RESEARCH NOTE



State reach and development in Africa since the 1960s: new data and analysis

Carl Müller-Crepon 匝

Department of Political Science and International Relations, University of Oxford, Oxford, UK Corresponding author. Email: carl.muller-crepon@politics.ox.ac.uk

(Received 14 December 2020; revised 9 June 2021; accepted 27 June 2021; first published online 10 November 2021)

Abstract

Prominent arguments hold that African states' geography limits state capacity, impedes public service provision, and slows economic development. To test this argument, I collect comprehensive panel data on a proxy of local state capacity, travel times to national and regional capitals. These are computed on a yearly 5×5 km grid using time-varying data on roads and administrative units (1966–2016). I use these data to estimate the effect of changes in travel times to capitals on local education provision, infant mortality rates, and nightlight emissions. Within the same location, decreases in travel times to its capitals are robustly associated with improved development outcomes. The article advances the measurement of state capacity and contributes to understanding its effects on human welfare.

Keyword: Data collection

1. Introduction

Adverse population distributions, arbitrary colonial borders, and deficient transport networks limit state capacity in Africa and are blamed to impede good governance and economic development on the continent (e.g., Herbst, 2000). However, empirical tests of this reasoning lag behind its frequent use as we lack geographically disaggregated and time-varying data on African states' capacity to estimate its effect on local development. This research note introduces such panel data on local state capacity proxied by travel times to regional and national capitals. With these data, I document a generally positive, robust association between improving local state capacity and development.

The new data of state capacity builds on previous measures that have, after an initial focus on the country-level (Hendrix, 2010), captured important subnational variation in state capacity (Boone, 2003). However, their reliance on either surveys (Wig and Tollefsen, 2016) or age-heaping in censuses (Lee and Zhang, 2017) limits spatio-temporal coverage and availability in low-capacity states. As a remedy, I develop a continent-wide, yearly proxy of local state capacity based on travel times to post-independence regional and national capitals between 1966 and 2016.¹ I compute these travel times for a 5×5 km grid of the continent using new panel data on administrative geographies and road networks digitized from the Michelin map corpus. The data are validated with information on local state capacity from the Afrobarometer (2018).

I use the new data to estimate the effect of local state capacity on development. Previous studies provide cross-sectional evidence that locations farther away from capitals experience more conflict (e.g., Tollefsen and Buhaug, 2015), suffer from corruption (Campante and Do, 2014; Krishna and Schober, 2014), and exhibit lower levels of development (Acemoglu *et al.*, 2015; Henn, 2020), an effect that is historically persistent (Pierskalla *et al.*,

¹Available at https://github.com/carl-mc/state_reach_africa.

[©] The Author(s), 2021. Published by Cambridge University Press on behalf of the European Political Science Association.

2017).² I address potential cross-sectional endogeneity of administrative geographies by focusing on variation within locations and country-years. I find that reductions in travel times to capitals are associated with increases in education and infant survival rates, as well as nightlight emissions. These results are not explained by spurious migration, differential pre-trends, endogenous road construction, improvements in economic market access, or ethnic favoritism and conflict.

2. State capacity and development

State capacity denotes states' ability to enforce their will by monitoring and steering citizens' behavior (Mann, 1984; Migdal, 1988). It consists of their administrative capacity, military strength, and ability to tax individuals (Hendrix, 2010). Physical access to citizens through transport infrastructure is a key determinant of state capacity (Boulding, 1962; Mann, 1984; Herbst, 2000; Acemoglu *et al.*, 2015).

Four mechanisms link physical accessibility via local state capacity to development. First, low transport costs between states' headquarters and the population enables bureaucrats and police officers to enforce law and order, thereby triggering the developmental effects of centralized state institutions (Huntington, 1968; Campante and Do, 2014). Second, states' provision of services such as health care and education depends on their capability to monitor demand and control agents, both of which increases with local state capacity (Krishna and Schober, 2014; Henn, 2020). In addition, citizens' access to specialized services (e.g., hospitals and courts) increases near the administrative capitals that typically harbor them. Third, smooth local accessibility increases citizens knowledge of and ability to tap into private rents from the state such as public sector jobs or subsidies (Ades and Glaeser, 1995; Banerjee *et al.*, 2018).

A fourth mechanism that links state capacity with citizens' welfare consists of taxation. As states become locally more capable, they are able to collect more taxes, which will, ceteris paribus, decrease citizens' material welfare. With that, the net effect of state capacity depends on whether the exchange of taxes for state-provided good is generally beneficial for citizens (Timmons, 2005) or not (Scott, 2017). While unable to disentangle the above theorized four mechanisms, the empirical analysis below sheds light on the overall effect of state capacity on local development.

3. Approximating state reach

Because physical accessibility is a necessary (but not sufficient) condition for state capacity, I proxy local state reach with the weighted sum of travel times to a locations' national and regional capitals:³

total state reach_{*p,t*} =
$$\sum_{u=1}^{U} \omega_u$$
 state reach_{*p,u,t*} (1)

$$=\sum_{u=1}^{U} -\omega_{u} \ln\left(1 + d_{t}(p, C_{p,u,t})\right)$$
(2)

where state reach on a level of administrative hierarchy u toward point p at time t is calculated as the travel time (in hours) on the shortest path between p and its capital $C_{p,u,t}$ on the road network at time t.⁴ The log-transform⁵ captures the convex relation between travel times to capitals and state

²Similarly, historical post offices fostered development in the US (Rogowski et al., 2021).

³No panel data exist on 2nd or 3rd level administrative units.

⁴The relation of travel times with state capacity likely varies by country. This can be reflected by normalizing travel times by country characteristic x_c . The country-year fixed effects in the analysis constitute such a normalization because $\log((1 + d_t(p, C_{p,u,t}))/x_c))$ decomposes to $\log(1 + d_t(p, C_{p,u,t})) - \log(x_c)$.

⁵Adding 1 hour prevents taking the log of 0 in capitals.

capacity (Figure 2). Weights ω_u denote the impact of times to capitals on level u. Because ω_u likely varies across different state activities, I estimate ω_u for each development outcome separately. I measure state reach_{p,u,t} with time-varying data on regional and national administrative geographies and road networks.⁶

First, I collect comprehensive panel data on the boundaries and capitals of first-level (regional) administrative units since African countries' independence. Drawing on diverse secondary sources, the dataset covers 1763 unique region-periods between independence and 2016 (Appendix A). Data on national borders and capitals come from Cshapes (Weidmann and Gleditsch, 2010). Changes in administrative unit designs and capital locations lead to temporal variation in state reach_{p,u,t}.

Second, I transform the corpus of Michelin road maps into a digital road atlas akin to a timevarying Google Maps.⁷ Scanned maps are available at a scale of 1:4 million⁸ for 23 years between 1966 and 2014. Following Müller-Crepon *et al.* (2021), pixels that depict roads are classified with a fully convolutional neural network, vectorized, and transformed into a planar graph that covers the African landmass at a resolution of 0.0417 decimal degrees (\approx 5 km; Appendix B). Each edge on the network is associated with a travel time derived from the quality of roads observed on the Michelin maps. With these data on roads and administrative units, I can compute the travel time on the shortest path from each location in Africa to its corresponding national and regional capitals in a given year. Doing so results in the measure state reach_{*p*,*u*,*t*} for the grid cells of an Africa-wide raster with a 0.0417 decimal degree resolution for every year between 1966 and 2016 as depicted in Figure 1.

I validate the data as a proxy for local state capacity using a state capacity index constructed from data on Afrobarometer's (2018) enumeration areas. The survey provides geocoded data on the presence of the police and military, post offices, schools, and hospitals, as well as public services such as water, sewage, and electricity. I aggregate this information via the first component of a Principal Component Analysis, capturing 36.2% of the data's variation. Figure 2 shows that travel times to capitals correlate strongly with the resulting state capacity index. Travel times also fit the index better than mere geodesic, as-the-crow-flies distances (Appendix C.1). In sum, travel times to capitals are a valid proxy of local state capacity.

3.1 State reach in Africa: weak, growing, and unequal

African states' reach has increased since their independence (Figure 3). Travel times between national capitals and citizens have decreased from a 1966 average of 11.7 hours to 9.3 hours in 2016, a change of 20.3 percent. Travel times to regional capitals decreased from 5.1 to 3.5 hours.

Three trends underlie this development (Appendix C.2). First, administrative geographies have changed. Côte d'Ivoire, Nigeria, and Tanzania have relocated their national capitals⁹ and most governments have created new administrative regions (Grossman and Lewis, 2014; Grossman *et al.*, 2017).¹⁰ Second, transport infrastructure has improved, with quality-weighted mileage increasing by about 50 percent since independence. Lastly, many citizens moved closer to capitals as urbanization rates doubled from 20 to 40 percent (World Bank, 2018).

The maps in Figure 1 show that travel times to capitals exhibit substantive variation within and across countries. In contrast to small Rwanda, travel times to the DR Congo's capital Kinshasa

⁶Surpassing rail or air transport, road transport is the dominant motorized transport mode in Africa (Jedwab and Storeygard, 2020).

⁷Jedwab and Storeygard (2020) use the Michelin maps to measure economic market access.

⁸This corresponds to 1 mm per 4 km and puts an upper precision limit on the data.

⁹Abidjan to Yamoussoukro (1983), Lagos to Abuja (1991), and Dar es Salaam to Dodoma (1974, but ministries remain in Dar es Salaam).

¹⁰Similarly, capitals of independent Eritrea and South Sudan moved closer to citizens.



Figure 1. State reach in Africa, 1966-2016.

vary between 0 and 50 hours, with a high median of 34.3 hours, reflecting the country's "difficult geography" (Herbst, 2000) and sparse infrastructure. Changes in travel times since 1966 are marked by similar variation. States and regions that seceded or relocated their capital experienced the largest improvements. These changes exhibit substantive spatial variation because changes in administrative geographies and road networks have spatially varying effects. My empirical strategy exploits this fact.

4. Data on local development

The empirical analysis examines whether changes in African states' reach affected local development, measured with data on education, infant mortality, and nightlight emissions.¹¹ To measure primary education, I rely on the Demographic and Health Survey (DHS, 2018). The DHS encodes the educational achievement of all members of sampled households. Assuming that they live where they were raised, I model the level of education of 2.1 million individuals older than 15 as having been influenced by the travel time to their capitals at age 6. An "attended primary school" dummy serves as the main outcome, capturing the arguably most important component of educational services and avoiding the skewed distribution of individuals' education levels.¹²

The DHS also measures the under-1 mortality of 2.9 million infants born to women aged 15– 49. I model infants' deaths as depending on the travel time to their mothers' capitals at their birth. Infant mortality and education rates exhibit local temporal variation because children and

¹¹Appendix D.2 provides summary statistics.

¹²Appendix E.6 shows robustness across different education levels.



Figure 2. Travel times to national and regional capitals and state capacity index. Values are demeaned by country × survey round and averaged within 40 *x*-axis quantiles.



Figure 3. Decreasing travel times to regional and national capitals in Africa 1960–2015. Population-weighted averages based on HYDE population estimates (Goldewijk *et al.*, 2010) and travel times to capitals. (a) Average time to national capital. (b) Average time to regional capital.

household members sampled in the same place have been born and schooled at different times. In modeling this variation, I assume that migration decisions do not respond to changes in state reach in a manner that correlates with the observed outcomes and show that results hold among non-migrants.

To mitigate the uneven survey coverage and counter concerns about migration biases, I also proxy local development with nightlight emissions (Weidmann and Schutte, 2017) derived from satellite images (1992–2013; National Geophysical Data Center, 2014). I measure nightlights within centroidal Voronoi cells $(400 \text{ km}^2)^{13}$ that are nested within administrative regions (Figure 4 and Appendix D.1). In contrast to quadratic cells that frequently overlap with

¹³Results are robust to varying cell sizes (Appendix E.7).



Figure 4. Voronoi cells, regions, and nightlights in Uganda 1992–2013. Panel (a) plots all regional borders between 1992 and 2013 (red) and Voronoi cells (black). Unpopulated cells in (b) in grey.

(changing) administrative borders, travel times can be consistently aggregated to the Voronoi cells. Nightlights are log-transformed after adding a constant of 0.001.¹⁴

5. Empirical strategy

The empirical analysis estimates the effect of travel times to capitals on local development. To account for any cross-sectional endogeneity from, e.g., strategic road, capital, and border placements and geographical confounders, I only study *temporal* variation in travel times and development within the same location:

$$Y_{i,p,c,t,s} = \alpha_p + \lambda_{c,t} + \mu_s + \beta_1 \text{ time to nat. cap.}_{p,t} + \beta_2 \text{ time to reg. cap.}_{p,t} + \delta X_i + \epsilon_{i,p,c,t},$$
(3)

where β_1 and β_2 are the effects associated with travel times to point p's regional and national capitals at time t. In parallel to weights ω_u in Equation 1, the sum of β_1 and β_2 proxies the total effect of state capacity. The units of analysis i are individuals in the education and infant mortality analyses and Voronoi cells in the examination of nightlight emissions. DHS respondents i are nested in location p and year t.¹⁵ p is synonymous with i where Voronoi cells are the units of analysis. The model controls for all constant attributes of locations/Voronoi cells, country-years, and surveys through fixed effects α_p , $\lambda_{c,b}$ and μ_s . Individual-level controls X_i in the education models consist of respondents' sex, age, and its square. Infant mortality models include mothers' age at birth and its square, infants' birth-order, and its square, as well as female and twin dummies. I add no time-varying covariates to the baseline nightlight model. I cluster standard errors on the point and country-year levels.

¹⁴I drop cells with oil wells that cause bright flares of burning gas (Lujala et al., 2007).

¹⁵In the individual-level analyses, locations p are the grid cells from the 5 × 5 km travel time raster. The DHS's random displacement of clusters by up to 10 (2) kilometers in rural (urban) areas adds noise.

	Table 1.	Changes i	in time t	o national/	regional	capital	and local	develo	pmer
--	----------	-----------	-----------	-------------	----------	---------	-----------	--------	------

	Primary educ. (0/100)	Infant mort. (0/100)	Light/capita (log)
	(1)	(2)	(3)
Time to nat. capital (log)	- 2.667***	0.878***	- 0.052**
	(0.400)	(0.233)	(0.026)
Time to reg. capital (log)	- 1.253***	0.203	- 0.037**
	(0.305)	(0.145)	(0.019)
$\beta_1 + \beta_2$:	- 3.92***	1.082***	- 0.089***
	(0.465)	(0.262)	(0.028)
Point FE: Country-year FE: Survey FE: Controls: Mean DV: Observations Adjusted <i>R</i> ²	yes yes yes 70 2,091,736 0.452	yes yes yes 9.6 2,895,564 0.051	yes yes - -6.5 1,507,013 0.836

Notes: OLS linear models. Individuals are the units of the education and infant mortality analyses, Voronoi cells those of the nightlight analyses. Control variables for models with primary education as the dependent variable consist of responents' age and age squared, as well as a female dummy. Where infant mortality is the dependent variable, models include an infant's mother's age at birth and its square, the birthorder and its square, as well as a female and twin dummy. Standard errors clustered on the point and country-year levels. Significance codes: *p < 0.1; **p < 0.05; ***p < 0.01.

Relying on this two-way fixed-effect estimator,¹⁶ I assume that changes in travel times to capitals are exogenous to local development outcomes observed thereafter. Robustness checks account for potential violations of this assumption, in particular non-parallel trends, endogenous roadbuilding, spurious correlations with changes in economic market access, or potential omitted variables such as ethnic inclusion or civil war.

6. Results

The results show that reductions in travel times toward regional and national capitals are, first, robustly associated with increasing primary education rates. Second, infant mortality rates improve as travel times to national, but not regional capitals decrease. Lastly, nightlight emissions significantly increase with declining travel times to regional capitals. Their relation to changes in the time to national capitals is of roughly similar size but statistically more unstable. Interpreting the sums of β_1 and β_2 as total effect of state capacity reveals its consistently positive, meaningful, and statistically significant association with local development.

Table 1 presents the baseline results. Interpreting coefficients in substantive terms, Model 1 indicates that a reduction of travel times to national (regional) capitals from 2 to 1 hours is associated with a precisely estimated increase in primary education rates by 1.1 (0.54) percentage points. Model 2 shows that infant mortality rates decrease by 0.38 percentage points when the time to the national capital decreases from 2 to 1 hours. In contrast, there is no evidence for a meaningful association of infant mortality with regional capital accessibility. The respective coefficient is close to zero and statistically insignificant.

Distinguishing the effects of relocations of national capitals¹⁷ from those of road network changes reveals that national capital relocations primarily drive the effects on education and infant mortality rates (Appendix E.1). Changes due to road networks have a smaller, but statistically significant effect on primary education and no effect on infant mortality.

¹⁶The "treatments" are continuous, time-variant travel times and do not suit a binary difference-in-difference design. ¹⁷There are very few regional capital relocations.



Figure 5. Robustness checks. Coefficient estimates with 95 percent Cls.

Lastly, Model 3 yields a significant association of travel times to regional capitals with nightlight emissions. As the local distance to the regional capital decreases from 2 to 1 hours, per-capita nightlight emissions increase by 1.8 percent. The same change in the distance to its national capital is associated with a similar increase in nightlights, which is however affected by differential pre-trends (Appendix E.3). The estimate's noisiness is related to the short coverage of the data (1992–2013) with changes of national capitals in newly independent South Sudan (2013) and Eritrea (1992), each with only one pre-/post-treatment year. Thus, changes in road networks with their more diffuse impact on development drive the estimate.

Figure 5 presents the results of robustness checks fully discussed in Appendix E. First, I find that the effects of travel times to capitals hold among non-migrant DHS respondents. Second and evidence against reverse causality, I find no significant effect of differential pre-trends, except for those that affect the nightlight model. Third, the results are not caused by potentially endogenous road building accounted for by time-varying local road density-measures or by computing travel times on time-invariant road data from 1966. Fourth, a potentially spurious correlation of travel times to capitals with economic market access to the 1530 biggest African cities does not drive the results. Fifth, the results are not due to potential omitted variables, in particular ethno-political inclusion, exposure to ethnic civil wars, regional and national capital dummies, and population density controls. Lastly, Appendix E.6 demonstrates robustness across alternative outcome specifications.

These analyses suggest generally positive developmental effects of increases in state capacity as proxied by estimated aggregate effects of travel times to national and regional capitals. Where national and regional administrations move closer to the citizens they govern, education rates improve. Infant mortality rates increase with travel times to national but not regional capitals. Lastly, nightlights become brighter when regional capitals move closer. Pre-trends partially drive the developmental effects of closer national capitals on nightlights, raising concerns about reverse causality.

7. Conclusion

One important constraint of states' capacity and their ability to foster development is their physical access to the population (Herbst, 2000). Measuring African states' varying success in making their citizens accessible, this research note has introduced new spatio-temporal data on travel times between administrative capitals and citizens as a proxy for post-colonial state capacity. The new data show a generally positive association between increases in local state capacity on various indicators of citizens' wealth and well-being. Difficult geographies with hard-to-access populations thus impede governance and economic development on the continent. At the same time, the results reject geographic determinism since states can and do improve their reach and thereby foster local development.

Offering new data and empirical results, this research note opens up avenues for future research. A first set of questions concerns the effects of state capacity on different state-provided goods and under different institutional settings. Second, there might be important heterogeneity in the mechanisms through which and when local state capacity matters most for citizens' welfare. Crises such as droughts, floods, or violent conflicts may exacerbate the effects of weak state-hood. Lastly, researchers could use the new measure of state capacity to study state-society interactions more generally, ranging from resource extraction, over accountability, to conflict processes.

Acknowledgements. I thank Lukas Dick and Benjamin Füglister for excellent research assistance. I am grateful to two anonymous reviewers, Lars-Erik Cederman, Robert H. Bates, Simon Hug, Philipp Hunziker, Horacio Larreguy, Clara Neupert-Wentz, Yannick Pengl, Ryan Saylor, and audiences at the APSA Annual Meeting 2019, the AFK workshop in Hamburg, and ETH Zurich for very helpful comments and feedback. I acknowledge support from the Swiss National Science Foundation under grant P0EZP1_165233.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/psrm.2021.60.

References

Acemoglu D, García-Jimeno C and Robinson JA (2015) State capacity and economic development: a network approach. American Economic Review 105, 2364–2409.

Ades AF and Glaeser EL (1995) Trade and circuses: explaining urban giants. *Quarterly Journal of Economics* 110, 195–227. Afrobarometer (2018) Afrobarometer data. Available at http://www.afrobarometer.org.

Banerjee A, Hanna R, Kyle J, Olken BA and Sumarto S (2018) Tangible information and citizen empowerment: identification cards and food subsidy programs in Indonesia. *Journal of Political Economy* 126, 451–491.

Boone C (2003) Political Topographies of the African State. Cambridge: Cambridge University Press.

Boulding K (1962) Conflict and Defense. Lanham: University Press of America.

- Campante FR and Do QA (2014) Isolated capital cities, accountability, and corruption: evidence from US States. American Economic Review 104, 2456–2481.
- DHS (2018) Demographic and Health Surveys. Integrated Demographic and Health Series (IDHS), version 2.0, Minnesota Population Center and ICF International. Available at http://idhsdata.org.
- Goldewijk KK, Beusen A and Janssen P (2010) Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1. *The Holocene* 2010, 1–9.

Grossman G and Lewis JI (2014) Administrative unit proliferation. American Political Science Review 108, 196-217.

- Grossman G, Pierskalla JH and Dean EB (2017) Government fragmentation and public goods provision. *Journal of Politics* **79**, 823–839.
- Hendrix CS (2010) Measuring state capacity: theoretical and empirical implications for the study of civil conflict. *Journal of Peace Research* 47, 273–285.
- Henn S (2020) Complements or Substitutes? How Institutional Arrangements Bind Traditional Authorities and the State in Africa. Unpublished working paper. http://soerenhenn.com/files/Henn_Chiefs.pdf.
- Herbst J (2000) States and Power in Africa. Princeton: Princeton University Press.
- Huntington SP (1968) Political Order in Changing Societies. New Haven: Yale University Press.
- Jedwab R and Storeygard A (2020) The Average and Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa 1960–2010. NBER working paper 27670.
- Krishna A and Schober G (2014) The gradient of governance: distance and disengagement in Indian villages. *Journal of Development Studies* 50, 820–838.
- Lee MM and Zhang N (2017) Legibility and the informational foundations of state capacity. *The Journal of Politics* **79**, 118–132.
- Lujala P, Rød JK and Thieme N (2007) Fighting over oil: introducing a new dataset. *Conflict Management and Peace Science* 24, 239–256.

- Mann M (1984) The autonomous power of the state: its origins, mechanisms and results. *European Journal of Sociology* 25, 185–213.
- Migdal JS (1988) Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World. Princeton: Princeton University Press.
- Müller-Crepon C, Hunziker P and Cederman L-E (2021) Roads to rule, roads to rebel: relational state capacity and conflict in Africa. *Journal of Conflict Resolution* 65, 563–590.
- National Geophysical Data Center (2014) DMSP-OLS Nighttime Lights Time Series, Version 4. Electronic resource. Available at http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html.
- Pierskalla J, Schultz A and Wibbels E (2017) Order, distance, and local development over the long-run. *Quarterly Journal of Political Science* 12, 375–404.
- Rogowski JC, Gerring J, Maguire M and Cojocaru L (2021) Public infrastructure and economic development: evidence from postal systems. *American Journal of Political Science*, Online First.
- Scott JC (2017) Against the Grain. A Deep History of the Earliest States. New Haven: Yale University Press.

Timmons JF (2005) The fiscal contract: states, taxes, and public services. World Politics 57, 530-567.

Tollefsen AF and Buhaug H (2015) Insurgency and inaccessibility. International Studies Review 17, 6-25.

Weidmann N and Gleditsch KS (2010) Mapping and measuring country shapes: the cshapes package. *R Journal* 2, 18–24.
Weidmann NB and Schutte S (2017) Using night light emissions for the prediction of local wealth. *Journal of Peace Research* 54, 125–140.

Wig T and Tollefsen AF (2016) Local institutional quality and conflict violence in Africa. *Political Geography* 53, 30–42. World Bank (2018) World Development Indicators. Accessed on 2019/03/20 from https://databank.worldbank.org/data.

Cite this article: Müller-Crepon C (2023). State reach and development in Africa since the 1960s: new data and analysis. *Political Science Research and Methods* **11**, 663–672. https://doi.org/10.1017/psrm.2021.60