The majority of entries for ethnic groups in Alesina et al. (2003) came from the *Encyclopedia Brittanica*, and a large plurality of the rest came from the CIA *World Factbook*.

However, there is one problem related to the different purposes of the "fractionalization" exercise and our goal here of constructing a "partitioned" variable. What is politically salient within each nation and what is relevant for international "partition" are different in one obvious respect. There are some groups that are meaningful minorities within a nation but are too broad and general (and not identified with any nation in particular) at the international level to be considered partitioned. The litmus test is whether a group could have conceivably made up a nation if it had not been "partitioned." So, for example, "blancos" were listed as a separate ethnic group in Bolivia, where the ethnic differences between whites and indigenous people are politically salient, but obviously it would not be meaningful to talk about "partition of whites" across nations. The same would go for groups such as "Spanish speakers," "English speakers," "mestizos" or "blacks." Hence, we omitted all such general groups in our measurement of "partitioned." This was not a subjective decision that would contaminate our results, because the judgment was made only on the basis of what makes sense as "international partition" and was not influenced by development outcomes in any given nation. Our coding was done before running the first regression.[10] Details of the constitution of the model are available from the authors.

A related problem is deciding when a group in two neighboring nations is, in reality, the same ethnic group split by a border. Again, we went for mechanical criteria so our results would not be contaminated by subjective judgments influenced by outcomes. Any two groups from Alesina et al. (2003)

[9] This index is then different (i.e., more refined) than one based only on language. Here is an example: a purely language-based index would group together African Americans and whites in the US. The ethnic criterion would not.

[10] After we distributed this paper as a working paper, we rechecked the coding for mistakes, and there were a handful, but the corrections made no difference to the results.

that have the same group name and are in adjacent nations are classified as partitioned (with the exception of excessively general groups, as defined in the preceding paragraph). This excludes some groups that speak the same (or nearly the same) language. So, for example, the Flemish in Belgium are not taken to be Belgian "Dutch," because Flemish and Dutch are different names in the dataset. However, French, German, and Italian speakers in Switzerland are considered to be partitioned from their respective neighbors, because the same names is used for these groups as in the home nation. The names come from the original sources (such as *Encyclopedia Brittanica* and CIA *World Factbook*) for Alesina et al. (2003), which we interpret as passing judgment on whether these sources see a politically salient group in one nation as being the same as in a neighboring nation. This naming convention very likely misclassifies some cases, but our review of the data finds that most cases are uncontroversial. Most importantly, once again we prefer orthogonal measurement errors to an error term that will be correlated with the error term in the outcome variable if subjective judgments are made based on outcomes.

There is one other potentially serious problem with our "partitioned" measure - the discontinuity at 100 percent ethnic share in one nation and zero in another nation, implying zero partitioning, versus 99.999 percent partitioning in the first nation and .001 percent in another nation, implying 99.999 percent partitioning in the first nation (of course, the shares will not necessarily sum to one because the total populations of the two nations will be different, but we assume the same population for purposes of illustration). However, this knife-edge property is not a serious problem in our data in practice. It is ruled out by our stipulation (originally determined in Alesina et al. (2003)) that the ethnic groups be "politically salient." An ethnic group that was either such an overwhelming majority of one nation or an extremely tiny minority of another nation is unlikely to be politically salient in either case. To check on this, we looked for examples where our exercise yielded an ethnic group that was (1) greater than 70 percent in one nation and (2) less than 30 percent in other nations (a generous test of this lopsidedness problem, since 70 percent is a lot less compelling a problem than 99.999 percent). There was only one ethnic group that met these two conditions: the Shona people are 76 percent of Zimbabwe, 11 percent of Mozambique, and 12 percent of Botswana. Although Zimbabwe comes the closest to the knife-edge problem, it is still a long way from the knife edge. Zimbabwe is actually a good example of the importance of partition, since sanctuaries for Shona rebels in Mozambique were important in the war that changed Rhodesia into Zimbabwe, and Zimbabwean refugees have fled to their cross-border co-ethnics during the current political crisis. Admittedly, the knife-edge feature is an unappealing theoretical property and would limit the application of this measure to all possible cases. However, we have seen that our emphasis on political salience of ethnic groups (following Alesina et al. 2003) rules out cases anywhere near the knife edge. Having only one case that comes even remotely close to the knife-edge problem (and having even that case be an arguably strong exhibit of the effects of partition) makes us think that this potentially serious theoretical problem with our measure simply does not occur in our data.

One possible objection to this variable is mobility of people. If members of the same ethnic groups wanted to be together, they could move into the same country. However, mobility of people is often not free, and many countries may prevent entry (or, in some cases, exit). The partitioned variable is calculated for 131 countries, and the indices $FRACTAL$ and $PARTITIONED$ are both available for 117 countries.

### 3.4 Other pre-existing measures

The literature of ethnolinguistic fractionalization has normally focused on one index of fractionalization, the Herfindhal index, which captures the probability that two randomly drawn individuals from the population of the country belong to different groups.[11] The original index was based on a linguistic classification of groups from a Soviet source (the *Atlas Narodov Mira*). It was originally used in the economic development literature by Mauro (1995) and Easterly and Levine (1997), and it is often referred to as Elf (Ethnolinguistic fractionalization) index. We label it $Elf1$ in the present paper. Alesina and al. (2003) proposed another index that, in addition to linguistic differences, includes differences based on other characteristics such as skin color, as we described above. (See Alesina and al. (2003) for more discussion about the construction of this variable.) They label it Fract, but to avoid confusion with our fractal variable, we label it $Elf2$ in the present paper.

### 3.5 Data and sources

Data for determining the fractal dimension for each country's political boundary comes from the GIS (Geographic Information Systems) format data set World Vector Shoreline. This data set is the largest-scale digital data set of political boundaries available today. The data is based on work done by the U.S. military in the early 1990's. The non-coastline borders for each country are isolated using ArcGIS software[12]. This data is then changed to a raster (digitized) format and then to a ".tif" format. With a few minor modifications, the software program ImageJ[13] calculates the box-count/box-size data for twelve different box-sizes; the smallest box-size corresponds to the smallest scale of the raster data exported from GIS (approximately 0.001 degrees latitude or longitude). A fractal dimension is calculated for each country using this data, ranging from 1.000 to 1.100. Finally, we take logs of the fractal dimension to achieve a fractal index, which ranges from 0 to 0.10. Table 2 displays all the sources for our variables. Table A1 in the appendix lists the value of $PARTITIONED$ and $FRACTAL$ for every country in our sample.

[11] Another index frequently used is a polarization index suggested by Montalvo and Raynal Querol (2995). We do not find any difference whne we use polarization insetad of fractionalization.

[12] ArcGIS 9.0 Desktop software from ESRI; www.esri.com.

[13] Available online at http://rsb.info.nih.gov/ij/download.html and at http://rsb.info.nih.gov/ij/developer/index.html.

### 3.6 Basic Correlations

Table 4 shows the correlation of our measures with other country characteristics related to statehood. The two measures of ethnic heterogeneity ($ELF1$ and $ELF2$) are highly correlated with each other and show a similar correlation pattern with other variables. We decide to concentrate hereafter on $ELF1$ because it has been more widely used in the previous literature.

Pre-existing ethnic heterogeneity would make state formation more difficult. Never having been colonized increases the likelihood of natural state formation, since colonizers usually determine what constitutes a nation. A long history of precolonial statehood, as summarized in the work of Bockstette, Chanda, and Putterman (2002), would also make natural state formation more likely. It is reassuring that there are high correlations between these measures and our new measures of artificial states.

$ELF1$ is strongly correlated with $PARTITIONED$, as expected. At the same time, the correlation falls enough below unity to make these potentially distinct measures capturing different phenomena (high diversity over a given unit of territory regardless of the borders vs. having ethnic groups split by borders). A long history of pre-colonial statehood is correlated with $FRACTAL$ but not with $PARTITIONED$, which provides further evidence that $FRACTAL$ is a general measure of natural statehood that captures other dimensions besides ethnic issues. $NON-COLONIAL$ has strong predictive power for both $PARTITIONED$ and $FRACTAL$, suggesting there was something in the very nature of colonization that increased the likelihood of former colonies being artificial states.

This pattern suggests to us that the two measures are capturing different dimensions of artificiality, with $PARTITIONED$ reflecting the ethnic dynamics of artificial statehood, while $FRACTAL$ captures non-ethnic features of artificiality. Both could be important in the artificial states phenomenon. For example, we hypothesize that $PARTITIONED$ captures the damaging effect of conflicting group loyalties between ethnic groups and nations. $FRACTAL$ may be capturing a dimension of artificiality that would exist even if the population in the nation were ethnically homogeneous and not split between neighboring countries. For example, as discussed above, the population in an ethnically homogeneous $FRACTAL$ artificial state may emphasize a narrow radius of group interest reaching only to the extended family or village, with little sense of common national purpose to finance national public goods. We will explore these two different dimensions of artificial statehood further using principle components analysis below.

## 4 Empirical results

### 4.1 Which states are "artificial?"

In order to illustrate which states are most artificial according to both measures, we took countries that were in the top third of $PARTITIONED$ and in the bottom third of $FRACTAL$ (the straightest borders). Given the weak correlation between the two measures, there were not that many countries in both – 13 to be exact. These "most artificial" states are Chad, Ecuador, Equatorial Guinea, Eritrea, Guatemala, Jordan, Mali, Morocco, Namibia, Niger, Pakistan, Sudan, and Zimbabwe. Many observers would not be suprised by this list , as our previous historical discussion illustrates.

What about the US and Canada? Their border is a straight line for most of its length; are they artificial states? According to our measures, yes; they do score relatively in terms of how artificial they are, which is certainly not consistent with a view of artificial as failed states. One may note that this a case in which borders were drawn before many people actually moved in. In many ways, the same applies to US states: in the West, borders that were drawn when the population density was still extremely low are often straight lines. On the contrary, borders of East coast states, drawn earlier amidst higher population density, are less straight. The US and Canada are included in our regressions below. This is a good example of the use of judgment calls that we did *not* make. One may argue that the knowledge of the history of how the US/Canada border was drawn should lead us to exclude it from our classification of artificiality. We have chosen not to use judgment calls of this type (which makes our empirical results less strong), because they would compromise the objectivity of our measures. An interesting line of future research would be to investigate this observation about when borders were drawn more systematically because the case of Canada and the US may not be unique.

## 4.2 Economic Success

We now turn to verifying whether these new measures of artificial states are correlated with economic success. There are a variety of outcomes that artificiality might affect. Unfortunately, these outcomes are all heavily correlated with each other and so don't convey much independent information. We consider the most obvious variable that measures economic success: per capita income (specifically, its log value in 2002). Table 5A reports our first set of regressions. Standard errors are in parentheses. Columns (1) and (2) include only $PARTITIONED$ and $FRACTAL$ as regressors. Their coefficients have the expected sign and are statistically significant at the 1% and 5% levels, respectively. Remember that we always expect the opposite sign of the coefficient on the two variables of interest, since higher values of $PARTITIONED$ and lower values of $FRACTAL$ are associated with more artificial states. Column (3) includes these two variables together; they both remain significant, but $FRACTAL$ is slightly below the 5% level.

We next start a long series of horse races between alternative hypotheses explaining GDP per capita and our artificial states measure, as well as other obvious robustness checks. Column (4) adds the measure of fractionalization $Elf1$ and a widely used variable for tropical climate zone. These are tests of whether our variables are proxying for other previous hypotheses, such as the ethnic fractionalization story of Easterly and Levine (1997) and Alesina et al. (2002). The tropical climate variable stands in for two additional alternative hypotheses: the direct effect of climate hypothesis of Gallup and Sachs (1999) and the indirect effect of climate through institutions (Acemoglu, Johnson and Robinson (2001), Easterly and Levine (2003), Rodrik, Subramanian and Trebbi (2004). Note that neither ethnic fractionalization nor tropical climate has a coefficient significant at the 5% level, but both our artificial state variables continue to have statistically significant coefficients of 5% or more.

Table 5b presents more sensitivity analyses to other geographic variables. In general, the variable $PARTITIONED$ is very robust, while $FRACTAL$ is less so. The coefficient on $FRACTAL$ always has the expected sign, but it is significant in only four of six regressions. The variables to which $FRACTAL$ is sensitive are either desert or "hot and dry climate," which are obviously closely related concepts. The desert variable is not itself significant. What seems to be going on is that desert countries are also more likely to have straight-line borders, so by dummying out these countries, we reduce the amount of variation in $FRACTAL$, and it becomes more difficult to detect its effect. For completeness, we also consider two other features of countries, land area and population density (continuing to include a geographic variable - the tropical climate zone measure). Borders might not matter as much in very large countries or in lightly settled countries. However, these controls are insignificant, while both $PARTITIONED$ and $FRACTAL$ remain significant.

Table 5c shows sensitivity to including continent effects. This is a stringent robustness check, since it removes all of the cross-continent variation in income (59 percent of the total variation in per capita income). The first column shows that the variable $FRACTAL$ does not survive the inclusion of a full set of regional dummies, but the coefficient on the variable $PARTITIONED$ remains significant at the 5% level. When we consider the continent dummies one at a time, both Africa and Europe reduce the significance of $FRACTAL$ to the 10% level. The other continent dummies leave $FRACTAL$ significant at 5%.

Table 5d returns to horse races with other hypotheses for long-run development. It obviously would be problematic to introduce endogenous variables to represent the other hypotheses, so we opt for a reduced form of income regressed on the instruments used for endogenous variables in other hypothesized determinants of development. Frankel and Romer (1999) use the geographic determinants of a country's trade share to instrument for trade opennness in a level regression for per capita income. We enter their geographic trade propensity variable and find that it leaves both $FRACTAL$ and $PARTITIONED$ significant at 5%. Hall and Jones (1999) define a variable called "social infrastructure" as a weighted average of openness and institutional quality and instrument for it with share of the population speaking English or share of the population speaking any European language. The English-speaking variable is insignificant, and our artificial state measures retain significance. The European language variable is significant and reduces the significance of $FRACTAL$ to 10%, while leaving $PARTITIONED$'s significance unchanged.

Lastly, we introduce the measure of years of experience with statehood shown to be important by Bockstette, Chanda, and Putterman (2002). This variable is significant and eliminates the significance of $FRACTAL$. $PARTITIONED$ is robust. In the next section, we consider the idea that history of statehood is one of several measures that jointly captures state artificiality and attempt to identify the principal components of these different measures.

In summary, the $PARTITIONED$ measure of artificial statehood is robust to all the tests we run here of alternative hypotheses and other plausible controls. Our $FRACTAL$ measure seems to have some information content, as it is significant in most of our regressions. However, its significant is fragile - it doesn't survive continent effects, some geographic variables, and a measure of years of experience with pre-colonial states. Since the $PARTITIONED$ and $FRACTAL$ measures are uncorrelated with each other and are both plausible measures of artificial statehood, we want to better understand what dimension of statehood is captured by each variable. We explore this further in the next section using principal components.

## 4.3 Principal Components of Artificial Statehood

We use principal components to try to get further insight into the different dimensions of artificial states that might affect development. We have our two measures of artificiality, $PARTITIONED$ and $FRACTAL$, and three other variables that we discussed for Table 4 as related to artificial state outcomes, $ELF1$, $NON-COLONIAL$, and $STATEHISTORY$. We run principal components on these five variables and examine the first two components, which account for most of the variance. The principal components method is usually employed as a variable reduction technique when there are multiple measures of the same concept or concepts. We will employ it like this is by using the first two principal components in place of the five underlying variables in our regressions. However, of more interest to us is what the exercise can tell us about the dimensions that these five variables are capturing. Table 6a shows that the first principal component is simply a weighted average of all five measures in the direction of "natural states," with $PARTITIONED$ and $ELF1$ entering negatively and $FRACTAL$, $NON-COLONIAL$, and $STATEHISTORY$ entering positively, as expected. This composite measure gives a decent weight to $PARTITIONED$ but not much to $FRACTAL$. The second principal component, by contrast, weights $FRACTAL$ more heavily than any other variable, and $PARTITIONED$ and $ETHNIC$ change signs. To get some insight into this second component, we can rewrite it as: $.46*(FRACTAL-(-PARTITIONED))+.16*(FRACTAL+NON-COLONIAL+STATEHISTORY)-(-PARTITIONED-ELF1))+smallresidualterms.$ Taking the negative of $PARTITIONED$ and $ELF1$ transforms them into measures where an increase means an increase in the direction of "natural statehood." Hence, this measure is capturing the difference between the effects of the non-ethnic measures of "naturalness" ($FRACTAL$, $NON-COLONIAL$, and $STATEHISTORY$) and the ethnic measures $-PARTITIONED$ and $-ELF$, with an extra weight on the difference between the two new measures introduced in this paper, $FRACTAL$ and $PARTITIONED$.

We enter the two principal components into all of the same regressions done in previous tables, except that we omit the five variables underlying the principal components from all of the regressions. The first component is largely robust, only failing to be significant in one regression with latitude (in which the second principal component is highly significant). The second component is more fragile, as $FRACTAL$ was. It is not significant in the regression with all the continent dummies, although it is not driven out by any one continent dummy when they are entered one at a time. The other occasions in which it fails are when either of two closely related variables are include: (1) hot, dry climate or (2) $DESERT$. These latter variables are not themselves significant, so it seems like the regression is unable to distinguish between the effects of the second component and desert climate.

The principal components exercise is not necessarily to be preferred over the conventional regressions above. We conclude from the exercise that the $FRACTAL$ variable appears to capture a second dimension of artificial statehood (correlated with colonial and state history) that has an additional effect on development outcomes, even though the effect is somewhat fragile.

# 5 Conclusions

The borders of many countries have been the result of processes that have little to do with the desire of people to be together or not. In some cases, groups who wanted to be separate have been thrown into the same political unit; others have been divided by artificial borders. Former colonizers have been mainly responsible for such mistakes, but the botched agreements after the two major wars of the last century have also played a role.

The main contribution of this paper is to provide two new measures meant to capture how "artificial" political borders are. One measure considers how straight land borders are, under the assumption that straight borders are more likely to be artificial and less likely to follow geographic features or the evolution of hundreds of years of border design. The second measure focuses on ethnic or linguistic groups separated by borders. We have then investigated whether these variables are correlated with the political and economic success of various countries. Even after controlling for a host of geographical variables and colonial history, the variable that captures the partition of groups is a significant determinant of GDP per capita. The variable capturing the straightness of borders is less robust, but it is also correlated to GDP per capita in most specifications.

Probably the single most important issue that we have not addressed is that of migrations. One consequences of artificial borders is that people may want to move, if they can. Often movement of peoples is not permitted by various governments, but migration certainly occurs. In some cases, migrations that respond to artificial borders may be partly responsible for economic costs, wars, dislocation of people, refugee crises and a hots of undesirable circumstances. Thus, the need to migrate created by the wrong borders may be one reason why artificial borders are inefficient. But sometimes the movement of people may correct for the artificial nature of borders. This dynamic aspect of movement of people and migrations (and, for that matter, changes of borders) is not considered in this paper, in which we view a static picture of the world.

Finally, note that the methodology used to construct our $FRACTAL$ variable can be usefully applied to other types of boundaries, such as sub-national administrative units or electoral districts. In the latter case, a squiggly border may be an indication of active gerrymandering.

# References


[1] Acemoglu, Daron, Simon Johnson and James Robinson (2001), "The Colonial Origin of Comparative Development: An Empirical Investigation." *American Economic Review.* v91:5, p1369-401.

[2] Ade, Adefuye (1985). "The Kakwa of Uganda and the Sudan: The Ethnic Factor in National and International Politics," in A.I. Asiwaju (ed.), *Partitioned Africans: Ethnic Relations across Africa's International Boundaries, 1884-1984*, St. Martin's Press: New York, 1985.

[3] Alcala, Francisco and Antonio Ciccone (2004), "Trade and Productivity." *Quarterly Journal of Economics.* v119:2, p613-46.

[4] Alesina, Alberto, Reza Baqir and William Easterly (1999), "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics*, v114:4, p1243-84.

[5] Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg (2003), "Fractionalization." *Journal of Economic Growth.* v8:2, p155-94.

[6] Alesina, Alberto and Enrico Spolaore (1997), "On the Number and Size of Nations." *Quarterly Journal of Economics.* v112:4, p1027-56.

[7] Alesina, Alberto and Enrico Spolaore (2003), *The Size of Nations.* Cambridge MA: MIT Press.

[8] Alesina, Alberto, Enrico Spolaore and Romain Wacziarg (2000), "Economic Integration and Political Disintegration." *American Economic Review.* v90:5, p1276-96.

[9] Bockstette, Valerie, Areendam Chanda and Louis Putterman (2002), "States and Markets: The Advantage of an Early Start." *Journal of Economic Growth.* v7:4, p347-69.

[10] Bourke, P. (1993), *Fractal Dimension Calculator User Manual, Online.* Available: http://astronomy.swin.edu.au/~pbourke/fractals/fracdim/fdc_orig/

[11] Bruk, S. I., and V.S. Apenchenko (eds.) (1964), *Atlas Narodov Mira.* Moscow: Glavnoe Upravlenie Geodezii i Kartografii.

[12] ATLAS NARODOV MIRA [ATLAS OF THE PEOPLES OF THE WORLD]. Moscow: Glavnoe Upravlenie Geodezii i Kartografii, 1964.

[13] Easterly, William (2006), *The White Man's Burden: Why the West's Efforts to Aid the Rest Did So Much Harm and So Little Good.* New York: The Penguin Press.

[14] Easterly, William and Ross Levine (1997), "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics.* v112:4, p1203-50.

[15] Englebert, Peter, Stacy Tarango and Matther Carter (2002), "Dismemberment and Suffocation: A Contribution to the Debate on African Boundaries." *Comparative Political Studies.* v35:10, p1093-1118.

[16] Flaherty, Mark S. and William W. Crumplin (1992), "Compactness and Electoral Boundary Adjustment: An Assessment of Alternative Measures." *The Canadian Geographer.* v36:2, p159-171.

[17] Frankel, Jeffrey A. and David Romer (1999), "Does Trade Cause Growth?" *American Economic Review*, v8:3, p379-99.

[18] Gallup, John Luke and Jeffrey D. Sachs with Andrew D. Mellinger (1999), "Geography and Economic Development," in Pleskovic, Boris and Joseph E. Stiglitz (eds.), *World Bank Annual Conference on Development Economics 1998.* The World Bank: Washington, DC, p127-8.

[19] Glaeser, Edward, Rafael La Porta, Florencio Lopez de Silanes and Andrei Shleifer (2004), "Do Institutions Cause Growth?" *Journal of Economic Growth.* v9:3, p271-303.

[20] Hall, R. and C. Jones (1999). "Why Do Some Countries Produce So Much More Output per Worker Than Others?" *Quarterly Journal of Economics*, v114:1, p83-116. Available: http://elsa.berkeley.edu/~chad/papers.html.

[21] Herbst, Jeffrey (2002), *State and Power in Africa.* Princeton, NJ: Princeton University Press.

[22] La Porta, Rafael, Florencio Lopez de Silanes, Andrei Shleifer and Robert Vishny (1999), "The Quality of Government." *Journal of Law, Economics and Organization.* v15:1, p315-388.

[23] Mandelbrot, Beniot (1967), "How Long is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension." *Science.* New Series v156:3763, p636-638.

[24] MacMillan, Margaret (2002), *Paris 1919: Six Months That Changed the World.* New York: Random House.

[25] Mauro, Paolo (1995), "Corruption and Growth." *Quarterly Journal of Economics.* v110:3, p681-712.

[26] Niemi, Richard G., Bernard Grofman, Carl Carlucci, and Thomas Hofeller (1990), "Measuring Compactness and the Role of a Compactness Standard in a Test for Partisan and Racial Gerrymandering." *Journal of Politics.* v52:4, p1155-1181

[27] Peitgen, Heinz-Otto, Hartmut Jurgens and Dietmar Saupe (1992), *Chaos and Fractals; New Frontiers of Science.* New York: Springer-Verlag.

[28] Putterman, Louis (2007), State Antiquity Index Dataset, Version 3, Brown University.

[29] Van Der Veen, Roel (2004), *What Went Wrong With Africa: A Contemporary History.* Amsterdam: KIT Publishers.

[30] Winn, Peter (1992), *Americas: The Changing Face of Latin America and the Caribbean*. New York: Pantheon Books.

**Figure 1 – Artificial versus Organic boundaries – Sudan and France**

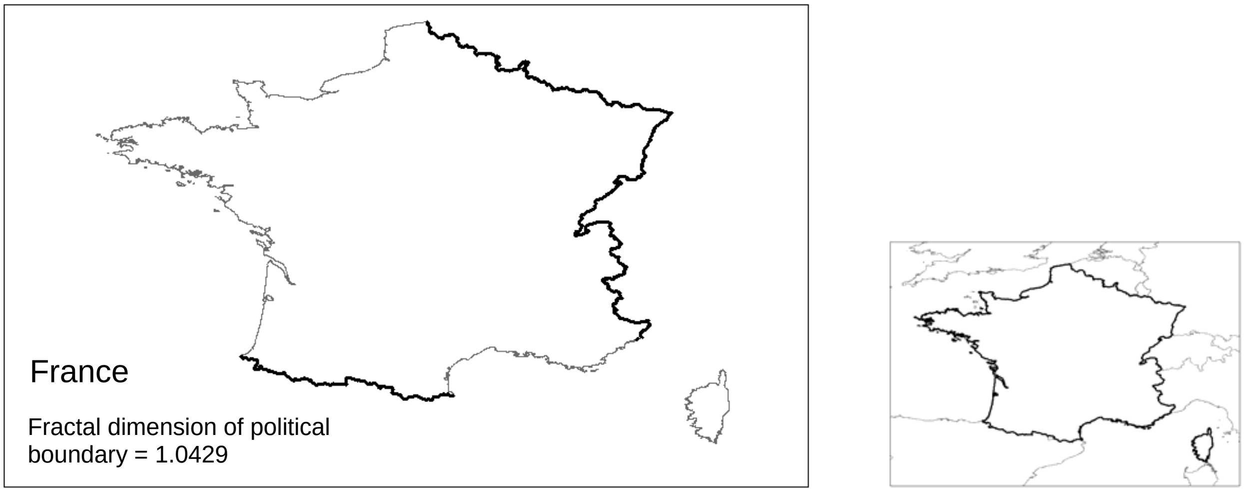



Figure 1a – France, with political boundaries highlighted at left

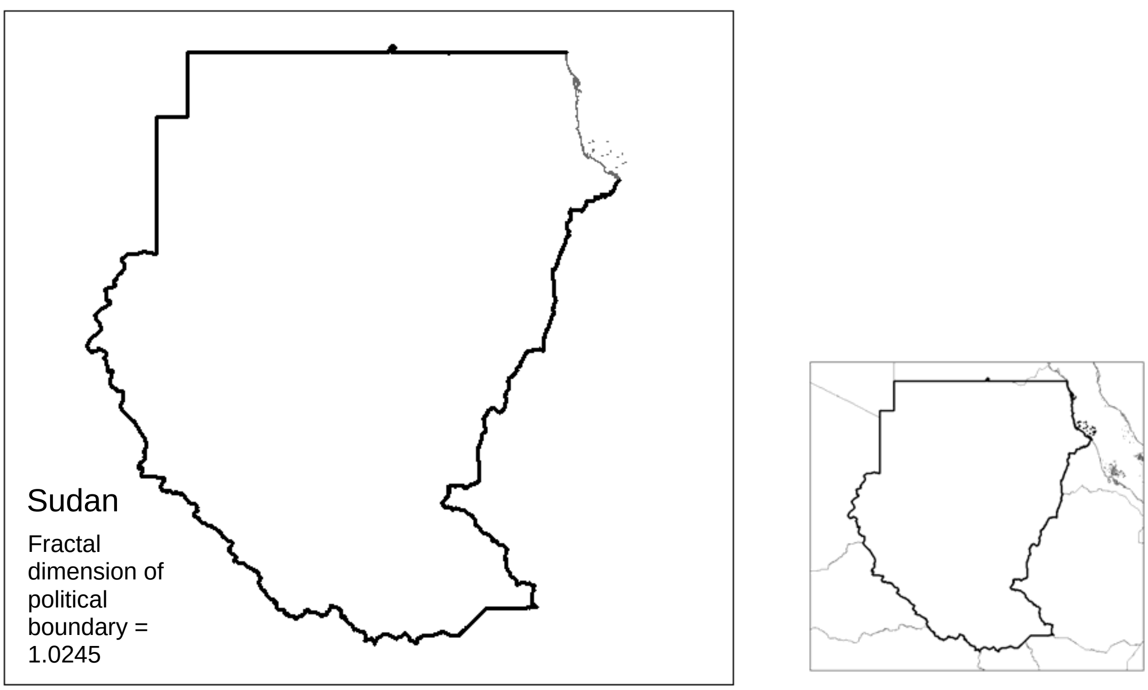



Figure 1b – Sudan, with political boundaries highlighted at left

## Figure 2a – 2d – Projections of two grids of different sizes onto Straight and Very Squiggly lines

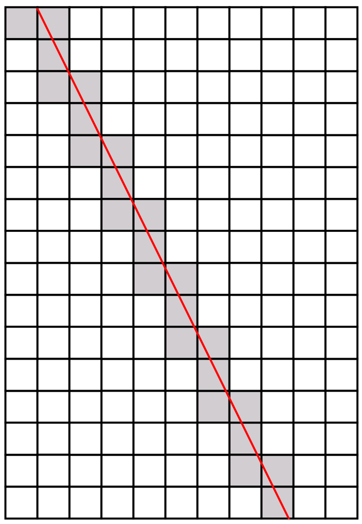


Figure 2a:
Box size = 2;  Box count = 24

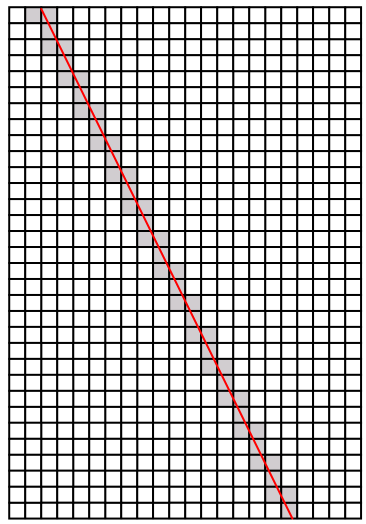


Figure 2b:
Box size = 1;  Box count = 48

**Straight line**

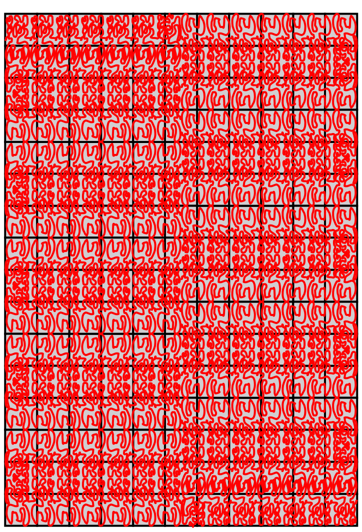


Figure 2c:
Box size = 2;  Box count = 176

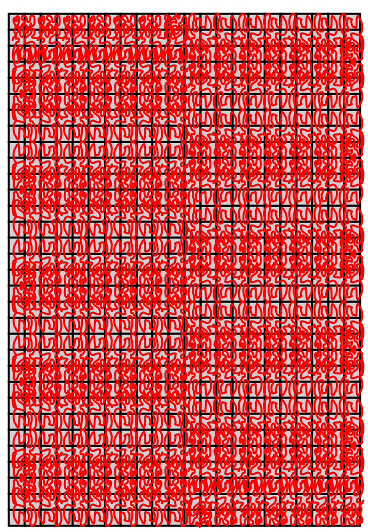


Figure 2d:
Box size = 1;  Box count = 704

**Very squiggly line**
(fills the whole page)

**Figure 2e – 2g – Projections of two grids of different sizes onto a Somewhat Squiggly line; Calculation of Fractal Dimension**

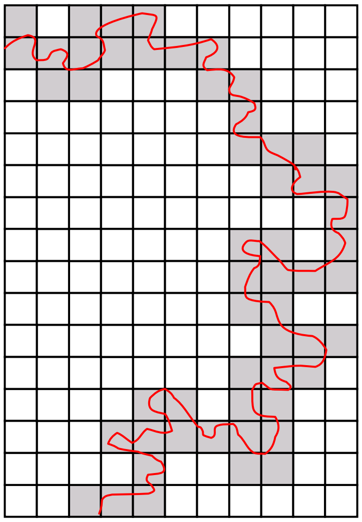


Figure 2e:
Box size = 2; Box count = 54

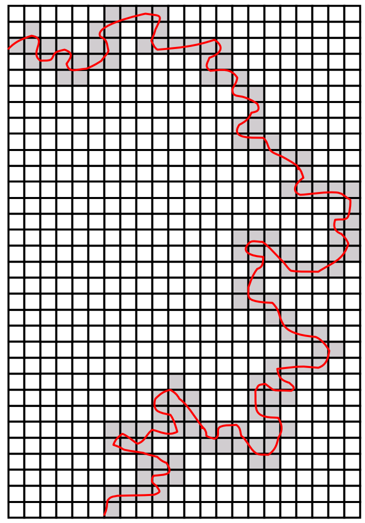


Figure 2f:
Box size = 1; Box count = 130

**Squiggly line**

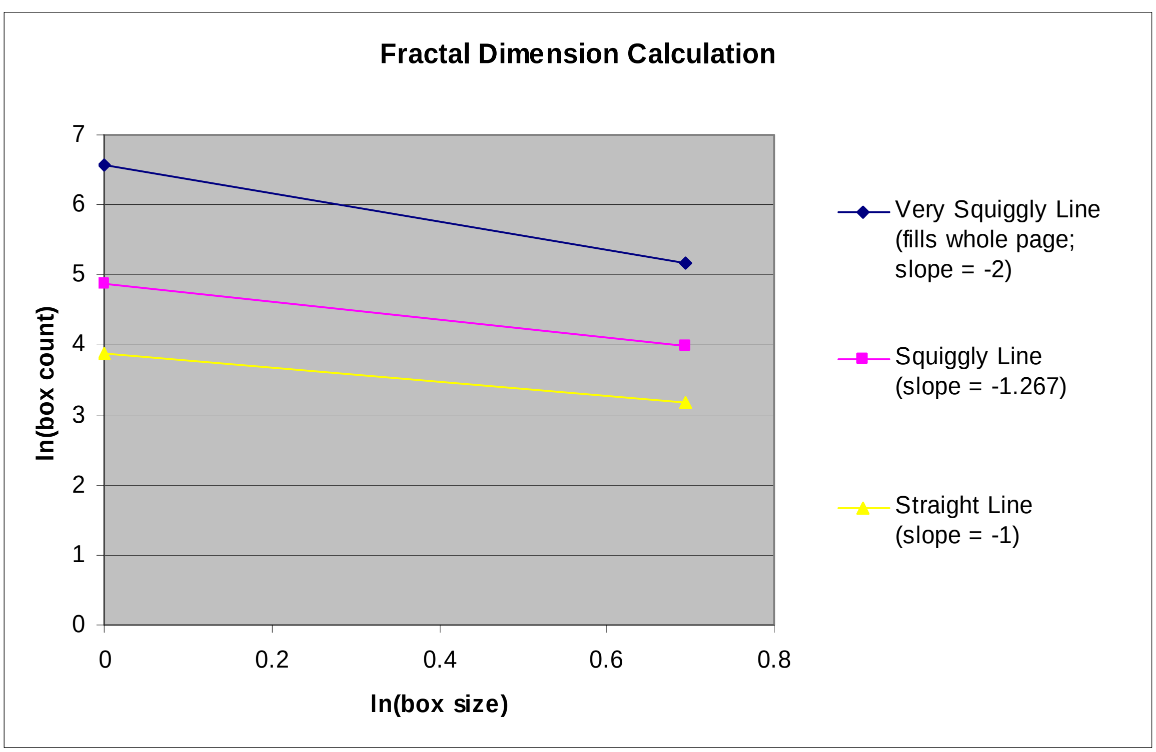
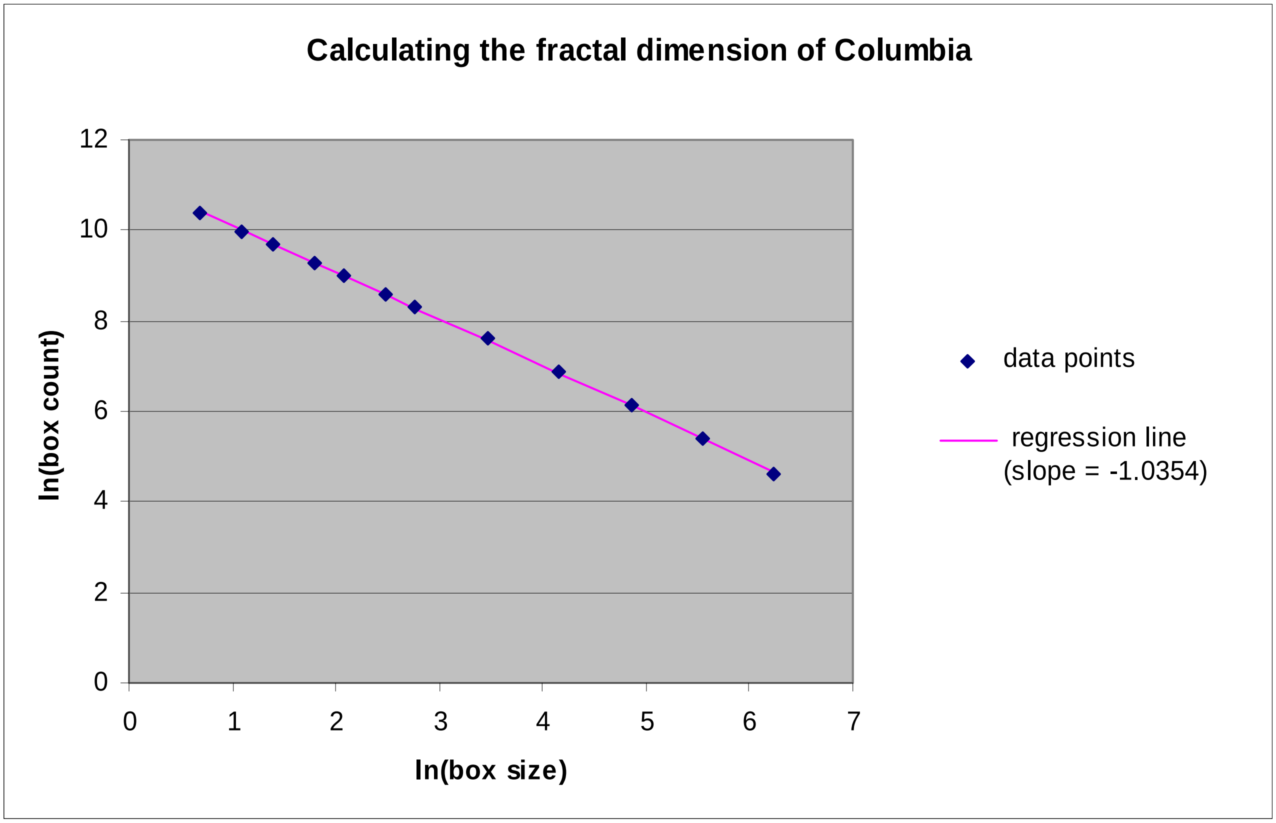



Figure 2g: Fractal Dimension Calculation

**Figure 3 – Calculation of the Fractal Dimension of Columbia's Border**



# Table 1 – Fractal Dimension Calculation

Straight Line (Figures 1a and 1b)

| box size | box count | ln (box size) | ln (box count) |
|---|---|---|---|
| 1 | 48 | 0.000 | 3.871 |
| 2 | 24 | 0.693 | 3.178 |

Regression coeff: -1.000

Fractal Number: 1.000

Very Squiggly Line (Figures 1c and 1d)

| box size | box count | ln (box size) | ln (box count) |
|---|---|---|---|
| 1 | 704 | 0.000 | 6.557 |
| 2 | 176 | 0.693 | 5.170 |

Regression coeff: -2.000

Fractal Number: 2.000

Squiggly Line (Figures 1e and 1f)

| box size | box count | ln (box size) | ln (box count) |
|---|---|---|---|
| 1 | 130 | 0.000 | 4.868 |
| 2 | 54 | 0.693 | 3.989 |

Regression coeff: -1.267

Fractal Number: 1.267

## Table 2 – Variable Definitions and Data Sources

| Variable Name | Description | Source |
|---|---|---|
| Partitioned | Fraction of country's population belonging to groups partitioned by a boarder. | Authors; based on ethnic data from Alesina et al 2003 |
| Fractal | Log of fractal dimension of country's non-coastline boarder (12 box sizes used). | Authors; based on World Vector Shoreline Dataset (GIS format) |
| ELF1 | Ethno-linguistic fractionalization index | 1960 data from the Atlas Narodov Mira |
| ELF2 | Ethno-linguistic fractionalization index | Alesina et al 2003 |
| Non-colonial | Dummy variable is one if the country never was a colony | Authors |
| Years of Ancient State History | Log of Discounted Sum of Years of Ancient Statehood, 1 AD to 1950 AD | Putterman (2007) |
| Log GDP per capita | Log of per capital GDP in 2002 | Summers-Heston, updated with World Bank per capita growth rates |
| Climate Zone A | Percent of cultivated land in Koppen-Geiger climate zone A (humid climate with no winter) | Gallup and Sachs (1999) |
| Climate Zone B | Percent of cultivated land in Koppen-Geiger climate zone B (dry climate with no winter) | Gallup and Sachs (1999) |
| Tropics % | Percentage of total land area in the tropics (between the Tropics of Capricorn and Cancer) | Gallup and Sachs (1999) |
| Distance to Equator | Distance from the middle of the country to the equator. | Gallup and Sachs (1999) |
| Desert % | Percent of total land area in Koppen-Geiger climate zone BW (desert) | Gallup and Sachs (1999) |
| Land Area | Log of total land area in kilometers squared | Gallup and Sachs (1999) |
| Population Density | Log of population density as experienced by the typical citizen (sum of pop density of many small regions, weighted by the pop of each region) | Gallup and Sachs (1999) |
| Frankel-Romer trade propensity | Log of predicted trade share based on gravity equation featuring only | Frankel and Romer 1999 |

| | | |
|---|---|---|
| | geographic variables (distance between bilateral trade partners and size) | |
| Share of population speaking English | Fraction of population that speaks English | Hall and Jones (1999) |
| Share of population speaking European language | Fraction of the population that speaks any European language | Hall and Jones (1999) |

## Table 3 – Summary Statistics

| Variable | Number of Obs | Mean | Std Dev | Min | Max | Number of Zeros |
|---|---|---|---|---|---|---|
| Partitioned | 132 | 27.247 | 28.617 | 0 | 100 | 20 |
| Fractal | 144 | 0.037 | 0.020 | -0.006 | 0.104 | 0 |
| ELF1 | 113 | 41.469 | 29.787 | 0 | 93 | 1 |
| ELF2 | 187 | 0.439 | 0.258 | 0 | 0.930 | 1 |
| Non-colonial | 209 | 0.301 | 0.460 | 0 | 1 | 146 |
| Log GDP per capita | 164 | 8.503 | 1.141 | 5.423 | 10.710 | N/A |
| Rule of Law | 208 | 0 | 1 | -2.31 | 2.01 | 0 |
| Government Effectiveness | 209 | 0 | 1 | -2.32 | 2.25 | 0 |
| Climate Zone A | 156 | 0.324 | 0.414 | 0 | 1 | 80 |
| Climate Zone B | 156 | 0.181 | 0.298 | 0 | 1 | 88 |
| Tropics % | 208 | 0.548 | 0.499 | 0 | 1 | 94 |
| Distance to Equator | 191 | 0.265 | 0.179 | 0.003 | 0.710 | N/A |
| Desert % | 160 | 0.156 | 0.305 | 0 | 1 | 109 |
| Log Land Area | 164 | 12.211 | 1.702 | 7.900 | 16.625 | N/A |
| Log Frankel-Romer trade propensity | 145 | 2.994 | 0.798 | 0.833 | 5.639 | #N/A |
| Share of population speaking English | 146 | 0.098 | 0.271 | 0.000 | 1.000 | 111 |
| Share of population speaking European language | 146 | 0.271 | 0.405 | 0.000 | 1.064 | 78 |
| Log Years of Ancient State History | 148 | 5.772 | 0.774 | 2.931 | 6.758 | #N/A |
| Log Population Density | 164 | 5.172 | 1.489 | 0.455 | 9.045 | N/A |

## Table 4 – Correlations of different ethnic and artifical state measures

| | PARTITIONED | FRACTAL | ELF1 | ELF2 | Non-colonial |
|---|---|---|---|---|---|
| **FRACTAL** | 0.071 | | | | |
| **ELF1** | 0.492** | 0.100 | | | |
| **ELF2** | 0.482** | -0.217** | 0.766** | | |
| **Non-colonial** | -0.228** | 0.249** | -0.404** | -.253** | |
| **Years of Ancient State History** | -0.095 | 0.205* | -0.270** | -0.262** | 0.401** |

+,* and ** refer to 10%, 5%, and 1% significance levels, respectively.

## Table 5a – Impact on Log GDP per Capita[1]

| Coefficient on: | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| PARTITIONED | **-0.020**** | | **-0.018**** | **-0.014**** |
| | **(0.003)** | | **(0.003)** | **(0.004)** |
| FRACTAL | | **11.49*** | $10.49^{+}$ | **9.535*** |
| | | **(5.563)** | (5.859) | **(4.316)** |
| ELF1 | | | | -0.0028 |
| | | | | (0.0039) |
| NON-COLONIAL | | | | **1.258**** |
| | | | | **(0.233)** |
| CLIMATE - ZONE A | | | | $-0.530^{+}$ |
| (Hot, rainy) | | | | (0.288) |
| CONSTANT | 9.005** | 7.952** | 8.541** | 8.507** |
| | (0.121) | (0.237) | (0.269) | (0.211) |
| OBSERVATIONS | 116 | 129 | 106 | 73 |

[1] Standard errors in parenthesis. $^{+}$,* and ** refer to 10%, 5%, and 1% significance levels, respectively.

## Table 5b – Impact on Log GDP per Capita[1]

| Coefficient on: | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| PARTITIONED | **-0.013**** | **-0.012**** | **-0.012**** | **-0.014**** | **-0.013**** | **-0.013**** |
| | **(0.004)** | **(0.003)** | **(0.003)** | **(0.004)** | **(0.004)** | **(0.004)** |
| FRACTAL | 5.709 | 6.795[+] | **8.184*** | 6.900 | **8.927*** | **13.56**** |
| | (4.948) | (3.871) | **(3.860)** | (5.399) | **(4.338)** | **(4.801)** |
| ELF1 | -0.0039 | -0.0014 | -0.00075 | -0.0037 | -0.0022 | -0.0050 |
| | (0.0044) | (0.0037) | (0.00316) | (0.0043) | (0.0040) | (0.0038) |
| NON-COLONIAL | **1.502**** | **1.098**** | **0.560*** | **1.531**** | **1.238**** | **1.294**** |
| | **(0.211)** | **(0.202)** | **(0.247)** | **(0.217)** | **(0.234)** | **(0.250)** |
| CLIMATE - ZONE A (Hot, rainy) | | | | | -0.508[+] | -0.475 |
| | | | | | (0.287) | (0.286) |
| CLIMATE - ZONE B (Hot, dry) | -0.079 | | | | | |
| | (0.307) | | | | | |
| TROPICS % | | **-0.777**** | | | | |
| | | **(0.203)** | | | | |
| DISTANCE TO EQUATOR | | | **3.139**** | | | |
| | | | **(0.658)** | | | |
| DESERT % | | | | 0.149 | | |
| | | | | (0.310) | | |
| LAND AREA | | | | | 0.071 | |
| | | | | | (0.079) | |
| POPULATION DENSITY | | | | | | 0.118 |
| | | | | | | (0.075) |
| CONSTANT | 8.442** | 8.775** | 7.566** | 8.369** | 8.090** | 6.908** |
| | (0.236) | (0.215) | (0.253) | (0.272) | (0.502) | (1.070) |
| OBSERVATIONS | 73 | 73 | 73 | 72 | 73 | 73 |

[1] Standard errors in parenthesis. [+],* and ** refer to 10%, 5%, and 1% significance levels, respectively.

# Table 5c – Impact on Log GDP per Capita with continent dummies[1]

| Coefficient on: | 1[2] | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| PARTITIONED | **-0.00845*** | **-0.00902**** | **-0.0125**** | **-0.0135**** | **-0.0138**** | **-0.0143**** | **-0.0124**** |
| | **(0.00339)** | **(0.00311)** | **(0.00386)** | **(0.00375)** | **(0.00376)** | **(0.00371)** | **(0.00373)** |
| FRACTAL | 5.311 | 6.245+ | **8.900*** | **9.419*** | 8.171+ | **10.27*** | **9.927*** |
| | (3.957) | (3.601) | **(4.289)** | **(4.31)** | (4.685) | **(4.48)** | **(4.308)** |
| ELF1 | -0.00148 | 0.00142 | -0.00017 | -0.00347 | -0.0031 | -0.00217 | -0.00465 |
| | (0.0032) | (0.003) | (0.00356) | (0.00395) | (0.0039) | (0.00383) | (0.00389) |
| NON-COLONIAL | 0.514 | **0.999**** | **1.479**** | **1.232**** | **0.934**** | **1.288**** | **1.301**** |
| | (0.323) | **(0.215)** | **(0.285)** | **(0.24)** | **(0.213)** | **(0.24)** | **(0.235)** |
| CLIMATE - ZONE A | -0.431 | **-0.583*** | **-0.611*** | -0.575+ | -0.506+ | -0.473 | -0.424 |
| | (0.286) | **(0.263)** | **(0.295)** | (0.29) | (0.291) | (0.294) | (0.284) |
| AFRICA | **-2.327**** | **-0.937**** | | | | | |
| | **(0.265)** | **(0.214)** | | | | | |
| LATIN AMERICA | **-1.600**** | | 0.419+ | | | | |
| | **(0.245)** | | (0.221) | | | | |
| ASIA & OCEANIA | **-1.315**** | | | 0.467 | | | |
| | **(0.385)** | | | (0.321) | | | |
| EUROPE | **-0.899*** | | | | 0.390+ | | |
| | **(0.35)** | | | | (0.209) | | |
| MIDDLE EAST | **-1.253**** | | | | | 0.47 | |
| | **(0.272)** | | | | | (0.301) | |
| NORTH AMERICA | | | | | | | **2.008**** |
| | | | | | | | **(0.263)** |
| CONSTANT | 10.28** | 8.769** | 8.249** | 8.514** | 8.551** | 8.411** | 8.453** |
| | (0.259) | (0.198) | (0.253) | (0.211) | (0.215) | (0.225) | (0.206) |
| OBSERVATIONS | 73 | 73 | 73 | 73 | 73 | 73 | 73 |
| R-SQUARED | 0.767 | 0.742 | 0.668 | 0.664 | 0.657 | 0.662 | 0.69 |

[1] Standard errors in parenthesis. +,* and ** refer to 10%, 5%, and 1% significance levels, respectively.
[2] Omitted region is North America.

## Table 5d: Impact on Log GDP per capita controlling for variables used in other levels regressions[1]

| Coefficient on: | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| PARTITIONED | **-0.0140**** | **-0.0136**** | **-0.0128**** | **-0.0119**** | **-0.0146**** |
| | **(0.0037)** | **(0.0038)** | **(0.00358)** | **(0.00338)** | **(0.0037)** |
| FRACTAL | **9.535*** | **10.89*** | **10.26*** | 8.303$^{+}$ | 7.337 |
| | **(4.316)** | **(4.58)** | **(4.271)** | (4.421) | (5.317) |
| ELF1 | -0.0028 | -0.0032 | -0.00397 | -0.000494 | -0.00187 |
| | (0.0039) | (0.0039) | (0.00369) | (0.00359) | (0.00369) |
| NON-COLONIAL | **1.258**** | **1.290**** | **1.261**** | **1.343**** | **1.127**** |
| | **(0.233)** | **(0.253)** | **(0.232)** | **(0.239)** | **(0.251)** |
| CLIMATE - ZONE A (HOT, RAINY) | -0.530$^{+}$ | -0.525$^{+}$ | -0.558$^{+}$ | **-0.599*** | -0.384 |
| | (0.288) | (0.292) | (0.29) | **(0.286)** | (0.296) |
| TRADE PROPENSITY | | -0.106 | | | |
| | | (0.148) | | | |
| ENGLISH SPEAKING SHARE | | | 1.457 | | |
| | | | (0.984) | | |
| EUROPEAN LANGUAGE SPEAKING SHARE | | | | **0.597**** | |
| | | | | **(0.218)** | |
| YEARS OF ANCIENT STATE HISTORY | | | | | **0.277*** |
| | | | | | **(0.134)** |
| CONSTANT | 8.507** | 8.753** | 8.472** | 8.235** | 6.952** |
| | (0.211) | (0.399) | (0.203) | (0.21) | (0.838) |
| OBSERVATIONS | 73 | 73 | 73 | 73 | 71 |

[1] Standard errors in parenthesis. $^{+}$,* and ** refer to 10%, 5%, and 1% significance levels, respectively.

**Table 6a: First Two Principal Components of Five Ethnic and Artificial State Measures**

| Coefficient on: | First Principal Component | Second Principal Component |
|---|---|---|
| Eigenvalue | 2.035 | 1.273 |
| Difference | 0.763 | 0.456 |
| Proportion | 0.407 | 0.255 |
| Cumulative | 0.407 | 0.662 |
| | | |
| Coefficients on: | | |
| PARTITIONED | -0.382 | 0.589 |
| FRACTAL | 0.126 | 0.728 |
| ELF1 | -0.569 | 0.160 |
| NON-COLONIAL | 0.557 | 0.255 |
| STATE HISTORY | 0.452 | 0.182 |

# Table 6b – Impact on Log GDP per Capita of First Two Principal Components Controlling for Continent Dummies [1]

| Coefficient on: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| FIRST PRINCIPAL COMPONENT | **0.525**** | **0.417**** | **0.227*** | **0.231*** | **0.401**** | **0.415**** | **0.372**** | **0.416**** | **0.463**** |
| | **(0.0844)** | **(0.107)** | **(0.103)** | **(0.106)** | **(0.108)** | **(0.108)** | **(0.105)** | **(0.108)** | **(0.101)** |
| SECOND PRINCIPAL COMPONENT | 0.236$^{+}$ | **0.330*** | 0.158 | **0.266*** | **0.343*** | **0.322*** | 0.277$^{+}$ | **0.338*** | **0.297*** |
| | (0.135) | **(0.144)** | (0.135) | **(0.114)** | **(0.144)** | **(0.142)** | (0.158) | **(0.146)** | **(0.145)** |
| CLIMATE - ZONE A (hot, rainy) | | -0.538$^{+}$ | -0.378 | **-0.663*** | -0.594$^{+}$ | -0.566$^{+}$ | -0.458 | -0.5 | -0.402 |
| | | (0.296) | (0.311) | **(.270)** | (0.296) | (0.311) | (0.27) | (0.301) | (0.319) |
| AFRICA | | | **-2.546**** | **-0.839**** | | | | | |
| | | | **(0.193)** | **(0.219)** | | | | | |
| LATIN AMERICA | | | **-1.872**** | | 0.135 | | | | |
| | | | **(0.233)** | | (0.17) | | | | |
| ASIA & OCEANIA | | | **-1.672**** | | | 0.249 | | | |
| | | | **(0.391)** | | | (0.279) | | | |
| EUROPE | | | **-1.214**** | | | | 0.377 | | |
| | | | **(0.311)** | | | | (0.288) | | |
| MIDDLE EAST | | | **-1.579**** | | | | | **0.385*** | |
| | | | **(0.268)** | | | | | **(0.191)** | |
| NORTH AMERICA | | | | | | | | | **2.270**** |
| | | | | | | | | | **(0.171)** |
| CONSTANT | 8.548** | 8.726** | 10.58** | 9.079** | 8.711** | 8.716** | 8.603** | 8.697** | 8.651** |
| | (0.0851) | (0.137) | (0.105) | (0.165) | (0.138) | (0.139) | (0.182) | (0.145) | (0.124) |
| OBSERVATIONS | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 |
| R-SQUARED | 0.612 | 0.636 | 0.75 | 0.703 | 0.638 | 0.639 | 0.643 | 0.64 | 0.688 |

[1] Standard errors in parenthesis. $^{+}$,***** and ** refer to 10%, 5%, and 1% significance levels, respectively.

## Table 6c – Impact on Log GDP per Capita of First Two Principal Components Controlling for Other Development Determinants [1]

| Coefficient on: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| FIRST PRINCIPAL COMPONENT | **0.525**** | **0.312**** | 0.19 | **0.541**** | **0.445**** | **0.406**** | **0.421**** | **0.439**** | **0.366**** |
| | **(0.084)** | **(0.111)** | (0.114) | **(0.096)** | **(0.099)** | **(0.123)** | **(0.105)** | **(0.097)** | **(0.112)** |
| SECOND PRINCIPAL COMPONENT | 0.212 | **0.287*** | **0.299*** | 0.204 | **0.341*** | **0.329*** | **0.336*** | **0.316*** | **0.323*** |
| | (0.155) | **(0.131)** | **(0.124)** | (0.17) | **(0.138)** | **(0.144)** | **(0.143)** | **(0.139)** | **(0.131)** |
| CLIMATE - ZONE A (Hot, rainy) | | | | | -0.457+ | -0.538+ | -0.530+ | -0.554+ | **-0.640*** |
| | | | | | (0.273) | (0.296) | (0.289) | (0.273) | **(0.296)** |
| CLIMATE - ZONE B (Hot, dry) | -0.158 | | | | | | | | |
| | (0.324) | | | | | | | | |
| TROPICS % | | **-0.823**** | | | | | | | |
| | | **(0.252)** | | | | | | | |
| DISTANCE TO EQUATOR | | | **3.035**** | | | | | | |
| | | | **(0.701)** | | | | | | |
| DESERT % | | | | -0.15 | | | | | |
| | | | | (0.325) | | | | | |
| LAND AREA | | | | | 0.116 | | | | |
| | | | | | (0.072) | | | | |
| POPULATION DENSITY | | | | | | 0.022 | | | |
| | | | | | | (0.092) | | | |
| TRADE PROPENSITY | | | | | | | -0.106 | | |
| | | | | | | | (0.153) | | |
| ENGLISH SPEAKING SHARE | | | | | | | | 1.833+ | |
| | | | | | | | | (1.035) | |
| EUROPEAN LANGUAGE SPEAKING SHARE | | | | | | | | | **0.490*** |
| | | | | | | | | | **(0.222)** |
| CONSTANT | 8.574** | 8.982** | 7.708** | 8.567** | 7.237** | 8.601** | 9.017** | 8.692** | 8.617** |
| | (0.105) | (0.162) | (0.212) | (0.1) | (0.924) | (0.497) | (0.468) | (0.119) | (0.122) |
| OBSERVATIONS | 71 | 71 | 71 | 70 | 71 | 71 | 71 | 71 | 71 |
| R-SQUARED | 0.613 | 0.681 | 0.717 | 0.605 | 0.654 | 0.636 | 0.639 | 0.672 | 0.66 |

[1] Robust standard errors in parenthesis. +,* and ** refer to 10%, 5%, and 1% significance levels, respectively.

## Table 7 – Partitioned and Fractal Variables

| Country | Partitioned | Fractal | Country | Partitioned | Fractal |
|---|---|---|---|---|---|
| Afghanistan | 74.7 | 0.0387784 | Laos | – | 0.0443039 |
| Albania | 10.7 | 0.0549994 | Latvia | 42.4 | 0.0323801 |
| Algeria | 20 | 0.013163 | Lebanon | 18 | 0.0682286 |
| Andorra | 55.1 | – | Lesotho | 15 | 0.0400953 |
| Angola | 20.77 | 0.0165523 | Liberia | 28.9 | 0.0459384 |
| Argentina | 0.15 | 0.0344787 | Libya | 0 | 0.0029258 |
| Armenia | 0 | 0.0627588 | Liechtenstein | 6.2 | 0.0495613 |
| Australia | 0 | – | Lithuania | 17.5 | 0.0483234 |
| Austria | – | 0.0555577 | Luxembourg | 22.6 | 0.10436 |
| Azerbaijan | 7.8 | 0.0576272 | Macedonia | 30.9 | – |
| Bahrain | 0 | – | Malawi | 66 | 0.0412763 |
| Bangladesh | – | 0.0803165 | Malaysia | 28.1 | 0.0532849 |
| Belarus | 18.4 | 0.0502272 | Mali | 89 | 0.025853 |
| Belgium | 63.9 | 0.0708405 | Malta | 0 | – |
| Belize | 6.6 | 0.0205279 | Mauritania | – | 0.015962 |
| Benin | 37.4 | 0.0365344 | Mexico | 0 | 0.0243315 |
| Bhutan | – | 0.02744 | Moldova | 28.9 | 0.0808332 |
| Bolivia | 60.76 | 0.028772 | Monaco | 63 | – |
| Bosnia-Herzegovina | 54.1 | 0.0694794 | Mongolia | 5.9 | 0.0194791 |
| Botswana | 17.1 | 0.0294423 | Morocco | 40 | 0.0056441 |
| Brazil | 0.1 | 0.0415929 | Mozambique | 75 | – |
| Brunei | 100 | – | Myanmar | 1.4 | 0.0480375 |
| Bulgaria | 20.7 | 0.0319928 | Namibia | 64.6 | 0.0237068 |
| Burkina Faso | 73.1 | 0.036303 | Nepal | – | 0.0417655 |
| Burundi | 97.4 | 0.0467118 | Netherlands | 6.1 | 0.0929256 |
| Cambodia | 10.9 | 0.0473892 | New Zealand | 0 | – |
| Cameroon | 37.3 | 0.0351837 | Nicaragua | 0 | 0.0379603 |
| Canada | 0 | 0.0120471 | Niger | 73.4 | 0.0176729 |
| Central African Republic | 43.8 | 0.0418999 | Nigeria | 58 | 0.0316343 |
| Chad | 51.4 | 0.0175157 | Norway | 1 | 0.016572 |
| Chile | 0.5 | 0.0309559 | Oman | 0 | 0.0151743 |
| China | – | 0.036679 | Pakistan | 57 | 0.0266516 |
| Colombia | 1 | 0.0348071 | Panama | 1.4 | – |
| Congo | – | 0.0261842 | Papua New Guinea | – | -0.0056359 |
| Congo, Dem Rep (Zaire) | – | 0.0241071 | Paraguay | 5.2 | 0.0331543 |
| Costa Rica | 0 | 0.0461199 | Peru | 52.5 | 0.0476466 |
| Croatia | 15.4 | 0.0618193 | Poland | 0.1 | 0.0341985 |
| Cyprus | – | 0.0435001 | Portugal | 0.95 | 0.0351065 |
| Czech Republic | – | 0.066471 | Qatar | 0 | 0.0331155 |
| Denmark | 0.7 | 0.0120669 | Romania | 15.73 | 0.0451932 |
| Djibouti | 66.7 | 0.0313146 | Russia | 5.4 | 0.0357435 |
| Dominican Republic | – | 0.0557279 | Rwanda | 100 | 0.0720322 |
| Ecuador | 40 | 0.0212623 | San Marino | 12 | – |
| Egypt | 0 | 0.0105541 | Saudi Arabia | 0 | 0.00028 |
| El Salvador | 16 | 0.027693 | Senegal | 91.1 | 0.0464541 |
| Equatorial Guinea | 80 | 0.0073728 | Serbia And Montenegro | 23.1 | 0.0620166 |
| Eritrea | 83 | 0.0079781 | Sierra Leone | 33.7 | 0.0550279 |
| Estonia | 33.3 | 0.045948 | Singapore | 14 | – |
| Ethiopia | – | 0.0192632 | Slovak Republic | – | 0.0366983 |
| Finland | 6.11 | 0.0281986 | Slovenia | 6 | 0.0829063 |

| | | | | | |
|---|---|---|---|---|---|
| France | 5 | 0.0419574 | Somalia | 22 | 0.0029158 |
| French Guiana | – | 0.0250828 | South Africa | 21.98 | 0.0254923 |
| Gabon | 35.5 | 0.0297724 | Spain | 18.9 | 0.0245267 |
| Gambia | 62.9 | – | Sudan | 16.3 | 0.0241656 |
| Georgia | 19.7 | 0.0535029 | Suriname | – | 0.0378448 |
| Germany | 2.6 | 0.0573345 | Swaziland | – | 0.0263498 |
| Ghana | 84.7 | 0.0428391 | Sweden | 3.22 | 0.0210959 |
| Greece | 7.7 | 0.0568056 | Switzerland | 95.7 | 0.0816537 |
| Guatemala | 43.05 | 0.0253266 | Syria | 13.7 | – |
| Guinea | 88.4 | 0.0459671 | Tajikistan | 7.4 | 0.0581085 |
| Guinea-Bissau | 35.1 | 0.0327188 | Tanzania | – | 0.0250633 |
| Guyana | 6.8 | 0.0260477 | Thailand | 18.5 | 0.0371513 |
| Haiti | – | 0.0557279 | Togo | 0.5 | 0.0383742 |
| Honduras | 1.64 | 0.0356374 | Tunisia | 0 | 0.0223485 |
| Hungary | 5.3 | 0.0518711 | Turkey | 20 | 0.0443709 |
| India | – | 0.041401 | Turkmenistan | 19.4 | 0.0125608 |
| Indonesia | 17.6 | 0.0439307 | Uganda | 21.9 | 0.0236287 |
| Iran | 31 | 0.0310722 | Ukraine | 34.1 | 0.0514058 |
| Iraq | 20.2 | 0.023863 | United Arab Emirates | 0 | 0.0159324 |
| Israel | 19.47 | 0.0488663 | United Kingdom | – | 0.0747827 |
| Italy | 2.29 | 0.0545924 | United States | – | 0.0141888 |
| Ivory Coast | 59 | 0.0516242 | Uruguay | 0 | 0.0426762 |
| Jordan | 51 | 0.0160013 | Uzbekistan | 19.4 | 0.039586 |
| Kazakhstan | 41 | 0.0340632 | Venezuela | 0 | 0.0255508 |
| Kenya | 27 | 0.0133011 | Vietnam | 2.8 | 0.0521464 |
| Korea, North | 0 | 0.0277805 | West Bank | – | 0.0643072 |
| Korea, South | 0 | 0.0203319 | Yemen | – | 0.0117803 |
| Kuwait | – | 0.0114641 | Zambia | 31 | 0.0193908 |
| Kyrgyz Republic | 38.9 | 0.0486664 | Zimbabwe | 99 | 0.0276735 |