STEG WORKING PAPER

# PRIORITY ROADS: THE POLITICAL ECONOMY OF AFRICA'S INTERIOR-TO-COAST ROADS

Roberto Bonfatti, Yuan Gu, and Steven Poelhekke

OCTOBER 2022
STEG WP031

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

Roberto Bonfatti[†a], Yuan Gu[‡b], and Steven Poelhekke[§c]

[a]University of Padova and University of Nottingham

[b]APG Asset Management, Tinbergen Institute

[c]Vrije Universiteit Amsterdam, CEPR, and CESifo

September 27, 2022

**Abstract**

Africa's interior-to-coast roads are well placed to export natural resources, but not to support regional trade. Are they the best response to geography and comparative advantage, or the result of political distortions? To answer this question, we investigate the political determinants of road paving in West Africa in 1965-2014. Controlling for geography and comparative advantage, we find that autocracies more than democracies focused on connecting metal and mineral deposits to ports, resulting in more interior-to-coast networks. This deposit-to-port bias is driven by deposits located on the ruling elite's ethnic homeland. This suggests that Africa's interior-to-coast roads were at least in part the result of ethnic favoritism.

**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 thank Alexander Moradi and Hervé Gazel for kindly sharing their data, and Xuying Ma and Gian Luca Tedeschi for excellent research assistance. 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 Auckland, Nottingham, Tilburg, Rome Tor Vergata, St Andrews, Vrije Universiteit Amsterdam, Wageningen and Waikato, and conference participants at ETSG, CSAE, and the CEPR STEG workshop on Political Economy and Public Investment. All remaining errors are our own.

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

‡Email: yuan.gu@apg-am.nl.

§Corresponding author. Contact information: Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands. Email: steven.poelhekke@vu.nl

# 1 Introduction

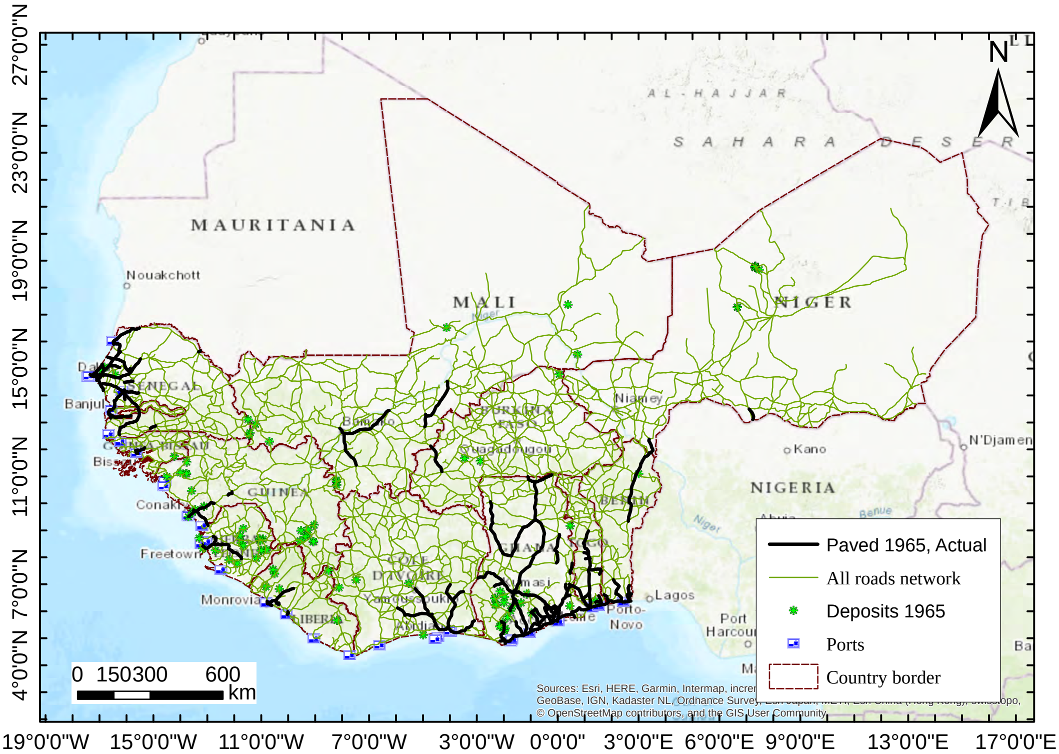
According to the institutional view of development, inclusive political institutions are required for sustained economic growth, as only these will deliver the public goods and market-supporting infrastructure that make growth possible. In contrast, extractive political institutions will focus on enriching the ruling elite and on making it hard for other groups to emerge (e.g. Acemoglu and Robinson, 2012). Indeed, Acemoglu et al. (2019) find that democracy has a positive effect on GDP per capita growth. One specific type of market-supporting infrastructure is the network of paved roads. In this paper, we empirically investigate the link between political institutions and the construction of such a network.

We focus on West Africa in the post-colonisation period (1965-2014). One striking fact about this region is that, following on a colonial pattern, paved roads have evolved mostly in an *interior-to-coast* direction (see Figure 1),[1] 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 1964, p. 23; Rodney 1972, 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 the actual networks should be compared with. Then, it is hard to rule out that the actual networks are the best response to the geography or comparative advantage of the West African countries.[2]

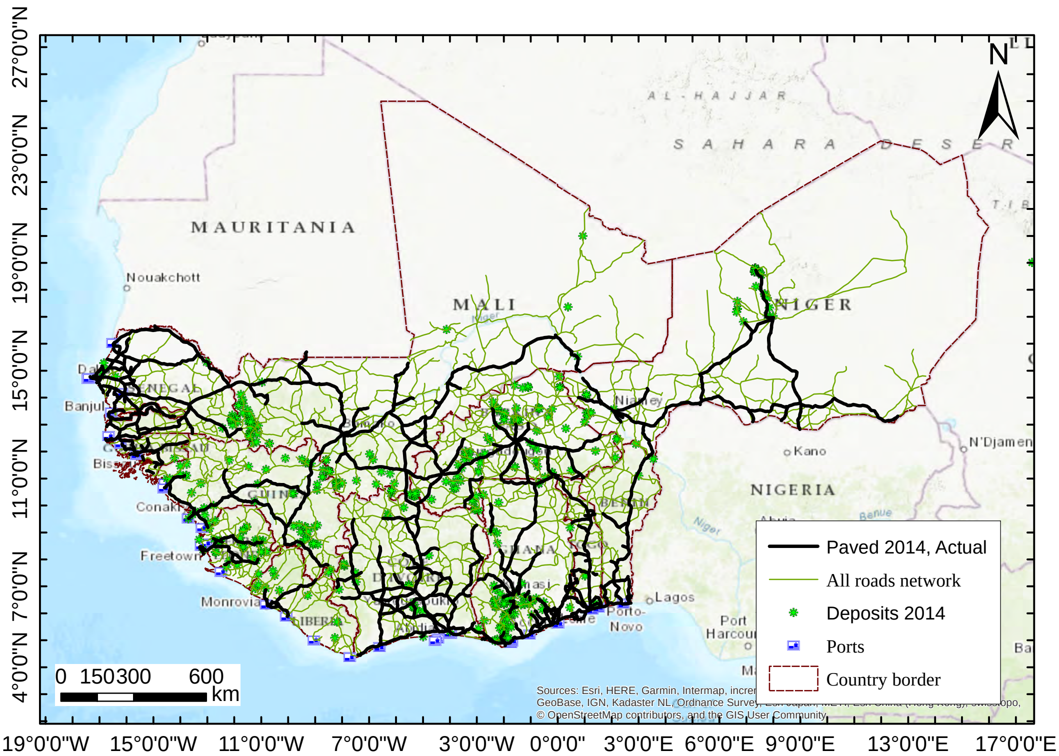
To make progress, we take an indirect approach and investigate the political circumstances under which these networks were built. Controlling for country fixed effects, and thus for geography and time-invariant comparative advantage, we document two facts. First, the West African paved road networks expanded in a more interior-to-coast fashion in periods of autocracy, relative to periods of democracy. In particular, for a given number of kilometers paved, autocracies relative to democracies focused more on paving connections between metal and min-

[1] In 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).

[2] When one draws the *entire* road networks of these countries (including unpaved roads and tracks), these do not look predominantly interior-to-coast. This casts doubt on the geography hypothesis, though of course geography could matter differently for paved and non-paved roads.




1965




2014

Figure 1: West African roads, paved and upaved, 1965 and 2014

eral deposits and ports. Since a substantial share of deposits are located in a country's interior,[3] this resulted in more interior-to-coast connections being built under autocracies. These findings are robust to controlling for proxies for time-varying comparative advantage (such as time fixed effects, measures of the stock of deposits, and country-specific commodity price indexes) and to instrumenting for autocracy as suggested by the most recent literature (using lags in space and time as in Acemoglu et al. 2019).

Second, we show that the *deposit-to-port bias* of autocracies was stronger, the more the deposits were located on the ruling elite's ethnic homeland. Indeed, in the case in which most of the deposits were located outside these homelands, the deposit-to-port bias disappears altogether: autocracies did not pave any more deposit-to-port roads than democracies. In other words, it was only when the ruling ethnic groups "owned" enough of the deposits, that autocracies appear to have had a greater appetite for deposit-to-port roads.

Seen through the lens of the institutional view, these results have a natural interpretation. Extractive political institutions, facing fewer constraints than inclusive ones, tend to select public policies which are more biased in favour of the ruling elite. But in ethnically fragmented societies such as those of West Africa - where the ruling elite prosper with the support of their ethnic groups (Padró i Miquel 2007) - this must result in stronger ethnic favoritism under extractive institutions than under inclusive ones. Indeed, in their study of road paving in Kenya that focuses on cities, Burgess et al. (2015) found that autocratic governments engaged in more ethnic favoritism than democratic ones.[4] After netting out geography and comparative advantage, deposit-to-port roads can be seen as targeted policies: they benefit the ethnic groups on whose territory the deposits are located, more than the rest of society.[5] Hence, you would expect deposit-to-port bias to be stronger when political institutions are extractive, and the deposits are located on the ruling elite's ethic homeland. This is precisely what we find.

In this interpretation, West Africa's interior-to-coast roads are at least in part the result of ethnic favoritism. Given the region's pervasive ethnic fragmentation, and its history of fre-

[3]At the start of our sample (1965), 36% of deposits were closer to the country's centroid than to the coast, with those closer to the coast being 145 kilometers inland on average. These numbers have increased to respectively 43% and 195 kilometers by the end of our sample (2014).

[4]Similar results are found by Hodler & Raschky (2014) and De Luca et al. (2018) for a large sample of countries.

[5]One might object that deposit-to-port roads should be valuable to an autocratic government independently on where they are located, since the government can appropriate the resource-revenues they generate through its control of the state apparatus. There are however several reasons, discussed in Section 5.3, to believe that parts of the value of these roads is only captured locally.

quent autocratic rule, we think this might be an important determinant of the interior-to-coast expansion of the West African networks since independence.

Our analysis is based on a newly-assembled dataset of roads, railways, metal and mineral deposits, ports and cities covering 12 West African countries over the period 1965-2014.[6] The data on roads and railways was obtained by digitising 23 successive editions of Michelin road maps of West Africa.[7] 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 that may overestimate the actual number of connections between grids, we take into account the actual links between roads, and digitally recreate the complete network and its evolution over time.[8] 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.[9]

Our first challenge is to measure the extent to which paving in any given country and period displayed a deposit-to-port bias - defined as the extent to which it focused on connecting metal and mineral deposits to port, as opposed to any other pair of locations. Building on the richness of our road data, we proceed in two steps. In the first step, for each country in our sample, we construct a counterfactual road expansion path, inspired by the work of Burgess et al. (2015) on Kenya. The counterfactual aims to approximate the way in which the paved roads would have expanded in 1965-2014, had the government only cared about connecting deposits to ports. More specifically, we take all deposit-port pairs that are not yet fully connected by paved roads, and rank them according to their "market potential" (the sum of port size and deposit value, divided by bilateral distance). In each period, we then re-allocate the actually paved kilometers to connecting the deposit-port pairs that are at the top of the ranking. Not surprisingly, in a

[6] We focus on roads and not on railways primarily because of lack of variation: almost no West African railway has been extended in our period of study. Furthermore, many of them have fallen in a derelict state. Nevertheless, in some of our specifications, we allow for the possibility that existing railways act as substitutes for roads.

[7] The map scans were kindly shared by Alexander Moradi, Free University of Bozen-Bolzano.

[8] We add a novel dataset to the literature: A comparable dataset based on the same underlying maps that covers more African countries is available from Jedwab & Storeygard 2022, but its units are 0.1 by 0.1 degree grids (of about 11 by 11 km) with dummies for road presence and quality. For our purposes, the precise shape of the network that tracks all actual connections is important, while adjacent grids may for example assume links where there are none, such as between parallel roads, across rivers, railways, mountain ranges or national borders.

[9] One noteworthy aspect of our approach is that although we have data at the level of individual locations (deposits), we then aggregate this up so that our unit of observation in the regressions is a country-period. This approach is adopted because 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 fashion). Clearly, the shape of a national road network is best investigated at the country level.

typical country and period, the counterfactual looks very interior-to-coast in shape. In the second step, we calculate the extent to which real paving in any given country and period overlapped with the counterfactual network for that country at the start of that period. Such country-specific and time-varying measure of overlap is a measure of deposit-to-port bias, since a large overlap denotes a type of paving that focused a lot on connecting deposits to ports.

We regress our measure of deposit-to-port bias on the level of autocracy in the country at the beginning of the period, controlling for country fixed effects, time fixed effects, country-specific trends, and a vector of country-level, time-varying controls. Our first result is that autocracies displayed a stronger deposit-to-port bias. This effect is large, with a one-standard deviation increase in autocracy resulting in half of a standard deviation increase in deposit-to-port bias. The effect is robust to instrumenting for autocracy using a lagged measure of autocracy in a country's region, as suggested by the most recent literature (Acemoglu et al. 2019). Moreover, it survives a large number of robustness checks, including measuring the deposit-to-port counterfactual in different ways and dropping countries one by one.

Having established that autocracies had a greater deposit-to-port bias than democracies, we next turn to consider possible mechanisms. Our results are not driven by autocracies paving more kilometers than democracies, by them being more in favor of trade openness, by them receiving more or less foreign aid or direct investment, by them being more involved in civil wars, or by the possibility that deposits were close to big cities, which autocracies might have had a bias for. In contrast, we find evidence that autocracies only had a deposit-to-port bias when a large enough share of the deposits to be connected were located on the ruling elite's ethnic homeland. As explained above, this suggests that West Africa's interior-to-coast roads are at least in part the result of ethnic favoritism.

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.[10] Second, internal trade costs remain very high in Africa,[11] suggesting that a great expansion of the African networks will be required in the future. Third, there are conflicting views on the type of links that such expansion should

[10] Jedwab & Storeygard (2022), p. 3, report that transport accounted for 14% of World Bank lending, and 22% of African Development Bank disbursements in 2012-2015.

[11] Atkin & Donaldson (2015) find that internal trade costs in Ethiopia and Nigeria are four to five times larger than in the US. Porteous (2019) finds slightly higher values for a sample covering all 42 Sub-Saharan African countries.

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.[12] On the other hand, China has been funding projects that essentially reinforce the existing interior-to-coast links (Bonfatti & 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 (Porteous 2019).

Our paper is most related to a small but growing literature on the political determinants of transport infrastructure. Burgess et al. (2015) study ethnic favoritism in the allocation of Kenyan roads in 1963-2011. They find strong evidence of favoritism, which however attenuates under democracy.[13] For a sample covering 39 Sub-Saharan African countries over the same period, Jedwab & Storeygard (2022) find that leaders favoured road building near cities within 150km of their hometown.Our original contribution to this literature is to document the impact of ethnic favouritism on the interior-to-port shape of the African networks, something which the literature on transport infrastructure has neglected but has been the subject of a long-standing debate in development economics.

This paper builds on earlier work by Bonfatti & Poelhekke (2017). They show that coastal African countries endowed with more deposit-to-port roads also feature national trade costs that are strongly biased in favor of trade with overseas countries, relative to neighboring countries. This effect is reversed for landlocked countries, presumably because the deposit-to-port roads can also be used to trade with transit neighbors. In conjunction with those earlier results, our current results suggest that the deposit-to-port bias associated with ethnic favoritism has had a large and long-lasting effect on the direction of trade in the African countries, and specifically one that has penalised regional integration.[14]

[12]The 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 et al. (2006), 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, Porteous (2019) 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.

[13]Alder et al. (2018) obtain results for China which resemble those of Burgess et al. (2015).

[14]In unreported regressions, we find no evidence that autocracies paved more kilometers in total. This suggests that autocracies' deposit-to-port bias resulted in them paving fewer roads of a different kind.

The paper is structured as follows. Section 2 describes the construction of our dependent variable (a measure of deposit-to-port bias), and Section 3 presents the empirical approach. Section 4 lists the data sources. Section 5 presents the main results, with additional results provided in Section 6 and in the Online Appendix. Section 7 concludes.

# 2 Construction of the dependent variable

To construct our dependent variable - a measure of the sitting government's preference for connecting metal and mineral deposits to ports, as opposed to any other pairs of locations - we proceed in two steps. First, for each country in our sample, we construct a counterfactual road expansion path. This counterfactual aims to approximate the way in which paved roads would have expanded in 1965-2014, had the government only cared about connecting deposits to ports. Our measure of deposit-to-port bias, which we construct in the second step, will be a measure of how close the actual expansion path was to the counterfactual. The two steps are described in sections 2.1 and 2.2 respectively.

## 2.1 A counterfactual road expansion path

For each country in our sample, we construct a counterfactual road expansion path inspired by the work of Burgess et al. (2015) for Kenya. As they do, 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 Michelin maps (we have 23 such maps, published in 1965-2014 and covering all countries in our sample).[15] We then re-allocate the quota of each period to paving different roads in that period, proceeding period after period as explained below. In other words, we imagine that the government would follow a different expansion path to the one it actually followed, leading to a different, counterfactual network by 2014.

Burgess et al. (2015) construct a counterfactual network expansion path which is meant to approximate the way in which paved roads would have expanded in Kenya since 1964, had the government only cared about connecting *cities* to one another. They do so by ranking all possible city pairs according to market potential (the sum of the cities' size, divided by their distance along

[15] A list of map years is reported in Section 4.1.

existing 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).[16] To take the topography into account, they construct the counterfactual by only paving unpaved roads along the shortest route between city pairs in the initial (1964) network. They can do so since all the roads paved during their period were initially unpaved, and 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 similar way, except that our counterfactual connects deposit-port pairs instead of city pairs. For each country independently, we prepare a list of all possible deposit-port pairs. We then rank them 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.[17] "Market potential" of a pair is defined as the sum of indices of deposit value and port size, divided by the distance along existing roads between them.[18] Large deposits are thus connected to large ports first, unless a deposit is very remote compared to smaller ones. We measure deposit value by multiplying the estimated size at discovery with the average price of the commodity being produced over our sample period.[19] We measure port size by their depth, on the logic that deeper ports can accommodate larger ships with more draft. 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. In case any unused quota remains 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.[20]

We start with deposits that had been discovered by 1965 and let additional deposits enter the ranking from the year that they were discovered, which shakes up the ranking over time through their location and size. We use year of discovery, as opposed to first year of production,

[16]When paving a connection between a pair, they start paving from the largest city in the pair.

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

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

[19]We compute four value categories by dividing the value distribution into four quartiles. See Section OA2.3 for more details.

[20]In robustness checks, we control for the total amount of paving in each period.

for two good reasons. First, the latter info is only available for 40% of our deposits. Second, the need for a paved road connection logically arises immediately after discovery. Indeed, the start of production may well be the effect, rather than the cause, of a paved road connection being set in place.[21]

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 2014. 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 2014.[22]

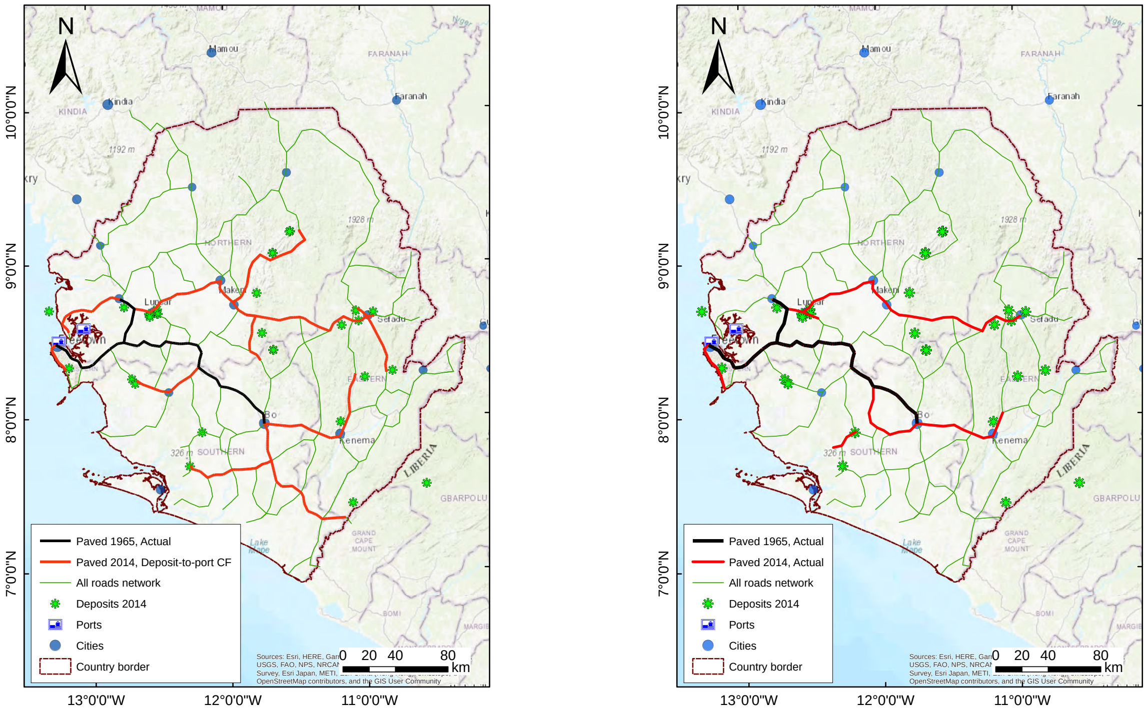
The deposit-to-port counterfactual for one country and period - Sierra Leone in 2012-13 - is illustrated in Figure 2, Panel I. In the panel, deposits are denoted by green asterisks, ports (including those on navigable rivers) by blue squares, the initial (1965) network by black lines, and the deposit-to-port counterfactual at the end of the period by red lines. For comparison, Panel II reports the actual network in 2014. As can be appreciated, the counterfactual network is quite similar to the actual one, suggesting that road paving in Sierra Leone has been placing a lot of emphasis on connecting deposits to ports.

In our baseline specification, we do not let the counterfactual network deteriorate over time, but this is something we allow for in robustness checks. Deterioration did happen in some countries and periods - perhaps as a result of lack of maintenance, or of traumatic events such as civil war - this may result in the counterfactual network being longer than the actual one.[23] The case of Sierra Leone, which we have just discussed, well illustrates this point. In robustness checks, described in Section 6.2, we consider several alternative ways to construct the counterfactual network, which take the issue of deterioration into account.

[21]For deposits for which this info is available, start of production came 6 years after discovery at the median (11 years on average).

[22]When no road of any quality exists to connect a deposit and a port (not even a track) we artificially connect the deposit to the road network by taking the shortest straight-line distance from the deposit location to the nearest road.

[23]It happens in 43% 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.




I: counterfactual, end of 2013

II: actual, end of 2013

Figure 2: Sierra Leone, deposit-to-port counterfactual paved road network (panel I) and actual paved road network, end of 2012-13 period (panel II)

## 2.2 The *deposit-to-port bias*$_{i,t}$ measure

Our dependent variable - a measure of deposit-to-port bias in country $i$ and period $t$ - is the share (in length) of roads paved in country $i$ and period $t$ which overlaps with the deposit-to-port counterfactual network at the end of period $t$. The intuition behind this definition is as follows. The counterfactual network at the end of period $t$ approximates the set of roads that a planner would have paved up to that point in time, had they cared only about connecting deposits to ports. By measuring the overlap between the roads actually paved in period $t$ and the above-mentioned set, we measure the extent to which paving decisions in period $t$ attempted to make the actual road network closer to the deposit-to-port counterfactual. We take this to be a measure of the sitting government's preference for connecting deposits to ports, as opposed any other pairs of locations.

Formally, the overlap between roads paved and the deposit-to-port counterfactual is calculated as follows (since this procedure applies to all countries, we omit country subscripts in the rest of this section). Let $A_{t,real}$ denote the actual road network in place at the end of period $t$, and $A_{t,cf}$ the counterfactual deposit-to-port network as it would have appeared at the end of the same period. 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 $t$.[24] Dividing the length of this overlap by the length of actual paving gives us the fraction of the actual paving in period $t$ that overlaps with the deposit-to-port counterfactual at the end of period $t$.

In summary, the deposit-to-port bias measure can be written as

$$\textit{deposit-to-port bias}_t = \frac{\mathcal{L}\left\{(A_{t,real} \setminus A_{t-1,real}) \cap A_{t,cf}\right\}}{\mathcal{L}\left\{A_{t,real} \setminus A_{t-1,real}\right\}}, \tag{1}$$

where the operator $\mathcal{L}\{S\}$ denotes the total length of roads contained in set $S$. Such measure ranges between 0 and 1, where a score of 1 denotes maximum deposit-to-port bias.

The construction of $\textit{deposit-to-port bias}_{i,t}$ can be illustrated using Figure 3. Let area $A_{0,real} = a$ represent the initial (1965) paved road network. Suppose 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 $b + c + d$. 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 a priority under the actual planner, but would have been a priority under the counterfactual planner. The overlap between the actual paving in period 1 and the counterfactual at the end of the period is then $d$, and $\textit{deposit-to-port bias}_1$ is equal to $d/(b+c+d)$. Suppose again that by the end of period $t = 2$ (Dec 31, 1968),[25] 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 $g + h + e$. 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 $A_{2,cfp} = a + c + d + e + f + h + i$. The overlap in our example is then $e + h$, and $\textit{deposit-to-port bias}_2$ is equal to $(e + h)/(e + g + h)$.[26]

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

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

[26] We could have alternatively considered as overlap the intersection of the actual and counterfactual *additional* paving, in which case the overlap in period 2 would be only area $h$, because area $e$ should have been paved one

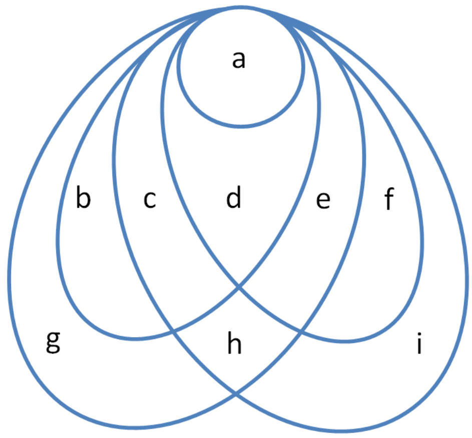
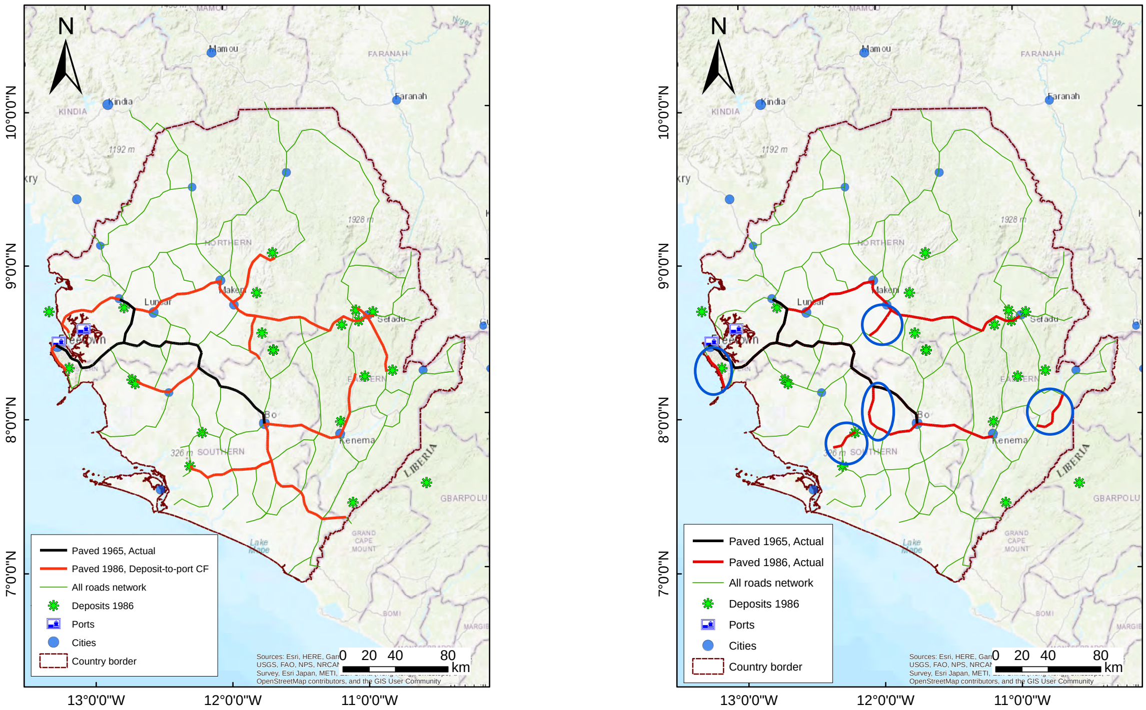



Figure 3: Construction of the deposit-to-port bias measure

Figure 4 further illustrates this procedure for the case of Sierra Leone in 1984-85 (period 10 in our sample). Panel I shows the deposit-to-port counterfactual as it appeared at the end of that period. 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. Panel II shows the actual paved network at the end of the period, with the roads paved in 1984-85 circled in blue. These circled additions measure a total of 158 kilometers in length. The measure of overlap for period 10, $\textit{deposit-to-port bias}_{10}$, compares the circled additions to the counterfactual network in Panel I, and finds that only one of them (south of Freetown) has a section overlapping with the counterfactual network. Dividing the length of this section by 158, we obtain $\textit{deposit-to-port bias}_{10} = 0.09$. This number means that only 9% of the roads that were paved in 1984-85 would have been paved in that period or earlier, had the government only cared about connecting deposits to ports.

It should be noted that the absolute value of our measure of deposit-to-port bias has no clear meaning. Thus, all we are going to be focusing on is changes in this measure over time (and thus over different types of government) within each country.

period earlier. However, this approach betrays the fact that area $e$ *was* paved in period 2, and it overlaps with the deposit-to-port 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-port 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.




I: counterfactual, end of 1985 II: actual paving, 1984-85

Figure 4: Sierra Leone, deposit-to-port counterfactual paved road network (panel I), end of 1984-85 period, and actual paving, 1984-85 period (panel II). The red lines in Panel II indicate the actual paved road network at the end of the 1984-85 period.

## 2.3 Summary statistics on the dependent variable

Table OA1 describes the distributions of $deposit\text{-}to\text{-}port\ bias_{i,t}$ across country-periods in our full sample. Our full sample is made of 12 countries which we observe for 22 periods, giving us 264 country-periods. However in country-periods in which no actual paving occurred, our dependent variable is zero by construction. To distinguish this kind of zeros from the the ones denoting a planner who paved roads that did not overlap with the deposit-to-port counterfactual, we drop the former type of zeros from the sample. We are thus left with 174 country-periods, which is also the maximum number of observations in our regressions.[27]

One might worry about deposits being disproportionately discovered along pre-existing paved roads. That however should compress $deposit\text{-}to\text{-}port\ bias_{i,t}$ towards zero, since it would make it more likely for the deposit-to-port counterfactual to be paved already. Reassuringly, Table 1 shows that $deposit\text{-}to\text{-}port\ bias_{i,t}$ has substantial average and standard deviation, at 0.26 and 0.34 respectively. That indicates that, in the average country-period, 26% of the kilometers that

[27]We get to 174 after dropping 10 observations for which no measure of autocracy is available.

Table 1: Summary statistics

| Variable | N | mean | s.d. | min | max |
|---|---|---|---|---|---|
| Autocracy, start of t | 164 | 3.43 | 5.35 | -8.00 | 9.00 |
| Autocracy, start of t (Acemoglu) | 154 | -0.16 | 0.36 | -1.00 | 0.00 |
| Deposit-to-port bias | 164 | 0.26 | 0.34 | 0.00 | 1.00 |
| Deposit-to-port bias, with diamonds | 164 | 0.26 | 0.34 | 0.00 | 1.00 |
| Deposit-to-port bias, allowing for deterioration (0.5) | 164 | 0.25 | 0.33 | 0.00 | 1.00 |
| Deposit-to-port bias, allowing for deterioration (1.0) | 164 | 0.24 | 0.32 | 0.00 | 1.00 |
| Deposit-to-port bias, main ports double weight | 164 | 0.26 | 0.34 | 0.00 | 1.00 |
| Deposit-to-port bias, excl. deposits near railways | 164 | 0.21 | 0.25 | 0.00 | 1.00 |
| City-to-city bias | 164 | 0.33 | 0.34 | 0.00 | 1.00 |
| Cumulative deposit overlap, end of t | 164 | 0.51 | 0.20 | 0.16 | 0.90 |
| Cumulative city overlap, end of t | 164 | 0.62 | 0.16 | 0.31 | 0.97 |
| Cum. deposit discoveries, start of t | 164 | 16.46 | 16.91 | 1.00 | 81.00 |
| Deposit discoveries, t-1 | 164 | 1.24 | 2.55 | 0.00 | 18.00 |
| Current paving (km) | 164 | 185.93 | 240.60 | 6.29 | 1794.12 |
| Current deterioration (km) | 164 | -47.53 | 87.48 | -622.60 | 0.00 |
| Price index, average t | 164 | 1.9e+10 | 5.0e+10 | 0.00 | 3.4e+11 |
| Log aid, average t | 164 | 18.61 | 1.21 | 15.36 | 21.17 |
| Log FDI, average t | 137 | 2.96 | 2.12 | -4.61 | 8.09 |
| Civil war, average t | 164 | 0.03 | 0.16 | 0.00 | 1.00 |
| Openness to trade, average t | 146 | 0.28 | 0.44 | 0.00 | 1.00 |
| Ethnic deposits | 164 | 0.43 | 0.42 | 0.00 | 1.00 |
| Lagged regional average of autocracy | 164 | 3.78 | 3.11 | -2.83 | 6.59 |

Note: This table provides summary statistics for all variables used in the analysis. See Section 4 for variable definitions and sources.

were newly paved in the period overlapped with the deposit-to-port counterfactual. Note also that *deposit-to-port bias*$_{i,t}$ has maximum and minimum values of 1 and 0, indicating that there are cases in which the new paving overlaps perfectly with the deposit-to-port counterfactual, and cases in which it does not overlap at all. This suggests that different governments in different country-periods took substantially different approaches to road paving, at least according to our measure.

Online Appendix Tables OA3 and OA4 provide country-by-country and period-by-period averages. One potential worry is that in "elongated" countries such as Benin or Togo the newly paved roads may have to necessarily overlap with the deposit-to-port counterfactual, which by construction has an interior-to-coast shape. That is not the case, however. Online Appendix Table OA3 shows that while *deposit-to-port bias*$_{i,t}$ was large on average in Benin (0.43), it was low on average in Togo (0.14). More importantly, the standard deviation (not reported in the table) was high in both countries, at 0.46 and 0.30 respectively. This suggests that there were periods in which the newly paved roads overlapped a lot with the deposit-to-port counterfactual, but also periods in which they overlapped little.

Finally, Online Appendix Table OA5 delves into the case of one country, Sierra Leone. In 7 out of 22 periods, no actual paving occurred in Sierra Leone, necessarily leading to $\textit{deposit-to-port bias}_{i,t}$ being equal to zero. Of the 15 periods with positive paving, 8 report a positive value for $\textit{deposit-to-port bias}_{i,t}$. Importantly, there is substantial variation in $\textit{deposit-to-port bias}_{i,t}$ over time, suggesting that different governments had different paving priorities.

# 3 Empirical approach

Our baseline specification tests for the hypothesis that autocracies, relative to democracies, focused more on connecting metal and mineral deposits to ports, as opposed to any other pairs of location. We run

$$\textit{deposit-to-port bias}_{i,t} = \beta\, A_{i,t} + \gamma\, X_{i,t} + f_i + f_t + trend_i + \epsilon_{it}, \qquad (2)$$

where $\textit{deposit-to-port bias}_{i,t}$ is our measure of deposit-to-port bias in country $i$ and period $t$ (as described in the previous section), $A_{i,t}$ is a measure of autocracy at the start of period $t$, $f_i$ and $f_t$ are respectively country and period fixed effects, $trend_i$ are country-specific linear trends, and $X_{i,t}$ denotes a vector of country-level, time-varying controls.[28] We focus on two-way fixed effect regressions as opposed to analyzing 'events' of democratization or its reverse, because countries often change slowly from autocracy to democracy and several change regime more than once.[29]

Our choice of periods is dictated by the years of publication of the Michelin maps of West Africa, which we use to detect the paving of roads. 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 23 maps published between 1965 and 2014, which results in 22 irregular time periods with an average length of 2.23 years (and a minimum and maximum length of 1 and 7 years). For example, our first three maps were published in 1965, 1967 and 1968. Thus, our first period

[28]We estimate our main results in levels and allow for deterministic trends by including $trend_i$. In Online Appendix Section OA4 we test for and reject non-stationarity of autocracy (since all countries have become more democratic over time) and re-estimate our main results in first differences.

[29]Online Appendix Figure OA1 shows the evolution of autocracy over time for each country.

spans 1965 and 1966, while our second period spans 1967 only.[30] A full list of map years and periods is provided in footnote 34.

Our main measure of autocracy, $A_{i,t}$, will be the inverse of the Polity IV index of democracy, which is available yearly. Thus, our measure will be equal to 10 for the strongest autocracy, and -10 for the weakest one (or the strongest democracy).

In our baseline specification, the vector $X_{i,t}$ contains three types of variables, the first of which controls for the extent to which a country's deposits have already been connected to ports. One potential worry is that $\textit{deposit-to-port bias}_{i,t}$ will be mechanically lower when a greater share of the counterfactual has already been paved, and this may confound our results.[31] To rule this out, we define an additional variable, $\textit{cumulative deposit overlap}_{i,t}$, which measures the share of the counterfactual that is actually paved at the start of period $t$, and always control for it in our specifications.

The second type of variables proxy for time-varying comparative advantage. These variables include the stock of mineral and metal deposits at the start of the period, the number of deposits discovered in the previous period, and country-specific, time-varying indexes of the world prices of metals and minerals exported by each country (more details on the construction of such index are provided in the Online Appendix).

The third type of variable controls for the extent to which the deposits to be connected were located on the ethnic homeland of the ruling elite. Previous research has uncovered the importance of ethnic favouritism in shaping road paving in Kenya (Burgess et al. 2015). To the extent that the autocracy happens to be correlated with the "ethnic salience" of the deposits to be connected, then our estimates of the effect of autocracy on road paving might be biased. To address this concern, we construct a new variable capturing the extent to which the deposits to be connected are located on the ruling elite's ethnic homeland. We begin by counting the total number of yet unconnected deposits located on the deposit-to-port counterfactual at the end of period $t-1$ (or equivalently, at the start of period $t$). These are the deposits that are not yet connected by a paved road at the end of period $t-1$, but should be connected according to the deposit-to-port counterfactual. The variable $\textit{ethnic deposit}_{i,t-1}$ measures the weighted (by

[30]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.

[31]For example, if the deposit-to-port counterfactual is entirely paved at the start of period $t$, then $\textit{deposit-to-port bias}_{i,t}$ must necessarily be zero.

deposit size) share of these deposits that are located on the ethnic homeland of the ruling elite at the end of period $t-1$.[32] We control for $\textit{ethnic deposit}_{i,t-1}$ in our baseline specification.

Finally, we will also run a version of (2) in which we interact $A_{i,t}$ with $\textit{ethnic deposit}_{i,t-1}$. This is to test for the hypothesis according to which, if autocracies had a stronger deposit-to-port bias than democracies, they did so because they focused more on enriching the ruling ethnic groups.

# 4 Data sources

We construct a panel covering 12 West African countries over 22 time periods between 1965 and 2014, 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 do not include Gambia and Nigeria, even though they are part of West Africa, because of lack of variation in deposit discoveries.[33]

In this section, we briefly describe the data, focusing on the most original ones (those on roads and deposits). A full description of the data is provided in the Online Appendix.

## 4.1 Roads

We have digitised 23 Michelin West Africa road maps covering the 12 West African countries listed at the start of this section, between 1965 and 2014. A "period", in our dataset, is defined as the time lapsing between two consecutive maps. 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 periods to start at the beginning of each publication year. For example, the first two maps were published in 1965 and 1967: our first period then starts on January 1, 1965 and ends on December 31, 1966. This gives us 22 periods for each country, for a theoretical maximum of 264 observations.[34]

[32] Results are robust to using the simple share. A deposit is coded to be on the elite's homeland if the group on whose homeland it is located is classified by the Ethnic Power Relationship (EPR) dataset as being "dominant", "senior partner" or "junior partner" in the national government, while it is not coded as such if the group is "powerless", "discriminated" or "irrelevant".

[33] In Nigeria, we could only identify the year of discovery for 4 out of 22 deposits, and these were all discovered before 1965, leaving no variation within our sample period. Similarly, Gambia has one deposit with unknown discovery date and is thus also excluded.

[34] The map editions are 1965, 1967, 1968, 1969, 1971, 1973, 1976, 1983, 1984, 1986, 1988, 1989, 1990, 1991, 1996, 1998, 2000, 2002, 2003, 2007, 2009, 2012 and 2014. Although we also have some earlier years, we start in

Roads are classified along two dimensions: their regional importance in terms of major, secondary and local roads, and their quality. In our analysis, we use all roads (major, secondary and local), and focus on the quality classification. There are six quality 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 other four types of worse-quality roads are also unpaved and not all-weather roads. We classify as "paved" the surfaced roads, and as "unpaved" all other roads.

In Appendix Table OA2 we track the evolution of roads in West Africa by comparing their quality at the start and end of our sample (1965 and in 2014). Most roads that were paved in 1965 were still paved in 2014 (9,496 out of 10,665 kilometers), with only 11% of them deteriorating to unpaved status. It was relatively rare for roads that did not exist in 1965 to appear as paved in 2014. The vast majority (88%) of "new" paved roads in 2014 were existing unpaved roads in 1965.[35] It is also evident that paved roads are less prone to deterioration than unpaved roads.

We infer each 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 1967 map but unpaved in the 1965 map, then we conclude that it was paved during the 1965-66 period (although we do not know exactly when).

Appendix Tables OA3 and OA4 report, respectively, average paving across countries for each period in our sample, as well as average paving across periods for each country. In the average country-period, 175 kilometers of roads were paved, and 47 kilometers deteriorated to unpaved status. The latter number indicates that deterioration played an important role in our sample, at least relative to paving. In our analysis, we will thus conduct several robustness checks to account for this. Paving continued rather uniformly throughout our sample period, with low points in the early 1990s and early 2000s. The country that paved the least was Liberia (63 kilometers per period) whereas the country that paved the most was Mali (294 kilometers).

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

[35] Of the 3,026 kilometers of roads which appear as paved in 2014 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 2014, before becoming paved roads by 2014.

## 4.2 Deposits

Our deposit data is the combination of three sources: the proprietary datasets 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,[36] for which we know the location, the year of discovery and the size of the deposit.[37]

Table OA6 reports descriptive statistics on deposits. A large majority of them were discovered after 1965 (304 out of 391). All countries had multiple deposits discovered after 1965, except for Guinea-Bissau (which had none), and Togo (which had one). There are 14 different types of metal and minerals in the sample, of which gold is by far the most frequent, followed by iron ore, uranium, diamonds[38] and bauxite. Finally, most deposits are of intermediate size (categories 3 and 2 listed above).

We estimate deposit value by multiplying the estimated size at discovery with the average price of the commodity being produced over our sample period.[39] The rank order of prices hardly ever changes over the period (and when it does, it is for commodities with comparable prices, such as Manganese and Zirconium) suggesting that it make sense to use average prices to infer the relative value of the different commodities. We then compute four value categories by dividing the value distribution into four quartiles.[40]

## 4.3 Other data

The data on ports comes from the 2016 World Port Index (WPI). This provides several port characteristics, including channel and pier depths in 12 categories. We use depth categories as categories of size, on the logic that deeper ports can accommodate larger ships with more draft. To get rid of small fishing ports, we exclude ports with less than the 5 meters of draft required

[36]Per source, 321 records are from MinEx, 54 from RMG, and 8 from MRDS.

[37]According 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".

[38]We drop diamond deposits in the baseline estimation, on the logic that diamonds are less likely to be transported via ports and bulk ships than to be airlifted. We include diamonds in robustness checks.

[39]For deposits that contain more than one metal or mineral we use the primary metal or mineral. MinEx states: "Primary metal refers to the main metal (in value-terms) contained in the deposit. For consistency, the value of any by-product credits (such as silver and gold) in the primary deposit have been converted into metal-equivalent." 84% of deposits do not any have by-products. Only 8% of deposits have by-products other than silver and gold.

[40]See Online Appendix Section OA2.3 for more details.

to accommodate bulk ships built from the 1950s onwards (Waters et al. 2000, Figure 4.46). This leaves us with nine categories of depth. We collapse those categories into four, which we use as weights in the ranking of which ports to connect first to deposits. We include all deep water ports that exist in 2016, 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.

As a measure of autocracy, we use either (minus) the "Polity" score provided by the Polity IV Project, or alternatively (minus) the democracy dummy constructed by Acemoglu et al. (2019). The former is a widely used indicator, which in our formulation ranges from +10 (strongly autocratic) to -10 (strongly democratic). The latter was constructed by combining the Polity IV index with several other independent measures of democracy. There is substantial variation in the Polity score, both across countries and within countries (see Online Appendix Figure OA1).[41] Most of our countries are autocratic in the midst of the sample period, some of them having experienced a short period of democracy after decolonisation. They then become less autocratic between 1990 and 2010, but the path towards this transition is quite different across countries.

Data on the ethnicity of deposits comes from a combination of Murdock's Ethnolinguistic Map of Africa (Murdock 1959) and the Ethnic Power Relationship Core Dataset 2019 (Vogt et al. 2015; EPR from now on). The first dataset allows us to identify the ethnic group on whose homeland the deposit is located. The second dataset gives information on the political power of such ethnic group, in each year. We match ethnic groups in the two datasets, listing as politically irrelevant those groups which appear in the Ethnolinguistic Map, but not in the EPR dataset (since the latter does not include ethnic groups that were never politically relevant). We then consider a deposit to be located on the ruling elite's homeland if the group on whose homeland the deposit is located is at that point in time "dominant", "senior partner", or "junior partner", and otherwise if the group is "powerless", "discriminated" or "irrelevant".[42] We use the EPR dataset, as opposed to datasets on the birthplace or ethic group of the current political leader (see Hodler & Raschky 2014 and De Luca et al. 2018), in that we favor its broader approach, acknowledging that ethnic favoritism may extend to junior partners in government.

[41] 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.

[42] The former three categories include, respectively, 4,566, 5,935 and 15,242 observations (deposit-year). The latter include 5,202, 775 and 11,668 observations. There are also 124 observations that are coded as "state collapse" in the EPR dataset, and for which we set the dummy equal to zero.

Data on the prices of our 14 metal and mineral commodities comes from UNCTAD (downloaded June 14 2017), Plunkert & Jones (1999), and Bazzi & Blattman (2014). We construct a country-specific, time-varying price index in which, in each country, each (time-varying) price is weighted by the total value of deposits of the underlying commodity (relative to the total value of deposits) in the middle of our sample period (1989).[43] In other words, the only time variation in the index comes from exogenous changes in prices.

Finally, data on aid, FDI, Civil War and openness come from the World Bank Development Indicators, the Correlates of War (COW) Project ("Intra-State War Data", 2010 version) and Wacziarg & Welch (2008). Average yearly 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. 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). Openness is measured by a dummy variable which is equal to zero if a country is "closed" to the world economy according to a set of criteria, and one if it is open. The dummy is available for all our countries up to the year 2005. For eight countries, it is equal to zero until it turns one in a year between 1984 and 2001, and remains one afterwards. For the remaining four countries (Liberia, Niger, Senegal and Togo) it is zero throughout.

# 5 Main results

We begin by providing our baseline results (Section 5.1), and we then show that they are robust to instrumenting for autocracy using a standard approach (5.2). We conclude by illustrating the results through a case study, and by discussing our favourite interpretation in light of the existing literature (5.3).

## 5.1 Baseline estimates

Table 2 presents our baseline results. In column (1) we simply regress *deposit-to-port bias*$_{i,t}$ on autocracy at the start of $t$, including country and period fixed effects and country-specific time-trends. The sample includes 12 West African countries over 22 periods, but we only include

[43]See Online Appendix Section OA2.3 for an explanation of how the value of deposit is estimated.

country-periods when new paving took place, and data on autocracy is available (leaving us with 174 observations). We correct standard errors for possible heteroskedasticity, and report wild bootstrap t-statistics in the bottom rows that allow for serial correlation by country.[44] We find a significant positive effect of autocracy on $\textit{deposit-to-port bias}_{i,t}$. This suggests that autocratic governments, relative to democratic ones, focused more on connecting deposits to ports, as opposed to any other pairs of locations.

We next show that this effect is robust to controlling for the share of the deposit-to-port counterfactual that has already been paved (column 2) for time-varying comparative advantage (column 3), and for a measure of the ethnic salience of the deposits to be connected (column 4). The cumulative deposit overlap variable, $\textit{cum. deposit overlap}_{i,t}$, measures the share of the deposit-to-port counterfactual that is already paved at the start of each period. It controls for the possibility that $\textit{deposit-to-port bias}_{i,t}$ may be mechanically lower when a greater share of the counterfactual has already been paved, and this may confound our results. As expected, the coefficient on this variable is negative (though not significant), however the coefficient on autocracy is unchanged.

Column 3 controls for three measures of time-varying comparative advantage: the stock of deposits at the start of each period ($\textit{cum. deposit discoveries}_{i,t}$), the number of discoveries in the previous period ($\textit{deposit discoveries}_{i,t-1}$), and a country-specific, time-varying price index, capturing the current world prices of the minerals produced in a country. One concern with our approach is that changes over time in the supply or demand or commodities, that are country specific and hence not captured by period fixed effects, may affect both political institutions and optimal paving decision, biasing our results. Indeed, the coefficient on $\textit{deposit discoveries}_{i,t-1}$ suggests that countries which have recently discovered deposits go on to pave more deposit-to-port roads. The coefficient on autocracy is unchanged, however. This provides reassurance that the effect we have identified is orthogonal to changes in comparative advantage.

Column 4 controls for a measure of the ethnic salience of the deposits to be connected. Previous research has shown that rulers of ethnically divided societies tend to reward their fellow ethnic groups with favourable government policies, including the allocation of roads (Hodler & Raschky 2014, Burgess et al. 2015). In our setting, to the extent that the autocracy happens to be correlated with the "ethnic salience" of the deposits to be connected, our estimates of the

[44] We come back to clustering in Section 6.2.

Table 2: Baseline model

| Dependent variable → | Deposit-to-port bias | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Autocracy, start of t | 0.022** | 0.024** | 0.025** | 0.031*** | 0.013 |
| | (0.009) | (0.010) | (0.010) | (0.009) | (0.013) |
| Cum. deposit overlap, start of t | | -0.393 | -0.220 | -0.358 | -0.284 |
| | | (0.333) | (0.340) | (0.333) | (0.324) |
| Cum. deposit discoveries, start of t | | | -0.004 | -0.003 | -0.001 |
| | | | (0.007) | (0.007) | (0.007) |
| Deposit discoveries, t-1 | | | 0.024* | 0.024** | 0.024* |
| | | | (0.013) | (0.012) | (0.012) |
| Price index, average t | | | -0.000 | -0.000 | -0.000 |
| | | | (0.000) | (0.000) | (0.000) |
| Ethnic deposits, start of t | | | | 0.476*** | 0.343** |
| | | | | (0.157) | (0.158) |
| Ethnic deposits * Autocracy | | | | | 0.028* |
| | | | | | (0.015) |
| Observations | 174 | 164 | 164 | 164 | 164 |
| R-squared | 0.430 | 0.434 | 0.452 | 0.503 | 0.515 |
| Country FE | Yes | Yes | Yes | Yes | Yes |
| Period FE | Yes | Yes | Yes | Yes | Yes |
| Country-specific trends | Yes | Yes | Yes | Yes | Yes |
| Wild bootstrap Autocracy t-stat | 1.864 | 1.990 | 1.982 | 3.196 | 2.617 |
| Wild bootstrap Autocracy p-value | 0.059 | 0.067 | 0.051 | 0.006 | 0.026 |
| Wild bootstrap interaction t-stat | | | | | 1.788 |
| Wild bootstrap interaction p-value | | | | | 0.081 |
| Marginal effects of autocracy for | | | | | |
| Median ethnicity | | | | | 0.022** |
| | | | | | (0.010) |
| Mean ethnicity | | | | | 0.025*** |
| | | | | | (0.010) |
| 75% percentile ethnicity | | | | | 0.040*** |
| | | | | | (0.011) |

*Notes*: This table shows OLS regressions of the effect of autocracy on the degree to which countries' actual paving overlaps with the counterfactual network that connects deposits to ports. All specifications include country and period fixed effects, and country-specific linear trends. Robust standard errors in parenthesis: *** p < 0.01, ** p < 0.05, * p < 0.10. Wild bootstrap statistics are clustered by country and are performed using the method of Roodman et al. (2019). See Section 4 for variable definitions and sources.

effect of autocracy on road paving might be biased. The variable *ethnic* $deposit_{i,t-1}$ controls for this, by measuring the share of the deposits to be connected that are on the ethnic homeland of the ruling elite at the end of period $t-1$ (or equivalently, at the start of period $t$; more details on this variable are provided in section 3). The positive coefficient on this variable suggests that, indeed, ethnic considerations were an important drivers of road paving in our sample. However, our coefficient of interest becomes larger and more significant, suggesting that, if anything, to omit this variable was biasing our results downwards.

The point estimate in column 4 suggests that the quality of institutions had a sizable effect on road-paving decisions. The average autocracy level in our sample is 3.43 with a large standard deviation of 5.35. This implies that a one-standard deviation increase in autocracy predicts a 0.17 point (or half of a standard deviation, c.f. Table 1) increase in *deposit-to-port* $bias_{i,t}$.

We next turn to interpreting our results. Our empirical specification alleviates the concern that results may be spuriously driven by geography, comparative advantage, or other important drivers of road paving decisions. Still, the finding that autocracies had a greater deposit-to-port bias than democracy could be interpreted in different ways. It could be that autocracies focused more on enriching the ruling groups, as in our favourite interpretation. Alternatively, it could be that autocracies simply paved more roads (mechanically leading to a greater overlap with the deposit-to-port counterfactual, and hence a higher deposit-to-port bias), or that they favoured more open-economy policies (leading to greater investment in export-oriented roads), or that they were more capable to attract mining-related FDIs (who were actually responsible for paving the roads).

We here present evidence in support of our favourite interpretation - that autocracies had a greater deposit-to-port bias because they focused more on enriching the ruling groups - and devote Section 6.1 to reviewing evidence against alternative interpretations. Our empirical strategy exploits variation in the location and discovery of the deposits to be connected. If our favourite interpretation is correct, then we would not expect autocracies to have the same deposit-to-port bias independently of the location of the deposits to be connected. On the contrary, we would expect such bias to be stronger if the deposits were located on the ruling groups' ethnic homeland, since the ruling groups would be better positioned to benefit from the deposit-to-ports roads in that case.

Column 5 of Table 2 adds to the baseline an interaction of autocracy with *ethnic* $deposit_{i,t-1}$.

Now, the coefficient on autocracy turns insignificant, while the one on the interaction term is positive and significant. Moreover, the sum of the two coefficients is estimated to be positive and significant at mean and median values of *ethnic* $deposit_{i,t-1}$, as shown at the bottom of the table.[45] This suggests that autocracies *did* have a deposit-to-port bias when the deposits to be connected were located on the ruling groups' ethnic homelands, but *did not* have such bias when the deposits were located in other parts of the country. This is *prima facie* evidence in support of our favourite interpretation.

## 5.2 IV results

Our baseline analysis has adopted a number of strategies to control for unobserved heterogeneity. Because we control for country fixed effects, our results are not driven by time-invariant features of a country such as geography, deep-seated comparative advantage, or pre-1965 historical experiences. Similarly, period fixed effects rule out that our results are driven by events such as the passing of time since decolonisation, or the end of the Cold War. Country-specific, linear time trends absorb factors such as long-run population or economic growth. Finally, we also include proxies for time-varying comparative advantage, and control for ethnic favouritism in the allocation of roads.

Nevertheless, to further alleviate the concern that our results may be driven by omitted factors, we deploy an instrumental variable approach. To identify an instrument - that is an exogenous variable that is correlated with autocracy, but not with a country's paving decisions - we follow the most recent literature on democracy and growth. Several recent papers have emphasised the fact that democratisation tends to spread across countries in the same region (Buera et al. 2011; Acemoglu et al. 2019). Building on this notion, Acemoglu et al. (2019) instrument for democracy in country $i$ and year $t$ using average democracy in $t-1$ in the same World Bank region that $i$ belongs to (but including only countries that had a similar political regime as $i$ at the start of the sample), and including time fixed effects.[46]

[45] The minimum value of *ethnic* $deposit_{i,t-1}$ for which autocracies are estimated to have a deposit-to-port bias (with 95% confidence) is 0.25. About 56% of our observations have a value of *ethnic* $deposit_{i,t-1}$ above this threshold.

[46] We thus assume that democratisation in other countries in the region has no direct effect on the overlap of paving with our counterfactual. Such a spatial instrument assumes in addition that there is no indirect effect of democratisation in the region on our outcome variable. That is, the share of overlap of paving with the counterfactual is independent of the share of overlap of paving in other countries in the region. Only landlocked countries may depend to some extent on paving in one specific other country in the region, which is their transit

Unfortunately, in adapting the Acemoglu et al. (2019) IV approach to our setting, we are constrained by the limited geographical coverage of our sample, and by the similar political history of our countries. All our countries are in the same region, which pre-empts us from exploiting any time variation in a specification with time fixed effects. Furthermore, only one of our countries (Sierra Leone) was not an autocracy at the start of the sample (its autocracy score is negative), limiting our capacity to instrument for democracy using different regional averages for different countries. Given these limits, the variation left would be entirely cross sectional, and it would be largely mechanical given that average regional democracy is calculated by excluding the country under consideration.[47]

To circumvent these issues, we replace the period effects with less demanding 5-year effects, and we use as an instrument average autocracy in West Africa, as opposed Africa as a whole. Because the 5-year effects typically span over multiple periods (recall that the average period lengths is 2.23 years), this allows us to exploit some of the time variation in the instrument.[48] To focus on West Africa as a reference region helps to make the instrument more precise. By West Africa, we mean the countries in our sample and their direct neighbors: Algeria, Cameroon, Cape Verde, Chad, Libya, Mauritania, and Nigeria.[49]

Results in our IV approach, presented in Table 3, confirm our baseline results. Column 1 reproduces our baseline. Column 2 shows that to replace the period effects with 5-year effects does not, in itself, make our results any stronger. Column 3 and 4 present, respectively, the first and second stage. The instrument is strong enough, with an F-test above 10: lagged average regional autocracy is a strong predictor of a country's own-autocracy.[50] In addition, following

neighbour. However, in our counterfactual paving network, landlocked countries pave to the border crossing that leads to the port in the transit country, which is independent of shape of the counterfactual network in the transit country. We show that results are robust to dropping landlocked countries.

[47]This generates a spurious, *negative* cross sectional correlation between own and average regional democracy, since countries with below-average own democracy increase the average when they are excluded, and countries with above-average democracy increase it.

[48]We create ten dummies for the years 1965-1969; 1970-1974; 1975-1979; etc. For each period, the dummy equals one when the year of the Michelin map at the end of the period falls within the 5-year period. For example, the second dummy is one only for the periods 1 Jan 1969 - 31 Dec 1970 and 1 Jan 1971 - 31 Dec 1972, and zero otherwise. This is because these are the only two periods for which the year of the map at the end of the period (respectively 1971 and 1973) falls within the 5-year period of the second dummy.

[49]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.

[50]We do not add more than one lag of democracy, as Acemoglu et al. (2019) do, because adding a second or third lag weakens the joint significance of the instruments, while reducing sample size. The second stage in that case is still significant but 0.079** and 0.067**, respectively.

Table 3: IV: Endogenous autocracy

| Dependent variable → | Deposit-to-port bias | | Autocracy | Deposit-to-port bias |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Autocracy, start of t | 0.031*** | 0.031*** | | 0.089** |
| | (0.009) | (0.010) | | (0.035) |
| Cum. deposit overlap, start of t | -0.358 | -0.398 | 7.558*** | -0.748* |
| | (0.333) | (0.319) | (2.585) | (0.401) |
| Cum. deposit discoveries, start of t | -0.003 | -0.004 | -0.128** | 0.003 |
| | (0.007) | (0.007) | (0.055) | (0.008) |
| Deposit discoveries, t-1 | 0.024** | 0.019 | 0.079 | 0.019 |
| | (0.012) | (0.013) | (0.120) | (0.013) |
| Price index, average t | -0.000 | -0.000 | 0.000 | -0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Ethnic deposits, start of t | 0.476*** | 0.406** | -1.557 | 0.490*** |
| | (0.157) | (0.159) | (1.357) | (0.147) |
| Lagged regional average of autocracy | | | 1.028*** | |
| | | | (0.340) | |
| Observations | 164 | 164 | 164 | 164 |
| Country FE | Yes | Yes | Yes | Yes |
| Period FE | Yes | No | No | No |
| Country-specific trends | Yes | Yes | Yes | Yes |
| 5-Year FE | No | Yes | Yes | Yes |
| Wild bootstrap Autocracy t-stat | 3.196 | 3.286 | | 4.187 |
| Wild bootstrap Autocracy p-value | 0.006 | 0.008 | | 0.000 |
| 1st stage F-test | | | | 10.954 |
| Anderson-Rubin weak IV robust 95% C.I. | | | | [0.025,0.193] |

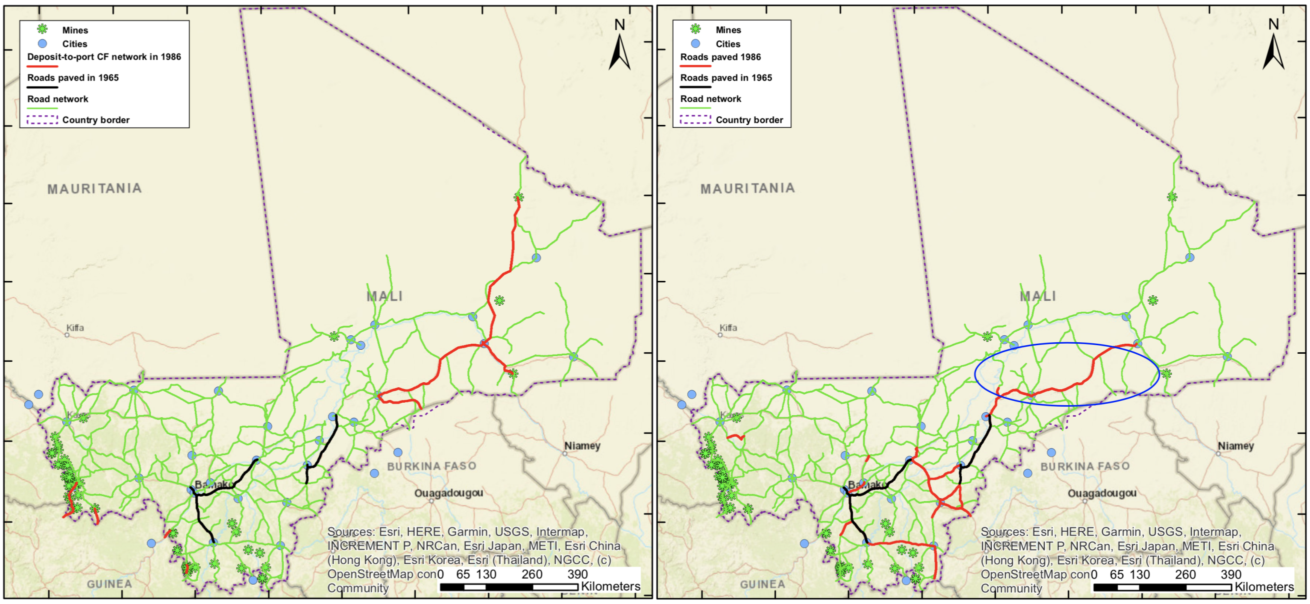
*Notes*: This table shows OLS (columns 1 and 2) and IV (columns 3 and 4) regressions of the effect of autocracy on the degree to which countries' actual paving overlaps with the counterfactual network that connects deposits to ports. All specifications include country and period fixed effects, and country-specific linear trends. Robust standard errors in parenthesis: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Wild bootstrap statistics are clustered by country and are performed using the method of Roodman et al. (2019). See Section 4 for variable definitions and sources.

Andrews et al. (2019), we provide Anderson-Rubin confidence intervals, which are robust to weak identification and are efficient in the one-instrument case, and which show that the coefficient of interest is comfortably above zero. We conclude that our baseline result, namely that autocracies have a greater deposit-to-port bias in road paving, is confirmed.

The fact that our IV coefficient is larger than our OLS coefficient may reflect a downward bias of OLS estimates, due to omitted factors or measurement error in the index of autocracy. The presence of omitted factors which bias OLS downwards is consistent with results discussed later (see Table 4), showing that to control for additional factors (such as foreign aid, FDIs and civil war) leads to a sizeable increase in the coefficient on autocracy. This may reflect the fact that, for example, autocracies attract less FDIs, which are also responsible for paving some of the deposit-to-port roads. It is also possible that the polity index provides only an imperfect measure of autocracy. Indeed, Acemoglu et al. (2019), studying the effect of democracy on growth, also obtain larger IV estimates compared to OLS, though the difference is smaller in their case.

### 5.3 Discussion

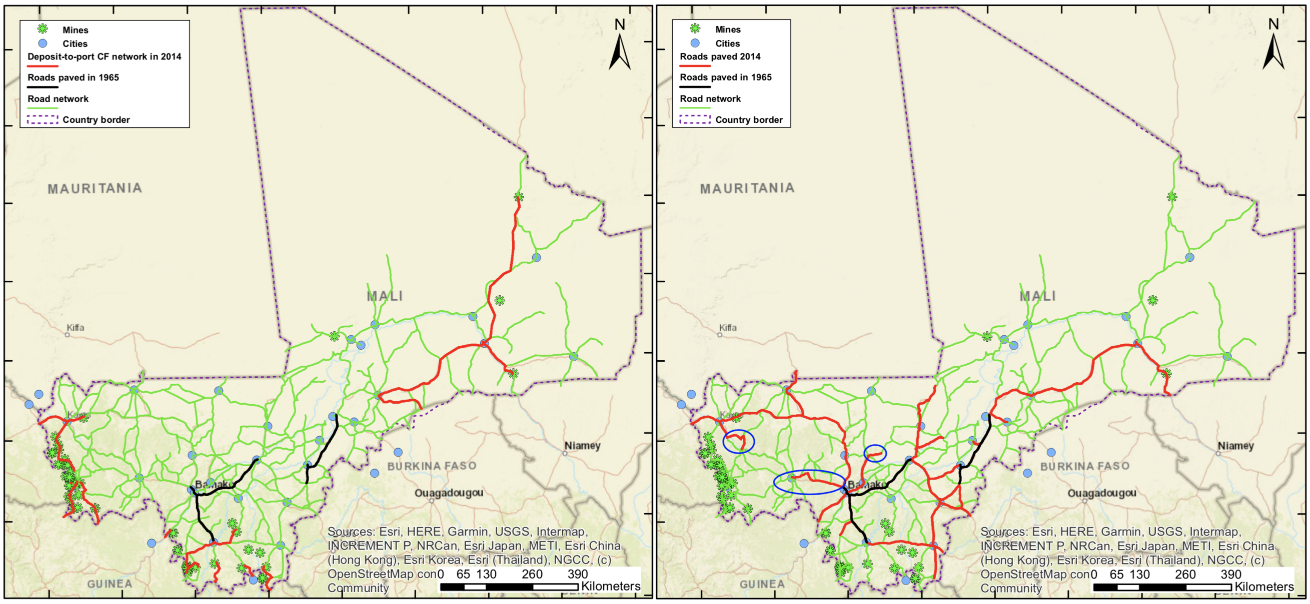
In this section, we first conduct a brief case study to illustrate the positive relationship between autocracy and deposit-to-port bias, and we then further discuss our favourite interpretation of the evidence provided in Section 5.1.




I: counterfactual, end of 1985 II: actual paving, 1984-85

Figure 5: Mali, deposit-to-port counterfactual (panel I) and actual (panel II) paved road network at the end of 1984-85. The blue circles indicate roads which were paved in 1984-85.

Figure 5 illustrates road paving in Mali in 1984-85, a country-period in which our dependent variable, *deposit-to-port bias*$_{i,t}$, was higher than average. Panel I illustrates the deposit-to-port counterfactual network at the end of the period. By then, a counterfactual planner only caring about connecting deposits to ports would have connected the mines along the country's Western and Southern border to the transit countries of Guinea and Ivory Coast. In addition, it would have paved the long road required to connect the three Easternmost mines of, from North to South, "In Darset", "Tilemsi" and "Ansongo". The latter two are still amongst Mali's most




I: counterfactual, end of 2013 II: actual paving, 2012-13

Figure 6: Mali, deposit-to-port counterfactual (panel I) and actual (panel II) paved road network at the end of 2012-13. The blue circles indicate roads which were paved in 2012-13.

important mines, producing respectively phosphate and manganese. Panel II represents the actual network at the end of the period, with blue circles indicating the 443 km of roads paved in 1984-85. Despite the fact that several cities around Bamako were not yet connected by paved road with the capital (which even then was by an order of magnitude bigger than any other city in Mali), the government of the day chose to pave the long eastward road to the Tilemsi and Ansongo mines instead. Sixty-eight percent of this paving overlaps with the deposit-to-port counterfactual, resulting in $\textit{deposit-to-port bias}_{i,t} = 0.68$ (as compared to an average for this variable of 0.26). Thus, this government is estimated to have had a stronger than average deposit-to-port bias.

Figure 6 illustrates road paving in 2012-13, when $\textit{deposit-to-port bias}_{i,t}$ was instead lower than average. Again, the two panels respectively illustrate the deposit-to-port counterfactual, and the actual network, at the end of the period, with blue circles indicating the 251 km of roads paved in 2012-13. Although parts of the counterfactual had not yet been paved (both in the North-East and in the West of the country), the government of the day chose to expand the network around Bamako, with connections paved to Kita and Kayes to the west of the capital, and to Sekou to the east. None of this paving overlaps with the deposit-to-port counterfactual, resulting in $\textit{deposit-to-port bias}_{i,t} = 0$. Thus, this government is estimated to have had a lower than average deposit-to-port bias.

It turns out that the government in charge in Mali in 1984-85 was much more autocratic than the one in power in 2012-13: thus, these two observations contribute to our finding that autocracies had a greater deposit-to-port bias than democracy. Mali endured 31 years of autocracy between 1960 and 1991. The ruler in charge in 1984-85, Moussa Traoré, was (unofficially or officially) in power between 1968 and 1991. Indeed, Mali's autocracy score is fixed at 7 (where 10 is most autocratic, and -10 is least autocratic) between 1960 and 1991. Political change only came in 1992, with the passing of a democratic constitution which laid out a competitive party system with a two-term limit for the president, and guaranteed freedom of expression. Democratic elections were held in 1992, 1997, 2002 and 2007, with the winner of the last two elections, Amadou Toumani Toure, peacefully succeeding to the winner of the first two. Although the years 2012-13 were marred by political and military turmoil due to a Tuareg rebellion in the North,[51] the next elections (held in 2013) were deemed credible by international observers (CIA Factbook).[52] Thanks to the 1992 constitution, Mali "... enjoys a quite vibrant, free, and pluralistic press on a continent where media rights cannot be taken for granted" (USAID 2011, p. 29). As of today, "Significant media freedom and largely functioning markets, particularly in informal sectors, as well as the urban culture of demonstrations all provide mechanisms of feedback to the government" (Bertelsmann Stiftung 2022, p. 12). Indeed, Mali's autocracy score has been fluctuating between the relatively low values of -7 and -5 since 1992.

We now further discuss our favourite interpretation of the evidence provided in the previous section. To recap, autocracies displayed a stronger deposit-to-port bias than democracies: they focused more on connecting deposits to ports, as opposed to any other pairs of locations. In our favourite interpretation, this was due to the fact that autocracies, facing fewer constraints (e.g. due to the lack of competitive elections and weaker scrutiny from the media and civil society), tend to select public policies which are more biased in favour of the ruling groups. The logic of this interpretation is that, after controlling for geography and comparative advantage, deposit-to-port roads can be seen as targeted policies: they benefit the ethnic groups on whose territory the deposits are located, more than the rest of society. Thus, if autocracies are as we

[51]Following the rebellion, which was aided by an influx of Al Qaeda fighters from Lybia, a military coup ousted Toure in March 2012. There followed a government of national unity, a french military interventions in the North, and eventually new elections in 2013.

[52]Available here: https://www.cia.gov/the-world-factbook/countries/mali/

have described them, then we would expect them to select more targeted policies such as deposit-to-port roads. Consistently with this interpretation, we find that, indeed, the deposit-to-port bias of autocracies was stronger when the deposits to be connected were located on the ruling groups' ethnic homeland.

Our interpretation rests on two notions, the first of which is that extractive political institutions, facing fewer constraints than inclusive ones, tend to select public policies that are more biased in favour of the ruling groups. This is one of the central arguments of the literature on institutions and development (e.g. Acemoglu et al. 2019), which has been further developed by Padró i Miquel (2007) for the case of ethnically fragmented societies. The latter paper has shown how autocratic leaders can persist in power, while extracting enormous rents, by biasing public policy in favour of their own ethnic group. That autocracies tend to engage in more ethnic favouritism than democracies has been shown empirically by a number of papers, including Hodler & Raschky (2014) and Burgess et al. (2015). The literature has also shown that greater resource rents tend to result in more corruption in autocracies, compared to democracies (Bhattacharyya & Hodler 2010 and Arezki & Gylfason 2013).

The second notion is that deposit-to-port roads are targeted policies which disproportionately benefit the ethnic groups on whose homeland the deposits are located. One might object that deposit-to-port roads should be valuable to an autocratic government independently on where they are located, since the government can appropriate the resource-revenues they generate through its control of the state apparatus.[53] There are however several reasons to believe that parts of the value of deposit-to-port roads are only captured locally. First, the mines may have positive local economic effects, at least in expectations. While the literature on the local effects of mines has found mixed results,[54] recent evidence has found that resource discoveries can have a very large positive effects on local expectations (Cust & Mensah 2020). Second, deposit-to-port roads might also be used to transport people and non-mineral goods, thus having a positive local economic effect on top of any effect via greater resource rents. For both of these reasons,

[53]Consistent with this, Berman et al. (2020) find that, in Africa, mines tend to increase feelings of deprivation within the ethnic groups hosting the mines, but only if those groups are powerless in national politics.

[54]Aragón & Rud (2013) found positive effects of a Peruvian gold mine on local incomes (channeled through an increase in the price of locally-produced non-tradeable products and services), but only after the company owning the mines adopted a policy directed at increasing local supply linkages. While resource extraction may have a negative dutch disease effect on local tradeable sectors, Pelzl & Poelhekke (2021) found that, in Indonesia, this was only true for labour-intensive extraction processes, while capital intensive processes benefited both tradeable and non-tradeable sectors. For Africa, Berman et al. (2020) find limited evidence of a positive local economic effect of mines.

autocrats may face requests for resource development from their fellow ethnic members. Finally, the exploitation of mineral deposits can have important consequences for the distribution of powers within an ethnically fragmented society. On the one hand, resource discoveries may encourage separatist movements by those who own the deposits. On the other, greater mineral rents may boost the military power of the ethnic groups on whose ethnic homeland the minerals are located (Berman et al. 2017). Anticipating this, autocrats may want to develop deposits in their own ethnic homeland more than in their rivals', thus mitigating the risk of rebellions and strengthening their hold on power.

# 6 Additional results

This section is divided in two parts. In the first, we further corroborate our interpretation of the evidence presented in Section 5.1, by ruling out alternative interpretations. In the second, we conduct a number of robustness tests.

## 6.1 Ruling out alternative mechanisms

Having shown that autocracies had a deposit-to-port bias in road paving, and having presented and discussed our favourite interpretation of this, we now turn to ruling out alternative interpretations in Table 4.

In column 1, we add the amount of current paving to our baseline (column 4 of Table 2). This variable is included to rule out a mechanical channel through which autocracies might affect overlap with the deposit-to-port counterfactual (other than through their intrinsic deposit-to-port bias). If autocracies pave more or less kilometers than democracies,[55] and different amounts of paving mechanically leads to more overlap with the counterfactual, then our results might overestimate the deposit-to-port bias of autocracies. We find that the coefficient on current paving is positive but insignificant, while the coefficient on autocracy is unchanged. This suggests that this mechanical channel is not important.

Column 2 adds the economy's degree of openness to our baseline. This is done to rule out the following alternative explanation for the greater deposit-to-port bias of autocracies.

[55] In an (unreported) regression of current paving on autocracy (with country and period fixed effects and country-specific trends) we find a positive but marginally insignificant relationship (p-value: 0.17).

If autocracies are more in favor of an open economy, then they may be more focused than democracies on exploiting a country's comparative advantage. To achieve that goal, they may decide to pave more deposit-to-port roads. Although a broad measure of openness may not be the perfect proxy for such attitudes, to control for it may at least give an indication as to whether such alternative channel is likely to be important. Reassuringly, the coefficient on autocracy is essentially unchanged. This makes sense retrospectively, given that the autocracies in our sample were often socialist or military regimes, with arguably weak preferences for economic openness.

Columns 3 and 4 control for the inflows of foreign direct investment, and foreign aid. It may be the case that foreign investors and donors have their own agenda on deposit-to-port roads. For example, investors in extractive industries may want more such roads, while donors may want more or less, depending on their view on optimal infrastructure development, or on self-interest.[56] To the extent that autocracies receive more or less investment or aid, this may bias our results. Results however tend to dispel this worry, since the coefficient on autocracy is again essentially unchanged.

Column 5 controls for the occurrence of civil wars, of which there were several in our sample.[57] Inter-ethnic fighting may make public policies even more parochial then they normally are. This may be especially true for road construction, since a balanced expansion of the network (connecting different parts of the country to one another) may simply not be an option during period of fighting. Indeed, the coefficient on civil war on deposit-to-port bias is estimated to be positive in column 5, and positive and significant in column 7 (where we include all the controls at once). Once again, however, the coefficient on autocracy remains in the same ballpark.

One final concern is that autocracies may appear to connect deposits to ports, not because they had a deposit-to-port bias, but because they wanted to connect cities, which happened to be located near deposits. Cities may tend to be located near deposits for two reasons. First, exploration and extraction costs may be lower near cities. Second, cities may endogenously emerge near deposits, due to the demand for local inputs generated by mines.

To address this concern, we construct a measure of a government's preference for connecting cities, and control for it in column 6. To construct this measure, we first construct a second

[56]Western donors have been advocating a balanced expansion of the African transport networks over the last two decades, while China has been sponsoring more deposit-to-coast roads. For some evidence in support of this claim, see Bonfatti & Poelhekke (2017).

[57]These included Côte d'Ivoire (2002-04), Guinea (2000-01), Liberia (1989-90; 1992-95; 2002-03) and Sierra Leone (1991-96, 1998-00), in addition to Guinea-Bissau (1998-99).

Table 4: Mechanisms

| Dependent variable → | Deposit-to-port bias | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Autocracy, start of t | 0.031*** | 0.034*** | 0.039*** | 0.030*** | 0.035*** | 0.028*** | 0.050*** |
| | (0.009) | (0.010) | (0.011) | (0.010) | (0.009) | (0.008) | (0.013) |
| Current paving, t | 0.000 | | | | | | 0.000* |
| | (0.000) | | | | | | (0.000) |
| Openness to trade, average t | | -0.085 | | | | | -0.055 |
| | | (0.140) | | | | | (0.162) |
| Log FDI, average t | | | 0.001 | | | | 0.013 |
| | | | (0.042) | | | | (0.039) |
| Log aid, average t | | | | -0.032 | | | 0.104 |
| | | | | (0.116) | | | (0.120) |
| Civil war, average t | | | | | 0.250 | | 0.486*** |
| | | | | | (0.200) | | (0.140) |
| City overlap | | | | | | 0.376*** | 0.432*** |
| | | | | | | (0.099) | (0.124) |
| Observations | 164 | 146 | 137 | 164 | 164 | 164 | 119 |
| R-squared | 0.509 | 0.498 | 0.550 | 0.504 | 0.513 | 0.586 | 0.692 |
| Standard controls included | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Period FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-specific trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Wild bootstrap Autocracy t-stat | 3.375 | 2.731 | 2.704 | 2.530 | 3.855 | 3.145 | 5.816 |
| Wild bootstrap Autocracy p-value | 0.007 | 0.017 | 0.021 | 0.031 | 0.001 | 0.007 | 0.001 |

*Notes*: This table shows OLS regressions of the effect of autocracy on the degree to which countries' actual paving overlaps with the counterfactual network that connects deposits to ports. All specifications include country and period fixed effects, and country-specific linear trends. Robust standard errors in parenthesis: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Wild bootstrap statistics are clustered by country and are performed using the method of Roodman et al. (2019). See Section 4 for variable definitions and sources.

counterfactual road expansion path, which aims to approximate the way in which paved roads would have expanded in 1965-2014, had the government only cared about connecting cities to one another.[58] In constructing this counterfactual, we follow Burgess et al. (2015). The idea of the counterfactual is that - unlike mineral deposits, which mostly need to be connected to a port for exporting purposes - cities need to be connected to one another for internal trade, as well as to a port for international trade. It is constructed by ranking all possible city pairs in a country by market potential (the sum of population divided by distance),[59] and then reallocate the paving quota starting from the top pairs in each period using a procedure that is identical to the one used for the deposit-to-port counterfactual. This procedure tends to connect cities to one another as well as to a port, given that in all our coastal countries except for Cote d'Ivoire, the capital city has a port. In fact, the large size of the capital city attracts lots of connections from interior cities, thus implicitly recognising the importance of the international trade which flows through the capital. We then construct a measure of overlap of current paving with the counterfactual, as explained in Section 2.2, but using the city-to-city counterfactual instead of the deposit-to-port counterfactual. The result is a measure of a government's preference for connecting cities to one another, as opposed to any other pairs of locations.

Controlling for such measure, the coefficient on autocracy reduces slightly in size, but remains strongly significant. This alleviates the concerns that our earlier results were driven by a greater interest of autocracy in connecting cities. Interestingly, the coefficient on $\textit{city-to-city bias}_{i,t}$ is positive and significant, suggesting that the deposit-to-port and city-to-city counterfactuals partially overlap. This indicates that, at least to some extent, deposits do tend to be located near cities.

## 6.2 Robustness

We now present a battery of tests that show the robustness of our main results.

[58]The data on cities was kindly shared by Hervé Gazel of the Africapolis project (Université Jean Moulin Lyon 3. See: `http://www.afd.fr/lang/en/home/publications/travaux-de-recherche/archives-anciennes-collections/NotesetEtudes/Africapolis` It contains the location and population of cities with at least 10,000 inhabitants in 2010, 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.

[59]Cities located in neighboring countries, but less than 50km away from the border, are included.

### ***Clustering***

So far, we have presented results using robust standard errors, and provided wild bootstrap statistics that account for possible autocorrelation. More standard is to cluster standard errors, but with only 12 countries our setting falls short of the rule-of-thumb that clustering requires at least 50 clusters (Cameron & Miller 2015). While it is not clear that 12 clusters is necessarily too few, clustering may then lead to overfitting and a downwards-biased cluster-robust variance matrix estimate, and potential over-rejection. Robustness Table 5 nevertheless presents our baseline regression and main result with clustered standard errors by country (in column 1) and by country and year (in column 2): the confidence level is very similar to the wild bootstrap and robust estimates.

The remainder of Robustness Table 5 presents results using different definitions of the counterfactual.

### ***Diamonds***

First, in column 3, we include deposits that contain diamonds, even though our prior is that these have such a high ratio of value to weight that they would not be transported via roads and bulk ships. Only 6% of deposits in our sample contain diamonds and including them does not change our results.

### ***Deterioration of roads***

Second, we investigate the effect of accounting for the deterioration of roads. In our baseline counterfactual we did not include deterioration: we simply do not let the baseline counterfactual networks deteriorate over time, because we only observe the implicit paving budget and not the budget for the maintenance of roads. However, in our sample, we observe deterioration as well as upgrading of roads. In particular, it happens in 43% of country-periods that the quality of at least one segment of road deteriorates (of typically 1 km length). At the median, 3% of the paved road network becomes unpaved in such an event (see also Table 1). This may be the result of lack of maintenance, or of traumatic events such as natural disasters or civil war such as in Sierra Leone between 1991 and 2002.

The first issue is that, in the presence of deterioration, paving in any given period may involve

Table 5: Robustness

| Dependent variable → | Deposit-to-port bias | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Counterfactual includes → | | | Diamonds | Deterioration of roads, with repaving costs equal to: | | Main ports double weight | Excl. deposits near railroads | Ranked by distance only | Ranked by value only |
| | | | | 0.5 * paving | 1.0 * paving | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Autocracy, start of t | 0.031*** | 0.031*** | 0.030*** | 0.028*** | 0.024*** | 0.031*** | 0.020*** | 0.035*** | 0.038*** |
| | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.007) | (0.009) | (0.009) |
| Cum. deposit overlap, start of t | -0.358 | -0.358 | | -0.310 | -0.279 | -0.358 | -0.337 | -0.549* | -0.438 |
| | (0.277) | (0.323) | | (0.362) | (0.362) | (0.333) | (0.272) | (0.307) | (0.271) |
| Observations | 164 | 164 | 164 | 164 | 164 | 164 | 164 | 165 | 165 |
| R-squared | 0.503 | 0.503 | 0.480 | 0.495 | 0.478 | 0.503 | 0.638 | 0.521 | 0.529 |
| Country FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Period FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-specific trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clustered on | country | country & year | | | | | | | |
| Wild bootstrap Autocracy t-stat | | | 2.684 | 2.307 | 2.015 | 3.196 | 3.500 | 3.574 | 4.576 |
| Wild bootstrap Autocracy p-value | | | 0.018 | 0.033 | 0.072 | 0.006 | 0.005 | 0.004 | 0.001 |

*Notes*: This table shows OLS regressions of the effect of autocracy on the degree to which countries' actual paving overlaps with the counterfactual network that connects deposits to ports. All specifications include country and period fixed effects, and country-specific linear trends. Robust standard errors in parenthesis: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Wild bootstrap statistics are clustered by country and are performed using the method of Roodman et al. (2019). See Section 4 for variable definitions and sources.

roads that have never been paved before, 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 counterfactual. We tackle this issue in two ways. First, we use the World Bank's 2008 ROCKS dataset (also used by Collier et al. 2016), containing more than 3,000 road construction projects in 89 low and middle-income countries. We find that the per-kilometer cost of paving an entirely new road is on average no different from repaving a previously paved road. This holds in the full sample and in Africa in particular.[60] This provides at least some reassurance that we are not making a first-order mistake in constructing the counterfactual the way we do.

To provide further robustness, we adjust the baseline counterfactual in columns 4 and 5 and proceed in steps to include deterioration. First, we collect for each country and year the actual deterioration of roads by comparing maps from year to year, similar to the way that we collected actual paving. In constructing the counterfactual roads we now let the previous' year network deteriorate by an amount equal to the observed deterioration since the previous year, and apply deterioration to the *lowest* ranked paved connection. This assumes that deterioration is (at least in part) a decision to divert scarce resources for road maintenance away from the least important connections, which captures the idea that governments can willingly cut back on maintenance or at least fail to implement repairs between years for which we observe road maps. It also ensures that the counterfactual expands less (or even contracts) in periods of deterioration, and that the counterfactual planner may spend part of the paving quota to maintain the existing network (since it will be used to offset the effect of the deterioration quota on the lowest-ranked pairs). The next step is then to pave roads by applying the actual paving budget to the *most* important connections. Depending on the routes, the paving can then pave roads that have never been paved before or repave roads that were paved sometime in the past but deteriorated over time. The final step is to vary the 'price' of repaving roads: in column 4 repaving 'costs' only half of what is costs to pave a road that was never paved before ($0.5 * paving$). In other words, 1 km of paving budget can *repave* 2 km of roads. Alternatively, in column 5 repaving costs the same

[60] The ROCKS dataset contains many different types of road construction activities. After careful reading we identify two categories, "Upgrading to Bituminous 2L" and "Reconstruction bituminous", as the best candidates to represent the paving of a new road, and the re-paving of a previously existing road. We are thus left with 1,124 projects, 333 of which are located in Africa. We regress cost per kilometer on "first paving" (an indicator variable for the first category), an interaction between first paving and the Africa dummy, country and year fixed effects, the length of the overall road project, and the type of terrain on which the road was built. The coefficient on first paving is negative and insignificant, for either values of the Africa dummy.

of what is costs to pave a road that was never paved before ($1.0 * paving$). Since we argued using road construction data that repaving roads is not cheaper than paving roads for the first time, the second version follows our preferred baseline specification in terms of how repaving is treated.

In both cases, the magnitude of the main effect is somewhat smaller, but equally significant.

### *Main ports*

In column 6 we double the weight of the port that is in or near a country's largest coastal city, or, in the case of landlocked countries, the port of the nearest largest coastal city in the transit country. In practice, this does not materially change the ranking of deposit-port pairs and thus yields the same estimate of autocracy on deposit-to-port bias.

### *Railroads*

In our main analysis, we focus on roads and not on railways primarily because of lack of variation: almost no West African railways has been extended in our period of study. Furthermore, many of them have fallen in a derelict state. A few railways are operational and some of those are owned and operated outright by mining companies. For example, the railway between Yekepa and the port of Buchanan in Liberia was built in the 1960s for the Tokadeh iron ore mine, was damaged during the civil war, and has been rebuilt and reopened by ArcelorMittal in 2011.[61] We explicitly include railways in column 7, and show that our main result is robust. Specifically, we find for every deposit whether the nearest mode of transport is a road or a railway. 22 deposits in seven countries are closer to a railway than to a road: we thus assume that these deposits were and are served by a railway and do not need to be connected to ports via paved roads.[62] We thus exclude these deposits from the ranking of deposit-port connections when constructing the counterfactual. Our main result is robust to this exercise.

[61] Source: `https://liberia.arcelormittal.com/what-we-do/rail.aspx`.

[62] It may strike the reader as unlikely that any deposit would not also have a road nearby. However, the scale of our maps (1:30,000) ensures that we analyse relatively main roads as opposed to local roads (such as logging or private roads).

### ***Distance versus value***

In columns 8 and 9 we change the ranking of deposit-to-port connections by ranking only based on the distance between them in column 8, or only by the sum of their size in column 9. Neither approach materially alters our main result.

### ***Alternative measure of autocracy***

In the Online Appendix Section OA5 we follow Acemoglu et al. (2019) and replace the inverse of the Polity IV score, defined between -10 and 10, with their dummy value of autocracy equal to one if a country is non-democratic overall and zero otherwise. 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. Online Appendix Figure OA1 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 below zero (except for the years 1968-1969 in Sierra Leone).

Tables OA10, OA11, OA12, and OA13 repeat all of our main results with the dummy version of autocracy. To compare magnitudes, we can use the difference between the average inverse polity score among democratic countries (which is -6) and the average autocracy score among autocratic countries (which is +6). We thus find that moving 12 units on the inverse polity score towards autocracy increases deposit-to-port bias by 0.26, while we find that the Acemoglu et al. dummy suggests that becoming autocratic increases deposit-to-port bias by a very similar 0.28. In summary, all of our results are highly robust to changing the definition of the main explanatory variable.

### **Dropping one country at a time**

In Online Appendix OA3 we test whether any single country is influential enough to drive our main results. To this end, we drop countries one by one from the sample and re-estimate the regression of Table 2, column 4. In addition, we drop all landlocked countries from the sample. In all cases the point estimates are significant and very similar to the average effect.

# 7 Conclusions

In this paper, we have used a newly-constructed dataset to document two new facts about the expansion of the West African road networks since independence. For a sample of twelve countries and spanning almost five decades, we have shown that, first, more autocratic governments displayed a stronger preference for paving interior-to-coast roads, and particularly deposit-to-port roads. Second, autocracies only had such a deposit-to-port bias when the deposits to be connected were located on the ruling elite's ethnic homeland.

The most plausible interpretation of these findings is that autocracies, facing fewer constraints than democracies, focused more on paving roads that would benefit the ruling groups, and less on paving other types of roads (inclusive those most required for sustained economic growth). For example, the lack of competitive elections may make it less important for autocracies to seek consensus through a balanced road investment program, while weaker scrutiny from the media and civil society 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 favoritism 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 is important to know, particularly at a time in which Chinese investments are helping to exacerbate the interior-to-coast pattern of the existing African networks.

In order to corroborate this interpretation, one would need to construct and estimate a quantitative model of politics, road building and trade for 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


Acemoglu, D., Naidu, S., Restrepo, P. & Robinson, J. A. (2019), 'Democracy does cause growth', *Journal of Political Economy* **127**(1), 47–100.

Alder, S., Kondo, I. et al. (2018), Political distortions and infrastructure networks in china: A quantitative spatial equilibrium analysis, *in* '2018 Meeting Papers. Society for Economic Dynamics'.

Andrews, I., Stock, J. H. & Sun, L. (2019), 'Weak instruments in instrumental variables regression: Theory and practice', *Annual Review of Economics* **11**(1), 727–753.
**URL:** *https://doi.org/10.1146/annurev-economics-080218-025643*

Aragón, F. M. & Rud, J. P. (2013), 'Natural resources and local communities: evidence from a peruvian gold mine', *American Economic Journal: Economic Policy* **5**(2), 1–25.

Arezki, R. & Gylfason, T. (2013), 'Resource rents, democracy, corruption and conflict: Evidence from sub-saharan africa', *Journal of African Economies* **22**(4), 552–569.

Atkin, D. & Donaldson, D. (2015), 'Who's getting globalized? the size and implications of intra-national trade costs', *NBER Working Papers* (21439).

Bazzi, S. & Blattman, C. (2014), 'Economic shocks and conflict: Evidence from commodity prices', *American Economic Journal: Macroeconomics* **6**(4), 1–38.

Berman, N., Couttenier, M. & Girard, V. (2020), 'Natural resources and the salience of ethnic identities', *Working Paper Series, Universidade Nova de Lisboa* .

Berman, N., Couttenier, M., Rohner, D. & Thoenig, M. (2017), 'This mine is mine! how minerals fuel conflicts in africa', *American Economic Review* **107**(6), 1564–1610.

Bertelsmann Stiftung (2022), *BTI 2022 Country Report — Mali*, Gütersloh: Bertelsmann Stiftung.

Bhattacharyya, S. & Hodler, R. (2010), 'Natural resources, democracy and corruption', *European Economic Review* **54**(4), 608–621.

Blimpo, M. P., Harding, R. & Wantchekon, L. (2013), 'Public investment in rural infrastructure: Some political economy considerations', *Journal of African Economies* **22**(suppl_2), ii57–ii83.

Bonfatti, R. & Poelhekke, S. (2017), 'From mine to coast: Transport infrastructure and the direction of trade in developing countries', *Journal of Development Economics* **127**, 91–108.

Breitung, J. (2001), The local power of some unit root tests for panel data, *in* B. Baltagi, T. Fomby & R. Carter Hill, eds, 'Advances in Econometrics, Vol. 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels', Emerald Group Publishing Limited, Bingley, pp. 161–178.

Buera, F. J., Monge-Naranjo, A. & Primiceri, G. E. (2011), 'Learning the wealth of nations', *Econometrica* **79**(1), 1–45.

Burgess, R., Jedwab, R., Miguel, E., Morjaria, A. & Padró i Miquel, G. (2015), 'The value of democracy: evidence from road building in kenya', *American Economic Review* **105**(6), 1817–51.

Buys, P., Deichmann, U. & Wheeler, D. (2006), *Road network upgrading and overland trade expansion in sub-Saharan Africa*, The World Bank.

Cameron, A. C. & Miller, D. L. (2015), 'A practitioner's guide to cluster-robust inference', *Journal of Human Resources* **50**(2), 317–372.

Collier, P., Soderbom, M. & Kirchberger, M. (2016), 'The cost of road infrastructure in low- and middle-income countries', *World Bank Economic Review* **30**(3), 522–548.

Cust, J. & Mensah, J. T. (2020), 'Natural resource discoveries, citizen expectations and household decisions', *World Bank, Policy Research Working Papers* .

De Luca, G., Hodler, R., Raschky, P. A. & Valsecchi, M. (2018), 'Ethnic favoritism: An axiom of politics?', *Journal of Development Economics* **132**, 115–129.

Hodler, R. & Raschky, P. A. (2014), 'Regional favoritism', *The Quarterly Journal of Economics* **129**(2), 995–1033.

Jedwab, R. & Storeygard, A. (2022), 'The average and heterogeneous effects of transportation investments: Evidence from sub-saharan africa 1960-2010', *Journal of the European Economic Association* **20**(1), 1–38.

Murdock, G. P. (1959), *Africa its peoples and their culture history*, McGraw-Hill book company.

Nkrumah, K. (1964), *Africa Must Unite*, Heineman.

Padró i Miquel, G. (2007), 'The control of politicians in divided societies: The politics of fear', *Review of Economic studies* **74**(4), 1259–1274.

Pelzl, P. & Poelhekke, S. (2021), 'Good mine, bad mine: Natural resource heterogeneity and dutch disease in indonesia', *Journal of International Economics* **131**, 103457.

Plunkert, P. A. & Jones, T. S. (1999), *Metal prices in the United States through 1998*, US Department of the Interior, US Geological Survey.

Porteous, O. (2019), 'High trade costs and their consequences: an estimated dynamic model of african agricultural storage and trade', *American Economic Journal: Applied Economics* **11**(4), 327–66.

Rodney, W. (1972), *How Europe underdeveloped Africa*, Bogle-L'Ouverture Publications.

Roodman, D., Ørregaard Nielsen, M., MacKinnon, J. G. & Webb, M. D. (2019), 'Fast and wild: Bootstrap inference in stata using boottest', *The Stata Journal* **19**(1), 4–60.

Sachs, J., McArthur, J. W., Schmidt-Traub, G., Kruk, M., Bahadur, C., Faye, M. & McCord, G. (2004), 'Ending africa's poverty trap', *Brookings papers on economic activity* **2004**(1), 117–240.

USAID (2011), *Democracy and Governance Assessment of Mali*.

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L.-E., Hunziker, P. & Girardin, L. (2015), 'Integrating data on ethnicity, geography, and conflict: The ethnic power relations data set family', *Journal of Conflict Resolution* **59**(7), 1327–1342.

Wacziarg, R. & Welch, K. H. (2008), 'Trade liberalization and growth: New evidence', *The World Bank Economic Review* **22**(2), 187–231.

Waters, J. K., Mayer, R. H. & Kriebel, D. L. (2000), 'Shipping trends analysis', *Report prepared for Institute for Water Resources, Alexandria, Virginia* .

# Online Appendix

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

*Roberto Bonfatti, Yuan Gu, Steven Poelhekke*

September 27, 2022

### Table of Contents

## OA1 Additional Descriptive statistics

Table OA1: Summary statitics, full sample

| Variable | N | mean | s.d. | min | max |
|---|---|---|---|---|---|
| Current paving (km) | 264 | 124.82 | 208.59 | 0.00 | 1,794.12 |
| Current deterioration (km) | 264 | -43.20 | 80.41 | -622.60 | 0.00 |
| Deposit-to-port bias | 264 | 0.18 | 0.30 | 0.00 | 1.00 |
| Deposit-to-port bias, with diamonds | 264 | 0.18 | 0.30 | 0.00 | 1.00 |
| Deposit-to-port bias, allowing for deterioration (0.5) | 264 | 0.17 | 0.29 | 0.00 | 1.00 |
| Deposit-to-port bias, allowing for deterioration (1.0) | 264 | 0.16 | 0.29 | 0.00 | 1.00 |
| Deposit-to-port bias, main ports double weight | 264 | 0.18 | 0.30 | 0.00 | 1.00 |
| Deposit-to-port bias, excl. deposits near railways | 264 | 0.14 | 0.22 | 0.00 | 1.00 |
| City-to-city bias | 264 | 0.22 | 0.32 | 0.00 | 1.00 |
| Cumulative deposit overlap, end of t | 264 | 0.53 | 0.22 | 0.00 | 1.00 |
| Cumulative city overlap, end of t | 264 | 0.64 | 0.18 | 0.00 | 1.00 |
| Ethnic deposits | 264 | 0.40 | 0.42 | 0.00 | 1.00 |

Note: This table shows summary statistics for the main variable of interest for all country-years (including those with zero paving).

## OA2 Data Appendix

We require data on autocracy, detailed road construction over time and space, the location, size and type of deposits, the location and size of ports and cities, the world prices of metals and minerals, foreign aid, FDI flows, periods of war, openness, a delineation of ethnic lands and a measure of the political influence of each ethnic group. In this appendix, we describe our sources in detail.

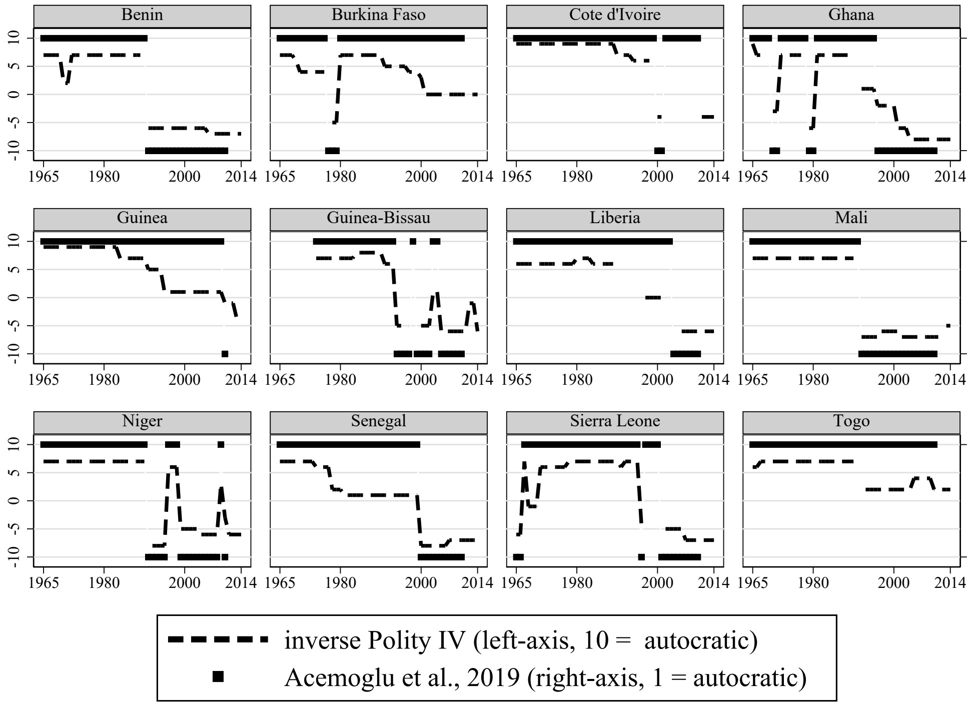
### OA2.1 Autocracy

Our main measure of autocracy is equal to minus the polity score provided by the Polity IV Project. This is a widely used score which ranges from +10 (strongly democratic) to -10 (strongly autocratic). Consequently, our measure will range from +10 (strongly autocratic) to -10 (strongly democratic). This measure is available in all years in our sample, excluding a few in which there was either a collapse in government authority, or a period of transitions (Polity IV Project codes -77 and -88 respectively). Since $A_{i,t}$ should measure the state of autocracy at the beginning of period $t$, and the Polity IV index measures political conditions on December 31 of each year (see Marshall et al., 2017, p. 12), we construct $A_{i,t}$ using the inverse of the Polity IV index for the last year before period $t$. For example, for the period that goes from January 1, 1965 till December 31, 1966, $A_{i,t}$ is constructed using the inverse of the 1964 Polity IV index, which measures the state of autocracy on December 31, 1964.

Figure OA1 shows that there is a lot of variation in the inverse polity score 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 autocracy equal to the inverse of the measure constructed by Acemoglu et al. (2019) by combining the Polity IV index with several other independent measures of democracy. This measure is available until 2012 and records 11 of our 12 countries as autocratic in 1965 and only 3 as autocratic by 2012. Our sample thus reveals the same unprecedented spread of democracy as highlighted in Acemoglu et al. (2019). 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 OA1: Autocracy scores




## OA2.2 Roads

We have digitised 23 Michelin West Africa maps covering the 12 West African countries listed at the start of this section over the period 1965 to 2014. 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 define our sample periods as starting at the beginning of each publication year. For example, our first two road maps were published in 1965 and 1967: our first period then starts on January 1, 1965 and ends on December 31, 1966. This gives us 22 periods for each country, for a theoretical maximum of 264 observations.[63]

The maps show colour- and pattern-coded roads, which we digitize using the ArcGIS platform.

[63]The map editions are 1965, 1967, 1968, 1969, 1971, 1973, 1976, 1983, 1984, 1986, 1988, 1989, 1990, 1991, 1996, 1998, 2000, 2002, 2003, 2007, 2009, 2012 and 2014. Although 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 unequal time span between years seems to suggest that a new map was only issued when enough changes to the road network were observed.

To raise the resolution of the maps 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. In our analysis, we use all roads (major, secondary and local), and focus on the quality classification. There are six quality 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 other four types of worse-quality roads are also unpaved and not all-weather roads. We classify as "paved" the surfaced roads, and as "unpaved" all other roads.

In Table OA2 we track the distribution of roads in West Africa by quality, in 1965 and in 2014. 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 10,665 kilometers of roads that were paved in 1965, 9,496 kilometers, or 89%, were still paved in 2014, while the remaining 11% deteriorated to unpaved status (possibly due to lack of maintenance and/or war). Of the latter, 187 kilometers deteriorated to the point of vanishing by 2014. Note that the paving of roads where no road was previously present is relatively rare: paving normally happens to existing roads or tracks.[64] Note also that compared to paved roads, unpaved roads are more prone to deterioration.

Figure OA2 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. 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, Guinea-Bissau and Togo also paved many roads, while Ghana, Guinea, Liberia, Mali, Niger and Sierra Leone paved relatively

[64]Note also that, of the 3,026 kilometers of roads which appear as paved in 2014 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 2014, before becoming paved roads by 2014.

Table OA2: Road quality changes between 1965 and 2014 in km

| Status in 1965 | | | Status by 2014 | | | |
|---|---|---|---|---|---|---|
| | | | No road | Unpaved (lowest quality) | Unpaved (medium quality) | Paved (good quality) |
| No road | 21,241 | → | 2,096 | 7,668 | 10,547 | 3,026 |
| Unpaved (lowest quality) | 27,729 | → | *1,482* | 15,812 | 7,607 | 4,310 |
| Unpaved (medium quality) | 57,423 | → | *1,679* | *8,697* | 30,620 | 18,124 |
| Paved (good quality) | 10,665 | → | *187* | *235* | *934* | 9,496 |

Note: The table lists changes in road quality across 13 West African countries as inferred from comparing the digitised Michelin maps of West Africa for 1965 and 2014. 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 2,096 km which were recorded as track or roads at some point between 1965 and 2014, but these no longer exist by 2014. Only paved (surfaced) roads can be used year-round. Road quality deteriorations are denoted in italics. For example, 187 km of paved roads disappeared by 2014, and 235 deteriorated to the "unpaved (track)" quality. For sources, see Section OA2.2.

few.

To construct our dependent variable, we need a measure of road paving in each country $i$ and period $t$. We construct such measure by comparing maps published at the beginning and end of the period. For example, if a road shows up as paved in the 1967 edition but unpaved in the 1965 edition, then we conclude that it was paved sometime between January 1, 1965 and December 31, 1966 (although we do not know exactly when). Tables OA3, OA4 and OA5 present the distribution of paving by country, by period, and by country-period for the case of Sierra Leone. The first two tables show that some countries paved more than Figure OA2 suggests, even though it was low compared to their surface area.

### OA2.3 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.[65]

We start from MinEx, which is a dataset of deposits and has the best coverage in terms of

[65] Per source, 321 records are from MinEx, 54 from RMG, and 8 from MRDS.

Table OA3: Road paving characteristics by country, averages across periods

| Counterfactual includes → | Current paving | Current deterioration | Deposit-to-port bias | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Diamonds | Deterioration of roads, repaving costs equal to: | | Main ports double weight | Excl. deposits near rail |
| Country ↓ | | | | | 0.5 ∗ paving | 1.0 ∗ paving | | |
| Benin | 99.95 | −4.98 | 0.43 | 0.43 | 0.43 | 0.41 | 0.43 | 0.26 |
| Burkina Faso | 189.94 | −18.88 | 0.16 | 0.16 | 0.16 | 0.16 | 0.16 | 0.17 |
| Cote d'Ivoire | 266.40 | −52.09 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.15 |
| Ghana | 177.10 | −139.45 | 0.52 | 0.52 | 0.52 | 0.51 | 0.52 | 0.56 |
| Guinea | 138.26 | −4.40 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.19 |
| Guinea Bissau | 63.60 | −16.02 | 0.18 | 0.18 | 0.14 | 0.14 | 0.18 | 0.23 |
| Liberia | 63.24 | −9.43 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.08 |
| Mali | 293.87 | −67.74 | 0.16 | 0.16 | 0.15 | 0.15 | 0.16 | 0.10 |
| Niger | 265.99 | −33.68 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.02 |
| Senegal | 206.19 | −62.49 | 0.25 | 0.25 | 0.24 | 0.22 | 0.25 | 0.18 |
| Sierra Leone | 86.49 | −37.36 | 0.27 | 0.32 | 0.20 | 0.15 | 0.27 | 0.32 |
| Togo | 121.82 | −65.26 | 0.14 | 0.14 | 0.11 | 0.11 | 0.14 | 0.12 |
| Sample average | 174.57 | −46.55 | 0.25 | 0.25 | 0.23 | 0.22 | 0.25 | 0.20 |

Note: This table shows means by country, for country-periods with positive paving. Paving and deterioration in km. See Section 4 for variable definitions and sources.

Table OA4: Road paving characteristics by period, averages across countries

| | Paving since last map | Deterioration since last map | Deposit-to-port bias | | | | | |
|---|---|---|---|---|---|---|---|---|
| Counterfactual includes → | | | | Diamonds | Deterioration of roads, repaving costs equal to: | | Main ports double weight | Excl. deposits near rail |
| ↓ Period ending Jan 1: | | | | | 0.5 ∗ paving | 1.0 ∗ paving | | |
| 1967 | 158.03 | −57.91 | 0.07 | 0.07 | 0.03 | 0.03 | 0.07 | 0.02 |
| 1968 | 137.70 | −82.93 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.19 |
| 1969 | 129.94 | −27.69 | 0.16 | 0.16 | 0.14 | 0.14 | 0.16 | 0.19 |
| 1971 | 211.79 | −69.32 | 0.23 | 0.23 | 0.23 | 0.22 | 0.23 | 0.20 |
| 1973 | 135.92 | −42.13 | 0.21 | 0.21 | 0.20 | 0.18 | 0.21 | 0.11 |
| 1976 | 258.05 | −30.24 | 0.41 | 0.39 | 0.31 | 0.31 | 0.41 | 0.30 |
| 1983 | 542.69 | −19.49 | 0.33 | 0.33 | 0.34 | 0.30 | 0.33 | 0.21 |
| 1984 | 183.26 | −34.00 | 0.12 | 0.12 | 0.12 | 0.12 | 0.12 | 0.19 |
| 1986 | 210.85 | −43.74 | 0.23 | 0.26 | 0.22 | 0.22 | 0.23 | 0.16 |
| 1988 | 113.74 | −91.65 | 0.37 | 0.37 | 0.35 | 0.33 | 0.37 | 0.18 |
| 1989 | 22.83 | −44.29 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.53 |
| 1990 | 74.06 | −28.57 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 | 0.24 |
| 1991 | 93.78 | 0.00 | 0.19 | 0.19 | 0.19 | 0.19 | 0.19 | 0.12 |
| 1996 | 182.62 | −74.40 | 0.44 | 0.44 | 0.34 | 0.34 | 0.44 | 0.21 |
| 1998 | 141.91 | −42.21 | 0.27 | 0.27 | 0.27 | 0.27 | 0.27 | 0.26 |
| 2000 | 212.86 | −26.75 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 | 0.25 |
| 2002 | 77.37 | −161.33 | 0.32 | 0.32 | 0.32 | 0.28 | 0.32 | 0.13 |
| 2003 | 47.06 | −16.50 | 0.19 | 0.19 | 0.19 | 0.19 | 0.19 | 0.23 |
| 2007 | 251.24 | −19.47 | 0.12 | 0.20 | 0.12 | 0.12 | 0.12 | 0.15 |
| 2009 | 109.22 | −10.54 | 0.33 | 0.33 | 0.27 | 0.24 | 0.33 | 0.28 |
| 2012 | 197.37 | −74.48 | 0.46 | 0.46 | 0.45 | 0.36 | 0.46 | 0.37 |
| 2014 | 81.72 | −50.41 | 0.24 | 0.24 | 0.24 | 0.24 | 0.24 | 0.30 |
| Sample average | 174.57 | −46.55 | 0.25 | 0.25 | 0.23 | 0.22 | 0.25 | 0.20 |

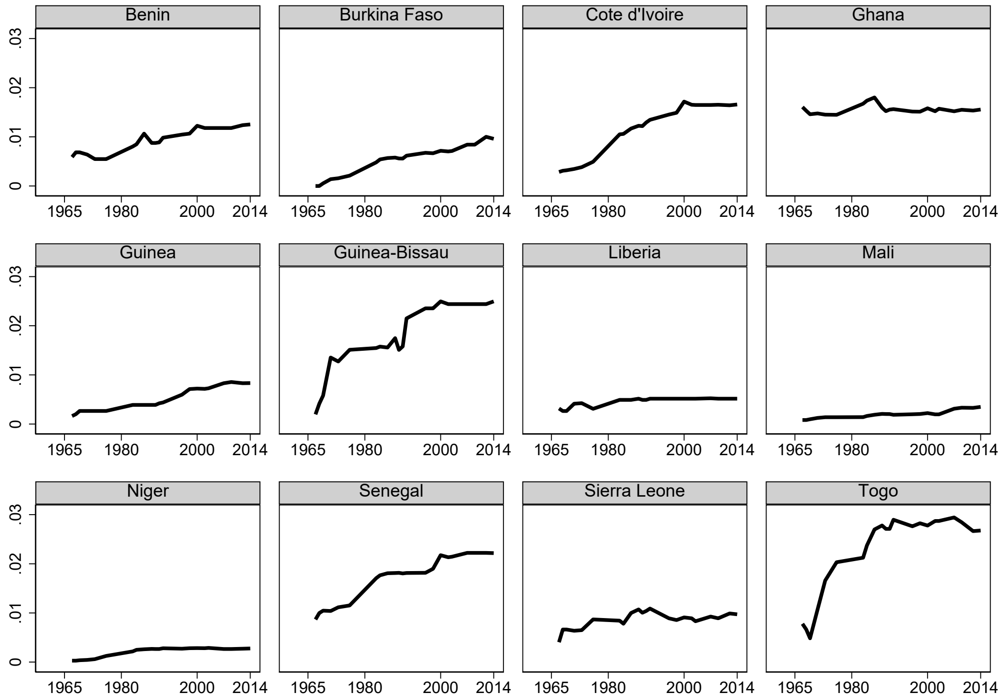
Note: This table shows means by period, for country-periods with positive paving. Paving and deterioration in km. See Section 4 for variable definitions and sources.

Table OA5: Road paving characteristics by period, Sierra Leone

| | Paving since last map | Deterioration since last map | Deposit-to-port bias | | | | |
|---|---|---|---|---|---|---|---|
| Counterfactual includes → | | | | Diamonds | Deterioration of roads, repaving costs equal to: | | Main ports double weight |
| ↓ Period ending Jan 1: | | | | | 0.5 ∗ paving | 1.0 ∗ paving | |
| 1967 | 81.16 | −60.62 | 0.31 | 0.31 | 0.08 | 0.08 | 0.31 |
| 1968 | 229.14 | −42.38 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 |
| 1969 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1971 | 110.52 | −128.48 | 0.45 | 0.45 | 0.43 | 0.43 | 0.45 |
| 1973 | 106.66 | −97.95 | 1.00 | 1.00 | 0.87 | 0.87 | 1.00 |
| 1976 | 193.56 | −36.07 | 0.72 | 0.57 | 0.24 | 0.24 | 0.72 |
| 1983 | 0.00 | −17.73 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1984 | 0.00 | −45.41 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1986 | 157.59 | 0.00 | 0.09 | 0.38 | 0.09 | 0.09 | 0.09 |
| 1988 | 52.62 | 0.00 | 0.56 | 0.56 | 0.41 | 0.20 | 0.56 |
| 1989 | 0.00 | −48.80 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1990 | 25.82 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1991 | 35.71 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1996 | 0.00 | −143.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1998 | 0.00 | −25.82 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2000 | 38.55 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2002 | 25.58 | −38.55 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2003 | 25.82 | −68.87 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2007 | 68.87 | 0.00 | 0.00 | 0.66 | 0.00 | 0.00 | 0.00 |
| 2009 | 0.00 | −25.58 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2012 | 133.05 | −61.89 | 0.81 | 0.81 | 0.81 | 0.36 | 0.81 |
| 2014 | 12.73 | −25.58 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Sample average | 58.97 | −39.41 | 0.18 | 0.22 | 0.14 | 0.11 | 0.18 |

Note: This table shows means by period, for Sierre Leone. Paving and deterioration in km. See Section 4 for variable definitions and sources.

Figure OA2: Stock of paved roads




year of discovery. 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.[66]

Summary statistics on deposits are presented in Table OA6. The database includes various kinds of metals and minerals: bauxite, chromium, copper, gold, iron ore, manganese and others. In the baseline estimation we focus on all metals and minerals except diamond deposits, since diamonds are less likely to be transported via ports and bulk ships than to be airlifted. Three-quarters of deposits were discovered after 1965, almost half are at least of major size, and gold is most common.

To calculate the value of deposits, we proceed in three steps. First, the deposit size is needed. For a number of mines, this figure is already provided in the MinEx database, consisting of the estimated deposit size at discovery. For the remaining group of mines, while an exact size is

[66] According 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".

Table OA6: Discoveries of deposits and their characteristics

| Country | Discoveries | | Metal/mineral | | Size | |
|---|---|---|---|---|---|---|
| | before or in 1965 | after 1965 | | | | |
| Benin | 2 | 1 | Gold | 257 | Super Giant | 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 | 163 |
| Guinea | 16 | 34 | Bauxite | 22 | Minor | 68 |
| Guinea-Bissau | 1 | 0 | Nickel | 7 | | |
| Liberia | 5 | 5 | Manganese | 6 | | |
| Mali | 8 | 46 | Phosphate | 6 | | |
| Niger | 9 | 27 | Titanium, Zirconium | 2 | | |
| 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 2014, the distribution of the main metal or mineral contained in them, and the distribution of deposit size. For sources, see Section 4.2.

not given, the MinEx database indicates the deposit's size category (supergiant, giant, major, moderate, and minor). We group giant and supergiant together because there is only one supergiant deposit (see Table OA6), and construct four categories (giant=4, major=3, moderate=2, minor=1).[67] We then employ the estimated weighted average size, which is provided by Minex for each mineral and size category, to fill in the deposit size of the remaining mines. For example, the estimated average size of an iron mine in the giant size category is 3,162 million tons of reserves.

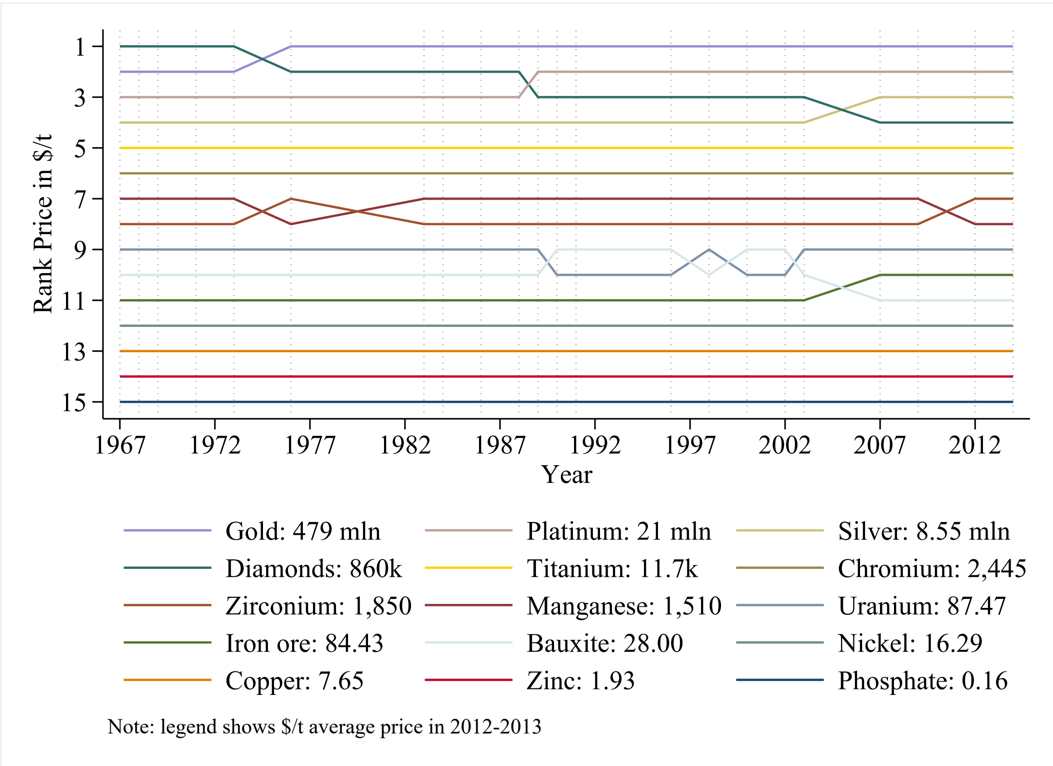
The second step involves the mineral prices by which the size of the deposits needs to be multiplied. To avoid values being influenced by price fluctuations, we collect, for each mineral, a time series of yearly commodity prices from 1960 to 2018 (sources indicated below). We therefore calculate each mine value by multiplying the deposit size with the average mineral price across the period under consideration. The rationale for using average prices is illustrated by Figure

[67]The 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 differ by a factor five). 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.

OA3 which ranks unit prices of all metals and minerals found in our sample of deposits over time. Using period-average unit prices (in USD per tonne), the figure shows that the rank-order is very stable over time. This implies, for example, that a gold mine of a certain size will always be worth more than a similarly-sized iron ore mine. Only diamonds have become relatively cheaper over time (possibly due to the advent of synthetic diamonds), but including or excluding diamonds does not affect our results. Other minerals that now and then switch rank have very similar unit prices.

Finally, based on the resulting mine values, we compute four value categories by dividing the value distribution into four quartiles.

Figure OA3: Unit price ranking of metals and minerals




## OA2.4 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 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 can accommodate larger ships with more draft, and 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.

### OA2.5 Cities

The data on cities was kindly shared by Hervé Gazel of the Africapolis project.[68] 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.[69] 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.[70]

### OA2.6 Prices

The data on the world price of 15 metal and mineral commodities comes from UNCTAD (downloaded June 14 2017), Plunkert & Jones (1999), Bazzi & Blattman (2014) and indexmundi.com.

### OA2.7 Aid, FDI, war and openness

We collected data for control variables on aggregate inflows of foreign aid and foreign direct investment, on the occurrence of war, and on openness to trade. 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. 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),

[68] Université Jean Moulin Lyon 3. See: `http://www.afd.fr/lang/en/home/publications/travaux-de-recherche/archives-anciennes-collections/NotesetEtudes/Africapolis`

[69] Table OA7 shows further details on the population distribution of these towns.

[70] Spearman's rank correlation test clearly rejects independence of 1960 and 2010 towns, with $\rho = 0.598$

Table OA7: 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 existed, 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.

Liberia (1989-90; 1992-95; 2002-03) and Sierra Leone (1991-00), in addition to Guinea-Bissau (1998-99). Finally, the data on openness comes from Wacziarg & Welch (2008), who constructed a dummy variable which is equal to zero if a country is "closed" to the world economy according to a set of criteria, and one if it is open. The dummy is available for all our countries up to the year 2005. For eight countries, it is equal to zero until it turns one in a year between 1984 and 2001, and remains one afterwards. For the other four countries (Liberia, Niger, Senegal and Togo) it is zero throughout.

### OA2.8 Ethnicity

For each deposit, we want to construct a measure of the extent to which the ethnic group on whose homeland the deposit is located is politically influential in each period. To this purpose, we combine data from Murdock's Ethnolinguistic Map of Africa (Murdock 1959) and from the Ethnic Power Relationship Core Dataset 2019 (Vogt et al. 2015; EPR from now on). The first dataset allows us to identify the ethnic group on whose homeland each deposit is located. The second dataset gives information on the political power of each ethnic group, in each year. More in detail, for each ethnic group that, in year $t$ or in any earlier year, was politically relevant, the EPR dataset classifies the group's influence on the central government into five categories: "dominant", "senior partner", "junior partner", "powerless" and "discriminated".[71] A sixth category, "irrelevant", is reserved for groups that are not politically relevant in period $t$ or in an earlier year, but will become so in a later year. Ethnic groups that were never politically relevant are not included. We matched ethnic groups in the two datasets, listing as irrelevant those groups which appear in the Ethnolinguistic Map, but not in the EPR dataset. We then create a dummy that is equal to one if the group's political status is dominant, senior partner, or junior partner, and zero if it is powerless, discriminated or irrelevant.[72] For each deposit, the dummy can switch on and off, as the group's influence can change over time. Since the EPR dataset measures political conditions on January 1 of each year, we measure a deposit's political influence at the start of period $t$ with the dummy for the first year of that period.

[71] The dataset also has category "monopoly", which however does not apply to any ethnic group in our sample.

[72] The former three categories include, respectively, 4,566, 5,935 and 15,242 observations (deposit-year). The latter include 5,202, 775 and 11,668 observations. There are also 124 observations that are coded as "state collapse" in the EPR dataset, and for which we set the dummy equal to zero.

## OA3 Dropping one country at a time

In this section we test for the sensitivity of our main results to dropping each country from the sample, one at a time. In Figure OA4 we rerun the regression of Table 2, column 4. Each dot is the point estimate of the regression that excludes the country that is listed on the left, and horizontal spikes show the 95% confidence intervals. We first repeat the average effect, then drop individual countries and finally, in the bottom of the Figure, we drop all landlocked countries. The panel on the left measures autocracy using the PolityIV database, while the panel on the right measures autocracy using the dummy in Acemoglu et al. (2019). In all cases the point estimates are significant and very similar to the average effect, suggesting that no single country is influential enough to drive the results.

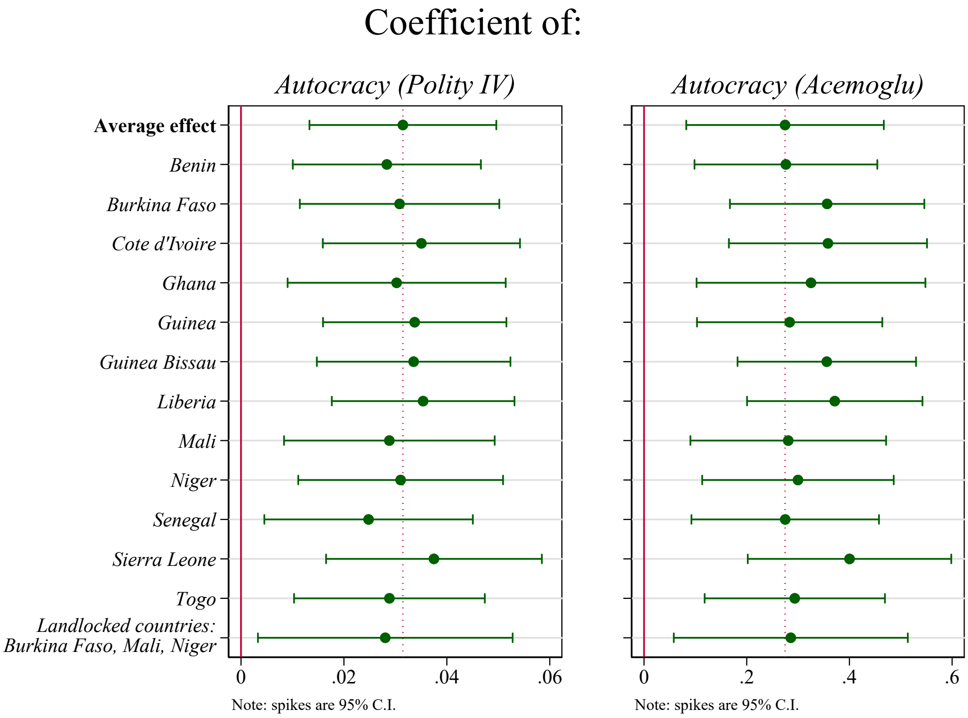



Figure OA4: Dropping one country at a time

## OA4 Testing for trends and first differenced specifications

We estimate our main results in levels and allow for deterministic trends by always including $trend_i$ in the regressions. Here we test formally for non-stationarity of autocracy, since all countries have become more democratic over time, and re-estimate our main results in first differences.

Table OA8 performs a range of Breitung unit root tests for autocracy in levels (in the first four rows), and in first differences (in the last four rows). The Breitung (2001) test has good power with small datasets. It regresses the first difference of the variable on 1 or 2 lags of itself for prewhitening of the series before computing the test based on the residuals of that regression. The test includes a trend (in the last two columns) and is robust to cross-sectional correlation. The drawback is that the test requires a balanced panel, which is why we construct balanced subsets of our data by taking only the years from 1975, or by excluding Guinea-Bissau. We do this for both the Polity IV and the Acemoglu version of autocracy.

The test indeed finds a unit root for the Polity IV version of democracy in levels with 1 or 2 lags, but when also a country-specific trend is included the tests comfortably reject a unit root. We thus find that including $trend_i$ in the regressions is important to take care of a possible trend in autocracy when usign the Polity IV version. First differencing the variable in the bottom rows shows that the tests reject higher order non-stationarity.

For completeness, Table OA9 performs our baseline regression in first differences with and without additional (two-way) fixed effects and trends, both for the Polity IV and the Acemoglu measure of autocracy. Our main result is robust to all these alternative specifications, even though we lose power due to the first difference transformation dropping the early observations.

Table OA8: Breitung unit root tests

| | H0: Panels contain unit roots<br>Ha: Panels are stationary | | | | | |
|---|---|---|---|---|---|---|
| Lags/trend → | (1) | (2) | (1)+trend | (2)+trend | N | T |
| Autocracy, post 1975 | 0.386 | 1.007 | -3.963 | -2.888 | 13 | 40 |
| | 0.242 | 0.439 | 0.000 | 0.002 | | |
| Autocracy, excl. Guinea-Bissau | 0.230 | 1.199 | -2.303 | -2.031 | 12 | 50 |
| | 0.123 | 0.424 | 0.023 | 0.002 | | |
| Autocracy (Acemoglu et al., 2019), post 1975 | -1.332 | -0.555 | -0.696 | -0.095 | 13 | 36 |
| | 0.028 | 0.153 | 0.003 | 0.043 | | |
| Autocracy (Acemoglu et al., 2019), excl. Guinea-Bissau | -1.544 | -1.152 | 0.515 | 0.468 | 12 | 46 |
| | 0.008 | 0.039 | 0.140 | 0.100 | | |
| Δ Autocracy, post 1975 | -9.903 | -5.063 | -9.296 | -7.196 | 13 | 39 |
| | 0.000 | 0.000 | 0.000 | 0.000 | | |
| Δ Autocracy, excl. Guinea-Bissau | -5.927 | -6.344 | -8.729 | -9.359 | 12 | 49 |
| | 0.001 | 0.000 | 0.000 | 0.000 | | |
| Δ Autocracy (Acemoglu et al., 2019), post 1975 | -4.278 | -10.550 | -2.241 | -2.422 | 13 | 35 |
| | 0.000 | 0.000 | 0.000 | 0.000 | | |
| Δ Autocracy (Acemoglu et al., 2019), excl. Guinea-Bissau | -6.363 | -11.306 | -1.906 | -1.603 | 12 | 45 |
| | 0.000 | 0.000 | 0.000 | 0.000 | | |

*Notes*: Table shows Breitung test statistics with p-values below. The Breitung test requires a balanced panel. It regresses the first difference of the variable on (1) or (2) lags of itself for prewhitening of the series before computing the test based on the residuals of that regression. The test includes a trend in columns 3 and 4 and is robust to cross-sectional correlation.

Table OA9: First differenced estimates of baseline regression

| Dependent variable → | Deposit-to-port bias | | | | Deposit-to-port bias | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Autocracy, start of t | 0.030* | 0.029* | 0.030* | 0.036** | | | | |
| | (0.016) | (0.018) | (0.018) | (0.018) | | | | |
| Autocracy, start of t (Acemoglu) | | | | | 0.589*** | 0.567*** | 0.608*** | 0.617*** |
| | | | | | (0.133) | (0.138) | (0.134) | (0.138) |
| Observations | 150 | 150 | 150 | 150 | 152 | 152 | 152 | 152 |
| R-squared | 0.053 | 0.193 | 0.266 | 0.328 | 0.170 | 0.323 | 0.392 | 0.416 |
| Country dummies | No | No | Yes | Yes | No | No | Yes | Yes |
| Period dummies | No | Yes | Yes | Yes | No | Yes | Yes | Yes |
| Country-specific trends | No | No | No | Yes | No | No | No | Yes |

*Notes*: This table repeats the baseline specification with all variables expressed in first differences.

## OA5 Main results using Acemoglu et al. (2019)'s measure of autocracy

Table OA10: Baseline model using Acemoglu et al. (2019)'s measure of autocracy

| Dependent variable → | Deposit-to-port bias | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Autocracy, start of t (Acemoglu) | 0.274*** | 0.287*** | 0.286*** | 0.322*** | -0.013 |
| | (0.102) | (0.097) | (0.097) | (0.088) | (0.142) |
| Cum. deposit overlap, start of t | | -0.249 | -0.234 | -0.322 | -0.283 |
| | | (0.316) | (0.320) | (0.330) | (0.307) |
| Cum. deposit discoveries, start of t | | 0.002 | 0.002 | 0.002 | 0.009 |
| | | (0.008) | (0.008) | (0.007) | (0.008) |
| Deposit discoveries, t-1 | | 0.041*** | 0.041*** | 0.038*** | 0.035** |
| | | (0.015) | (0.015) | (0.014) | (0.015) |
| Price index, average t | | | -0.000 | -0.000 | -0.000 |
| | | | (0.000) | (0.000) | (0.000) |
| Ethnic deposits, start of t | | | | 0.478*** | 0.546*** |
| | | | | (0.138) | (0.142) |
| Ethnic deposits * Autocracy | | | | | 0.506*** |
| | | | | | (0.178) |
| Observations | 173 | 162 | 162 | 162 | 162 |
| R-squared | 0.406 | 0.458 | 0.458 | 0.517 | 0.537 |
| Country FE | Yes | Yes | Yes | Yes | Yes |
| Period FE | Yes | Yes | Yes | Yes | Yes |
| Country-specific trends | Yes | Yes | Yes | Yes | Yes |
| Wild bootstrap Autocracy t-stat | 2.093 | 2.127 | 2.098 | 2.409 | 1.675 |
| Wild bootstrap Autocracy p-value | 0.023 | 0.031 | 0.034 | 0.017 | 0.143 |
| Wild bootstrap interaction t-stat | | | | | 2.640 |
| Wild bootstrap interaction p-value | | | | | 0.045 |
| Marginal effects of autocracy for | | | | | |
| Median ethnicity | | | | | 0.148 |
| | | | | | (0.102) |
| Mean ethnicity | | | | | 0.203** |
| | | | | | (0.092) |
| 75% percentile ethnicity | | | | | 0.479*** |
| | | | | | (0.102) |

*Notes*: This table shows OLS regressions of the effect of autocracy on the degree to which countries' actual paving overlaps with the counterfactual network that connects deposits to ports. All specifications include country and period fixed effects, and country-specific linear trends. Robust standard errors in parenthesis: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Wild bootstrap statistics are clustered by country and are performed using the method of Roodman et al. (2019). See Section 4 for variable definitions and sources.

Table OA11: IV: Endogenous autocracy using Acemoglu et al. (2019)'s measure of autocracy

| Dependent variable → | Deposit-to-port bias | | Autocracy | Deposit-to-port bias |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Autocracy, start of t (Acemoglu) | 0.322*** | 0.374*** | | 1.660*** |
| | (0.088) | (0.089) | | (0.588) |
| Cum. deposit overlap, start of t | -0.322 | -0.394 | 0.472* | -0.841* |
| | (0.330) | (0.327) | (0.271) | (0.491) |
| Cum. deposit discoveries, start of t | 0.002 | 0.001 | -0.008 | 0.012 |
| | (0.007) | (0.007) | (0.007) | (0.012) |
| Deposit discoveries, t-1 | 0.038*** | 0.020* | 0.014 | 0.010 |
| | (0.014) | (0.012) | (0.018) | (0.023) |
| Price index, average t | -0.000 | -0.000 | -0.000 | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Ethnic deposits, start of t | 0.478*** | 0.465*** | -0.120 | 0.612*** |
| | (0.138) | (0.151) | (0.119) | (0.199) |
| Lagged regional average of autocracy (Polity) | | | 0.073** | |
| | | | (0.032) | |
| Observations | 162 | 162 | 162 | 162 |
| Country FE | Yes | Yes | Yes | Yes |
| Period FE | Yes | No | No | No |
| Country-specific trends | Yes | Yes | Yes | Yes |
| 5-Year FE | No | Yes | Yes | Yes |
| Wild bootstrap Autocracy t-stat | 2.409 | 2.890 | | 3.049 |
| Wild bootstrap Autocracy p-value | 0.017 | 0.003 | | 0.018 |
| 1st stage F-test | | | | 5.557 |
| Anderson-Rubin weak IV robust 95% C.I. | | | | [0.798,...] |

*Notes*: This table shows OLS regressions of the effect of autocracy on the degree to which countries' actual paving overlaps with the counterfactual network that connects deposits to ports. All specifications include country and period fixed effects, and country-specific linear trends. Robust standard errors in parenthesis: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Wild bootstrap statistics are clustered by country and are performed using the method of Roodman et al. (2019). See Section 4 for variable definitions and sources.

Table OA12: Mechanisms using Acemoglu et al. (2019)'s measure of autocracy

| Dependent variable → | Deposit-to-port bias | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Autocracy, start of t (Acemoglu) | 0.327*** | 0.336*** | 0.380*** | 0.308*** | 0.327*** | 0.304*** | 0.354*** |
| | (0.086) | (0.090) | (0.111) | (0.104) | (0.090) | (0.082) | (0.113) |
| Current paving, t | 0.000* | | | | | | 0.000 |
| | (0.000) | | | | | | (0.000) |
| Openness to trade, average t | | -0.068 | | | | | -0.002 |
| | | (0.124) | | | | | (0.144) |
| Log FDI, average t | | | 0.017 | | | | 0.030 |
| | | | (0.036) | | | | (0.033) |
| Log aid, average t | | | | -0.048 | | | 0.045 |
| | | | | (0.114) | | | (0.113) |
| Civil war, average t | | | | | 0.232 | | 0.297** |
| | | | | | (0.169) | | (0.115) |
| City overlap | | | | | | 0.349*** | 0.442*** |
| | | | | | | (0.090) | (0.111) |
| Observations | 162 | 153 | 135 | 162 | 162 | 162 | 126 |
| R-squared | 0.527 | 0.514 | 0.557 | 0.518 | 0.527 | 0.592 | 0.666 |
| Standard controls included | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Period FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-specific trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Wild bootstrap Autocracy t-stat | 2.482 | 2.529 | 2.162 | 1.943 | 2.309 | 2.661 | 3.001 |
| Wild bootstrap Autocracy p-value | 0.014 | 0.018 | 0.062 | 0.060 | 0.025 | 0.015 | 0.027 |

*Notes*: This table shows OLS regressions of the effect of autocracy on the degree to which countries' actual paving overlaps with the counterfactual network that connects deposits to ports. All specifications include country and period fixed effects, and country-specific linear trends. Robust standard errors in parenthesis: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Wild bootstrap statistics are clustered by country and are performed using the method of Roodman et al. (2019). See Section 4 for variable definitions and sources.

Table OA13: Robustness using Acemoglu et al. (2019)'s measure of autocracy

| Dependent variable → | Deposit-to-port bias | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Counterfactual includes → | | | Diamonds | Deterioration of roads, with repaving costs equal to: | | Main ports double weight | Excl. deposits near railroads | Ranked by distance only | Ranked by value only |
| | | | | 0.5 * paving | 1.0 * paving | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Autocracy, start of t (Acemoglu) | 0.322** | 0.322** | 0.299*** | 0.324*** | 0.324*** | 0.322*** | 0.234*** | 0.367*** | 0.389*** |
| | (0.128) | (0.120) | (0.094) | (0.090) | (0.087) | (0.088) | (0.053) | (0.086) | (0.084) |
| Cum. deposit overlap, start of t | -0.322 | -0.322 | -0.581+ | -0.318 | -0.385 | -0.322 | -0.441+ | -0.556* | -0.419+ |
| | (0.269) | (0.287) | (0.366) | (0.360) | (0.360) | (0.330) | (0.268) | (0.308) | (0.274) |
| Observations | 162 | 162 | 162 | 162 | 162 | 162 | 162 | 163 | 163 |
| R-squared | 0.517 | 0.517 | 0.494 | 0.502 | 0.495 | 0.517 | 0.640 | 0.539 | 0.545 |
| Country FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Period FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country-specific trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clustered on | country | country & year | | | | | | | |
| Wild bootstrap Autocracy t-stat | | | 2.011 | 2.117 | 2.286 | 2.409 | 3.957 | 2.751 | 3.288 |
| Wild bootstrap Autocracy p-value | | | 0.060 | 0.038 | 0.022 | 0.017 | 0.003 | 0.012 | 0.003 |

*Notes*: This table shows OLS regressions of the effect of autocracy on the degree to which countries' actual paving overlaps with the counterfactual network that connects deposits to ports. All specifications include country and period fixed effects, and country-specific linear trends. Robust standard errors in parenthesis: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Wild bootstrap statistics are clustered by country and are performed using the method of Roodman et al. (2019). See Section 4 for variable definitions and sources.