

Priority Roads: the Political Economy of Africa's Interior-to- Coast Roads

*Roberto Bonfatti*¹

*Yuan Gu*²

*Steven (S.) Poelhekke*³

¹ University of Padova, Italy; University of Nottingham, UK

² Vrije Universiteit Amsterdam, The Netherlands

³ University of Auckland, New Zealand; Vrije Universiteit Amsterdam, The Netherlands

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

Priority Roads: the Political Economy of Africa's Interior-to-Coast Roads*

Roberto Bonfatti^{†a}, Yuan Gu^{‡b}, and Steven Poelhekke^{§c}

^aUniversity of Padova and University of Nottingham

^bVrije Universiteit Amsterdam, Tinbergen Institute

^cUniversity of Auckland, Vrije Universiteit Amsterdam and Tinbergen Institute

January 22, 2019

Abstract

Africa's interior-to-coast roads are well suited to export natural resources, but not to support regional trade. Are they the optimal response to geography and comparative advantage, or the result of suboptimal political distortions? We investigate the political determinants of road paving in West Africa across the 1965-2012 period. Controlling for geography and the endogeneity of democratization, we show that autocracies tend to connect natural resource deposits to ports, while the networks expanded in a less interior-to-coast way in periods of democracy. This result suggests that Africa's interior-to-coast roads are at least in part the result of suboptimal political distortions.

JEL codes: P16, P26, D72, H54, O18, Q32

Keywords: political economy, democracy, infrastructure, natural resources, development

*Bonfatti and Poelhekke gratefully acknowledge financial support from the International Growth Centre (IGC Cities Funded Project Nr 39304). We also thank Robin Burgess, Gabriel Felbermayr, Guy Michaels, Ameet Morjaria, Oliver Morrissey, Alex Trew for useful feedback, as well as seminar participants at the World Bank and the Universities of Nottingham, Tilburg, Rome Tor Vergata, St Andrews, Vrije Universiteit Amsterdam and Wageningen, and conference participants at ETSG and CSAE. All remaining errors are our own.

[†]Email: roberto.bonfatti@unipd.it. Also affiliated with CESifo, Munich.

[‡]Email: y.gu@vu.nl.

[§]Corresponding author. University of Auckland, Owen G. Glenn Building, 12 Grafton Rd., Auckland 1010, New Zealand. Email: steven.poelhekke@auckland.ac.nz. Also affiliated with CESifo, Munich.

1 Introduction

According to the institutional view of development, inclusive political institutions are required for sustained economic growth, as only these will deliver the public goods and market-supporting infrastructure that make growth possible. In contrast, extractive political institutions will focus on enriching the ruling elite and on making it hard for other groups to emerge (e.g. Acemoglu and Robinson, 2012). Indeed, Acemoglu et al. (2018) find that democracy has a positive effect on GDP per capita. One specific type of market-supporting infrastructure is the road network. In this paper, we empirically investigate the link between political institutions and the construction of the network of paved roads.¹

We focus on West Africa in the post-colonisation period (1965-2012). One striking fact about this region is that, following up on a colonial pattern, paved roads have evolved mostly in an *interior-to-coast* direction (see Figures 1 and 2),² which makes them potentially useful to export natural resources but relatively ill-suited to promote internal and regional connectivity. Although the shape of these roads has long been criticised by policy makers and development economists (e.g. Nkrumah, 1963, p. 23; Rodney, 1982, p. 209; Sachs et al., 2004, p. 182), a conclusive proof that they are indeed suboptimal has been hard to produce. This is because the complexity of the optimal network design problem together with limited data availability for this region make it hard to identify the optimal network that these networks should be compared with. Then, it is hard to rule out that they are not simply the optimal response to the geography or comparative advantage of the West African countries.³

[Figures 1 and 2 about here]

To make progress, we take an indirect approach and investigate the political circumstances under which these networks were built. Controlling for country fixed effects and thus for geography,

¹We focus on roads and not on railways because railways are in a derelict state due to poor maintenance, and almost none have been extended since independence, thus giving us no variation to analyze. Moreover, railways tend to have roads near them that run in parallel: based on Vmap0 (www.mapability.com), 86% of used and unused rail kilometers in Africa connect to roads, while at the median the remaining 14% have a road within 1 kilometer. See also Jedwab and Moradi (2016).

²In most countries in this region, paved roads tend to run from the interior to the coast. Very few roads run in the direction of connecting two interior-to-coast roads, two overland borders, or two neighboring countries (particularly after excluding the interior-to-coast roads of the landlocked countries).

³It is interesting to note that when one draws the *entire* road networks of these countries (that is including all unpaved roads and tracks), these do *not* look predominantly interior-to-coast. This casts doubt on the possibility that geography is the driving factor behind the interior-to-coast shape of the *paved* road network, though of course geography could matter differently for paved roads and lower quality roads.

we document that the West African paved road networks expanded in a more interior-to-coast fashion in periods of autocracy, relative to periods of democracy. In particular, keeping fixed the number of kilometers paved, autocracies relative to democracies focused more on connecting metal and mineral deposits to ports, and less on connecting cities to cities (though the latter result is less strong). The largest cities tend to be on or near the coast, and metal and mineral deposits tend to be scattered and thus more likely to be located in a country’s interior. Combined with the region’s history of frequent autocratic rule, this empirical result goes towards explaining the observed interior-to-coast shape of the network. These findings are robust to controlling for factors that may drive changes in both political institutions and comparative advantage over time (such as time fixed effects, the world prices of the specific metals and minerals that each country exports, and civil conflict) and to instrumenting for democracy as suggested by the most recent literature on democracy and growth (Acemoglu et al., 2018).

Seen through the lens of the institutional view, these results have a natural interpretation. Deposit-to-port roads will generate resource rents: since these are more easily captured by the ruling elite under autocracy than under democracy, one would expect autocracy to be more biased in favour of these roads. Conversely, city-to-city roads may favour more widespread economic development. But since the elite may have a lesser need to secure widespread consensus under autocracy,⁴ and may even fear more the political consequences of economic development, one would expect autocracy to be biased against such roads. In this interpretation, the interior-to-coast expansion of West African roads since independence was at least in part the suboptimal consequence of extractive political institutions.

Uncovering the political determinants of African infrastructure, as this paper tries to do, is important for a number of reasons. First, transport infrastructure is an important item of expenditure for many African governments and donors.⁵ Second, internal trade costs remain very high in Africa,⁶ suggesting that a great expansion of the African networks will be required

⁴For example, Padró i Miquel (2007) and Burgess et al. (2015) argue that, in ethnically fragmented Africa, autocrats were often able to survive by only securing the support of their own ethnic groups, which is typically concentrated in a small part of the country. This is not to deny that democracy may also neglect less political influential regions, as shown by Blimpo et al. (2013) for the case of Ghana, Senegal, Mali and Benin.

⁵Jedwab and Storeygard (2018, p. 3) report that transport accounted for 14% of World Bank lending, and 22% of African Development Bank disbursements in 2012-2015.

⁶Atkin and Donaldson (2015) find that internal trade costs in Ethiopia and Nigeria are four to five times larger than in the US. Porteus (2017) finds slightly higher values for a sample covering all 42 Sub-Saharan African countries.

in the future.⁷ Third, there are conflicting views on the type of links that such expansion should prioritise. On the one hand, the African Development Bank, backed by Western donors, has been calling for more interior-to-interior links of the kind that Africa most sorely misses.⁸ On the other hand, China has been funding projects that essentially reinforce the existing interior-to-coast links (Bonfatti and Poelhekke, 2017, p. 105), often in cooperation with governments of doubtful democratic credentials. Finally, these decisions are likely to have large welfare consequences. For example, it is widely accepted that high internal trade costs are a major obstacle to the modernisation of African agriculture (Blimpo et al., 2013, p. 62), whose low productivity and high labour share are found to explain most of the income gap between sub-Saharan Africa and the rest of the world (Porteus, 2017).

Our analysis is based on a newly-assembled dataset of roads, cities, metal and mineral deposits, and ports, covering 12 West African countries over the period 1965-2012. The data on roads was obtained by digitising 18 successive editions of Michelin road maps of West Africa. For each year of publication, we have information on the full network of paved and unpaved roads, so that we can track the expansion of the paved road network. Rather than working with gridded data, we take into account the links between roads and digitally recreate the complete network. The data on deposits, which we put together from several different sources, gives us the location and year of discovery of 391 metal and mineral deposits in West Africa, of which we also know the size and main metal or mineral extracted.

Our first challenge is to measure the extent to which paving in any given period displayed a *relative city bias* - defined as the extent to which it focused on connecting cities to one another, as opposed to connecting deposits to ports. We proceed in two steps. In the first step, for each country in our sample, we construct two counterfactual road expansion paths, both inspired by the work of Burgess et al. (2015) on Kenya. The first counterfactual aims to approximate the locations where unpaved roads would have been paved from period to period between 1965 and 2012, had the government only cared about connecting cities to one another (both domestic

⁷More infrastructure may spread growth more evenly as it facilitates the propagation of shocks (Amarasinghe et al., 2018).

⁸The Bank has been pushing for the completion of the Trans-African Highway network, a large part of which is made up of interior-to-interior links. This approach is backed up by Buys, Deichman and Wheeler (2010), who argue that to connect the major cities of Sub-Saharan Africa through primary roads would boost trade by US\$ 250bn at a relatively little cost. More recently, Porteus (2018) shows, using an estimated model of agricultural trade, that to reduce African trade costs to international standards along the Trans-African Highway network links would achieve a significant portion of the welfare gains achievable by reducing trade costs on a much larger set of links.

cities and foreign cities close to the country’s borders). More specifically, we take all city pairs that are not yet connected by paved roads, and rank them according to their market potential (sum of population size divided by bilateral distance). In each period, we then re-allocate the actually paved kilometers to connecting the city pairs that are at the top of the ranking. The second counterfactual approximates the location where unpaved roads would have been paved, had the government only cared about connecting deposits to ports in order to facilitate the mining of these deposits and the export of its metals and minerals. Here we rank deposit-port pairs by port-size and deposit-size and start paving from deep ports to nearby large deposits. The two counterfactuals turn out to be substantially different, and thus represent two very different approaches to road paving.

In the second step, we calculate the extent to which real paving in any given period overlapped with each of the two counterfactual networks. We then take the difference between the two measures of overlap. The resulting measure is an estimate of the planner’s relative city bias in that period, because a high score denotes a type of paving that focused a lot on connecting cities to one another, and very little on connecting deposits to ports.

We regress our measure of relative city bias in any given period on the level of democracy at the beginning of the period, controlling for country fixed effects, time fixed effects, time-specific trends, and a rich set of controls. Our main result is that, controlling for the number of kilometers paved in any given period, democracies display a stronger relative city bias. This effect is large, with a one-standard deviation increase in democracy resulting in half-standard deviation increase in relative city bias. The effect is robust to instrumenting for democracy using a lagged measure of democracy in a country’s region, as suggested by the most recent literature on democracy and growth (Acemoglu et al., 2018).

When we unpack our measure of relative city bias into its two component parts (overlap with the city-to-city counterfactual and overlap with the deposit-to-port counterfactual) we find that this is mostly driven by democracies paving fewer deposit-to-port kilometers than autocracies. Democracies do appear to also pave more city-to-city kilometers than autocracies, but only with some delay, which could be due to democracies being slower to get paving decision implemented.

Our results are robust to measuring democracy with the dummy variable proposed by Acemoglu et al. (2018), to controlling for ethnic favouritism in the allocation of paved roads, to calculating the counterfactual network using some of the alternative approaches proposed by

Burgess et al. (2015), and to using different definitions of the instrument.

One noteworthy aspect of our approach is that although we have data at the level of individual locations (cities and deposits), we then aggregate this up so that our unit of observation in the regressions is a country-period. This approach is adopted out of necessity, because of two reasons. First, any regression that had cities and deposits as units of observations would run into the problem that these are fundamentally different types of locations, making it hard to define control variables that are common to both. Secondly, and perhaps more importantly, our goal is to investigate how the *shape* of the national road networks evolved over time (specifically whether they evolved in an interior-to-coast versus interior-to-interior fashion), and not just whether it was cities or deposits that were connected. Clearly, the shape of a national road network is best investigated at the country level.

The paper is most closely related to a small but growing literature on the political determinants of transport infrastructure. Burgess et al. (2015) study ethnic favouritism in the allocation of Kenyan roads in 1963-2011. They find strong evidence of favouritism - implying a departure from an “ideal” city-to-city counterfactual constructed as described above - which however attenuates under democracy. For a sample covering 39 Sub-Saharan African countries over the same period, Jedwab and Storeygard (2018) find that democracies tend to pave more roads close to big cities.⁹ This is consistent with the findings by Burgess et al. (2015), since favoured ethnic groups will not necessarily be located close to the big cities. We weakly confirm these results on democracy’s “closeness to cities” for West Africa, though we establish it using a different, network-based measure of closeness. However, our main contribution is to document the bias of autocracies in favour of connecting deposits to ports, and relate our findings to the debate on the interior-to-coast shape of the African networks.

Alder and Kondo (2018) obtain results for China which resemble those of Burgess et al. (2015). They compare the expansion of Chinese motorways in 1992-2007 to an optimal network calculated by connecting the country’s largest cities through a heuristic algorithm. They find that a significant portion of the discrepancy between the two networks is due to an especially high weight being given to the birthplaces of officials in power at the time of expansion. Finally, Felbermayr and Tarasov (2015) derive optimal investment on transport infrastructure in a con-

⁹They also find that democracies tend to pave more roads. Inspired by this, and to avoid any mechanical effect of extra paving on our measure of overlap, we always control for the number of kilometers paved in each period.

tinuum of locations disposed along a line, and show (theoretically and empirically, using data on European roads) that if neighboring governments act non-co-operatively, then the equilibrium features too little investment close to the borders. While this may provide an additional explanation for the scarcity of *inter*-national interior-to-interior links in Africa,¹⁰ it cannot easily explain the scarcity of *intra*-national interior-to-interior links, nor the pattern of changes over time that we document in this paper.

The paper builds on earlier work by Bonfatti and Poelhekke (2017). They show that coastal African countries endowed with more mine-to-port roads also feature national trade costs that are strongly biased in favour of trade with overseas countries, relative to neighboring countries.¹¹ This effect is reversed for landlocked countries, presumably because the mine-to-port roads can also be used to trade with transit neighbors. In conjunction with those earlier results, the results in the present paper suggest that the bias in road investment decisions associated with extractive political institutions will have a large and long-lasting effect on the direction of trade in the African countries, and specifically one that penalises regional integration.¹²

The paper is structured as follows. Section 2 presents the basic empirical approach, and Section 3 lists the data sources. Section 4 describes the construction of the relative city bias measure, our measure of bias towards connecting cities to one another (as opposed to deposits to ports). Section 5 presents the main results, with robustness checks being provided in Section 6 and in the Online Appendix. Section 7 concludes.

¹⁰Note however that Porteus (2017) finds that cross-border trade costs in Africa are not correlated with differences in colonial heritage. This seems to suggest that, at least in colonial times, the scarcity of inter-national interior-to-interior links was not driven by a lack of co-ordination between neighboring governments. In a similar vein, Bonfatti and Poelhekke (2017) find no evidence that pairs of neighboring African countries sharing the same coloniser trade more with each other than other pairs of neighboring African countries.

¹¹Not observing trade costs directly, they infer this from observed import flows. Three facts suggest that what they identify is trade costs, and not simply the comparative advantage of metal- and mineral-endowed countries for trading with rich countries that tend to be located overseas. First, they show that the effect is opposite for coastal and landlocked countries. Second, they control for the difference in income between trading countries. And third, they show that metal- and mineral-endowed coastal countries also import more from overseas *African* countries.

¹²Our results are that, controlling for kilometers paved, autocracy builds more deposit-to-port roads relative to city-to-city roads, which by the earlier results would mean national trade costs that are more biased in favour of overseas countries relative to neighbouring countries (recall that, unlike deposit-to-port roads, city-to-city roads are allowed to connect locations in neighboring countries). In unreported regressions, we find no evidence that autocracies pave more kilometers. This suggests that autocracy is bad for city-to-city roads (and thus potentially for the cost of trading with neighboring countries) not only in relative terms but also in absolute terms.

2 Empirical approach

Let $city\ bias_{i,t}$ denote our measure of *relative city bias* - defined as the bias towards connecting cities to one other, as opposed to deposits to ports - in country i and period t (full details on how this measure is constructed will be provided in Section 4). Our hypothesis is that relative city bias will be more pronounced in periods of democracy, compared to periods of autocracy.

To test for this hypothesis, we regress $city\ bias_{i,t}$ on a measure of the state of democracy in country i at the beginning of period t , $D_{i,t}$:

$$city\ bias_{i,t} = \beta D_{i,t} + \gamma X_{i,t} + f_i + f_t + trend_i + \epsilon_{it}, \quad (1)$$

where we always include country fixed effects f_i , period fixed effects f_t , and country-specific trends $trend_i$, and $X_{i,t}$ denotes a vector of time-varying, country-level controls.

Our choice of periods is dictated by the years of publication of the Michelin maps of West Africa used to assess road paving. Such maps are not published every year, although for all years in which they are published they cover all countries in our sample. We observe 18 maps published between 1965 and 2012, which results in 17 irregular time periods with an average length of 2.76 years (and a minimum and maximum length of 1 and 6 years). Since Michelin maps are published at the beginning of a year and are supposed to capture the situation on the ground shortly before publication, we take our periods to start at the beginning of each publication year. For example, our first two road maps were published in 1965 and 1968: our first period then goes from January 1, 1965 till December 31, 1967.

Our dependent variable, $city\ bias_{i,t}$, is constructed using all new paving that occurred in period t . We infer the period's paving by comparing maps published at the beginning and end of the period. For example, if a road shows up as paved in the 1968 edition but unpaved in the 1965 edition, then we conclude that it was paved sometime between January 1, 1965 and December 31, 1967 (though we do not know exactly when).

Our main measure of democracy will be the Polity IV index, which is available yearly. Since $D_{i,t}$ is supposed to measure the state of democracy at the beginning of period t , and the Polity IV index measures the state of democracy on December 31 of each year (see Marshall et al., 2017, p. 12), we construct $D_{i,t}$ using the value of the Polity IV score the last year before period t . For example, for the period that goes from January 1, 1965 till December 31, 1967, $D_{i,t}$ is

constructed using the 1964 Polity IV score, which measures the state of democracy on December 31, 1964.

The vector $X_{i,t}$ contains variables which may affect road paving decisions and correlate with democracy. It includes stock variables measured at the beginning of each period (e.g. a country's stock of deposits), flow variables measured period by period (e.g. current paving, deposit discoveries), and flow variables measured yearly (commodity prices, foreign aid, FDI, and an indicator for civil war) for which we take the average over the period. To avoid confusion, the variables' names reported in the regression tables will always clearly specify the timing of measurement of each variable.

It is worth pointing out that the specification in (1) is quite demanding. Any time invariant country characteristic that may make it natural for a country to have a more or less interior-to-coast road network - be it geography, the time-invariant determinants of comparative advantage, colonial heritage, etc - will be absorbed by the country fixed effects. Any common time process that may matter for the way in which the shape of the networks evolve - e.g. the passing of time since independence, the end of the Cold War and the subsequent wave of democratisation, technological progress, price changes that apply to all metals and minerals - will be captured by the period fixed effects. Any mechanical effect driven by the fact that our periods are of irregular length will also be captured by the period fixed effects. Finally, any country-specific, linear time process - such as those driven by population or economic growth, or by price trends in the specific metals and minerals produced by each country - will be accounted for by the country-specific trends. We will, in summary, only be using the within-country variation which is left after removing all of these time processes, which we will then further discipline by adding a variety of country-specific, time varying controls.

To try and address any remaining endogeneity concerns, we will also instrument for democracy using the lagged regional average of democracy, as suggested by Acemoglu et al. (2018).

3 Data and descriptives

We construct a panel covering 12 West African countries over 17 time periods between 1965 and 2012, although not all variables are available for the entire sample. The sample includes Benin, Burkina Faso, Côte d'Ivoire, Ghana, Guinea, Guinea Bissau, Liberia, Mali, Niger, Senegal, Sierra

Leone and Togo (we are forced to drop Gambia due to the fact that it does not have any metal or mineral deposits). We require data on democracy, detailed road construction over time and space, the location, size and type of deposits, the location and size of cities and ports, the world prices of metals and minerals, foreign aid, FDI flows, periods of war, and a delineation of ethnic lands and the ethnicity of political leaders. We describe sources in detail in this section, and provide a table of summary statistics for the regression sample in Appendix Table A1.

3.1 Democracy

Our main measure of democracy is the widely used Polity IV score, which ranges from -10 (autocracy) to +10 (democracy). Figure 3 shows that there is a lot of variation between countries and within countries. Indeed, for most of the countries in our sample, the within-country variation is similar in magnitude to that of Kenya, the country for which Burgess et al. (2015) find a strong effect of democratisation on the allocation of roads. Overall, most countries in our sample are relatively autocratic up to the 2000s, with some countries (e.g. Ghana, Burkina Faso and Sierra Leone) also experiencing periods of democracy shortly after decolonisation. In robustness checks, we also employ a dichotomous measure of democracy, constructed by Acemoglu et al. (2018) by combining the Polity IV index with several other independent measures of democracy. According to this definition, only one of our 12 countries was a democracy in 1965, but nine countries had become democratic by 2012. Our sample thus reveals the same unprecedented spread of democracy as highlighted in Acemoglu et al. (2018). The use of multiple sources for the dummy potentially reduces measurement error but may come at the cost of reduced variation because half of the countries in our sample only change status once.

[Figure 3 about here]

3.2 Roads

We have 18 Michelin West Africa maps covering our 12 countries from 1965 to 2012. The map editions are 1965, 1968, 1969, 1971, 1973, 1976, 1983, 1984, 1986, 1989, 1990, 1996, 1998, 2002, 2003, 2007, 2009, and 2012.¹³ The maps show colour- and pattern-coded roads, which we digitize

¹³Although we also have some earlier years, we start in 1965 because Michelin changed the legend of their maps in that year. While we do not know for certain that no map was issued for the intermittent years, the

using the ArcGIS platform. To raise the resolution of the maps from 1:60,000 to 1:30,000 we start with a world transport map from the ArcGIS public online database and track the status of roads forwards and backwards in time, and update the base map where necessary. To allow for network analysis (see further down), we also make sure that all road segments connect.

Roads are classified along two dimensions: their regional importance in terms of major, secondary and local roads, and their quality. The road quality classification is broken down into six categories: “surfaced”, “improved”, “partially improved”, “earth roads”, “tracks”, and “others”. According to the Michelin legend, surfaced roads are paved with asphalt and/or concrete and are suitable for all-weather conditions and vehicles. In contrast, improved roads are unpaved even though they receive regular mechanical maintenance, and are only suitable for high speeds in certain sections. The four types of worse-quality roads also are unpaved and not all-weather roads. In our analysis, we focus on the quality classification and use all roads (major, secondary and local) changing status between unpaved (equal to any of the five lower quality categories) and paved (equal to “surfaced” and thus all-weather roads). However, we use the location of all unpaved roads that exist as of 2012 as the location of segments of road that may or may not be paved by the government in any earlier year.

In Table 1 we track the distribution of roads in West Africa by quality, in 1965 and in 2012. In the table, the five categories of unpaved roads are grouped in the following way: Unpaved (medium quality) = “improved” and “partially improved”; Unpaved (lowest quality) = “earth roads”, “tracks” and “others”. We see that roads are both upgrading and downgrading over time. For example, of the 10,477 kilometers of roads that were paved in 1965, 9,295 kilometers, or 90%, were still paved in 2012, while the remaining 10% deteriorated to unpaved status (possibly due to lack of maintenance and/or war). Of the latter, 113 kilometers deteriorated to the point of vanishing by 2012. Note that the paving of roads where no road was previously present is relatively rare: paving normally happens to existing roads or tracks.¹⁴ Note also that compared to paved roads, unpaved roads are more prone to deterioration.

[Table 1 about here]

unequal timespan between years seems to suggest that a new map was only issued when enough changes to the road network were observed.

¹⁴Note also that, of the 2,036 kilometers of roads which appear as paved in 2012 but do not appear at all in 1965, many (if not most) must have first come into existence as non-paved tracks or roads at some point between 1965 and 2012, before becoming paved roads by 2012.

Figure 4 shows for each country the evolution of the stock of paved roads since 1965, expressed in kilometers per surface area. There is substantial heterogeneity, with some countries showing almost no activity (Liberia, Mali, Niger) and others making a lot of progress. However, in Online Appendix Tables OA1 and OA2 we show details of the distribution by country and period, which shows that some countries paved more than Figure 4 suggests even though it was low compared to their surface area. The stock of paved roads is in general increasing over time. For example, Côte d’Ivoire increased its stock almost six-fold from 900 kilometers to 5,200 km. Similarly, Senegal, Niger and Mali also paved many roads, while Ghana, Sierra Leone, Guinea-Bissau, and Liberia paved relatively few.

[Figure 4 about here]

3.3 Deposits

Our deposit data is the combination of three sources: the proprietary data sets by MinEx Consulting and SNL Mining (formerly known as RMG), and the publicly available data from the United States Geological Survey (the Mineral Resources Data System or MRDS). Combined, these sources provide us with 391 records of West African deposits, for which we know the location, the year of discovery and the size of the deposit. We start from MinEx, which is a dataset of deposits and has the best coverage in terms of year of discovery.¹⁵ We then update missing dates using MRDS and RMG, which are datasets of mines, the distinction being that a deposit is a concentrated occurrence of metal or mineral resources, while a mine is an industrial facility used to mine the resources. We interpret the year of discovery listed for the mine as the year of discovery of the deposit in which the mine is located.¹⁶ The size of deposits is classified by MinEx as supergiant, giant, major, moderate, and minor. We group giant and supergiant together because there is only one supergiant deposit, and construct four categories (giant=4, major=3, moderate=2, minor=1), which we use as weights in the ranking of which deposits to connect first.¹⁷

¹⁵According to MinEx, “The discovery date refers to when the deposit was recognised as having significant value. This is usually set as the date of the first economic drill intersection”.

¹⁶Per source, 242 records are from MinEx, 131 from RMG, and 18 from MRDS.

¹⁷The original delineation is not linear in the metal- and mineral-content of ore, which depends on ore grade and on the metal or mineral. For example, in MinEx a giant gold mine has > 6Moz Au and a major one > 1Moz Au (which differ by a factor six), but a giant iron ore mine contains > 1000Mt and a major one > 200Mt (which

The database include various kinds of deposits: bauxite, chromium, copper, gold, iron ore, manganese and others. We include all metals and minerals except diamond deposits, since diamonds are less likely to be transported via ports and bulk ships than to be airlifted. Table 2 lists the number of discoveries by country and period and the distribution of metals, minerals and size. Three-quarters of deposits were discovered after 1965, almost half are at least of major size, and gold is most common.

[Table 2 about here]

3.4 Cities

The data on cities was kindly shared by Hervé Gazel of the Africapolis project.¹⁸ It contains the location and population of cities with at least 10,000 inhabitants in 2010, for 33 countries and for the years 1960, 1970, 1980, 1990, 2000 and 2010. We observe 822 cities within our sample. We use the 477 cities which exist in 1960 and use their population size to rank them. Table 3 shows the population distribution of these towns. We find that 1960 population explains 85% of the variation in 2010 population, and cities that exist in 1960 make up 88% of the population of cities that exist in 2010. New cities that were created after 1960 have on average only 18,000 inhabitants and never more than 96,000.¹⁹

[Table 3 about here]

3.5 Ports

We use the 2016 World Port Index (WPI) to identify ports. It provides the location, characteristics, channel and cargo pier depth (in 12 categories), shipping facilities, and available services of ports in the World. According to Waters et al. (2000, figure 4.46), dry bulk ships built from the 1950s onwards require at least 5m of draft. We therefore keep all ports (excluding marine

differ by a factor five). For our purpose, it is sufficient that the rank order from giant to moderate is preserved. A few deposits have other size measures which come from RMG and MRDS: for reference, we group one ‘medium’ mine into moderate and one ‘small’ mine into minor. Only a subset of deposits also list the size in terms of the estimated amount of metal- or mineral-containing ore.

¹⁸Université Jean Moulin Lyon 3. See: <http://www.afd.fr/lang/en/home/publications/travaux-de-recherche/archives-anciennes-collections/NotesetEtudes/Africapolis>

¹⁹Spearman’s rank correlation test clearly rejects independence of 1960 and 2010 towns, with $\rho = 0.598$

terminals used for oil) with channel and cargo pier depth of at least 5m. This way we exclude many fishing ports that are unlikely to be used for metal or mineral exports. We then group the remaining nine categories of depth into four: a-d (the deepest ports, weight=4); e-g (the second deepest ports, weight=3); h-k (weight=2); l-n (weight=1). Deeper ports will have priority to road connection in our counterfactual network. We include all deep water ports that exist in 2016, based on the logic that even if some of them did not exist in 1965, given the geography they could have been built where they are today.

3.6 Prices, aid, FDI and war

We collected data for control variables on the world price of metal and mineral commodities, on aggregate inflows of foreign aid and foreign direct investment, and on the occurrence of war. The data on the world price of 15 metal and mineral commodities comes from UNCTAD (downloaded June 14 2017), US Geological Survey (1999), and Bazzi and Blattman (2014). We constructed a country-specific, time-varying index in which each price is weighted by the share of deposits producing that metal or mineral in a given country in the middle of our sample period (1989).²⁰ The data on aid and FDI come from the World Bank Development Indicators. Average inflows of foreign aid and FDI in our sample is 173 and 105 million US\$, respectively, with Ghana and Côte d’Ivoire typically receiving most aid and Ghana receiving most FDI. Finally, the data on war come from the Correlates of War (COW) Project. There was no war between our countries in our period, however we identify several civil wars in the Mano River Region, including Côte d’Ivoire (2002-04), Guinea (2000-01), Liberia (1989-90; 1992-95; 2002-03) and Sierra Leone (1991-00), in addition to Guinea-Bissau (1998-99).

3.7 Ethnicity

Finally, we use data from Murdock’s *Ethnolinguistic atlas* of Africa (Murdock, 1959, 1967) to determine whether each city and deposit in country i and period t is co-ethnic with the current government leader, where “co-ethnic” means located in the ethnic homeland the leader. For any city or deposit, this indicator can switch on and off over time, as the ethnicity of the leader

²⁰In other words, the only time variation in the index comes from exogenous changes in prices. We construct the index in this way out of necessity, since we do not observe the exact size of each deposit, and thus cannot infer the way in which any new discovery will affect the overall value of the country’s deposits.

changes. To track the ethnicity of political leaders, we rely on wikipedia and other web sources.

4 Construction of the relative city bias measure

To construct the dependent variable, our measure of *relative city bias* - a measure of bias towards connecting cities to one another, as opposed to connecting deposits to ports - we proceed in two steps. First, for each country in our sample, we construct two counterfactual road expansion paths. The first path aims to approximate the way in which paved roads would have expanded in 1965-2012, had the government only cared about connecting cities. The second approximates the way in which they would have expanded had the government only cared about connecting deposits to ports. Our measure of relative city bias, which we construct in the second step, will be a measure of how close the actual expansion path was to either counterfactuals. The two steps are described in sections 4.1 and 4.2 respectively, and section 4.3 provides descriptive statistics on the relative city bias measure.

4.1 Construction of the counterfactual road expansion paths

To construct our two counterfactual road expansion paths, we follow Burgess et al. (2015)’s methodology for Kenya. We begin by calculating how many kilometers of roads were paved in each country in each period (the paving “quota” for that period), by comparing the observed status of roads in two subsequent map years. We then re-allocate the quota of each period to paving different roads in that period, proceeding period after period as each counterfactual requires. In other words, we imagine that the government would follow two different expansion path to the one it actually followed, leading to two different counterfactual networks by 2012.

The first counterfactual aims to approximate the way in which paved roads would have expanded in 1965-2012, had the government only cared about connecting cities to one another. We label this the “city-to-city counterfactual”. To construct this counterfactual, we follow Burgess et al. (2015) very closely. They construct an “optimal” counterfactual expansion path by ranking city pairs according to their market potential (the sum of the cities’ size, divided by their distance along roads), and by re-allocating the paving quota in each period to yet unpaved bilateral connections in the order in which city pairs show up in the ranking (continuing with

unfinished pairs from the previous period before moving down the ranking).²¹ They also take border crossings into account, by including foreign cities “near” borders in the ranking (roads to these cities are paved up to the border). The logic behind this procedure is that it should be optimal to first connect city pairs with a greater trade potential. This procedure also implicitly takes into account international trade, since the ports through which this is channeled typically have an important role in the initial paved network (the one that the counterfactual network expands on) and, being big cities, are likely to be assigned an important role in the counterfactual expansion.²² To take the topography into account, Burgess et al. (2015) construct their counterfactual by only paving along the shortest unpaved roads existing between cities in the initial (1964) network. They can do so since all the paved roads constructed during their period were initially unpaved, and were already in existence in 1964. The initial unpaved network, then, is a reasonable indication of the set of road segments that could potentially be paved.

We proceed in a the same way, except for two differences.²³ First, we define an objective criterion for including foreign cities: that they should be located within 50km of the border. The second difference is prompted by the fact that, differently from Burgess et al. (2015), some of the paved roads constructed in our sample did not exist at all in the initial (1965) networks. Thus, to use the initial unpaved networks to construct our counterfactuals would mean throwing away some of our knowledge of the topography, since we would be constraining ourselves to not build roads in places where we know roads could be built (as evidenced by the fact that they were built at a later stage). To circumvent this problem, we add to our 1965 networks all roads (paved or unpaved) which were built by 2012. This means that our counterfactual-building procedure is allowed to pave not only over roads that were unpaved in 1965, but also over territories where a road did not exist in 1965, but where a road would be built by 2012.

Figure 5 exemplifies this procedure for the case of Sierra Leone. Cities and towns are denoted by blue dots, and the initial (1965) road network by black lines. Foreign cities that are included in the ranking are also shown, as are the unpaved roads (by thin green lines). The red lines denote

²¹When paving a connection between a pair, they start paving from the largest city in the pair.

²²All coastal countries in our study except Côte d’Ivoire have capital cities which are also major port cities: Dakar in Senegal, Bissau in Guinea Bissau, Conakry in Guinea, Freetown in Sierra Leone, Monrovia in Liberia, Accra in Ghana, Lomé in Togo and Porto-Novo in Benin. All of these cities were already connected in 1965 (see Figure 1) and are large cities which are given an important role in the counterfactual expansion (for an illustration, see the example of Sierra Leone discussed below).

²³In robustness tests we experiment with Burgess et al. (2015)’s alternative ways of ranking city pairs, which are based on size only or on distance only.

the city-to-city counterfactual by 2012: this is our approximation of how paved roads would have evolved, had the government only cared about connecting cities to one another. Note that almost all domestic cities are well connected to one another by paved roads, and there is one such road that reaches out to the overland border with Guinea. Note also that Freetown has good connections with most cities in the country, reflecting its importance in the initial network, and its relative size (both of which reflect its importance in channeling international trade).

The second counterfactual approximates the way in which paved roads would have expanded had the government only cared about connecting deposits to ports. We label this the “deposit-to-port counterfactual”. To construct it, we proceed by analogy with the previous case. We rank deposit-port pairs by market potential, and re-allocate the paving quota in each period to connecting pairs in the order in which they show up in the ranking. We start with deposits that had been discovered by 1965 and let additional deposits enter the ranking from the year that they were discovered. We thus use the timing of discovery as exogenous variation. “Market potential” of a pair is defined as the sum of deposit size and port “size” (depth), divided by the distance along roads between them.²⁴ Large deposits are thus connected to deep ports first, unless a deposit is very remote compared to smaller ones. We always start paving from the port, because the port may have multiple uses so that even unfinished connections may be useful if they start from the port. Each time a new deposit is discovered, it enters the ranking at the start of the next period.²⁵ In case any quota is left over in a period (which may happen if there are few or no unconnected deposits at the beginning of the period), we drop it. This is based on the logic that a government who only cares about connecting deposits to ports would spend any surplus money on things other than paving.²⁶

The case of Sierra Leone is exemplified in Figure 6, where deposits are denoted by green asterisks, ports (including those on navigable rivers) by blue squares, and the deposit-to-port counterfactual by 2012 by red lines. It is important to note that the deposit-to-port counter-

²⁴When no road of any quality exists (not even a track) we take the shortest route from the deposit location to the nearest road. We also tried ranking by multiplying the deposit-size indicator and the port-depth indicator, but this does not change our results.

²⁵For landlocked countries, we pair each deposit to the closest port anywhere in West Africa, and pave that connection up to the country’s border.

²⁶Ades and Gleaser (1995) find that autocracies tend to have relatively larger capital cities than democracies, suggesting that more government spending is directed to the seat of government. We thus assume that unused quota is spent on the capital city rather than on road paving. In the estimation we will control for the total amount of paving in each period. Connecting cities always exhausts the paving quota because West African countries have not yet connected all cities.

factual in Figure 6 looks quite different from the city-to-city counterfactual in Figure 5. In particular, the former counterfactual features several interior-to-coast roads which are unconnected with each other and with the country’s borders, thus making it potentially more costly to travel in a North-South direction within Sierra Leone and to cross into neighboring Guinea in the North. Instead, the city-to-city counterfactual, with its more numerous North-South links, would make it potentially less costly to travel within Sierra Leone, and to cross into its neighbors. Of course the city-to-city counterfactual is still of limited length, because the simple re-allocation of roads cannot by itself tackle the issue that there are too few of them.

[Figures 5 and 6 about here]

Two important details in the construction of the counterfactuals must be flagged up. The first is prompted by the fact that in our sample we observe deterioration as well as upgrading of roads. In particular, it happens in 44% of country-periods that the quality of at least one segment of road deteriorates. At the median, 3% of the road network becomes unpaved in such an event. This may be the result of lack of maintenance, or of traumatic events such as civil war. In the presence of past deterioration, actual paving in any given period may involve roads that have never been paved, or roads that have already been paved in the past but have since deteriorated to unpaved status. If the two types of paving carry a different cost, then one should ideally give them different weights in calculating the quota used in the construction of the counterfactuals. Instead, we give them the same weight, implying that the construction of the counterfactuals does not distinguish between them. We do this out of necessity, since we do not observe the different cost of paving depending on whether a road was ever paved in the past.

Secondly, when constructing the counterfactuals, we only re-allocate the resources used to pave the road network (which we observe), and not those used to maintain it (which is not observed). This means that two periods with the same number of new kilometers paved, but with different number of kilometers deteriorated, are treated in the exact same way. Furthermore, when reallocating the paving quota, we do not let the counterfactual planner use the resources that are in reality used to pave new roads, to maintain the current counterfactual network instead. In fact, we simply do not let the counterfactual networks deteriorate over time. Again we do all this out of necessity, as we do not have a good proxy for the resources used to maintain roads, and we do not know how the cost of paving compares to the cost of maintaining a paved road

(though we can speculate that the latter will be an order of magnitude smaller). The implication is that the counterfactual networks are often longer than the actual networks towards the end of our period.

4.2 The relative city bias measure

Our measure of relative city bias - a measure of bias towards connecting cities to one another, as opposed to connecting deposits to ports - is calculated as follows:

$$city\ bias_{i,t} = city\ overlap_{i,t} - deposit\ overlap_{i,t} \quad (2)$$

where $city\ overlap_{i,t}$ and $deposit\ overlap_{i,t}$ measure the fraction of actual paving in country i and period t which overlaps with the city-to-city and deposit-to-port counterfactual networks for that country, as they would have appeared at the end of period t . These overlap measures range between 0 and 1, so that the relative city bias measure ranges between -1 and 1 . A score of 1 denotes the most extreme bias towards connecting cities to one another (as opposed to deposits to ports), while a score of -1 denotes the opposite.

The two overlap measures are calculated as follows. We only illustrate this for the case of $deposit\ overlap_{i,t}$, since the procedure to construct $city\ overlap_{i,t}$ is identical. Also, we omit the country subscript i for the rest of this section, but it should be clear that the same procedure applies to all countries in our sample. Let $A_{t,real}$ denote the real road network in place by the end of period t , and $A_{t,cf}$ the counterfactual deposit-to-port network as it would have appeared by the end of period t . The actual paving that occurred in period t is represented in set notation by $A_{t,real} \setminus A_{t-1,real}$, that is the difference between the actual road networks at the end of periods t and $t - 1$. Therefore, $(A_{t,real} \setminus A_{t-1,real}) \cap A_{t,cf}$ denotes the overlap between the actual paving in period t , and the counterfactual deposit-to-port network at the end of period t .²⁷ Dividing this overlap by $A_{t,real} \setminus A_{t-1,real}$ gives us the fraction of the actual paving in period t that overlaps with the deposit-to-port counterfactual at the end of period t , or our overlap measure:

$$deposit\ overlap_t = \frac{(A_{t,real} \setminus A_{t-1,real}) \cap A_{t,cf}}{A_{t,real} \setminus A_{t-1,real}}. \quad (3)$$

²⁷Obviously, such overlap can only include parts of the counterfactual network that had not been already paved at the start of period t .

To illustrate the method intuitively, we use Figure 7. Let area $A_{0,real} = a$ represent the initial (1965) paved road network. Assume that by the end of period 1 (Dec 31, 1967) the real network was extended towards the south-west, to cover area $A_{1,real} = a + b + c + d$. The actual paving in period 1 is thus $b + c + d$, which provides us with the paving quota for that period. Let area $A_{1,cf} = a + d + e + f$ towards the south-east represent the deposit-to-port counterfactual network as it would have appeared at the end of period 1. The reason for the discrepancy between counterfactual and actual network might be that there are deposits located in region f whose connection was not priority under the actual planner, but would have been priority under the counterfactual planner. The overlap between the actual paving in period 1 and the the counterfactual at the end of the period is then d , and *deposit overlap*₁ is equal to $d/(b + c + d)$.

[Figure 7 about here]

Suppose now that by the end of period $t = 2$ (Dec 31, 1968),²⁸ the real network was extended to cover area $A_{2,real} = a + b + c + d + e + g + h$. The actual paving in period 2 is thus $g + h + e$, which provides us with the paving quota for period 2. Under the deposit-to-port counterfactual, however, the additional paving in period 2 would have covered area $c + h + i$, such that the counterfactual network by the end of period 2 would have covered area $A_{2,cfp} = a + c + d + e + f + h + i$. At this point, we could proceed in two different ways.

On the one hand, if we considered as overlap the intersection of the actual and counterfactual *additional* paving, the overlap in period 2 would be only area h , because area e should have been paved one period earlier. However, this approach betrays the fact that area e *was* paved in period 2, and it overlaps with the deposit-to-coast counterfactual as it appeared in that period. The paving of e could well be the choice of a period 2 government who cares more about connecting deposits than did its predecessor in period 1, and wants to pave whatever is left unpaved of the deposit-to-coast counterfactual. To avoid unreasonably low overlaps because of this issue, we allow for the possibility that governments revert back to the entire (yet unpaved) counterfactual in any period. Consequently, we count as overlap in period 2 the overlap between the actual paving in that period, and the entire counterfactual network as it appears at the end of the period. The overlap in our example is then $e + h$, and *deposit overlap*₂ is equal to $(e + h)/(e + g + h)$.

²⁸As explained earlier, this choice of periods is dictated by the fact that the next Michelin map of West Africa was published in early 1969.

In summary:

$$A_{0,real} = a$$

$$A_{1,real} = a + b + c + d$$

$$A_{1,cf} = a + d + e + f$$

$$A_{2,real} = a + b + c + d + e + g + h$$

$$A_{2,cf} = a + c + d + e + f + h + i$$

$$deposit\ overlap_1 = \frac{(A_{1,real} \setminus A_{0,real}) \cap A_{1,cf}}{A_{1,real} \setminus A_{0,real}} = \frac{d}{b + c + d}$$

$$deposit\ overlap_2 = \frac{(A_{2,real} \setminus A_{1,real}) \cap A_{2,cf}}{A_{2,real} \setminus A_{1,real}} = \frac{e + h}{e + g + h}$$

The same rules apply for subsequent period, $t > 2$.

Figure 8 and 9 provide an example for Sierra Leone. Figure 8 shows the initial (1965) paved network in black segments, and the deposit-to-port counterfactual as it appeared at the end of period 8 (1984-1985) in red segments. The deposits used to construct the counterfactual (those discovered by the beginning of the period) are denoted by green asterisks, and the ports are denoted by blue squares. Figure 9 shows the actual paved network by the end of period 8, highlighting the actual paving in the period (that is the roads paved in 1984 or 1985) with red circles. These circled additions measure a total of 158 km in length, giving us the paving quota for the period. The measure of overlap for period 8, $deposit\ overlap_8$, compares the circled additions to the red network in Figure 8 (the deposit-to-port counterfactual as it appeared at the end of the period), and finds that three out of five of them have sections overlapping with the red network. Summing up the length of these sections and dividing by 158, we obtain $deposit\ overlap_8 = 0.35$. This number means that 35% of the roads that were paved in 1984 and 1985 would have been paved in that period or earlier, had the government only cared about connecting deposits to ports. In the same period, the overlap with the city counterfactual (not shown), $city\ overlap_8$, was 0.16, suggesting that only 16% of the roads paved in this period would have been paved, had the government only cared about connecting cities to one another. The resulting measure of relative city bias is $city\ bias_8 = 0.16 - 0.35 = -0.19$.

Strictly speaking, a negative value of *city bias* means that the government of Sierra Leone in 1984-85 had a bias against connecting cities to one another, and in favour of connecting deposits

to ports. Note however that we will not be focusing on the absolute value of our measure, but rather only on its changes over time (and thus over different types of government) within each country.

Online Appendix Table OA3 reports the value of *deposit overlap_t*, *city overlap_t* and *city bias_t* over time for Sierra Leone. In 6 out of 17 periods, no actual paving occurred in Sierra Leone, necessarily leading to *deposit overlap_t* = *city overlap_t* = *city bias_t* = 0. To make sure that these observations are not driving our results, we will drop them in our regressions. By far the most frequent case is one where the actual paving overlaps partially with both counterfactuals, resulting in values of *deposit overlap_t* and *city overlap_t* that are included between zero and one. The *city bias_t* measure is typically quite different from zero, confirming that the two counterfactuals are quite different from each other. Importantly, there is substantial variation in *city bias_t* over time, suggesting that different governments had different paving priorities. For example, the government in charge at the beginning of the 1984-85 period (an autocracy according to the Polity IV index) has a score of -0.19, while the government in charge at the beginning of the 1971-72 period (a democracy according to the Polity IV index) had a score of 0.27. This suggests that the latter government prioritised connecting cities with one another (as opposed to deposits to port) more than the former government. In a few cases where very few new kilometers were paved, the actual paving either perfectly overlaps with both counterfactuals, or does not overlap at all (resulting in *city bias_t* = 0 in both cases). To rule out that some mechanical link between the amount of current paving and *city bias_t* confounds our results, we will be controlling for the amount of current paving in our preferred specification.

[Figures 8 and 9 about here]

One potential worry is that *deposit overlap_{i,t}* and *city overlap_{i,t}* will be lower when a greater portion of the corresponding counterfactual has already been paved,²⁹ and this mechanical link may confound our results. To rule this out, we define two additional variables, *cumulative city overlap_{i,t}* and *cumulative deposit overlap_{i,t}*. To construct these, we measure in each country and period the total observed kilometers of paved roads that overlap with either counterfactual and divide by the total kilometers of paved roads of the counterfactual up to that

²⁹For example, if the deposit-to-port counterfactual is entirely paved at the start of period *t*, then *deposit overlap_{i,t}* must necessarily be zero.

year. Thus, rather than measuring the marginal changes in each period, we now compare the stock of paved roads with the stock of paved roads under either counterfactual. We will also control for these two variables in our preferred specification.

4.3 Descriptive statistics

Table 4 shows the distribution of paving, and the resulting overlap and relative city bias measures, across country-periods in our sample, with country-by-country and period-by-period details provided in Online Appendix Tables OA1 and OA2. Our sample is made up of 12 country which we observe for 17 periods, giving us 204 country-periods. After dropping the country-periods with no paving, we are left with 147 country-periods. Average paving in a country-period with positive paving was substantial (203 km) but with a high standard deviation (250 km).

Our sample of countries is characterised by a positive mean $city\ bias_{i,t}$ (0.12). Underlying this is the fact that in the average country-period, 19% of the kilometers that were newly paved in the period overlapped with the deposit-to-port counterfactual (the average value of $deposit\ overlap_{i,t}$ is 0.19), while 30% of them overlapped with the city-to-city counterfactual (the average value of $city\ overlap_{i,t}$ is 0.30). The mean of the three variables rises to 0.16, 0.26 and 0.42 when we exclude periods with no paving. The fact that the two overlap measures are on average high and have a maximum value of 1 confirms that the two counterfactuals are reasonable networks, that appropriately take into account the topography of each country. Note also that $city\ bias_{i,t}$ displays a substantial standard deviation (almost three times the mean) and has maximum and minimum values of 1 and -1 (indicating that there are cases in which the new paving overlaps perfectly with one counterfactual, but not at all with the other). This suggests that although the two counterfactuals may overlap, they are also substantially different. In other words, they represent two truly different approaches to road paving.

One potential worry is that in “elongated” countries such as Benin or Togo the two counterfactuals may both have a markedly interior-to-coast shape, resulting in a relative city bias measure that is compressed around zero. That is not the case, however. Table OA1 shows that while relative city bias is low on average in Benin (0.14), it is larger in Togo (0.49). More importantly, the standard deviation is high in both countries. This suggests that there were many periods in which the newly paved roads overlapped a lot with one counterfactual but little with the other, implying that the two are substantially different.

[Table 4 about here]

5 Main results

Table 5 builds up to our main estimate of the effect of democracy on the relative bias towards connecting cities to one another (as opposed to deposits to ports). In column (1) we simply regress our relative city bias measure on the democracy score at the start of each period t , including country and period fixed effects and country-specific time-trends. The sample only includes country-periods when new paving took place, and we correct standard errors for possible heteroskedasticity.³⁰ As explained above, periods are here defined as the time between the publication of two consecutive Michelin West Africa maps, which are the same for all our countries since they are all included in each map. We let democracy at the start of the period affect relative city bias during the rest of the period, as measured from the actual paving that took place during the period. We find a significant positive effect of democracy on relative city bias. This suggests that democracies tend to focus more on connecting cities to one another (as opposed to deposits to ports), relative to autocracies. The average democracy level in our sample is -3.7 with a large standard deviation of 5.2 . A one-standard deviation increase in democracy predicts a 0.2 point (or half a standard deviation, c.f. Table 4) increase in relative city bias. In other words, such an improvement in democracy would shape the road network in favour of cities and has a magnitude that is about equal to the difference between Burkina Faso (which on average scores 0.14 on $city\ bias_{i,t}$) and Senegal (which on average scores 0.37 on $city\ bias_{i,t}$).

In the second to fourth column we show that this effect is robust to adding additional controls. The cumulative deposit and city overlap variables, $cum.\ deposit\ overlap_{i,t}$ and $cum.\ city\ overlap_{i,t}$ measure the overlap of the entire stock of paved roads with either counterfactual network at the start of each period. They control for the possibility that democracies and autocracies may have different histories in terms of road paving, which may mechanically affect their current relative city bias. As one could expect, the two coefficients suggest a process of mean reversion over time: for example, a high score on $cum.\ deposit\ overlap_{i,t}$ (indicating that much of the deposit-to-port counterfactual has already been paved) predicts a low value of $deposit\ overlap_{i,t}$.

³⁰Because the polity score is missing for some country-period, the number of observations drops from 147 to 137. In subsequent columns, we lose another 10 observations due to the fact that the $cumulative\ city\ overlap_{i,t}$ and $cumulative\ deposit\ overlap_{i,t}$ are not defined in the first period.

(indicating that little of the deposit-to-port counterfactual is additionally paved during the period), thus leading to a higher relative city bias. The stock of deposits at the start of period (*cum. deposit discoveries_{i,t}*) controls for the possibility that countries with more deposits may both be more likely to be an autocracy, and to pave deposit-to-port roads. The amount of current paving during the period could mechanically drive our results if democracies systematically paved more or less kilometers than autocracies and the amount of paving had some mechanical relationship with our overlap measures (for example, it could lead to more overlap with the city-to-city counterfactuals, which are typically longer than the deposit-to-port counterfactuals). Note however that in an (unreported) regression of current paving on democracy (with country and period fixed effects and country-specific trends) we find a negative but very insignificant relationship.³¹ The lagged number of discoveries may similarly have a mechanical effect, if, for example, democracies invest less in exploration and discover fewer deposits, which then also leads to more paving budget for cities and a higher relative city bias score. These potentially confounding effects do not affect our results, however.

In column 4 we additionally control for a country-specific, time-varying price index which measures the current value of the country’s deposits, as well as for the inflow of foreign aid and FDI and for the incidence of civil wars during the current period. The worry is that these external and internal factors may matter for paving decisions, and may also be systematically related with democracy for unrelated reasons. None of these controls affect our main estimates, however. We note in passing that, although none of these coefficients is significant, the signs are what one could expect: higher prices of metals and minerals, more FDI, and more civil conflict are associated with relatively more deposit-to-port roads, while more foreign aid is associated with relatively more city-to-city roads. Finally, in column 5 we cluster standard errors by country to allow for serial correlation in the error term. This does not change our conclusion. Because the additional controls of column 4 severely limit the sample, and because the number of country-clusters is small, we favour specification 3 from now on.

[Table 5 about here]

In Table 6 we unpack the underlying determinants of the increase in relative city bias associ-

³¹With only country fixed effects we find weak evidence that democracies pave less, but this is not robust to adding period fixed effects. These results differ from Jedwab and Storeygard (2018), who in a sample of 43 sub-Saharan African in 1960-2015 find that democracy is positively correlated with investment in paved roads.

ated with democracy. Column 1 repeats the main estimate of column 3 in Table 5 (where control variables are included but not displayed). Columns 2 and 3 split the relative city bias measure into its two components: *deposit overlap*_{*i,t*}, the overlap of actual paving with the deposit-to-port counterfactual, and *city overlap*_{*i,t*}, its overlap with the city-to-city counterfactual. We see that an improvement in democracy mainly reduces the overlap between actual paving and the deposit-to-port counterfactual, while only barely increasing the overlap with the city-to-city counterfactual. Columns 4 to 6 add lagged democracy, to investigate whether there is a longer delay between a change in institutions (and thus in policy decisions) and actual implementation. This reveals that democracies also significantly increase the overlap of paving with the city-to-city counterfactual, but with a one-period lag. In summary, compared to autocracies, democracies seem to divert resources away from connecting deposits to ports, and with a delay towards connecting cities to one another. Such a delay might be explained by the fact that democracies are slower at getting paving decisions implemented, or by any technological reason that makes it harder to pave roads connecting cities.

[Table 6 about here]

One key worry is that democratisation may be endogenous to some underlying national process (for example, faster economic growth) which simultaneously triggers a different approach to road paving. In Table 7, following the recent literature on how democratisation itself may spread from country to country (Buera et al., 2011; Acemoglu et al., 2018), we instrument the democracy score with the lagged regional average of democracy scores in other countries in the region, which is composed of our sample of countries and their direct neighbors.³² For reasons that will be clear momentarily, we first show in columns 1 to 3 that our main results are robust to dropping the time effects (time fixed effects and country-specific trends). Column 4 then shows the first stage of the 2SLS regression.

Column 4 shows that improvements in democracy in a country’s region significantly predicts more democratisation at home. The instrument is very strong with a first stage F-test of 110. In Acemoglu et al. (2018), a country’s “region” is defined as the reference group of countries within a large geographical area (such as “Africa”) that had the same democracy score as the

³²Direct neighbors to our sample of countries are Algeria, Cameroon, Cape Verde, Chad, Libya, Mauritania, and Nigeria.

country in question at the start of the sample. In our setting only one country was a democracy in 1965, causing most countries to follow the same set of reference countries. This limits variation across space and time, which is why we favour dropping the time effects.³³ Our conclusion is that the predicted democracy score results in an estimate on *city bias_{i,t}*, *deposit overlap_{i,t}* and *city overlap_{i,t}* that is very similar to the estimates in Table 6.

[Table 7 about here]

6 Robustness

We now investigate whether our main finding holds up to a few important robustness checks. First, we follow Acemoglu et al. (2018) and replace the Polity IV score on democracy by a dummy equal to one if a country is democratic overall. More precisely, the dummy considers a country as democratic if during a given year after 1972 Freedom House codes it as “Free” or “Partially Free”, and Polity IV assigns it a positive score. If one of these two sources is not available, then the dummy is constructed using a variety of additional sources. Figure 3 shows the difference between the Polity IV score and the dummy. In our sample the dummy attains a value of 1 whenever the Polity IV score is above zero (except for the years 1968-1969 in Sierra Leone).

In Table 8, we repeat Table 5 using this alternative measure of democracy. Comparing the two tables, we see that the results are similar. One difference is that the coefficient on *city bias_{i,t}*, while still positive, is no longer significant in the full specification of column 4. We have experimented with moving from column 3 to column 4 gradually, adding the additional control variables one by one. The variable responsible for the loss of significance on *city bias_{i,t}* is FDI, whose negative coefficient is now significant. This seems to suggest that part of the effect of democracy on *city bias_{i,t}* is driven through FDI: presumably, autocracies attract more FDI,

³³Our second-stage results are robust to including the time effects. However, in that case we find a counter-intuitive result in the first stage: an improvement in democracy in a country’s region (over and above the time fixed effects and country-specific trends) *negatively* predicts democratization at home. This may suggest a process of mean reversion after an acceleration of democratization, or be due to collinearity between the regional instrument and the period fixed effects. The F-test in that case is still large and equal to 98. We also experimented with including all African countries in the regional average, and with splitting our countries along the median democracy score in 1965 to construct two separate reference groups. These instruments were much weaker with F-test scores below 6 and thus not suitable.

which push for more deposit-to-port roads being built. Note however that this inference is based on a smaller sample, since we lose one third of the observations when we move from column 3 to column 4. The estimated coefficient from our preferred specification (column 3) now suggests that a change from autocracy (0) to democracy (1) predicts a 0.44 point (or slightly more than a standard deviation) increase in relative city bias.

[Table 8 about here]

Table 9 repeats Table 6 (unpacking relative city bias in its two component parts) using the Acemoglu et al. (2018) dummy as a measure of democracy. It again finds that results are similar, except that the positive effect of lagged democracy on *city overlap*_{*i,t*} is no longer significant. Finally, in Online Appendix Table OA4 we also repeat the IV strategy for this alternative measure of democracy, finding that results are the same as in Table 7.

[Table 9 about here]

One potential problem with the Acemoglu et al. (2018) dummy is that it varies much less than the Polity IV score in our sample, particularly after controlling for country and time fixed effects. As can be seen from Figure 3, five of the twelve countries in our sample only change status once around the same time, and one country never changes status. Furthermore, the additional changes of status experienced by the other countries are typically very short-lived. This limited variation may contribute to explaining why some of our coefficients becomes insignificant when we measure democracy in this way.

A second important robustness consists in ruling out that our results are driven by ethnic favouritism. The literature has repeatedly found that public policy in Africa tends to favour groups that are “co-ethnic” with the president (using the terminology in Burgess et al., 2014), and particularly so in periods of autocracy.³⁴ If locations that are connected by the deposit-to-port counterfactual happened to be more co-ethnic (meaning located in the ethnic region of birth of the leader) than locations that are connected by the city-to-city counterfactual, then ethnic favouritism could explain why we find that autocracies tend to follow the former counterfactual.

³⁴For example, Hodler and Raschky (2014) find that politically less constrained leaders engage more in regional favoritism as measured by increasing night-time lights in their ethnic region of birth, but also that it is increasing in the time they are in power. Burgess et al. (2015) find that more is spent on roads and more kilometers of roads are paved in Kenya in the leader’s own ethnic region of birth if the leader is autocratic.

To address this, we construct a measure, *ethnic city bias*_{*i,t*} that compares, for each *i* and *t*, the number of co-ethnic cities connected by the city-to-city counterfactual to the number of co-ethnic deposits connected by the deposit-to-port counterfactual. The idea is that if, for example, the city-to-city counterfactual has few co-ethnic cities, but the deposit-to-port counterfactual has many co-ethnic deposits, then a more autocratic government could be more biased in favour of the deposit-to-port counterfactual, contributing to explaining our results. Once we include in the regressions *ethnic city bias*_{*i,t*} and its interaction with democracy, however, this spurious link between democracy and *city bias*_{*i,t*} should be controlled for.

More in detail, our measure is defined as follows:

$$\text{ethnic city bias}_{i,t} = \text{co-ethnic city}_{i,t} - \text{co-ethnic deposit}_{i,t}, \quad (4)$$

where *co-ethnic city*_{*i,t*} is the number of cities that are co-ethnic with the leader, are not connected to the paved road network in reality, but should be connected according to the city-to-city counterfactual. Similarly, *co-ethnic deposit*_{*i,t*} is the number of deposits that are co-ethnic with the leader, are not connected in reality, but should be connected according to the deposit-to-port counterfactual. We define a location as connected if its shortest route to the largest city (for cities) or main port (for deposits) is fully paved. We construct three different versions of *ethnic city bias*_{*i,t*}. The first version measures *co-ethnic city*_{*i,t*} and *co-ethnic deposit*_{*i,t*} using the simple count of the number of locations (cities and deposits) that satisfy the stated requirements. The second version scales the count of locations by the total number of locations (co-ethnic or not) that are not connected in reality, but should be connected according to the relevant counterfactual. Finally, the third version is similar to the second, but weighs locations (both in the numerator and in the denominator) by their size. Scaling captures the idea that a number of co-ethnic locations are more likely to bias the leader in favour of a counterfactual if they make up a large share of all locations that are unconnected in reality but connected according to the counterfactual. The two scaled versions are bounded by -1 and 1, while the count version is bounded by -22 and 8 in our data.

Table 10 presents the results. We add *ethnic city bias*_{*i,t*} and its interaction with democracy to our baseline specification. Regressions 1 to 3 use the polity score to measure democracy, while 4 to 6 use the Acemoglu dummy. The first row can be interpreted as the effect of democracy

on relative city bias when $ethnic\ city\ bias_{i,t} = 0$. If anything, controlling for ethnic favouritism strengthens our results. Neither $ethnic\ city\ bias_{i,t}$ nor its interaction with democracy are ever significant. This should not be taken as a sign that ethnic favouritism did not matter, however: since our measure only captures the *relative* co-ethnicity of the two counterfactuals, lack of significance does not mean that ethnic favouritism was not important in absolute terms.

[Table 10 about here]

In the Online Appendix we provide several other robustness checks. In Tables OA5 and OA6 we provide two variations on the IV strategy of Table 7. First, we redefine the instrument so that it only captures the lagged level of the Polity IV democracy score in the set of countries with which a country shares a border. Table OA5 shows that this instrument also predicts democratization, but also that it is a much weaker instrument. As a result, the effect on relative bias, deposit overlap and city overlap is estimated with less precision. Second, in Table OA6 we show that the IV strategy with the original instrument is robust to a smaller sample where we can include all of the control variables, even though the instrument is again weaker.

Finally, following Burgess et al. (2015), we experiment with the two alternative ways of constructing the counterfactual road expansion paths. In the first alternative, we rank city pairs and deposit-port pairs only by the inverse of their distance. We then recalculate the counterfactuals by connecting closer pairs first, irrespectively of their size. In the second alternative, we rank pairs only by the sum of their size, irrespectively of their distance. Table OA7 reproduces our preferred specification (Table 5, column 3) using these alternative counterfactuals. Results are robust to using the distance-only counterfactuals but not to using the size-only counterfactuals, even though the coefficients and standard errors are very similar in the two cases. While the general pattern is the same across the three measures, we believe that a measure that combines distance and importance by size is a more relevant input into decision making on road paving.

7 Conclusions

In this paper, we have used a newly-constructed dataset to establish a new fact about the way in which the West African road networks have expanded since independence. In a sample of twelve countries and spanning almost five decades, more autocratic governments have displayed

a stronger preference for paving interior-to-coast roads, and particularly deposit-to-port roads. Importantly, this finding is robust to controlling for country fixed effects (and thus for cross sectional variation in geography and comparative advantage), as well as for external factors (such as the prices of minerals) that may affect the evolution of both democracy and comparative advantage over time. It is also robust to instrumenting for democracy as suggested by the most recent literature on democracy and growth.

One possible interpretation of this finding is that autocracies attach a greater weight than democracies to rent extraction, which has led them to bias their paving decisions in favour of the extractive industries and against economic development more in general. For example, weaker electoral concerns may make it less required of autocracies to seek consensus through a balanced road investment program, while at the same time weaker media and civil society control may make it easier for them to hide a biased program. This interpretation is consistent with earlier results by Burgess et al. (2015) on Kenya, showing that democracies reduced ethnic favouritism in the allocation of roads. It is also consistent with the institutional view of economic development, according to which it is only inclusive political institutions that will deliver the market-supporting policies required to generate sustained economic growth.

In this interpretation, the notorious “interior-to-coast” nature of the African road networks is not (or not only) the natural result of geography or the natural choice of countries whose comparative advantage lies in exporting natural resources, but also the result of political distortions. The immediate policy implication is that, as advocated by many, the African countries should indeed focus on re-balancing their networks: this seems important to know, particularly at a time in which China’s investments in Africa are, if anything, helping to exacerbate the traditional interior-to-coast pattern.

In order to corroborate this interpretation, one would need to construct and estimate a quantitative model of road building, trade and growth in one or more African countries, which one could then use to make counterfactual welfare calculations for a case in which the country had been a democracy ever since independence. We leave this for future research.

References

- [1] Acemoglu, Daron and James A. Robinson (2012). *Why Nations Fail*, New York, Crown Publishers.
- [2] Acemoglu, Daron, Suresh Naidu, Pascual Restrepo and James A. Robinson (2018). “Democracy Does Cause Growth”, forthcoming, *Journal of Political Economy*.
- [3] Ades, Alberto F. and Edward L. Glaeser (1995). “Trade and Circuses: Explaining Urban Giants”, *Quarterly Journal of Economics*, Vol. 110, no. 1, pp. 195-227.
- [4] Alder, Simon and Illenin Kondo (2018). “Political Distortions and Infrastructure Networks in China: A Quantitative Spatial Equilibrium Analysis”, *mimeo*.
- [5] Amarasinghe, Ashani, Roland Hodler, Paul A. Raschky and Yves Zenou (2018). “Spatial Diffusion of Economic Shocks in Networks”, *mimeo*.
- [6] Bazzi, Samuel and Christopher Blattman (2014). “Economic Shocks and Conflict: Evidence from Commodity Prices”, *American Economic Journal: Macroeconomics*, vol. 6, no. 4, pp. 1-38.
- [7] Blimpo, Moussa P., Robin Harding and Leonard Wantchekon (2013). “Public Investment in Rural Infrastructure: Some Political Economy Considerations”, *Journal of African Economies*, Vol 2, pp. 57-83.
- [8] Bonfatti, Roberto, and Steven Poelhekke (2017). ”From mine to coast: transport infrastructure and the direction of trade in developing countries” *Journal of Development Economics* 127, 91-108.
- [9] Buera, Francisco J., Alexander Monge-Naranjo, Giorgio E. Primiceri (2011). Learning the Wealth of Nations, *Econometrica*, 79 (1), 145.
- [10] Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, Gerard Padr i Miquel (2015). The Value of Democracy: Evidence from Road Building in Kenya. *American Economic Review*, vol. 105, no. 6, pp. 1817-51.

- [11] Buys, Piet, Uwe Deichmann, David Wheeler (2010). Road Network Upgrading and Overland Trade Expansion in Sub-Saharan Africa. *Journal of African Economies*, Volume 19, Issue 3, Pages 399-432.
- [12] Donaldson, Dave (2015). The Gains from Market Integration. *Annual Review of Economics*, 7:619-647.
- [13] Felbermayr, Gabriel and Alexander Tarasov (2015). Trade and the Spatial Distribution of Transport Infrastructure/ CESifo Working Paper Series 5634, CESifo Group Munich.
- [14] Holdler, Roland and Paul A. Raschky (2014). Regional Favoritism. *The Quarterly Journal of Economics* 129(2) 995-1033.
- [15] Jedwab, Remi and Alexander Moradi (2016). The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa. *The Review of Economics and Statistics* 98(2) 268-284.
- [16] Jedwab, Remi and Adam Storeygard (2018). The Average and Heterogeneous Effects of Transportation Investments: Evidence from sub-Saharan Africa 1960-2010, mimeo.
- [17] Marshall, Monty G., Ted R. Gurr and Keith Jagers (2017). *Polity IV Project, Dataset Users' Manual*, Center for Systemic Peace.
- [18] Murdock, George P (1959). *Africa: its peoples and their culture history*, New York, McGraw-Hill.
- [19] Murdock, George Peter (1967). *Ethnographic atlas*, Pittsburgh, University of Pittsburgh Press.
- [20] Nkrumah, Kwame (1963). *Africa must unite*, New York, Praeger.
- [21] Padro i Miquel, Gerard (2007). "The Control of Politicians in Divided Societies: the Politics of Fear", in *Review of Economic Studies*, Vol. 74, No. 4, pp. 1259-1274.
- [22] Porteus, Obie (2017). "High Trade Costs and Their Consequences: An Estimated Dynamic Model of African Agricultural Storage and Trade", *mimeo*.

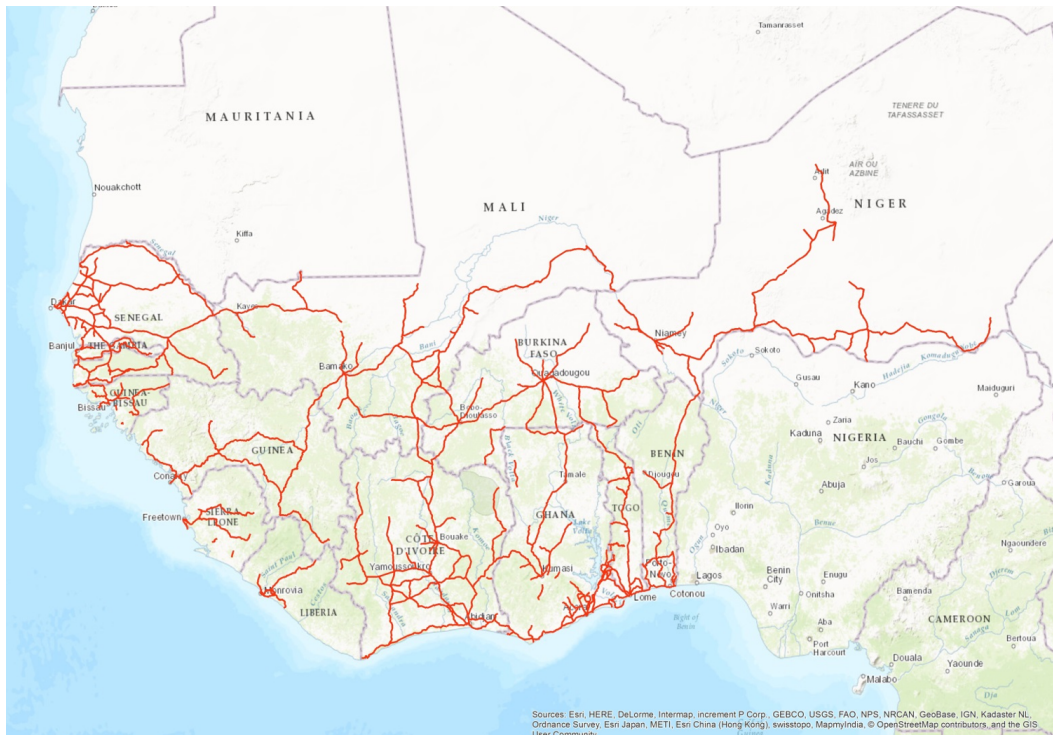
- [23] Rodney, W. (1982). *How Europe Underdeveloped Africa*, Washington DC, Howard University Press.
- [24] Sachs, J. D., J. W. McArthur, G. Schmidt-Traub, M. Kruk, C. Bahadur, M. Faye and G. McCord (2004). “Ending Africa’s Poverty Trap”, in *Brookings Papers on Economic Activity* No. 1 (2004), pp. 117-226.
- [25] US Geological Survey (1999). *Metal Prices in the United States Through 1998*, Washington DC, US Government Printing Office.
- [26] Waters, Jennifer K., Robert H. Mayer and David L. Kriebel (2000). *Draft Final Report - Shipping Trends Analysis*, Annapolis (MD), US Naval Academy Department of Naval Architecture and Ocean Engineering.

Figures

Figure 1: Paved roads in 1965



Figure 2: Paved roads in 2012



Note: Paved roads as appearing in the 1965 and 2012 Michelin road maps of West Africa. For a full discussion of sources, see Section 3.2.

Figure 3: Democracy scores

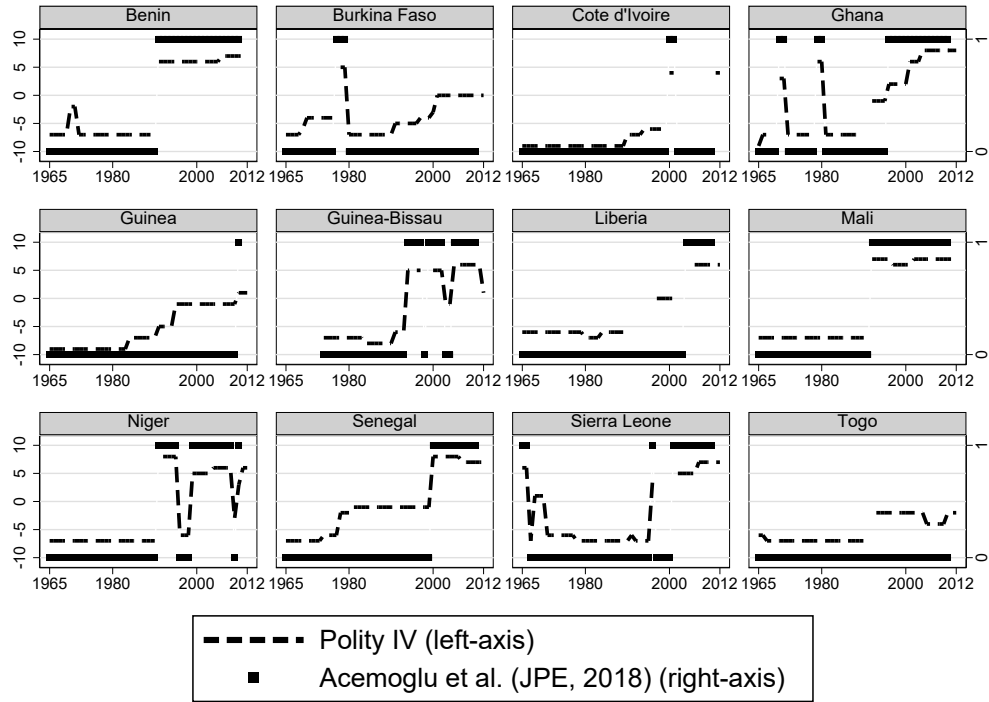
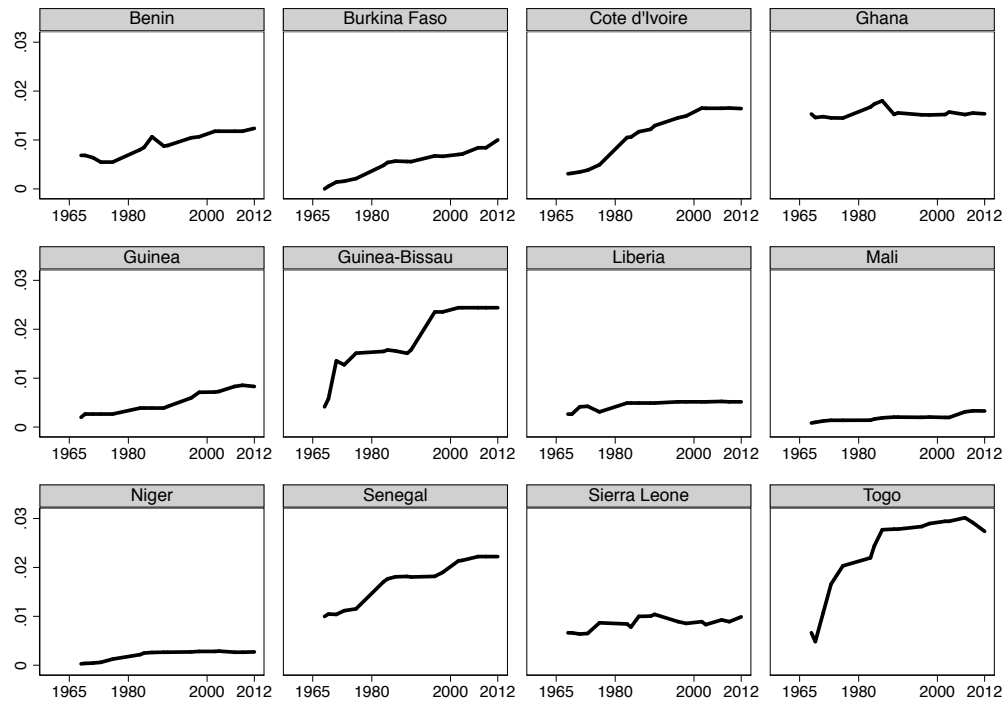


Figure 4: Cumulative paving since 1965, in km per square km of land area



Note: Paved roads as appearing in subsequent Michelin road maps of West Africa. For a full discussion of sources, see Section 3.2.

Figure 5: City-to-city counterfactual in 2012 Figure 6: Deposit-to-port counterfactual in 2012

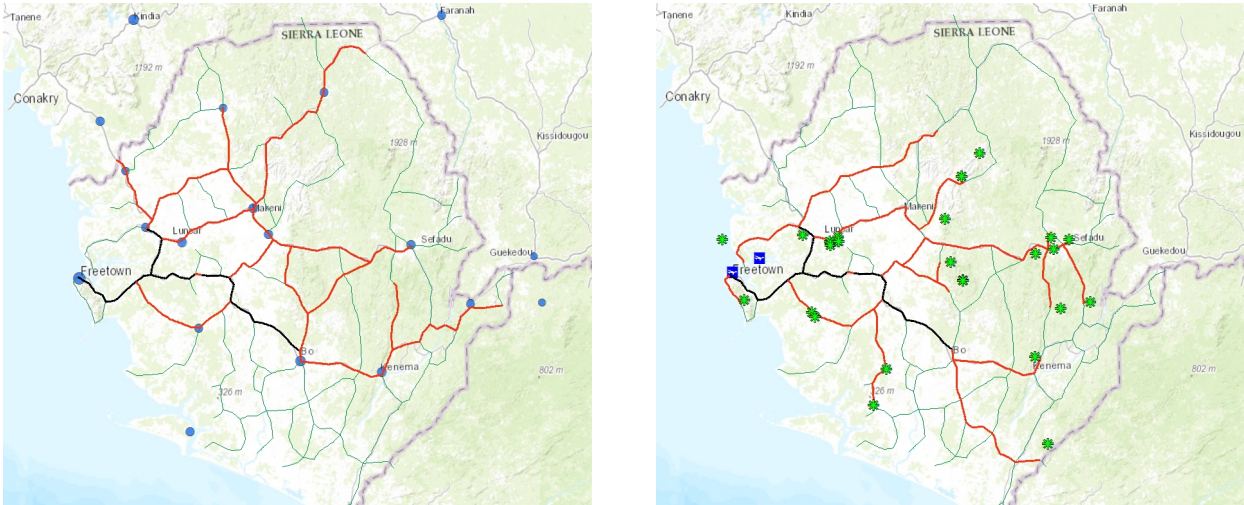


Figure 7: Overlap diagram

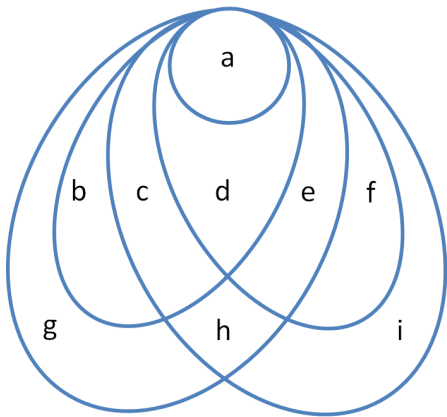
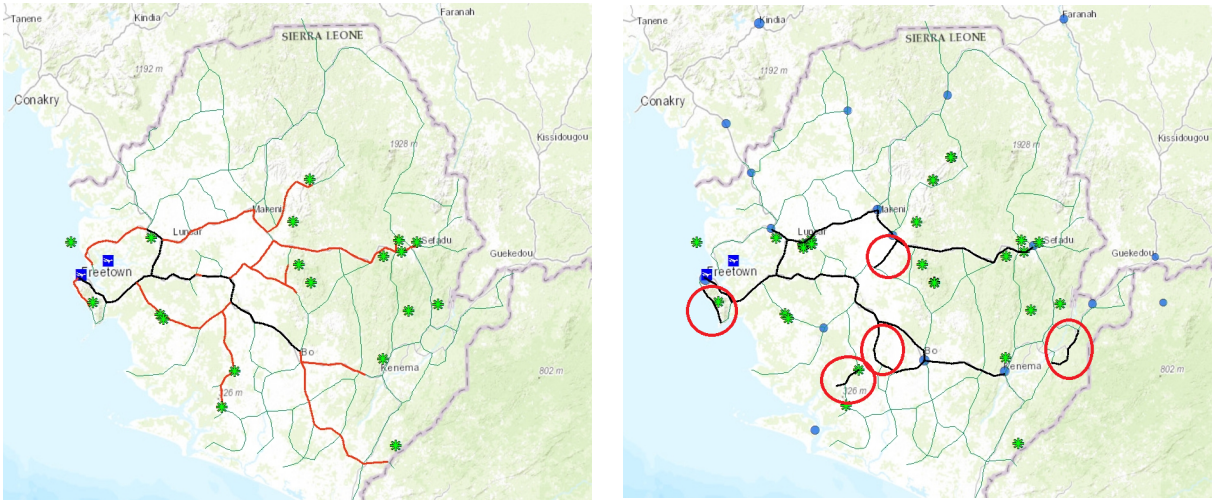


Figure 8: Deposit-to-port counterfactual in 1986

Figure 9: Actual network in 1986



Tables

Table 1: Road quality changes between 1965 and 2012 in km

Status in 1965			Status by 2012			
			No road	Unpaved (lowest quality)	Unpaved (medium quality)	Paved (good quality)
No road	16,678	→	719	4,938	8,985	2,036
Unpaved (lowest quality)	24,666	→	<i>1,103</i>	12,711	7,259	3,593
Unpaved (medium quality)	50,372	→	<i>687</i>	<i>5,414</i>	28,093	16,178
Paved (good quality)	10,477	→	<i>113</i>	<i>117</i>	<i>952</i>	9,295

Note: The table lists changes in road quality across 12 West African countries as inferred from comparing the digitised Michelin maps of West Africa for 1966 and 2012 (for a full discussion of sources, see Section 3.2). The three quality categories used in the table match to the six quality categories used by Michelin in the following way: Paved (good quality) = “surfaced”; Unpaved (medium quality) = “improved” and “partially improved”; Unpaved (lowest quality) = “earth roads”, “tracks” and “others”. “No road” are sections where a road is observed in at least one year after 1965. For example, there are 719 km which were recorded as track or roads at some point between 1965 and 2012, but these no longer exist by 2012. Only paved (surfaced) roads can be used year-round. Road quality deteriorations are denoted in italics. For example, 113 km of paved roads disappeared by 2012, and 117 deteriorated to the “unpaved (track)” quality. For sources, see Section 3.2.

Table 2: Discoveries of deposits and their characteristics

Country	Discoveries		Metal/mineral		Size	
	before 1965	after 1965				
Benin	1	2	Gold	257	SuperGiant	1
Burkina Faso	3	84	Iron ore	33	Giant	38
Côte d'Ivoire	4	20	Uranium	29	Major	119
Ghana	21	53	Diamonds	23	Moderate	164
Guinea	16	34	Bauxite	22	Minor	69
Guinea-Bissau	1	0	Nickel	7		
Liberia	5	5	Manganese	6		
Mali	8	46	Phosphate	6		
Niger	9	27	Titanium, Zirconium	3		
Senegal	3	14	Silver	1		
Sierra Leone	15	18	Chromium	1		
Togo	1	1	Copper	1		
			Platinum	1		
			Zinc	1		
Total	87	304	Total	391	Total	391

Note: The table describes the discoveries of deposits in 12 West African countries through 2012, the distribution of the main metal or mineral contained in them, and the distribution of deposit size. For sources, see Section 3.3.

Table 3: Distribution of cities

Country	Cities		Year	Population			
	in 1960	in 2010		mean	s.d.	min	max
Benin	24	81	1960	10,227	27,761	126	337,800
Burkina Faso	18	82	1970	18,149	55,951	249	624,091
Côte d'Ivoire	115	166	1980	29,878	98,508	1,067	1,251,272
Ghana	141	173	1990	45,168	157,047	1,560	2,107,460
Guinea	24	39	2000	64,498	230,042	5,610	3,043,326
Guinea-Bissau	2	6	2010	94,285	347,402	10,012	3,966,553
Liberia	5	23					
Mali	27	71					
Niger	25	51					
Senegal	31	57					
Sierra Leone	14	19					
Togo	51	54					
Total	477	822					

Note: This table shows the distribution of cities and population across countries and time. Between 2010 and 1960, 345 new cities have formed that have a population of at least 10,000 in 2010. For the years before these new cities exist, the population statistics in the right columns treat their population as missing. We use the location and population of cities in 1960 to generate the counterfactual paving schedules. Source: Africapolis project.

Table 4: Road paving characteristics

All country-periods	N	mean	s.d.	min	max
Current paving (km)	204	146.42	231.23	0.00	1794.12
Current deterioration (km)	204	-42.71	87.99	-759.17	0.00
City bias	204	0.12	0.34	-1.00	1.00
City overlap	204	0.30	0.36	0.00	1.00
Deposit overlap	204	0.19	0.31	0.00	1.00
Cumulative deposit overlap	203	0.52	0.22	0.16	1.00
Cumulative city overlap	203	0.69	0.18	0.21	1.00
Country-periods with new paving	N	mean	s.d.	min	max
Current paving (km)	147	203.20	250.45	6.29	1794.12
Current deterioration (km)	147	-47.21	96.91	-759.17	0.00
City bias	147	0.16	0.39	-1.00	1.00
City overlap	147	0.42	0.36	0.00	1.00
Deposit overlap	147	0.26	0.34	0.00	1.00
Cumulative deposit overlap	147	0.51	0.22	0.16	0.90
Cumulative city overlap	147	0.68	0.18	0.21	0.97

Note: This table shows summary statistics for the main variable of interest for a sample of country-years in which new paving occurs (bottom panel) and all country-years (top panel).

Table 5: City bias increases in democracy

Dependent variable →	City bias				
	(1)	(2)	(3)	(4)	(5)
Democracy, start of t	0.032 (0.015)	0.041 (0.018)	0.038 (0.017)	0.040 (0.021)	0.038 (0.019)
Cum. deposit overlap, start of t		1.708 (0.586)	1.862 (0.626)	2.068 (0.811)	1.862 (0.597)
Cum. city overlap, start of t		-1.875 (0.622)	-2.054 (0.638)	-2.466 (0.696)	-2.054 (0.332)
Cum. deposit discoveries, start of t		-0.019 (0.013)	-0.018 (0.012)	0.017 (0.024)	-0.018 (0.013)
Current paving, t			-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Deposit discoveries, t-1			-0.002 (0.024)	0.003 (0.036)	-0.002 (0.024)
Price index, average t				-0.000 (0.000)	
Log aid, average t				0.081 (0.163)	
Log FDI, average t				-0.062 (0.042)	
Civil war, average t				-0.868 (0.718)	
Country FE	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes
Country-specific trends	Yes	Yes	Yes	Yes	Yes
Observations	137	127	127	87	127
R-squared	0.526	0.603	0.608	0.738	0.608

Note: This table shows OLS regressions of the effect of democracy on countries' relative bias towards connecting cities (city bias). All specifications include country and period fixed effects and country-specific linear trends. Robust standard errors in parenthesis. Table A1 contains summary statistics. See Section 3 for variable definitions and sources.

Table 6: City bias and paving overlap with city and deposit counterfactual

Dependent variable →	City bias	Deposit overlap	City overlap	City bias	Deposit overlap	City overlap
	(1)	(2)	(3)	(4)	(5)	(6)
Democracy, start of t	0.038 (0.017)	-0.027 (0.013)	0.011 (0.015)	0.030 (0.019)	-0.030 (0.014)	0.000 (0.015)
Democracy, start of t-1				0.027 (0.020)	0.011 (0.015)	0.038 (0.016)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-specific trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	127	127	127	125	125	125
R-squared	0.533	0.564	0.576	0.547	0.603	0.672

Note: This table shows OLS regressions of the effect of democracy on countries' relative bias towards connecting cities (city bias) in columns 1 and 4, on the overlap of actual road paving with the deposit-to-port counterfactual in columns 2 and 5, and on the overlap of actual road paving with the city-to-city counterfactual in columns 3 and 6. The baseline controls are those included in Table 5, column 3. All specifications include country and period fixed effects and country-specific linear trends. Robust standard errors in parenthesis. Table A1 contains summary statistics. See Section 3 for variable definitions and sources.

Table 7: Instrumenting democracy with regional waves of democratization

Dependent variable →	City bias	City bias	City bias	Democracy, City bias start of t		Deposit overlap	City overlap
				1st stage	2nd stage		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Democracy, start of t	0.038 (0.017)	0.031 (0.015)	0.026 (0.011)		0.028 (0.014)	-0.029 (0.009)	-0.001 (0.012)
Cum. deposit overlap, start of t	1.862 (0.626)	0.342 (0.395)	-0.197 (0.289)	0.541 (2.168)	-0.186 (0.280)	-0.184 (0.213)	-0.370 (0.234)
Cum. city overlap, start of t	-2.054 (0.638)	-0.937 (0.506)	-0.585 (0.515)	-11.718 (3.607)	-0.561 (0.489)	0.135 (0.336)	-0.425 (0.396)
Cum. deposit discoveries, start of t	-0.018 (0.012)	-0.005 (0.009)	-0.004 (0.006)	0.044 (0.042)	-0.004 (0.006)	0.012 (0.004)	0.008 (0.005)
Current paving, t	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.001 (0.001)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Deposit discoveries, t-1	-0.002 (0.024)	-0.012 (0.024)	-0.012 (0.024)	0.010 (0.115)	-0.012 (0.022)	0.015 (0.013)	0.003 (0.013)
Lagged regional average of democracy				1.364 (0.138)			
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	No	No	No	No	No
Country-specific trends	Yes	No	No	No	No	No	No
1st stage F-test				110.07			
Observations	127	127	127	127	127	127	127
R-squared	0.533	0.398	0.245		0.114	0.132	0.095

Note: This table shows OLS regressions of the effect of democracy on countries' relative bias towards connecting cities (city bias) in columns 1-3. Column 4 is the first stage regression of a 2SLS specification where we instrument democracy with its lagged regional average. Columns 5-7 are the second stage regressions for three different dependent variables. All specifications include country fixed effects. Columns 1 and 2 include period fixed effects, and column 1 country-specific linear trends. Robust standard errors in parenthesis. Table A1 contains summary statistics. See Section 3 for variable definitions and sources.

Table 8: City bias increases in democracy (Acemoglu et al., 2018 definition of democracy)

Dependent variable →	City bias				
	(1)	(2)	(3)	(4)	(5)
Democracy, start of t (Acemoglu)	0.363 (0.151)	0.450 (0.155)	0.437 (0.152)	0.250 (0.207)	0.437 (0.121)
Cum. deposit overlap, start of t		1.372 (0.564)	1.455 (0.585)	1.955 (0.778)	1.455 (0.583)
Cum. city overlap, start of t		-2.185 (0.596)	-2.294 (0.612)	-2.722 (0.673)	-2.294 (0.313)
Cum. deposit discoveries, start of t		-0.013 (0.012)	-0.012 (0.012)	0.013 (0.027)	-0.012 (0.008)
Current paving, t			-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Deposit discoveries, t-1			-0.008 (0.024)	0.011 (0.040)	-0.008 (0.026)
Price index, average t				-0.001 (0.000)	
Log aid, average t				0.153 (0.153)	
Civil war, average t				-0.617 (0.650)	
Log FDI, average t				-0.073 (0.040)	
Country FE	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes
Country-specific trends	Yes	Yes	Yes	Yes	Yes
Observations	142	131	131	88	131
R-squared	0.496	0.585	0.587	0.723	0.587

Note: This table shows OLS regressions of the effect of Acemoglu et al.'s democracy dummy on countries' relative bias towards connecting cities (city bias). All specifications include country and period fixed effects and country-specific linear trends. Robust standard errors in parenthesis. Table A1 contains summary statistics. See Section 3 for variable definitions and sources.

Table 9: City bias and paving overlap with city and deposit counterfactual (Acemoglu et al., 2018 definition of democracy)

Dependent variable →	City bias	Deposit overlap	City overlap	City bias	Deposit overlap	City overlap
	(1)	(2)	(3)	(4)	(5)	(6)
Democracy, start of t (Acemoglu)	0.437 (0.152)	-0.256 (0.142)	0.181 (0.156)	0.418 (0.165)	-0.294 (0.136)	0.125 (0.162)
Democracy, start of t-1 (Acemoglu)				0.056 (0.187)	0.094 (0.158)	0.150 (0.186)
Baseline controls						
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-specific trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	131	131	131	130	130	130
R-squared	0.587	0.725	0.820	0.591	0.740	0.844

Note: This table shows OLS regressions of the effect of Acemoglu et al.'s democracy dummy on countries' relative bias towards connecting cities (city bias) in columns 1 and 4, on the overlap of actual road paving with the deposit-to-port counterfactual in columns 2 and 5, and on the overlap of actual road paving with the city-to-city counterfactual in columns 3 and 6. The baseline controls are those included in Table 5, column 3. All specifications include country and period fixed effects and country-specific linear trends. Robust standard errors in parenthesis. Table A1 contains summary statistics. See Section 3 for variable definitions and sources.

Table 10: Controlling for ethnic city bias

Dependent variable →	City bias					
	Polity score			Acemoglu dummy		
	count	count, scaled	size, scaled	count	count, scaled	size, scaled
Ethnic city bias definition →	(1)	(2)	(3)	(4)	(5)	(6)
Democracy, start of t	0.037 (0.017)	0.047 (0.019)	0.048 (0.019)	0.476 (0.155)	0.559 (0.188)	0.548 (0.179)
Ethnic city bias	-0.008 (0.013)	-0.068 (0.243)	-0.095 (0.242)	-0.023 (0.022)	-0.213 (0.292)	-0.290 (0.312)
Democracy, start of t * ethnic city bias	0.003 (0.002)	0.021 (0.033)	0.029 (0.034)	0.033 (0.023)	0.440 (0.318)	0.491 (0.326)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-specific trends	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	127	109	109	131	113	113
R-squared	0.542	0.558	0.560	0.525	0.539	0.541

Note: This table shows OLS regressions of the effect of democracy on countries' relative bias towards connecting cities (city bias), where columns 4-6 use Acemoglu et al.'s democracy dummy. We control for *Ethnic city bias_{i,t}*, which is defined as city ethnicity minus deposit ethnicity, where city ethnicity is the number of cities that are located in the ethnic region of the current ruler, are not connected to the capital in reality, but should be connected according to the city counterfactual. Deposit ethnicity is the number of deposits that are located in the ethnic region of the current ruler, are not connected in reality, but should be connected according to the deposit counterfactual. We define a location as connected if its shortest route to the largest city (for cities) or main port (for deposits) is fully paved. *Count* denotes where we count the number of deposits or cities to measure ethnicity. *Count, scaled* denotes where we scale the ethnicity count by the number of cities or deposits in the whole country that are not connected in reality but should be connected according to the relevant counterfactual. *Size, scaled* denotes where we measure ethnicity as the size-weighted sum of deposits or cities and scale by the size-weighted sum of cities or deposits in the whole country that are not connected in reality but should be connected according to the relevant counterfactual. Scaling reduces the number of observations because the denominator is zero in some cases. The baseline controls are those included in Table 5, column 3. All specifications include country and period fixed effects and country-specific linear trends. Robust standard errors in parenthesis. Table A1 contains summary statistics. See Section 3 for variable definitions and sources.

Appendix

Table A1: Summary statistics

variable	N	mean	p50	s.d.	min	max
City bias	137	0.17	0.05	0.40	-1	1
City overlap	137	0.43	0.39	0.36	0	1
Deposit overlap	137	0.26	0.08	0.34	0	1
Democracy (Polity IV)	137	-3.88	-7	5.19	-9	8
Democracy (Acemoglu)	137	0.17	0	0.38	0	1
Deposit discoveries	137	1.99	0	3.79	0	22
Cumulative deposit discoveries	137	15.12	9	15.52	1	72
Current paving	137	213.54	148.84	255.98	6.29	1,794.12
Cumulative deposit overlap	127	0.52	0.53	0.22	0.16	1.00
Cumulative city overlap	127	0.68	0.71	0.19	0.21	1.00
Regional average of democracy (Polity IV)	128	-4.07	-5.60	2.94	-6.59	2.83
Neighbors' reg. av. of democracy (Polity IV)	128	-4.23	-5.80	3.75	-8.33	6.50
Regional average of democracy (Acemoglu)	128	0.20	0.13	0.18	0.06	0.69
Price index, period average	87	258.82	111.98	385.55	11	1,951.31
log Aid, period average	87	18.48	18.60	1.03	15.63	20.41
log FDI, period average	87	2.23	2.53	1.84	-4.61	5.93
Civil war, period average	87	0.02	0	0.12	0	0.83
Ethnic city bias, count	137	-0.57	0	5.23	-22	8
Ethnic city bias, count, scaled	115	0.06	0	0.34	-0.85	0.80
Ethnic city bias, size, scaled	115	0.07	0	0.34	-0.83	0.83

Note: This table provides summary statistics for all variables used in the analysis. See Section 3 for variable definitions and sources.

Online Appendix

Table OA1: Road paving characteristics by country, period averages

Period	Current paving (km)	Current deterioration (km)	Deposit overlap	City overlap	City bias	
					mean	s.d.
Benin	118.97	0.00	0.40	0.54	0.14	0.68
Burkina Faso	210.30	−12.75	0.20	0.34	0.14	0.49
Côte d’Ivoire	320.86	−42.74	0.11	0.32	0.21	0.29
Ghana	199.40	−178.54	0.64	0.75	0.11	0.19
Guinea	189.41	−2.72	0.24	0.22	−0.03	0.35
GuineaBissau	62.49	−9.93	0.18	0.31	0.13	0.30
Liberia	72.89	−5.06	0.29	0.40	0.11	0.20
Mali	307.95	−63.58	0.05	0.18	0.13	0.28
Niger	296.80	−18.72	0.17	0.25	0.08	0.47
Senegal	243.12	−61.10	0.21	0.57	0.37	0.42
Sierra Leone	102.89	−42.41	0.53	0.50	−0.03	0.13
Togo	141.73	−58.87	0.18	0.67	0.49	0.38
Sample average	203.20	−47.21	0.26	0.42	0.16	0.39

Note: This table shows means (and standard deviation in the last column) by country across periods, for country-periods with positive paving. See Section 3 for variable definitions and sources.

Table OA2: Road paving characteristics by period, country averages

Period	Current paving (km)	Current deterioration (km)	Deposit overlap	City overlap	City bias	
					mean	s.d.
1965-1967	223.85	-88.02	0.03	0.10	0.08	0.18
1968	141.89	-31.64	0.17	0.29	0.12	0.18
1969-1970	211.79	-69.32	0.21	0.28	0.07	0.13
1971-1972	135.92	-42.13	0.19	0.33	0.14	0.15
1973-1975	258.05	-30.24	0.36	0.37	0.01	0.15
1976-1982	579.33	-21.26	0.32	0.44	0.12	0.26
1983	183.26	-34.00	0.10	0.42	0.33	0.29
1984-1985	210.85	-43.74	0.23	0.34	0.11	0.47
1986-1988	102.56	-128.58	0.28	0.48	0.20	0.39
1989	74.06	-28.57	0.35	0.59	0.23	0.35
1990-1995	220.46	-44.08	0.44	0.50	0.06	0.44
1996-1997	141.91	-42.21	0.20	0.56	0.36	0.33
1998-2001	171.06	-49.90	0.22	0.61	0.38	0.59
2002	47.06	-16.50	0.60	0.41	-0.19	0.68
2003-2006	261.03	-21.63	0.16	0.47	0.31	0.54
2007-2008	109.22	-10.54	0.32	0.30	-0.02	0.39
2009-2011	197.37	-74.48	0.42	0.87	0.45	0.47
Sample average	203.20	-47.21	0.26	0.42	0.16	0.39

Note: This table shows means (and standard deviation in the last column) by period, for country-periods with positive paving. See Section 3 for variable definitions and sources.

Table OA3: Road paving by period in Sierra Leone

Period	Stock of paved roads, end of period (km)	Current paving (km)	Current de- terioration (km)	Deposit overlap	City overlap	City bias
1965-1967	478.36	254.68	-47.38	0.20	0.00	-0.20
1968	478.36	0.00	0.00	0.00	0.00	0.00
1969-1970	460.40	110.52	-128.48	0.57	0.43	-0.15
1971-1972	469.10	106.66	-97.95	0.66	0.94	0.27
1973-1975	626.59	193.56	-36.07	0.51	0.56	0.04
1976-1982	608.86	0.00	-17.73	0.00	0.00	0.00
1983	563.44	0.00	-45.41	0.00	0.00	0.00
1984-1985	721.03	157.59	0.00	0.35	0.16	-0.19
1986-1988	724.85	29.64	-25.82	0.73	0.65	-0.08
1989	750.67	25.82	0.00	1.00	1.00	0.00
1990-1995	643.20	0.00	-107.47	0.00	0.00	0.00
1996-1997	617.38	0.00	-25.82	0.00	0.00	0.00
1998-2001	642.97	25.58	0.00	0.00	0.00	0.00
2002	599.91	25.82	-68.87	1.00	1.00	0.00
2003-2006	668.79	68.87	0.00	0.00	0.00	0.00
2007-2008	643.20	0.00	-25.58	0.00	0.00	0.00
2009-2011	714.36	133.05	-61.89	0.81	0.81	0.00

Note: This table shows data points per period for Sierra Leone. See Section 3 for variable definitions and sources.

Table OA4: Instrumenting democracy with regional waves of democratization: alternative measure of democracy (Acemoglu et al., 2018 measure of democracy)

Dependent variable →	City bias	City bias	City bias	Democracy, start of t (Acemoglu)	City bias	Deposit overlap	City overlap
	1st stage			2nd stage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Democracy, start of t (Acemoglu)	0.437 (0.152)	0.275 (0.131)	0.269 (0.137)		0.415 (0.223)	-0.491 (0.184)	-0.076 (0.218)
Cum. deposit overlap, start of t	1.455 (0.585)	0.327 (0.366)	-0.239 (0.282)	0.180 (0.202)	-0.226 (0.269)	-0.066 (0.206)	-0.292 (0.222)
Cum. city overlap, start of t	-2.294 (0.612)	-1.179 (0.482)	-0.770 (0.488)	-0.599 (0.336)	-0.684 (0.460)	0.156 (0.326)	-0.527 (0.392)
Cum. deposit discoveries, start of t	-0.012 (0.012)	-0.004 (0.008)	-0.000 (0.006)	0.004 (0.005)	-0.003 (0.007)	0.012 (0.005)	0.009 (0.005)
Current paving, t	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Deposit discoveries, t-1	-0.008 (0.024)	-0.014 (0.023)	-0.012 (0.023)	-0.004 (0.016)	-0.010 (0.022)	0.014 (0.013)	0.004 (0.013)
Lagged regional average of democracy				1.140 (0.252)			
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	No	No	No	No	No
Country-specific trends	Yes	No	No	No	No	No	No
1st stage F-test				29.433			
Observations	131	131	131	131	131	131	131
R-squared	0.511	0.388	0.230		0.083	0.022	0.070

Note: This table shows OLS regressions of the effect of Acemoglu et al.'s democracy dummy on countries' relative bias towards connecting cities (city bias) in columns 1-3. Column 4 is the first stage regression of a 2SLS specification where we instrument the democracy dummy with its lagged regional average. Columns 5-7 are the second stage regressions for three different dependent variables. All specifications include country fixed effects. Columns 1 and 2 include period fixed effects, and column 1 country-specific linear trends. Robust standard errors in parenthesis. Table A1 contains summary statistics. See Section 3 for variable definitions and sources.

Table OA5: Instrumenting democracy with *neighbors'* change in democratization

Dependent variable →	Democracy, start of t	City bias	Deposit overlap	City overlap
	1st stage	2nd stage		
	(1)	(2)	(3)	(4)
Democracy, start of t		0.016 (0.017)	-0.029 (0.013)	-0.012 (0.015)
Cum. deposit overlap, start of t	-1.825 (2.726)	-0.247 (0.290)	-0.182 (0.224)	-0.430 (0.247)
Cum. city overlap, start of t	-12.514 (4.056)	-0.700 (0.486)	0.138 (0.350)	-0.562 (0.401)
Cum. deposit discoveries, start of t	0.140 (0.051)	-0.000 (0.006)	0.012 (0.005)	0.012 (0.005)
Current paving, t	-0.002 (0.001)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Deposit discoveries, t-1	-0.022 (0.132)	-0.013 (0.022)	0.015 (0.013)	0.002 (0.013)
Lagged neighbors' regional average of democracy	0.799 (0.115)			
Country FE	Yes	Yes	Yes	Yes
Period FE	No	No	No	No
Country-specific trends	No	No	No	No
1st stage F-test	43.615			
Observations	127	127	127	127
R-squared		0.106	0.132	0.053

Note: This table shows the first stage regression of a 2SLS specification where we instrument democracy with its lagged regional average, counting only *neighboring* countries as a country's region, where neighbors also include countries that are not part of the regression sample. Columns 2-4 are the second stage regressions for three different dependent variables. All specifications include country fixed effects. Robust standard errors in parenthesis. Table A1 contains summary statistics. See Section 3 for variable definitions and sources.

Table OA6: Instrumenting democracy with regional waves of democratization: additional controls

Dependent variable →	Democracy, start of t	City bias	Deposit overlap	City overlap
	1st stage	2nd stage		
	(1)	(2)	(3)	(4)
Democracy, start of t		0.052 (0.024)	-0.046 (0.013)	0.006 (0.021)
Cum. deposit overlap, start of t	9.119 (3.614)	0.351 (0.466)	0.161 (0.292)	0.512 (0.362)
Cum. city overlap, start of t	-15.669 (3.958)	-0.700 (0.620)	0.053 (0.395)	-0.647 (0.476)
Cum. deposit discoveries, start of t	0.044 (0.136)	0.003 (0.017)	0.007 (0.010)	0.010 (0.013)
Current paving, t	-0.001 (0.001)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Deposit discoveries, t-1	-0.015 (0.242)	-0.009 (0.021)	0.019 (0.013)	0.010 (0.016)
Price index, average t	-0.009 (0.003)	-0.000 (0.000)	-0.000 (0.000)	-0.001 (0.000)
Log aid, average t	1.321 (0.699)	0.198 (0.077)	0.029 (0.064)	0.227 (0.069)
Log FDI, average t	0.133 (0.260)	-0.111 (0.034)	0.051 (0.031)	-0.060 (0.026)
Civil war, average t	0.193 (2.342)	-0.529 (0.365)	0.989 (0.224)	0.460 (0.234)
Lagged regional average of democracy	1.448 (0.236)			
Country FE	Yes	Yes	Yes	Yes
Period FE	No	No	No	No
Country-specific trends	No	No	No	No
1st stage F-test	33.404			
Observations	87	87	87	87
R-squared		0.290	0.265	0.285

Note: This table shows the first stage regression of a 2SLS specification where we instrument democracy with its lagged regional average. Columns 2-4 are the second stage regressions for three different dependent variables. All specifications include country fixed effects. Robust standard errors in parenthesis. Table A1 contains summary statistics. See Section 3 for variable definitions and sources.

Table OA7: Redefining counterfactuals: market potential, distance, and size

Dependent variable → Location ranking method →	City bias		
	market potential	distance	size
	(1)	(2)	(3)
Democracy, start of t	0.038 (0.017)	0.027 (0.015)	0.026 (0.017)
Cum. deposit overlap, start of t	1.862 (0.626)	0.893 (0.524)	1.547 (0.561)
Cum. city overlap, start of t	-2.054 (0.638)	-1.466 (0.601)	-1.729 (0.600)
Cum. deposit discoveries, start of t	-0.018 (0.012)	-0.018 (0.011)	-0.017 (0.011)
Current paving, t	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Deposit discoveries, t-1	-0.002 (0.024)	0.002 (0.020)	0.002 (0.021)
Country FE	Yes	Yes	Yes
Period FE	Yes	Yes	Yes
Country-specific trends	Yes	Yes	Yes
Observations	127	127	127
R-squared	0.533	0.535	0.479

Note: This table shows OLS regressions of the effect of democracy on city bias, where we change its definition. The locations used to calculate city bias (cities, deposits and ports) are ranked in three different ways: on the basis of market potential (the baseline), on the basis of distance only (column 2), and on the basis of size (column 3) where we use population for cities and the size index for deposits and ports. All specifications include country and period fixed effects and country-specific linear trends. Robust standard errors in parenthesis. Table A1 contains summary statistics. See Section 3 for variable definitions and sources.