

# **Trade and economic performance: does Africa's fragmentation matter?**

Paul Collier, Anthony J. Venables

29<sup>th</sup> Dec 2008.

## **Abstract**

The population of South Asia lives almost entirely in one mega-country or two large ones. In contrast, the rather smaller population of sub-Saharan Africa is spread across some fifty countries. Does this political fragmentation have economic consequences? We suggest that both private economic activity and the provision of public goods benefit from powerful scale economies that confer advantages on the South Asian model. Paradoxically, although Africa has a greater need than other regions for supra-national power structures, it has made far less progress towards regional unity.

\* This work was supported by the BP funded Oxford Centre for the Analysis of Resource Rich Economies and by the Centre for the Study of African Economies. Views expressed are those of the authors and do not represent institutional positions. Thanks to Mauro Caselli, Lisa Chauvet, Anke Hoefler, Hyesung Kim and Jean-Louis Warnholz for statistical analysis.

## **Authors' Addresses:**

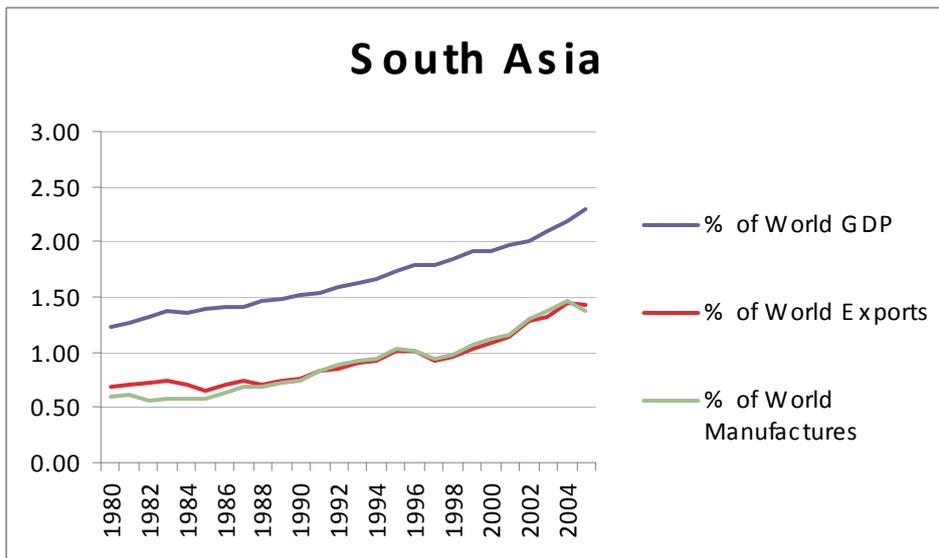
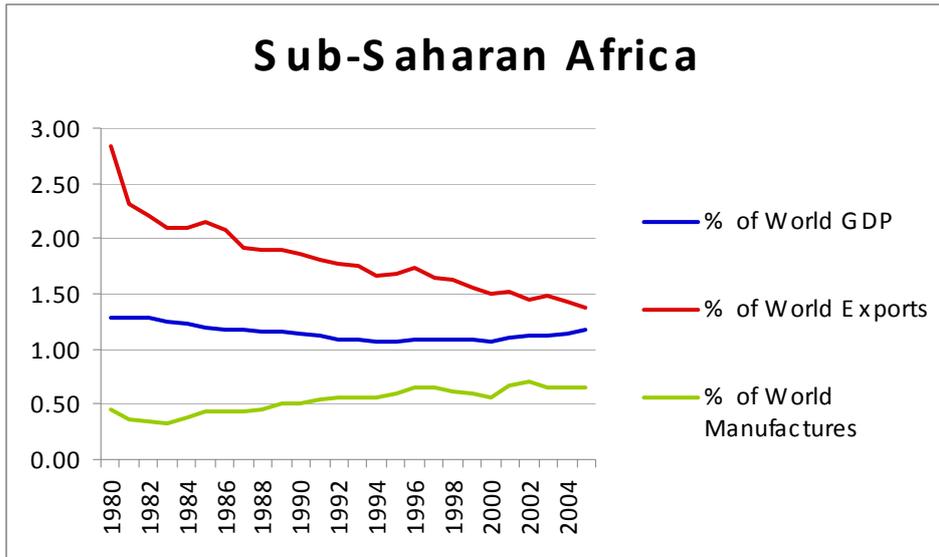
P. Collier  
CSAE  
Department of Economics  
Manor Road  
Oxford  
OX1 3UQ  
paul.collier@economics.ox.ac.uk

A.J. Venables  
Oxcarre  
Department of Economics  
Manor Rd  
Oxford  
OX1 3UQ  
tony.venables@economics.ox.ac.uk

## 1. Introduction

Africa and S. Asia have many features in common, and many that are strikingly different. They each contain roughly 1 billion people, and per capita income and human development indicator levels that were – in 1980 – broadly similar. But S. Asia is dominated by India, a unified state of 1,130 million people, and Pakistan and Bangladesh each have more than 150 million. By contrast, Africa, with a total population that is 60% that of South Asia, is divided into 54 states. Of these, even the largest, Nigeria, is smaller than the smallest of the big three South Asian countries, accounting for a mere 14% of Africa's population, whereas India accounts for 74% of that of South Asia. Some 95% of South Asians live in the big three, whereas the three largest countries of Africa account for only 28% of Africa's population. Indeed, the average African state has a population of only 17 million people, one-sixty-sixth the size of India. Does this matter?

The economic performance of the two regions has diverged sharply since 1980, as illustrated in the figures below. The regions had similar size GDPs in 1980, but whereas Africa's share of world GDP declined until the late 1990s, S. Asia's share nearly doubled. Africa's share of world exports was nearly four times that of S. Asia in 1980, but since then S. Asia's share has more than doubled and Africa's halved they are now of similar sizes. And whereas Africa's share of world manufacturing exports stagnated, S. Asia's share more than doubled. The Sub-Saharan African data is dominated by S. Africa, a single country accounting for around 20% of S.S Africa's GDP, and the performance of the rest of Africa is, on average, worse.



There are numerous reasons for this divergence, but in this paper we focus on a single set of issues; has Africa’s fragmentation into numerous small states contributed to its relatively poor performance?

The relationship between country size and economic performance has been analysed by a number of authors, with mixed findings. Recent work by Rose (2006) on a sample of 208 ‘countries’ (not all independent, eg the Isle of Man) leads him to the conclusion that ‘a country’s population has no significant consistent impact on its well-being’. In contrast, the growth literature has come up with evidence for a positive relationship between country size and growth. Key findings come from

Alesina, Spolaore and Wacziarg (and their coauthors, and see also Alcala and Ciccone 2003). The handbook article by Alesina et al (2005) provides a succinct summary of findings. They hypothesise that economic growth should be positively associated with country size and with openness, and negatively associated with their interaction. On a sample of 104 countries they find that these relationships are present and statistically significant. Furthermore, they are quantitatively significant: “for a country at the median level of openness (S. Korea) the effect of multiplying the country’s size by 10 would be to raise annual growth by 0.33 percentage points” (p1530).

The lack of consensus coming from these cross-country studies is perhaps unsurprising. On various measures of per capita income small countries rank both top (Luxembourg or Lichtenstein) and bottom (Burundi) of the list. Our approach is not to undertake aggregate cross-country studies, but rather to investigate the impact of fragmentation into separate states in a more micro-founded way. We will focus on Africa, and often compare it with S. Asia and with India in particular. Why is it that a fragmented (sub-) continent might be at a disadvantage relative to a more unified one?

Conceptually, there are three distinct mechanisms that generate costs of fragmentation. The first is fragmentation implies that natural advantages are likely to be unevenly distributed between countries. The second concerns the loss of scale economies, at the level of the firm, city, and country as a whole. The third concerns the loss of public goods as the scale of political cooperation is reduced.

Natural advantages, such as oil deposits or natural harbours, are unequally distributed across space. If a continent is politically fragmented, this increases the likelihood that natural advantages are unevenly distributed between countries. Evidently, this implies inequalities between countries. However, of greater pertinence for the present inquiry is that these inequalities are likely to imply inefficiencies: *average* income is reduced by fragmentation. This occurs if there are diminishing returns to having a natural advantage, in which case the aggregate benefits from nature would be greater the more equally shared are these advantages. Natural advantages may be fixed or transient, and we consider each of them. As examples of fixed differences in natural advantage we apply Collier’s classification of countries into resource-rich/ resource-scarce, and coastal/ landlocked. Transient differences arise as countries experience idiosyncratic shocks, with some countries benefiting, others possibly losing. We consider the exposure to export price shocks generated

by the structure of commodity exports and the fact that political fragmentation might reduce risk pooling. Section 2 covers these issues.

Whereas the first mechanism arose from diminishing returns and underlying or ‘first nature’ unevenness, the second is to do with increasing returns and the losses attributable to the inability of small countries to gain sufficient scale to work efficiently. One context in which this matters is urbanisation and the failure of Africa to develop highly productive urban centres of economic activity. Another is thick market effects, and the failure of many small African economies to have the scale to offer predictable economic environments. Section 3 covers these issues.

The third issue concerns the provision of public goods by means of politically organized collective action. The free-rider problem is frequently so acute that it can only be overcome by the coercive power of a government to tax its citizens, thereby generating the finance for public goods. Fragmentation of a continent into countries is, first and foremost, *political* fragmentation. As such it increases the costs of providing public goods, implying that provision will be both less adequate and more expensive. This is the subject of Section 4.

## **2. Uneven distribution of natural endowments**

### ***2.1 Fixed differences in natural advantage***

Fragmentation of a continent into countries means that geographically concentrated natural endowments – such as mineral resources, coasts or rivers – are likely to be unevenly distributed between countries, and so it turns out to be in Africa. Table 1 gives export earnings (as % GDP and per capita) from natural resources together with Collier’s (2008) classification of countries according to whether they are resource rich, landlocked, or coastal.

**Table 1: African first nature geography**

	<b>Fuels, Ores &amp; Metals: Export Value % GDP</b>	<b>Fuels, Ores &amp; Metals: Export Value per capita (\$)</b>	<b>Collier classification:</b>	<b>Year</b>
Equatorial Guinea	93.92	14591	Resource rich	2005
Angola	72.16	1471	Resource rich	2005
Congo, Rep.	71.46	1182	Resource rich	2005
Gabon	55.90	4071	Resource rich	2006
Chad	44.47	258	Landlocked	2005
Nigeria	40.94	214	Resource rich	2004
Botswana	34.74	1977	Resource rich	2005
Guinea	24.4	88	Resource rich	2006
Congo, Dem. Rep.	24.34	34	Resource rich	2006
Mauritania	19.79	123	Newly resource rich	2005
Mozambique	18.71	62	Newly resource rich	2005
Zambia	18.32	116	Resource rich	2005
Sudan	13.50	102	Newly resource rich	2005
Cote d'Ivoire	12.52	108	Coastal	2005
Mali	10.65	46	Landlocked	2004
Cameroon	9.4	89	Coastal (Ex-res-rich)	2005
Sierra Leone	9.4	20	Coastal	2005
South Africa	6.81	351	Coastal	2005
Zimbabwe	5.75	15	Landlocked	2005
Niger	4.55	12	Landlocked	2005
Kenya	4.52	21	Coastal	2004
Senegal	4.3	30	Coastal	2005
Togo	3.9	13	Coastal	2004
Namibia	3.11	97	Coastal	2006
Ghana	2.50	12	Coastal	2005
Madagascar	1.70	5	Coastal	2006
Central African Rep	1.62	5	Landlocked	2005
Tanzania	1.58	5	Coastal	2005
Cape Verde	0.87	19	Coastal	2006
Swaziland	0.81	9	Landlocked	2002
Uganda	0.66	2	Landlocked	2005
Ethiopia	0.54	0.7	Landlocked	2004
Burundi	0.34	0.3	Landlocked	2005
Rwanda	0.33	0.8	Landlocked	2005
Burkina Faso	0.31	1.2	Landlocked	2004
Mauritius	0.17	9	Coastal	2005
Benin	0.11	0.6	Coastal	2005
Malawi	0.05	0.1	Landlocked	2005
Liberia			Coastal	
Guinea-Bissau	0.44	0.9	Coastal	1995
Gambia, The	0.05	0.1	Coastal	2003
Comoros	0.03	0.1	Coastal	1997
Djibouti	0.01	0.1	Coastal	1990

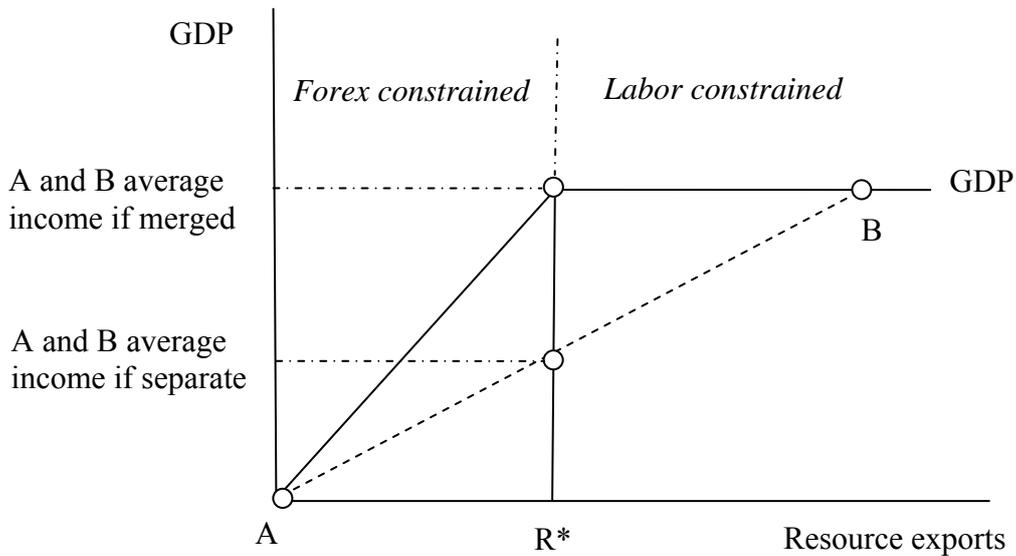
**Notes.** Diamond exports included in calculation for Angola, Botswana, DRC, Guinea and Zimbabwe, and gold exports included for DRC, Guinea and Ethiopia. For Mali, Fuels, Ores & Metals export consists solely of gold. For Sierra Leone, Fuels, Ores & Metals export consists solely of diamonds.

Resource export revenues per African citizen average range from several thousand dollars per head and more than half of GDP, to zero in resource scarce – and also landlocked – countries such as Malawi, Rwanda, Burundi or Uganda.

This unequal distribution matters for two reasons. The first is that it maps into an unequal distribution of resource rents per capita. Within a country resource rents (or at least those that find their way into public accounts) are likely to be spent throughout the country. The spending may not be spatially uniform; producing regions may be favoured, as in Nigeria, where 2005 Federal transfers from oil revenue amounted to \$210 per capita in oil producing states, and \$70 per capita in the North West of the country. But this is a far wider distribution of rents across intra-country boundaries than occurs across international boundaries (ie, zero).

It is not only equity that is damaged by unequal distribution of resource revenues. Since the economic impact of resource revenues is likely to be subject to diminishing returns their unequal distribution also leads to an efficiency loss. A simple economic model makes the point. Suppose that every country consumes and produces a single good that is non-tradable. Production of the good uses foreign exchange (imported oil or equipment) and domestic labour in fixed proportions. The only source of foreign exchange is resource revenues, and labour is in fixed supply. Real income in such an economy is illustrated in figure 1, in which resource revenue is measured on the horizontal axis. If resource exports are less than  $R^*$ , then production is foreign exchange constrained, and real income is given by the upwards sloping section of the line (with slope equal to the foreign exchange content per unit GDP). If natural resource earnings are greater than  $R^*$ , then the economy is labour constrained, this fixing income; further resource earnings beyond this point are simply accumulated as foreign assets. As a simplest case, suppose that one economy has no resource revenue (so is at point A) and another has resource revenue and is at point B. Average income – of the two separate countries – is the midpoint between A and B. Merger of the two economies would exactly double income, as illustrated.

**Figure 1: Income loss from uneven distribution of resources**



This is a very clear cut example – what insights does it provide to reality? There are two key elements to the argument. The first is that shortage of foreign exchange constrains production in economies without resource earnings. Many of the resource scarce and landlocked African economies have extremely low shares of exports in GDP – less than 15% of GDP for 8 such countries. Accessing world markets has been particularly difficult for these countries, and they are heavily aid dependent. If they were located within a single country, then such areas would earn the resources to finance ‘imports’ by intra-country trade. The argument must therefore turn on the fact that barriers to trade created by international borders are an order of magnitude greater than within-country trade barriers, and there is plenty of evidence that this is the case. The classic studies of the barriers created by international borders, compared to within country trade costs, are based on trade between Canada and the US. McCallum (2005) and Helliwell (1997) show that exports of Canadian provinces to other Canadian provinces are some twenty times larger than their exports to US states at the same distance. According to one study the US – Canada border is 7,000 miles wide, in the sense that it chokes off trade as much as would 7,000 miles of border-less distance. African borders are generally very much more difficult to navigate than is the border between the US and Canada. Limao and Venables (2001) find that poor infrastructure is particularly important in

choking off trade between African countries. The implication is that resource scarce land-locked regions face more acute problems financing imports as separate nations than they would as regions of a larger country.

The second part of the argument contained in figure 1 is that, at some point, there are diminishing returns to resource earnings. In the example of the figure this was the economy hitting full employment, so no more labour is available to produce further income, but the argument is more general. What are the sources of diminishing returns to the value of resource revenues in resource-rich African economies? There are often constraints on the supply of particular non-tradable services, such as construction services or particular labour skills. Spending then bids up the price of these inputs but does not buy additional real services. More generally, spending from resource revenues will be met by a combination of increased output, and crowding out of other expenditures. What are the expenditures that are crowded out? It may be exports, this giving rise to the 'Dutch disease'. Alternatively, monetary and exchange rate policy might be used to mitigate the Dutch disease, in which case crowding out will impact domestic activities, quite likely investment. If these activities are particularly valuable (as they would be if they are initially operating at a sub-optimal level) then crowding them out may actually reduce income.

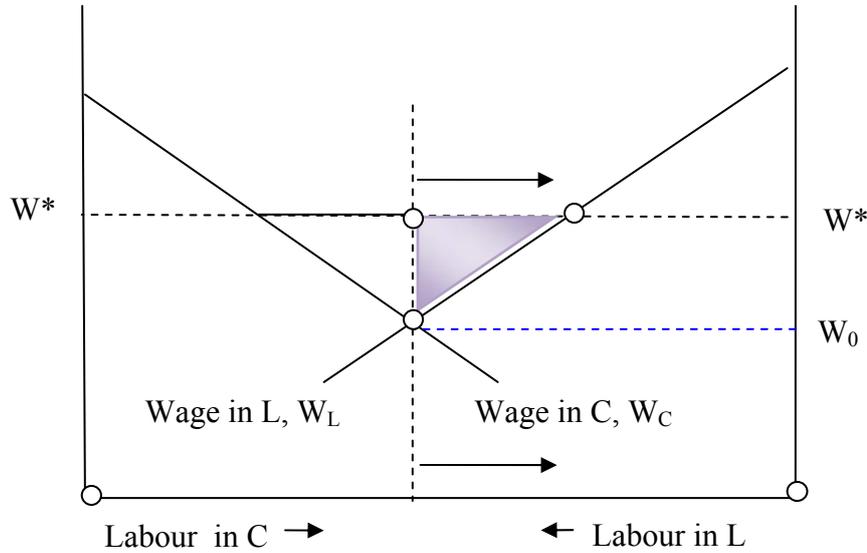
The key point for the present argument is that the smaller are resource revenues relative to the economy as a whole the more favourable will be the balance between increasing income vs crowding out other expenditures. If two countries are merged, the supply curve of merged country is the horizontal sum of the supply curve of each separately, so a given size increase in demand will lead to a larger quantity increase and smaller price increase in the merged economy than in a separate one. While this argument focuses on income and expenditure, other aspects of the 'resource curse' (eg political economy issues) may also exhibit increasing marginal cost, possibly meaning that benefits of resource revenue do not just flatten out with respect to resource revenue, but turn negative at the margin. In this case the citizens of *both* countries would gain from merger and a sharing of the economic impact of revenues.

It is not just natural resources that are unequally divided between countries, but also access to the coast. Coastal economies in developing countries are much better placed to engage in producing manufactures for world markets than are landlocked countries, and such activities are important drivers of growth (Jones and

Olken 2007, Spence 2008). These differential opportunities mean once again that the fragmentation of a region into separate countries may create both spatial inequalities and efficiency losses. A simple economic framework is helpful again. Suppose that economies can produce a good in which there are diminishing returns to labour. This might be agriculture, in which a fixed supply of land generates diminishing returns. Figure 2 illustrates two such economies. The total labour force is the length of the horizontal axis, and workers are equally divided between countries, with workers in coastal economy C measured from the left hand origin and those in landlocked economy L measured from the right. The value marginal product of workers in each country is given by the downward sloping lines, the slope reflecting the diminishing returns to labour. If this was all there was, wages in each country would be equal (and low,  $W_0$ ) at the intersection of these lines.

Suppose however that C, the coastal economy can also undertake an export activity for the world market that does *not* run into diminishing returns. Given productivity levels, this activity can pay wage  $W^*$ . The coastal economy will then have some of its labour move to this sector, and wages rise until they reach  $W^*$ . However, country L is landlocked and unable to access this source of employment, so is left with wages  $W_0$ . Fragmentation into two political units means a fixed division of the labour force, and the lack of migration between C and L creates inequality (the gap between  $W_0$  and  $W^*$ ) and also a loss of efficiency and real income. Within a single economy there would be internal migration of labour from L to C, so that the division of the total work force would move in line with the horizontal arrows illustrated. Migrants would gain, and would also bid up wages for those remaining, until all workers receive  $W^*$ . While landowners (owners of the sector specific factor) in L would lose, there is an overall real income gain from the migration equal to the shaded triangle. Furthermore, if the export activity had increasing returns to scale rather than constant returns then this expansion in employment would increase productivity and the wage  $W^*$ , raising income in both country C and country L.

**Figure 2: Coastal and landlocked economies**



What insight does this offer? Some Asian economies have witnessed massive migration to those regions (usually although not necessarily coastal) that have been successful in building up sectors supplying world markets. The best example is West to East migration in China. Internal migration is also important (if poorly documented) in India, both as seasonal migration, and as part of rapid urbanisation.

Africa has not yet developed such magnets of employment, but the analysis points to the fact that if it were to do so, fragmentation would prevent it from fully realising the benefits. In Africa there have been substantial migration flows, but they have often been problematic. Even where governments permit international migration, the non-citizen status of immigrants creates opportunities for the politics of xenophobia which increase as the stock of immigrants accumulates. Since attractive political niches at some point attract politicians willing to occupy them, the opportunity for xenophobia is unlikely to remain unexploited. In turn this political response exposes immigrants to violence and expulsion. The clearest instance of this depressing sequence arose from the sharp difference in natural advantage between Cote d'Ivoire and Burkina Faso. With its radically better economic opportunities due to its coast and its rainfall, Cote d'Ivoire attracted massive immigration from Burkina Faso. Indeed, at one stage around 40% of the labour force in Cote d'Ivoire was Burkinabe. During the 1990s this was exploited by populist politicians and was

instrumental in triggering the political meltdown into coups and civil war during which much of the migration was reversed.<sup>1</sup> Nigeria during the 1970s provides a second instance of the same sequence. The oil discovery in Nigeria created a sharp distinction in natural advantage vis-à-vis Ghana. This generated mass migration from Ghana: at one stage around a fifth of the Ghanaian population had emigrated. This in turn induced xenophobia in Nigeria: once economic conditions deteriorated the Ghanaians were formally expelled. South Africa witnessed its own backlash against immigration in 2008.

In addition to constraining migration, a further adverse effect of fragmentation can be illustrated with reference to figure 2. Owners of the specific factor in the diminishing returns sector in the coastal country (C) suffer an income loss as production of the export activity takes-off. The diminishing returns sector is most naturally thought of as agriculture, and as manufacturing expands the agricultural sector has to compete for labour and faces higher wages. Agriculture in C contracts, returns to land in country C fall, and C comes to import agricultural goods (from country L and the rest of the world). There is a likelihood that this creates a lobby for tariffs on agricultural imports. While this would raise the real income of landowners, it would have a negative on total income in the coastal economy. However, the political economy of protectionism is well-understood to favour small but cohesive interests over the general interest. The import tariff would also choke off manufacturing production in the coastal economy, possibly with damaging long run effects on development. And if the coastal economy were the primary market for the landlocked country's agricultural exports, then this would also damage country L. The point is of course, that the opportunity for using an import tariff – at least against imports from country L – exists only if C and L are separate countries. Merger therefore removes a policy instrument that is likely to be misused in response to political economy pressures. An example of such manipulation of trade barriers as a result of political fragmentation is the banning of food imports from Uganda by President Moi of Kenya in the early 1990s in response to lobbying from business interests who were holding large stocks of food.

---

<sup>1</sup> See Collier (2009) for a detailed account of this sequence.

## *2.2 Time-Variant differences in natural advantage: export volatility*

The economic advantages and disadvantages of a particular location at a particular date are shaped partly by fixed geography, and partly by the impact of short run shocks. The arguments we have developed above apply also to these transient effects, as well as the permanent ones. Shocks create unevenness between areas, and the ability to spread their impact (i.e. to pool risk) – with both distributional and efficiency implications – is impeded by fragmentation into national units.

Collier and Goderis (2008, 2008a, 2008b) investigate the consequences of commodity export shocks for GDP valued at constant prices. This abstracts from the income effect accruing directly from changes in the terms of trade and focuses on the consequences for real output. They find that the effects of shocks on growth are asymmetrical: adverse shocks significantly and substantially reduce output, whereas positive shocks do not have significant effects. The effect of adverse shocks is substantial. For example, for a typical African country whose commodity exports are initially around 35% of GDP, the consequence of a 30% fall in export prices would be to reduce growth in the following year by 3.6 percentage points. More generally, given the frequency of shocks and their cost, it is possible to estimate the discounted present value of the output losses that they generate. Using a 3% discount rate, for individual countries the cost is sometimes large: for example, for Nigeria it is 13% of one year's GDP. However, summed over the entire continent, country-by-country, the cost is modest at around four percentage points of one year's GDP.<sup>2</sup>

For this paper we have investigated whether had Africa been divided into fewer polities these costs would have been reduced.<sup>3</sup> We investigate both the consequences of a United Africa, which provides an upper bound to the analysis, and of political regrouping into four regional blocks. A United Africa would have reduced the costs by a mere 0.6 percentage points of the region's GDP. The gains from regional groupings are more variable, the largest being from a United West Africa, where they would amount to 1.7 percentage points of GDP.

---

<sup>2</sup> These costs are likely to be under-estimates because they omit longer term effects. Collier and Goderis find that the long term growth effects of dependence upon commodity exports are adverse unless governance is good. The mismanagement of volatility may be one of the routes by which poor governance has these adverse long-term effects.

<sup>3</sup> We would like to thank Benedikt Goderis for the substantial work involved in these recalculations.

The main reason why the effect is so limited is that the scope for risk-pooling in Africa is modest. Virtually all countries are commodity exporters. Not only are the prices of most commodities quite highly correlated, but a single commodity, oil, dominates Africa's exports. For a United Africa, around 65% of commodity exports would be oil. Hence, pooling has little impact on the size of the average shock. The key potential gain comes therefore not from reducing the average size of shocks but in changing their distribution. Most African countries do not export oil, so that a United Africa would have fewer enormous shocks and more moderate sized shocks than the distribution of country-specific shocks. Whether such redistribution matters depends upon the precise relationship between the size of shocks and their costs. The original specification of Collier and Goderis is linear: the costs of a shock are thus required to be proportionate to the size of the shock. In this case, redistributions that do not affect the average make no difference.

Hence, for this paper we investigated whether the true structure of costs was non-linear, rising more than proportionately with the size of the shock as seems inherently likely. We found little basis for a non-linear relationship. For example, when the square of the size is added to the regression, although it is negative it is not close to significance. Hence, we conclude that *given Africa's export structure*, the scope for reducing the cost of shocks through greater political unity is modest. However, as we will now argue, Africa's export structure may itself be a consequence of political fragmentation. Manufacturing and service exports benefit from scale economies that are less important in primary commodities. Hence, if political fragmentation has frustrated these scale economies it may have locked Africa into dependence upon primary commodity exports.

### **3. Scale, density and increasing returns**

The previous section pointed to diminishing returns as a source of loss for a fragmented continent with uneven distribution of resources or uneven impact of shocks. We now turn to increasing returns, and the cost of opportunities foregone by small countries. We first focus on a microeconomic mechanism that we think is particularly important in many small African countries: the fact that markets are too thin to be competitive. This constitutes a major obstacle to investment both through

high prices of capital goods and more fundamentally, through vulnerability to opportunistic behaviour. We then turn to how the microeconomic advantages of scale combine to produce a positive relationship between city size and productivity, and suggest that fragmentation has created an African city structure that is ill-suited to reap these productivity effects.

### ***3.1 Competition and the operation of markets***

Small economies are likely to have high levels of monopoly power. This makes them bad places for new investments to take place. Incumbent firms have little incentive to expand output or to innovate, and strong incentives to deter entry and innovation by newcomers. Monopoly raises the price so many intermediate goods as well as final output, thereby raising costs. And most importantly, monopoly creates the potential for opportunistic behavior in transactions and thereby creates a difficult business environment, even for the monopolists themselves.

Evidently, a small market is likely to be less competitive than a large one as – given some firm level economies of scale – fewer firms will operate. The effect will be particularly pronounced in sectors that are closed to trade. For example, the typical African economy has a very highly concentrated banking sector: often four banks dominate lending and this is a sufficiently small number to enable collusive oligopoly. The limited nature of the market also leads to a concentration of risks: banks are exposed to a high covariance of the risk of default.

Transport is another sector that is sheltered from international competition and that is often highly cartelized. A recent study of transport in Africa finds that the real costs of transport services are not abnormally high, but that trucking firms are able to charge exceptionally high prices (Teravaninthorn and Raballand 2008). Average prices per ton kilometre are \$0.02 in Pakistan, \$0.05 in China, \$0.08 for the Mombasa-Kampala run and \$0.11 for Doula-N'Djamena. Many African economies have restrictive regulatory regimes and transport cartels, and de-regulation of the trucking industry in Rwanda has been estimated to have reduced transport prices by 75%. For some countries, a major factor supporting cartelization is a treaty structure between countries designed to protect the national trucking industry from competition from neighboring countries.

Monopoly power raises relative prices, and a key relative price in an economy seeking to grow is the price of investment relative to the price of GDP as a whole. Given the level of saving, the higher is the relative price of investment, the less physical equipment it will purchase. This effect is quantitatively important, as the price of investment (relative to GDP as a whole) can be three or four times higher in developing countries than it is in high income countries. Recent work by Caselli and Feyrer (2007) shows that much of the variation in the marginal physical product of capital across countries is in fact due to this price effect.

Why is there this price difference? Part of it is attributable to the Balassa-Samuelson effect; investment has quite a high import content, and prices of tradable goods are relatively high in low income countries. Part of it may also be due to thin markets and monopoly power in supply of equipment and investment goods, and this may be a function of country size. To investigate this, we have explored the impact of GDP per capita (for the Balassa-Samuelson effect) and country size (for the market power effect) on this relative price. The evidence is given in table 2, which reports the results of regressing the price of investment relative to the price of GDP on real GDP per worker and on the number of workers, as a measure of economic size. The extremely strong dependence on output per worker is clear, and so too is the scale effect. Increasing the labour force by a factor of 20 reduces the relative price of investment by 13%. Restricting the sample to countries with a labour force of less than 20 million produces a quantitatively larger effect. Recalling that India's population is 66 times larger than the average African country, which has a workforce of less than 10 million, the implications for differences in the price of capital are substantial.

**Table 2: The relative price of investment** (all variables in logs)

	Price of investment	Price of investment
Real GDP per worker	-0.27 (-12.7)	-0.30 (-8.8)
Number of workers	-0.046 (-3.5)	-0.071 (-2.6)
N	N = 163	Population < 20 million, N=83
R <sup>2</sup>	0.50	0.50

Source, Caselli and Feyrer (2007) from Penn World Tables.  
t-statistics in parentheses

Thin markets and the resulting monopoly power increase the price of capital goods, but they have further pernicious effects. One is that they create an incentive for incumbent firms to actively pursue strategies that deter entry of new firms. From the perspective of firms that are already operating in a sector, if one firm devotes resources to keeping new entrants out of the sector, this is a public good. The strategies of entry-deterrence may be the use of predatory pricing, or the purchase of political influence. Evidently, from the perspective of society as a whole such behavior is undesirable. In an industry with many existing firms the free-riding problem implies that the returns to any one firm from such anti-social behavior are limited. But in a small market with an incumbent monopolist all the benefits to the existing industry are internalized and so the incentive to act to keep out new competitors is maximized.

Small and thin markets are also unattractive places to invest because investors are vulnerable to ‘hold-up’ – opportunistic behaviour by other firms with which they have to transact. Hold-up refers to the possibility that, once an investment has been sunk, the investor faces a monopsonistic purchaser of the output of the investment. Even if the purchaser and investor entered an agreement before the investment is undertaken, *ex post* the purchaser may act opportunistically, breaking the agreement and only offering a lower price. The investor will anticipate this possibility of hold-up, so may not make the investment in the first place. How is the hold-up problem overcome? One way is by making the *ex ante* contract legally binding but, even in countries with strong legal systems, it is often impossible to write a contract with the degree of completeness that will rule out such opportunistic behavior. The other is to

make sure that there are many alternative uses for the output of the project. This is partly a matter of the specificity of the investment (it might be a machine for making specialist parts demanded by a single manufacturer), and partly a matter of the size of the market in which the output is sold. Hold-up is more likely the fewer people are competing for the output.

This suggests that in small economies the threat of hold-up may be a major deterrent to investment. In agriculture, returns to investment are reduced if there is a monopsonistic grain merchant. In manufacturing, few potential purchasers of output deters investment. This gives rise to coordination failure – there is no incentive to enter on one side of the market until the other side has got more firms, and vice versa. And, in a small economy, even the return to the worst option, liquidating the investment, may be reduced by thin markets for second hand capital equipment. Distress sales are likely to be more coincident because smaller economies are less diversified, further depressing the expected price. The hold-up phenomenon applies not only to goods markets, but also to labour markets; the incentive to take training is reduced if the skill acquired can only be sold to one employer.

These arguments point to the fact that smallness does not just create static monopoly/ monopsony power, but also creates a fundamentally more risky business environment. Entry of new producers will be deterred by predatory behavior of incumbents and by the scarcity of outside options and consequent vulnerability to predatory and opportunistic behavior.

### ***3.2 Productivity and City Size***

Productivity tends to be higher in large (and/or dense) clusters of economic activity. This is the reason why cities form. Firms and workers locate to gain the benefits of this productivity advantage, despite the congestion costs and other diseconomies associated with large cities. A number of mechanisms drive this productivity effect. Some are narrowly technical, for example the fact that dense activity economises on transport costs, and improves communications (and possible learning externalities) between firms, and between firms and workers. Others are to do with the impact of size on market structure and the intensity of competition, as discussed in the preceding subsection. Still others are to do with political economy; a city with a large business sector is likely to have a strong business lobby, this producing a business-

friendly investment climate. The quantitative evidence of the productivity effect of city size comes largely from studies of cities in developed countries. Rosenthal and Strange (2004) report a consensus view that doubling city size is associated with a productivity increase of some 3 – 8%. This is a large effect – moving from a city of 100,000 workers to one of 3 million is predicted to increase productivity by more than 30%. Au and Henderson (2006) find even larger results for Chinese cities, where they estimate that moving from a city of 100,000 workers to one of 1.3 million workers raises productivity by 80%, although beyond this scale weak diminishing returns cut in.

In this section we investigate the effects of Africa's fragmentation on its city structure and argue that fragmentation has a negative impact on city size producing a city structure that is weak compared to an integrated country, such as India. We have argued elsewhere that Africa's failure to develop large clusters of economic activity has had major implications for its economic performance. Africa's failure to enter world markets for manufactured exports is best understood in terms of the location of productive clusters of activity (Collier and Venables 2007). Many Asian cities have already gained a head start in these sectors and grown highly productive clusters, this creating a barrier for new entrants. Asia's initial advantage over Africa when Asia first penetrated the global market in manufactures in the 1980s might have been modest. Quite probably the reasons for Africa's initial disadvantage have evaporated: for example, much of coastal Africa was beset by conflict (such as Mozambique) or poor economic policies (such as Ghana). Yet, as clusters have developed in Asia the resulting scale economies have given Asia a new and more formidable advantage: Africa may have missed the boat on industrialization unless OECD trade policies artificially create an offsetting temporary advantage for African manufactures which pump-primes the formation of clusters.

Does the political fragmentation of Africa have any bearing on this? The obvious fact is that small countries generally have smaller cities, so some of these productivity benefits are foregone. We first investigate this by a cross-country regression to explore the determinants of city size. We use a large data city of world cities, and take as dependent variable the population of the  $j$ th ranked city in country  $i$ . We work with the top five ranked cities in each country, giving – with some missing values for very small countries – 521 observations. The explanatory variables are

country population, country area, country per capita income, and the rank of the city in the country, i.e.,

$$\ln(\text{population}_{ij}) = a + b_1 \ln(\text{population}_i) + b_2 \ln(\text{area}_i) + b_3 \ln(\text{income pc}_i) + b_4 \ln(\text{rank}_j) \dots$$

Results are given in table 3 for specifications with and without regional fixed effects, for the world as a whole and for countries with per capita income of less than \$10,000. As is well known, national per capita income has a positive effect on city population. The city's within country rank has a negative effect, as it must by construction. The estimated parameter in our central specification is -1.08, close to Zipf's law (the rank size rule, stating that within countries the size of each city is inversely proportional to its rank in the city size distribution, see Gabaix and Ioannides 2004).

**Table 3: City population:** (all variables in logs)

	City population	City population	City population
Country population	0.639 (23.3)	0.70 (18.9)	0.731 (13.7)
Country area	0.169 (7.1)	0.107 (3.9)	0.085 (2.1)
GDP pc	0.27 (7.6)	0.27 (7.6)	0.102 (1.8)
City rank	-1.08 (-25.0)	-1.1 (-28.0)	-1.21 (-25.0)
Region fixed effects	No	Yes	No
N	521	521	Income per capita < \$10k. N = 325.
R2	0.80	0.83	0.83

Of most interest for our purposes is that fact that both country population and country area are highly significant determinants of city size. The sum of the coefficients on these two variables is 0.8, indicating that a merger of two similar size countries – ie a doubling of population and area – would lead to a 75% increase in the

size of the largest city. To see the quantitative implications of this, suppose that initially there are 10 separate countries, in each of which the largest city has a population of 3 million people. Combining these countries and letting city sizes adjust in line with the regularities given in table 3, produces a largest city of 19 million, and a size distribution of city populations given (for the first 10 cities) by; 19mn, 9.5mn, 6.3mn, 4.7mn, 3.8mn, 3.1mn, 2.7mn, 2.4mn, 2.1mn, 1.9mn...).

It is interesting comparing these calculations with the actual city size distributions of Africa as compared to India (table 4). Both the calculations and the actual data for India suggest that, compared to Africa, a large integrated country has much larger cities at the top of the rank; slightly fewer upper middle ranking (compare the ranks of the Indian and African cities with population around 3 millions as given in table 4); and many more mid to large cities (eg number of cities with population between 1 and 2 million).

This analysis suggests that the smaller size of African cities is due, in large part, to the fragmentation of countries. Our preceding analysis and the evidence from developed country studies (as summarised by Rosenthal and Strange 2004, cited above) suggests that this may have had an adverse effect on the productivity of African manufacturing. While size is not the only determinant of urban productivity, Africa's fragmentation and consequent urban structure may have impeded the development of major international manufacturing centers of the type that contribute to the performance of high growth economies.

**Table 4: Cities with population greater than 1 million: India and Africa.**

Mumbai	21,600,000	Lagos	Nigeria	10,100,000
Delhi	21,500,000	Kinshasa	Congo DR	8,200,000
Kolkata	15,700,000	Johannesburg	South Africa	7,800,000
Chennai	7,850,000	Al-Kharṭūm	Sudan	5,450,000
Bangalore	7,350,000	Abidjan	Côte d'Ivoire	4,225,000
Hyderābād	7,150,000	Durban	South Africa	3,600,000
Ahmadābād	5,650,000	Kano	Nigeria	3,600,000
Pune	4,625,000	Cape Town	South Africa	3,400,000
Sūrat	3,875,000	Accra	Ghana	3,350,000
Kānpur	3,475,000	Ibadan	Nigeria	3,200,000
Jaipur	3,050,000	Nairobi	Kenya	3,175,000
Lucknow	2,800,000	Adis Abeba	Ethiopia	3,100,000
Nāgpur	2,700,000	Dar es Salaam	Tanzania	3,000,000
Patna	2,350,000	Luanda	Angola	2,875,000
Indore	1,870,000	Dakar	Senegal	2,550,000
Vadodara	1,870,000	Pretoria	South Africa	2,450,000
Coimbatore	1,820,000	Harare	Zimbabwe	2,200,000
Bhopāl	1,810,000	Douala	Cameroon	2,000,000
Ludhiāna	1,730,000	Maputo	Mozambique	1,820,000
Āgra	1,700,000	Antananarivo	Madagascar	1,760,000
Kochi	1,660,000	Bamako	Mali	1,730,000
Visākhapatnam	1,610,000	Lusaka	Zambia	1,720,000
Meerut	1,600,000	Yaoundé	Cameroon	1,610,000
Asansol	1,580,000	Conakry	Guinea	1,600,000
Bhubaneswar	1,560,000	Kaduna	Nigeria	1,590,000
Nāshik	1,550,000	Kumasi	Ghana	1,520,000
Chandīgarh	1,520,000	Kampala	Uganda	1,490,000
Vārānasi	1,470,000	Lubumbashi	Congo (DR)	1,450,000
Kolhāpur	1,460,000	Muqdisho	Somalia	1,410,000
Jamshedpur	1,350,000	Brazzaville	Congo (Rep.)	1,330,000
Madurai	1,350,000	Lomé	Togo	1,320,000
Rājkot	1,320,000	Ouagadougou	Burkina Faso	1,260,000
Jabalpur	1,300,000	Benin City	Nigeria	1,180,000
Dhanbād	1,290,000	Port Harcourt	Nigeria	1,170,000
Amritsar	1,270,000	Port Elizabeth	South Africa	1,150,000
Allahābād	1,230,000	Freetown	Sierra Leone	1,110,000
Vijayawāda	1,220,000	Cotonou	Benin	1,090,000
Srīnagar	1,180,000	Maiduguri	Nigeria	1,040,000
Shambajinagar	1,170,000			
Solāpur	1,100,000			
Thiruvananthapuram	1,100,000			
Rānchi	1,090,000			
Jodhpur	1,040,000			
Guwāhātī	1,030,000			
Tiruchirāppalli	1,010,000			
Gwalior	1,000,000			

Source: <http://www.citypopulation.de/World.html>

#### 4. The Political Economy of Public Goods

Above we have considered those advantages of being a large country that arise from the greater diversity and geographic concentration of private economic activities. We now consider those that arise from government activities. A core function of a government is to supply public goods. By definition, public goods are subject to scale economies. The distinctive aspect of public goods that *ensures* scale economies is that consumption is non-rival: one person's consumption does not reduce that of another. However, many public goods also have more conventional scale economies: their production technology has high fixed costs which can be spread over more consumers as scale is expanded. A radio station displays both types of scale economy: listening is non-rival while, once the fixed costs of transmission have been incurred the number of hours per day broadcast is subject to much lower variable costs.

As the number of consumers is progressively increased to the point at which it includes all the inhabitants of the world, relatively few goods still have unrealized economies of scale. At some point consumption ceases to be non-rival, and technological scale economies are exhausted. Conceptually, we can order all the goods that might potentially be provided publicly according to the minimum size at which all scale economies are realized. Along this ranking we find first those goods that are only non-rival within the locality, then the nation, then the region, and finally those that are global.

The supply of public goods generates acute collective action problems which, except in a few cases, require coercive powers of taxation to overcome effectively. The highest level at which such powers are found is the nation. Hence, in practice, those scale economies that occur beyond the level of the nation state are generally not reaped.

The failure to reap such economies is widely lamented in discussions of the under-provision of global public goods (Barrett, 2007) and regional public goods (Cook and Sachs 1999). However, it has a powerful corollary through its implications for differences in the size of countries. Since there are some public goods that are global, *a fortiori*, there must be many more for which the minimum efficient size lies somewhere within the huge population range implied by the difference between the smallest African country and India. Within this range the smaller is the country the

fewer goods can efficiently be supplied at the level of the nation state. This consideration is reinforced once population is replaced as the metric of size by GDP. Public goods are economic activities and the relevant metric for scale is more likely to be the size of the economy than the size of the population. Since Africa is now the poorest region, its typical national economy is even tinier than suggested by the size of its population. The economy of Luxembourg is roughly the same size as that of the five countries of the East African Community combined.

It might seem that nations could be too large as well as too small for the efficient supply of public goods. If different communities have distinctive preferences for public goods, provision might work better if political decision units are small and so can reflect these differences. However, there is an asymmetry. Once the optimal scale has been reached, a large state can always replicate this scale by decentralizing supply to sub-national authorities: in other words, once all the scale economies have been reaped, further expansion can be under conditions of constant returns to scale. States that are too small do not have an equivalent option: a national government can choose to pass authority down, but it does not have the power to pass authority up. A large nation thus has an advantage over a small nation, and this may become very pronounced by the time we reach the tiny states that are common in Africa. We now analyze the provision of three public goods that are fundamental to prosperity: security, economic policy, and infrastructure.

#### ***4.1 Security***

Security is the clearest case of a public good that is subject to scale economies far beyond the size of the typical small African state. Like radio, security benefits from both types of scale economies. Over a wide range defense is non-rival: the same army that defends one community from rebellion can defend a proximate community. Over a very wide range it benefits from scale economies: big armies usually defeat little armies, a proposition formalized in contest success functions. The sheer power of scale economies in security has repeatedly been revealed in the expansion of empires. Once one power gets a military advantage over its neighbors, it can expand almost without limit if it chooses to do so. Rome, the Mongol Empire, Russia, and the nineteenth century European empires demonstrate that big is safe.

The incidence of warfare in Africa has been far higher than that in India and differences in scale in part account for this. Evidently, political union would have reduced the incidence of international war, but almost all of Africa's wars have been internal and so the key issue is how union would have affected this risk. Statistical analysis of the risk of civil war finds that while population significantly increases the risk, the effect is substantially less than proportionate: a territory under a single polity has a lower risk than the combined risk from two polities were it split in half (Collier, Hoeffler and Rohner, 2009). However, the case for scale is complicated by the trade-off with ethnic diversity. In general, in order to make a polity larger it is necessary to take in additional social groups and so diversity increases. Unfortunately, diversity increases the risk of civil war. Were Africa to have been split into fewer countries this adverse effect might have more than offset the benefit from greater scale. The issue has recently been analyzed by Wigstrom (2008) who has carefully investigated how mergers between neighboring African countries would affect ethnic diversity. It transpires that in some cases Africa's borders are so arbitrary that ethnic diversity would actually be *reduced* by merger. In these instances the scale and the diversity effects of political union work in the same direction, reducing the risk of civil war. Even were diversity to increase, he finds that over a wide range had Africa been divided into fewer independent polities it would have had a lower risk of civil war. While his analysis is of course hypothetical and cannot take into account many aspects of politics, it omits the consequences for peace of the economic benefits of scale discussed above. Since both the level and growth of income significantly reduce the risk of civil war, these economic effects of political union would have reinforced the effects discussed here.

The small scale of African polities not only increases the incidence of war, it increases military spending while at peace. Because large armies tend to defeat small armies, small states tend to compensate by spending a higher fraction of GDP on the military. Further, while a country can increase its own security by spending more on its army, this reduces the security of neighboring countries that feel threatened. Evidently, in response to a perceived increase in threat, the neighboring country will need to increase its own military spending, producing an arms race. The essence of an arms race is that if we think of military spending as producing security, an increase in spending by one country reduces the average productivity of its neighbor's spending, but increases the marginal productivity. Collier and Hoeffler (2007) establish that

neighborhood arms races have been common around the developing world. For the present paper we have used their results to simulate the reduction in average African military spending had there been a United Africa. We set the incidence of civil war equal to that estimated by Wigstrom (2008), eliminate the effect of neighborhood arms races, and set all other variables at their average actual values for the continent. The predicted level of military spending falls by a quarter from 3.2% of GDP to 2.4%.<sup>4</sup>

#### ***4.2 Good economic policies and governance***

A second fundamental public good is the provision of good economic policies and accountable government. The choice between good policies and bad policies is typically determined by a complex mix of influences. Among them are the interests of the elite, the political power of ordinary citizens, the degree to which both elites and citizens understand basic economic issues and so grasp how their interests are best served by policies, and finally the capacity of the civil service to design and implement policies.

Scale may enter here through various routes. One is that the quality of the civil service can be higher in a larger society simply because it can be more selective. It is around 50 times more competitive to become Permanent Secretary to the Ministry of Finance in India than in Africa and so the quality will on average be higher.

A second scale effect is that paradoxically a larger society can be better informed about economic issues than a small society. The key reason is that there are scale economies in the commercial media: radio, television, newspapers and magazines. A large market will permit more of these to exist than a small market. The serious discussion within a society of economic issues is highly dependent upon the existence of specialist media. India has such media whereas in Africa only South Africa comes anywhere close to providing a market in which specialist journals are viable. For example, the Indian newspaper *The Economic Times* has a circulation of 1.2 million, which is sufficient to finance a staff of economically qualified journalists. With the same density of circulation an economics newspaper in Zambia would have

---

<sup>4</sup> We would like to thank Anke Hoeffler for undertaking this calculation.

a circulation of under 10,000 and so would not be viable. Without a specialist media discussion in the society is likely to be less sophisticated and so the pace at which social learning takes place will be slower. In effect, the society needs a critical mass of educated citizens before social learning can be rapid. This may help to explain why India reformed its economic policies ahead of Africa.

A distinct reason why larger states may be able to reform faster is a corollary of their greater need to decentralize authority: more public goods reach the level at which decentralization is the efficient form of organization. Such decentralized authority introduces variation in strategies and this in turn provides a source of learning. Small societies can, of course, choose to replicate the same degree of decentralization, but the increased opportunities for learning come at the expense of forgone scale economies in provision. India, with its federal structure, has clearly had a very wide range of experimentation, some states pioneering in the provision of social services and others in encouraging foreign investment. Indeed, the equivalent of learning can occur even if successful experiments are not copied. People and firms will choose to relocate to attractive areas and this gradually shifts the weighted average of policies across the nation towards the most successful. Potentially, Africa's equivalent is that because it is divided into so many nations it enables policy variation at the national level. However, it is not clear that a small country is in any better position to learn *from other nations* than is a large country. India clearly learned a lot from China.

A related scale effect is the switch from discretion to rules in decision taking. At its best an intimate organization can function by tailoring each decision to the needs of the individual and the circumstance: decisions can be personalized. As the organization becomes larger this style of decision-taking breaks down because micro-management becomes overburdened, and is replaced by rule-based procedures. Rule-based decisions can seldom be as good as first-best discretionary decisions, but they are far better than either patronage-driven or idiosyncratic decisions. They also enable the government to have a credible commitment technology which may even dominate the best discretionary policies by providing an escape from the time consistency problem. Hence, we might expect that public decisions in large societies lie within a narrower range than those in small societies. This is closely analogous to the difference between autocracy and democracy. Autocrats have the discretion either to be very good or very bad, whereas democracies are rule-bound. Besley and

Kamatsu (2006) have recently compared the economic outcomes for these two types of government and indeed find that democracy truncates the distribution at both extremes.

While the above effects of scale are plausible, is there any evidence that they actually matter? Chauvet and Collier (2008) analyze the preconditions for policy turnaround in countries which initially have very poor economic policies and governance. They measure such turnarounds using the World Bank measure of economic policy and governance, the Country Policy and Institutional Assessment. This is a subjective, ordinal rating and so has some obvious drawbacks. However, most of the controversies in economic policy concern the higher range of these ratings: policies and governance that are very poor are often unmistakable. They define a turnaround as the passage from below one very low threshold to above one that while substantially higher is still quite modest relative to the rating of most developing countries. Their universe is all low-income countries over the period 1977-2005. From this universe they select those countries which had a period of at least four consecutive years during which economic policies and governance were below the threshold. They find that population size is one of the preconditions for a turnaround that is statistically significant: the larger the population the higher is the probability that a country that is initially below the threshold will achieve a sustained turnaround. Equivalently expressed, large countries appear to have more rapid social learning.

For this paper we have used the coefficients of the Chauvet-Collier model to estimate how long India, on the one hand, and the average African country on the other, would have taken to reform from a common initial position of poor policies to a common improved position. We create two artificial countries, 'India' and 'Zambia'. Both are identical in all characteristics other than their populations: indeed, all the other characteristics are set at the sample mean. 'India' has the population of India and 'Zambia' has the population of the average African country, both entered into the regression as logs. We set their initial CPIA score equal at 2.5, indicating very poor policies and pose the question how long would it take to reach a score of 3.5. Since the regression model is log-linear the consequence of the huge differences in population is likely to be exaggerated: the regression line is trying to explain the countries within the range and is likely to fail at the extremes. Nevertheless, the predicted difference in the pace of reform is striking: 'India' is predicted to bounce

out of bad policies into reasonable ones in only six years, whereas ‘Zambia’ is predicted to take around sixty years.<sup>5</sup>

### ***4.3 Infrastructure***

Transport and power infrastructure are public goods with such strong scale economies that the typical African polity is too small to exhaust them. Indeed, Africa is still dependent upon the transport infrastructure created during the period when its present polities were united into a few empires. Quite possibly, from the perspective of transport infrastructure the key feature of colonialism was not that the empires were ruled externally but that they temporarily united Africa into a few large polities. The most obvious problem generated by political division is the many countries that are landlocked. As shown by Limao and Venables (2001), the transport costs faced by landlocked countries are strongly affected by the infrastructure spending of their coastal neighbors. Evidently, these benefits to the landlocked are externalities from the perspective of the coastal countries: they are not internalized into the decision calculus and so spending is sub-optimal. However, the failure to internalize costs and benefits extends far more widely than the plight of the landlocked. The recent discovery of large iron ore deposits in Guinea by Rio Tinto Zinc provides a telling example. The exploitation of the deposits evidently requires an investment in a mine, but the pertinent issue is the consequential investments in transport infrastructure. Fortuitously, there is already a railway linking the deposit to a deepwater port, Buchanan, this being a legacy from the age of empires. However, Buchanan is in Liberia and the government of Guinea does not want to find itself subject to the hold-up problem vis-a-vis the government of Liberia. It has therefore insisted that the transport link be entirely within Guinea. This requires the construction of a new dedicated railway, and of a new deepwater port. This decision has more than doubled the total investment needed for the project, adding around \$4 billion. Evidently, these additional costs will be fully passed on. The government has agreed with Rio Tinto Zinc that it will absorb them through a reduced flow of royalty payments. Hence, the costs are ultimately borne by the people of Guinea. The decision is also costly to the

---

<sup>5</sup> We would like to thank Lisa Chauvet for providing this calculation.

people of Liberia: in particular, the port of Buchanan loses what may have been its key opportunity for scale economies.

The generation of electric power is also more costly if the market for power is politically segmented. Not only is the generation of power subject to scale economies, but non-coincident peaks in demand cannot be pooled. The resulting volatility in demand leads to both the installation of capacity that is idle for most of the time, and to energy rationing. Energy is a fundamental input into both resource extraction and manufacturing. The recent power cuts in South Africa are reducing investment in resource extraction, and high energy costs risk making manufacturing uncompetitive.

However, potentially the highest costs of political division arise from the interactions between transport and power. The resource extraction sector is highly intensive in both, especially in Africa where mineral deposits are often far from the coast. If, as in Guinea, the ore is exported unprocessed, it has a low value-to-weight ratio and so transport costs are high. Processing would reduce weight and so transport costs, but requires very large inputs of energy. In Africa the obvious source of non-exportable energy is hydro since the key input is rainfall on high ground, something which the continent has in abundance. Potentially, this non-tradable energy can be transmitted to resource extraction sites and used to process ore which is then cheaply transported to the coast. The exploitation of such synergies may yield huge pay-offs, but they would involve huge investments. Unfortunately, almost all such opportunities in Africa involve crossing frontiers between sovereign states. This returns us to the hold-up problem discussed above, but an added dimension. The hold-up problem within a state can at least in part be addressed by law: the difficulty is that of writing a contract that is sufficiently complete to cover all eventualities. The hold-up problem between states is radically more severe because the whole domain of international law is fragile: essentially, the concept of national sovereignty constitutes a barrier to the enforcement of any contract entered into by states.

## **5. Policy Implications: the Need for Integration**

In aggregate, Africa is less populous and poorer than India, yet it is sub-divided into around 50 independent states. In this paper we have suggested that this radical political sub-division of an already small economy has inflicted a wide range of costs

on African citizens. The benefits of smoothing are foregone – smoothing with respect to both temporary shocks and underlying natural differences. In the private economy manufacturing and services have the potential for large scale economies which are frustrated by political fragmentation. This skews Africa’s comparative advantage towards those sectors and modes of production where scale is less important, most notably peasant agriculture. In the public sector the lack of scale raises the cost of a wide range of public goods and so accentuates the problem of undersupply intrinsic to low-income. These losses to the private and public economies are mutually reinforcing: low private incomes reduce state revenues and so compound the under-provision of public goods, while the lack of public goods further reduces private incomes.

The evident implication is that Africa needs a process of political integration. Such processes have been common in other regions: over the last two centuries many states have chosen to create legal structures at the regional level that curtailed their sovereignty. During the 19<sup>th</sup> century the states within the United States of America gradually shifted power from each state to the federation. This shift in the locus of decision is even detectable in language: before the American Civil War the term ‘The United States of America’ was treated as a plural noun, whereas afterwards it became singular. In the 1940s the territories of British India decided that upon Independence they would divide into only two large polities, India and Pakistan, instead of back into the many small states that had preceded colonization. During the last half-century the European Union has gradually combined 25 states which have agreed to limit sovereignty across a wide range of economic decisions. Unlike Indian politicians, African politicians chose to dissolve the federations forged by the empires: the colonial map of Africa was far less fragmented than the current configuration with Nigeria (Africa’s largest country) being the sole exception. Post-colonial political fragmentation enormously increased the opportunities for the political class, multiplying by fifty the number of ministers required for government of the territory.

African governments have launched many initiatives aimed at greater regional integration, including at the political level the Organization for African Unity and its successor the African Union, and at the economic level an array of sub-regional trade arrangements too numerous to list. Yet the practical achievement in terms of economic integration falls far short of other regions. This is brought out by a comparison of California, Maharashtra, and Germany with Burundi. California,

Maharashtra and Germany are each over a hundred times as large as Burundi viewed as economic units. Yet in terms of autonomy in fiscal policy, monetary policy, trade policy, exchange rate policy, and the scope for judicial appeal against government, the government of Burundi has radically more power than the governments of California, Maharashtra or Germany. Since Burundi is such a small economic unit its citizens have radically greater need for a supra-national authority than Californians, Mahatashtrians, or Germans, yet they have radically less resort to one.

One reason why African efforts at greater integration have to date yielded so little may be that the efforts at political and economic integration have themselves not been integrated. The approach to political integration has been by means of a Pan-African entity, the African Union, which is so diverse as to be unwieldy: for example, it includes North Africa which identifies more closely with the Middle East than with Africa, and requires agreement among 53 sovereign governments for any action.

In contrast, the approach to economic integration has been too limited, being predominantly by means of sub-regional trade deals. These deals are so numerous and uncoordinated that they are mutually incompatible: governments have signed up to commitments which simply cannot all be met. This generates confusion and undermines the credibility of trade policies. More fundamentally, sub-regional trade deals between low-income countries generate economic forces which are liable to create severe political tensions. As Venables (2003) shows, they generate economic divergence, the poorest members of the integration scheme losing relative to the least-poor members. This is in stark contrast to regional integration arrangements between high-income countries, which generate forces for convergence. Indeed, even the analysis of Venables most probably underestimates the forces of divergence unleashed by regional integration between low-income countries since it rests only on the implications of comparative advantage. The forces unleashed by the scale economies discussed in Sections 2 and 3 imply further forces for divergence, for example with those cities that are initially economically largest, such as Nairobi, Johannesburg and Lagos, gaining at the expense of initially smaller cities in their respective regions. Hence, the politics of Africa's sub-regional trade schemes are almost inevitably going to be fraught.

A more promising alternative would be to learn from three aspects of successful integration in other regions. First, political and economic integration should go hand-in-hand: supra-national entities have to acquire real sovereignty over

particular domains of economic activity. For example, currently, despite the many sub-regional trade arrangements, each African country has been negotiating individually both in the WTO and in the Economic Partnership Agreements (EPAs) with the European Union. In contrast, the trade policies of each member country of the European Union are genuinely locked in common and so negotiation at the WTO and for the EPAs is done by the European Commission rather than by each member country individually.

Second, integration is easier if it grows from a small core of states of which none is too dominant. In Europe integration started with just six countries and has progressively grown to 25. In the USA union grew from a core of 13 founding states to its present total of 51. In Africa the most promising such nucleus is the East African Community (EAC), which has a core of three similarly sized states and has already added two new members. The EAC also has the advantage that it is building political institutions which might acquire real sovereignty over some aspects of policy, alongside steps towards economic integration.

Third, the economic agenda should be considerably broader than trade policy. There is scope for common rules in a wide range of economic policies, for example, common rules on investment and taxation which would enhance credibility. There is scope for the provision of common infrastructure: the East African Community used to run an integrated rail system, and power generation and distribution would be better handled at the sub-regional level. Unlike trade agreements, these other areas of policy cooperation are much more likely to generate mutual gains, so that the economic consequences will reinforce rather than undermine the political process.

Finally, we might note that the political stresses produced by poverty have tended to be perverse. Whereas what is needed is a stronger impetus towards unity, poverty is a fertile breeding ground for xenophobia and division. During 2008 there were riots in Johannesburg against immigration from Zimbabwe and *de facto* ethnic partition in Kenya. The vision of African Unity, pioneered by Nkrumah and Nyerere, is in need of serious revival.

## References:

- Alcala, F. and A. Ciccone (2003) 'Trade, extent of the market, and economic growth; 1960-66', unpublished.
- Alesina, A., E. Spolaore and R. Wacziarg (2004) 'Trade, growth and the size of countries' in P. Aghion and S. Durlauf (eds) *Handbook of Economic Growth*, North Holland, Amsterdam.
- Au, C-C and J.V. Henderson (2006) 'Are Chinese Cities Too Small?' *Review of Economic Studies*, 73, 549-576.
- Barrett, S. 2007, *Why Cooperate?* Oxford University Press.
- Besley, T. and M. Kudamatsu, 2007, 'Making Autocracy Work', CEPR discussion paper no. 6371.
- Caselli, F. and J. Feyrer, (2007) 'The Marginal Product of Capital' *Quarterly Journal of Economics*.
- Chauvet, L. and P. Collier, 2008, What are the Preconditions for Turnarounds in Failing States? *Journal of Conflict Management and Peace Science*, 25,332-348.
- Collier, P., 2009, *Wars, Guns and Votes: Democracy in Dangerous Places*, HarperCollins, New York.
- Collier, P. and B. Goderis, forthcoming, 'Does Aid Mitigate External Shocks?' *Review of Development Economics* (forthcoming).
- \_\_\_\_\_, 2008a, 'Structural Policies for Shock-Prone Commodity Exporters', mimeo, CSAE, Oxford.
- \_\_\_\_\_, 2008b, Commodity Prices, Growth, and the Natural Resource Curse, mimeo, CSAE, Oxford.
- Cook, L.D. and J. Sachs (1999) 'Regional public goods in international assistance' in I. Kaul, I. Grunberg and M. Stern (eds) *Global public goods; international cooperation in the 21<sup>st</sup> century*, UNDP, OUP, Oxford
- Gabaix, X. and Ioannides, Y. (2004), 'The evolution of city size distributions', in *Handbook of Urban and Regional Economics*, Vol. 4, Henderson, V. and Thisse, J. F., eds. (NorthHolland, Amsterdam, 2004)
- Limao, N. and A.J. Venables (2001) "Infrastructure, geographical disadvantage, transport costs and trade", *World Bank Economic Review*, 15, 451-479.
- Rose, A.K. (2006) 'Size really doesn't matter; in search of a national scale effect', processed UC Berkeley.

Rosenthal, S.S. and W.C. Strange (2004) 'Evidence on the nature and sources of agglomeration economies', in *Handbook of Urban and Regional Economics*, Vol. 4, Henderson, V. and Thisse, J. F., eds. (NorthHolland, Amsterdam, 2004)

Teravaninthorn, S. and Raballand, G. (2008), *Transport prices and cost in Africa: A review of the Main International Corridors*, World Bank: Washington D.C.

Venables, A.J. (2003) 'Winners and losers from regional integration agreements', *Economic Journal*, 113, 747-761.

Wigstrom, C., 2008, 'Does Size Matter?' M.Phil. Thesis , Oxford University.