

Does Local State Capacity Foster Development in Africa? New Data and Analysis

Carl Müller-Crepon*

ETH Zurich

December 7, 2019

Abstract

The colonial making of African states' geographies has limited their capacity and caused currently low levels of development on the continent. To test this prominent argument, I provide comprehensive panel data on local state capacity and estimate the effect of state capacity on local development. I proxy African states' time-varying capacity through local travel times to national and regional administrative capitals. Travel times are computed on a yearly 5×5 km grid with new data on roads and administrative units and capitals (1966–2016). With these data, I estimate the effect of changes in travel times to capitals on local education and infant mortality rates as well as nightlight emissions. Within the same location, development outcomes generally improve as travel times to its capitals decrease. The data and evidence presented in this letter improve the measurement of local state capacity and contribute to the understanding of its effects on human welfare.

*email: carl.mueller-crepon@icr.gess.ethz.ch. ETH Zurich; International Conflict Research; Haldeneggsteig 4; CH-8092 Zurich. I thank Lukas Dick and Benjamin Füglistler for excellent research assistance. I am grateful to Lars-Erik Cederman, Robert Bates, Simon Hug, Philipp Hunziker, Horacio Larreguy, Clara Neupert-Wentz, Yannick Pengl, Ryan Saylor, and participants at the APSA Annual Meeting 2019, the AFK workshop in Hamburg, and seminars at ETH Zurich for helpful comments and feedback.

Introduction

Adverse population distributions, colonial borders, and deficient transport networks limit state capacity in Africa and are blamed to cause misgovernance and low levels of development on the continent (e.g. [Herbst 2000](#)). However, empirical tests of this reasoning lag behind its frequent use. In particular, we lack geographically disaggregated and time-varying data on African states' capacity that allows for estimating its effect on local development. This letter presents new panel data on local state capacity proxied by travel times to post-independence regional and national capitals. With these data, I estimate generally positive effects of increases in state capacity on local development.

The new measure of state capacity adds to previous efforts of measurement that have, after an initial focus on the country-level ([Hendrix 2010](#)), turned to disaggregated data that captures spatial variation in developing states' capacity ([Boone 2003](#)). However, their reliance on either surveys ([Wig and Tollefsen 2016](#)) or age-heaping in censuses ([Lee and Zhang 2017](#)) limits temporal coverage and availability in low-capacity states.¹ As a remedy, I develop and validate a continent-wide, yearly proxy of local state capacity based on travel times to regional and national administrative capitals between 1966 and 2016. I compute these travel times for a 5×5 km grid of the continent using newly collected panel data on (1) administrative geographies and (2) road networks digitized from the Michelin map corpus.

With the new data, I estimate the effect of local state capacity on development. Previous studies provide cross-sectional evidence that locations farther away from administrative 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, García-Jimeno and Robinson 2015; Henn 2018](#)).² I improve identification and address the potential endogeneity of administrative geographies by holding constant all variation within points and country-years. The analysis shows that reductions in travel times to capitals generally come with increases in education and infant survival rates, as well as nightlight emissions. These effects are not explained by, *inter alia*, spurious migration, differential pre-trends, or improvements in economic market access.

¹Surveys also cover space incompletely and age-heaping may proxy education rates rather than state capacity ([A'Hearn, Baten and Crayen 2009](#)).

²Similarly, US counties with a post office in 1896 developed quicker ([Rogowski et al. 2019](#)).

State capacity and development

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

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 and trigger the developmental effects of centralized institutions promoted by the state (Campante and Do 2014; Huntington 1968). Second, states provide multiple public services such as education and health care. Such service provision depends on states' capability to monitor demand and control agents, therefore increasing in the level of local state capacity (e.g. Krishna and Schober 2014; Henn 2018). In addition, citizens' access to specialized services (i.e. hospitals, courts) increases with access to administrative capitals where these are typically located. Third, smooth local accessibility increases the possibilities for citizens to tap into private rents from the state such as public sector jobs or subsidies (Ades and Glaeser 1995).³

These three development-inducing effects are counteracted by a fourth mechanism, direct and indirect taxation. Local state capacity facilitates tax collection, which has, in and off itself, a negative effect on citizens' welfare. With that, the net effect of state capacity depends on whether the exchange of taxes for public services is generally beneficial for citizens (e.g. Timmons 2005) or not (e.g. Scott 2017). The empirical results below shed light on this question.

Approximating state reach

Because physical accessibility is a precondition for local state capacity, I proxy local state reach via the weighted sum of travel times from a location to its various (national, re-

³Part of this mechanism runs through information on these rents available to citizens (see e.g. Banerjee et al. 2018).

gional, etc.) administrative capitals:

$$\text{total state reach}_{p,t} = \sum_{u=1}^U \omega_u \text{state reach}_{p,u,t} \quad (1)$$

$$= \sum_{u=1}^U -\omega_u \ln(1 + d_t(p, C_{p,u,t})) \quad (2)$$

where the reach of the state on a level of administrative hierarchy u towards point p at time t is calculated as the travel time (in hours) on the shortest path between p and the capital $C_{u,t}$ on the road network at time t . The log-transform⁴ captures the convex relation between increasing travel times to capitals and decreases in state capacity (Figure 2). Weights ω_u denote the impact of times to capitals on each level u . I estimate ω_u in an outcome-specific manner below. To measure $\text{state reach}_{p,u,t}$, I collect time-varying data on administrative geographies at the national and regional level and road networks.

First, I create the first comprehensive geographic panel data on the boundaries and capitals of first-level (regional) administrative units since African countries' independence. Drawing on diverse sources, the dataset covers 1763 unique region-periods that span the time from independence to 2016 (see Appendix A). Data on national borders and capitals comes from Cshapes ([Weidmann and Gleditsch 2010](#)).

Second, I transform the Michelin road map corpus into a digital road atlas akin to a time-varying Google Maps. Scanned maps are available at a scale of 1:4 million⁵ for 23 years between 1966 and 2014. Pixels that depict roads are classified with a fully convolutional neural network (author and co-authors, 2019), vectorized, and transformed into a planar, georeferenced graph that covers the African continent at a resolution of .0417 decimal degrees ($\approx 5\text{km}$; Appendix B). Each edge on the network is associated with a travel time derived from the quality of roads observed on the original maps. With these data on roads and administrative units, I calculate $\text{state reach}_{p,u,t}$ for the grid cells of a raster with a .0417 decimal degree resolution for every year between 1966 and 2016 (Figure 1).

I validate the utility of the data as a proxy for local state capacity with information on the presence of the state in enumeration areas (EA) of the Afrobarometer ([2018](#)) surveys. Figure 2 shows that travel times to capitals correlate strongly with an index of binary indicators for whether an EA features the presence of the police and military, a school and hospital, as well as a number of government-provided public services such as water,

⁴The constant of 1 hour prevents taking the log of 0 in capitals.

⁵This corresponds to 1mm per 4km and puts an upper limit to the precision of the resulting data.

4

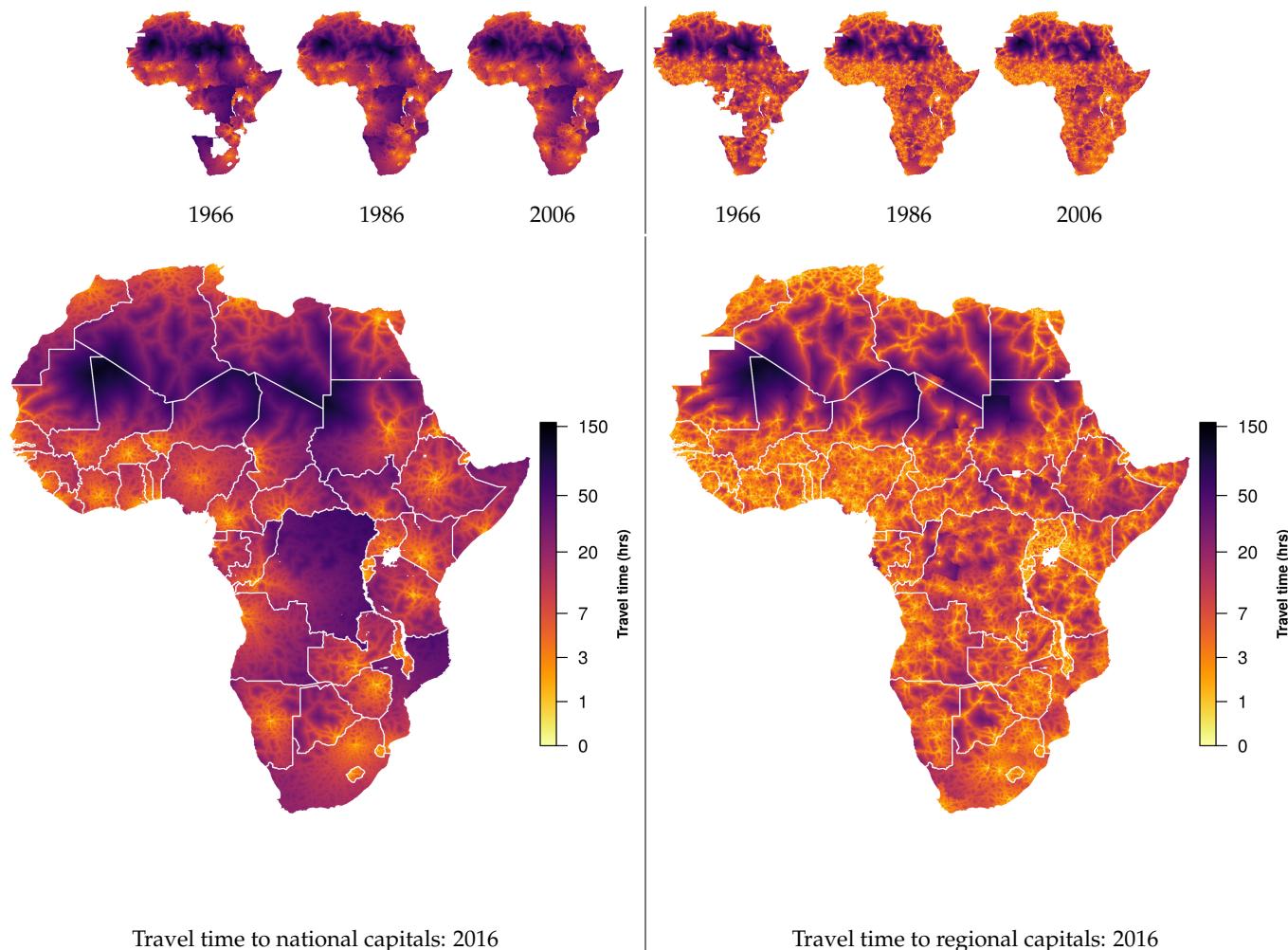


Figure 1: State reach in Africa, 1966–2016.

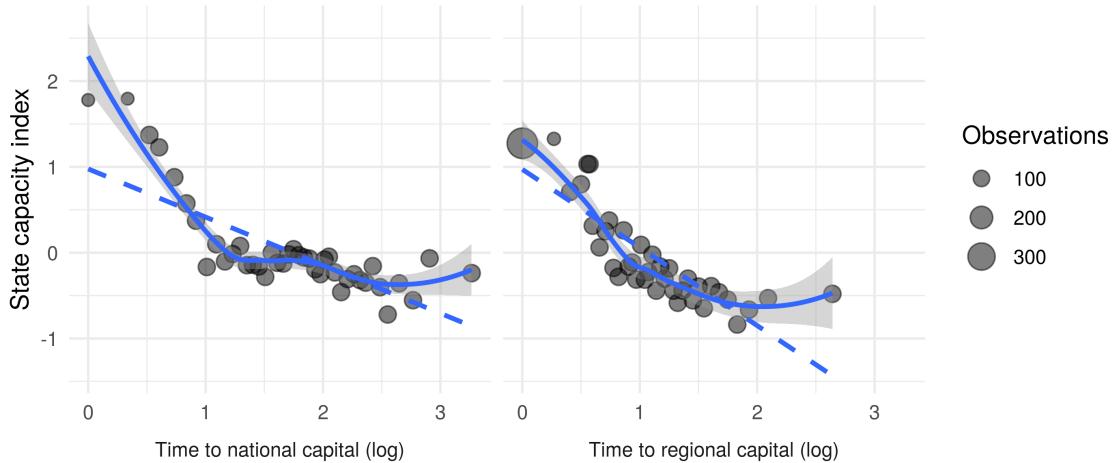


Figure 2: Travel times to national and regional capitals and state capacity index
Values are demeaned by country \times survey round and averaged within 40 x-axis quantiles.

sewage, and electricity. Travel times also fit these geographic trends better than mere geodesic, as-the-crow-flies distances (Appendix C.1).

State reach in Africa: weak, growing, and unequal

Since their independence, African states' reach has increased. Plotted in Figure 3, travel times between national capitals and citizens have decreased from an average of 11.7 hours in 1966 to 9.3 hours in 2016, a change of 20.8 percent. Times to regional capitals decreased by 32.2 percent from 5.1 hours in 1966 to 3.5 hours in 2016.

Three trends have caused this decrease (Appendix C.2). First, administrative geographies have changed. Côte d'Ivoire, Nigeria, and Tanzania have relocated their national capitals⁶ and most governments have created additional administrative regions (Grossman and Lewis 2014; Grossman, Pierskalla and Dean 2017).⁷ Second, transport infrastructure has improved, with quality-weighted mileage increasing by about 50 percent since independence. Lastly, the doubling of urbanization rates from 20 to 40 percent (World Bank 2018) has further decreased travel times to capitals between 1960 and 2016.

The maps in Figure 1 show that travel times to capitals exhibit substantive variation within and across countries. Travel times to the capital of a state like the DR Congo with a "difficult geography" (Herbst 2000) and poor infrastructure exhibit a wide range from 0

⁶Abidjan to Yamoussoukro (1983), Lagos to Abuja (1991), and Dar es Salaam to Dodoma (1974, but all Tanzanian ministries are still located in Dar es Salaam).

⁷Similarly, the independence of Eritrea and South Sudan has brought capitals closer to their citizens.

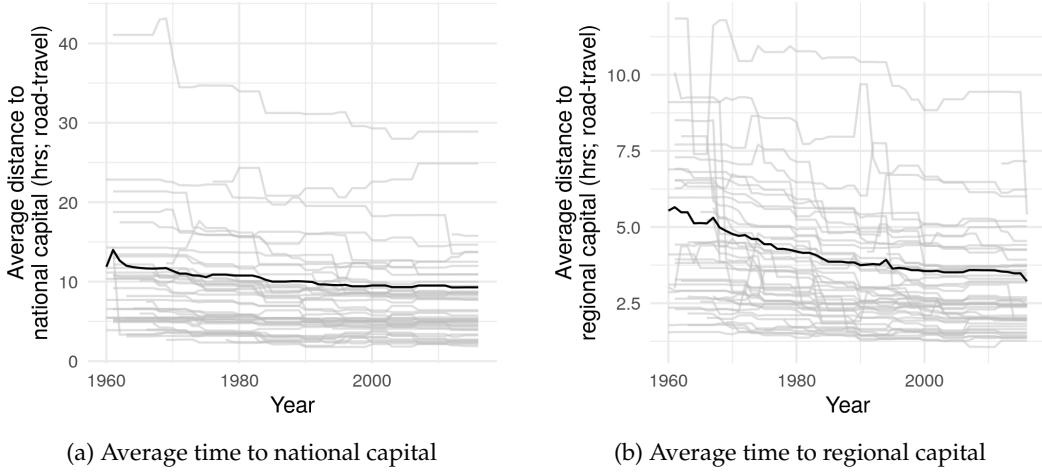


Figure 3: Decreasing travel times to regional and national capitals in Africa 1960–2015.

Population-weighted averages based on HYDE population estimates ([Goldewijk, Beusen and Janssen 2010](#)) and travel times to capitals.

to 50 hours, with a high median of 34.3 hours. In Mali, traveling from Bamako to the median citizen takes 7.2 hours and up to 28 to the far North. Capitals of small countries such as Rwanda naturally are closest to their median citizen. Similar variation marks changes in travel times since 1966. Here, states that seceded (e.g. South Sudan) and relocated their capital (e.g. Nigeria) experienced the largest improvements. These changes exhibit, again, substantive spatial variation because changes in administrative geographies, road networks, and population distributions do not equally affect all locations in a country. My empirical strategy exploits this fact.

Data on local development

States' reach in Africa shows substantial spatio-temporal variation over the past 50 years. I examine whether this variation affected local development with data on local education, infant mortality, and nightlight emissions.⁸

To measure the first two outcomes, primary education and infant mortality rates, I rely on the Demographic and Health Survey ([DHS 2018](#)). The DHS encodes the educational achievement of all members of sampled households. Assuming that household members live where they were raised, I model the level of education of 1.9 million individuals older than 15 as having been influenced by the travel time towards their capitals at age 6. A primary education dummy serves as the main outcome.

⁸See Appendix D.2 for summary statistics.

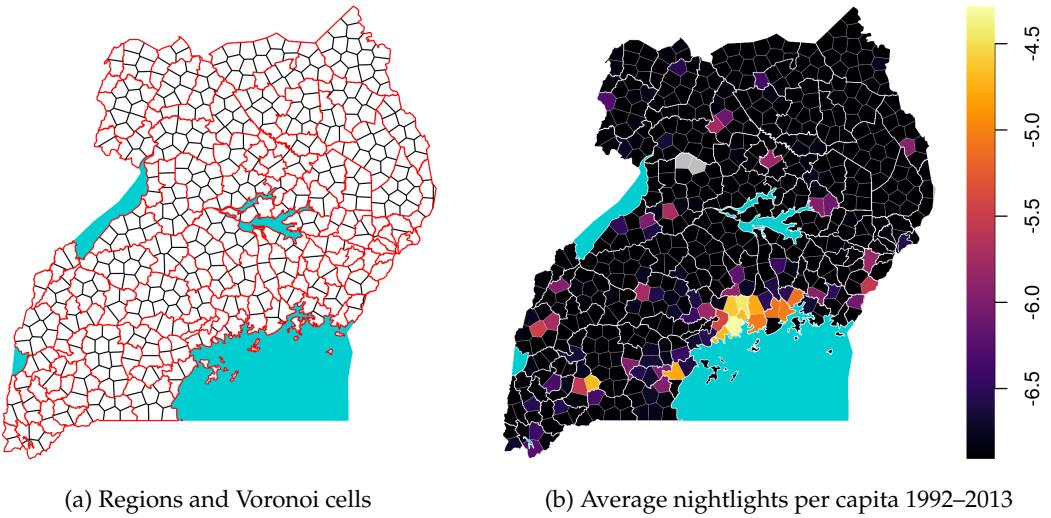


Figure 4: Voronoi cells, regions, and nightlights in Uganda 1992–2013.

Panel (a) plots all regional borders between 1992 and 2013 (red) and Voronoi cells (black). Unpopulated Voronoi cells in Panel (b) in grey.

The DHS also provides information about the under-1 mortality of infants born to women aged 15–49. Using this information on 2.6 million infants, 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 household members sampled in the same place are born and schooled at different points in time. 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. I investigate whether effects differ between migrants and non-migrants below.

To mitigate the uneven coverage of the DHS surveys and counter concerns about migration biases, I analyze local nightlight emissions per capita as a third measure of development ([Weidmann and Schutte 2017](#)) derived from satellite images (1992–2013; [National Geophysical Data Center 2014](#)). I aggregate these data up to the level of centroidal Voronoi cells with a size of 400 km^2 that are nested within all regions observed between 1992 and 2013 (Figure 4 and Appendix D.1). In contrast to quadratic cells that frequently overlap with (changing) administrative borders, I can consistently aggregate the travel time data to these Voronoi cells. Nightlights are log-transformed after adding a constant of .001.⁹

⁹I drop cells with oil wells ([Lujala, Rød and Thieme 2007](#)) that spoil the measure with bright flares of burning gas.

Empirical strategy

The empirical analysis aims to identify the effect of travel times to capitals on development despite endogenous processes such as strategic road, capital, and border placements. To account for any cross-sectional endogeneity, I only exploit *temporal* variation in travel times and development within the same location:

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

where β_1 and β_2 capture the effects of the travel times to point p 's regional and national capitals at time t . In parallel to EQ 1, the sum of β_1 and β_2 proxies the total effect of state capacity. DHS respondents i are spatially matched to grid-cell p in the travel time rasters of year t .¹⁰ p is synonymous with i where Voronoi cells are the units of analysis. The model controls for all constant attributes of points/units and country-years through fixed effects α_p and $\lambda_{c,t}$. Individual-level controls X_i in the education models consist in respondents' sex, age and its square. Infant mortality models include the mother's age at birth and its square, the infant's 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 p and country-year levels.

With that, I assume that changes in travel times to capitals are exogenous to local development outcomes observed thereafter.¹¹ Robustness checks account for potential violations of this assumption, such as reverse causation or spurious correlations of changes in times to capitals with changes in economic market access.

Results

The results show that, first, reductions in travel times towards regional and national capitals are robustly associated with higher 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 joint proxy for state

¹⁰Noise is added by the DHS's random displacement of clusters by up to 10 (2) kilometers in rural (urban) areas.

¹¹The setting does not allow for a binary difference-in-difference design because the treatments – travel times – are continuous variables with multiple, positive and negative changes.

capacity reveals a consistently positive, meaningful, and statistically significant effect on local development.

Table 1: Changes in time to national/regional capital and local development

	Primary educ. (0/100) (1)	Infant mort. (0/100) (2)	Light/capita (log) (3)
Time to nat. capital (log)	-2.662*** (0.421)	0.932*** (0.240)	-0.052* (0.027)
Time to reg. capital (log)	-1.325*** (0.323)	0.138 (0.160)	-0.037** (0.019)
$\beta_1 + \beta_2$:	-3.987*** (0.486)	1.069*** (0.273)	-0.089*** (0.029)
Point FE:	yes	yes	yes
Country-year FE:	yes	yes	yes
Survey FE:	yes	yes	yes
Controls:	yes	yes	-
Mean DV:	70	9.9	-6.5
Observations	1,893,067	2,634,209	1,506,991
Adjusted R ²	0.445	0.051	0.836

Notes: OLS linear models. Control variables for models with primary education as the dependent variable consist of respondents' 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

Table 1 presents the baseline results. Interpreting coefficients in substantive terms, Model 1 indicates that a decrease in 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 (.54) percentage points. Model 2 shows that infant mortality rates decrease by .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 impact of changes in the time to regional capitals on infant mortality rates. The respective coefficient is close to zero and statistically insignificant. This coincides with findings by [Grossman, Pierskalla and Dean \(2017\)](#) who report inconclusive effects of district splintering in Malawi, Nigeria, and Uganda.

Distinguishing the effects of relocations of national capitals from those of changes in road networks reveals that the effects of travel times to national capitals on education and mortality rates are primarily driven by relocations of capitals (Appendix E.1). Changes due to road networks have a smaller, but statistically significant effect on primary education rates and no effect on infants' mortality.

Lastly, Model 3 yields a significant association of reductions in travel times to regional capitals with higher nightlight emissions. As the average distance to the regional capital

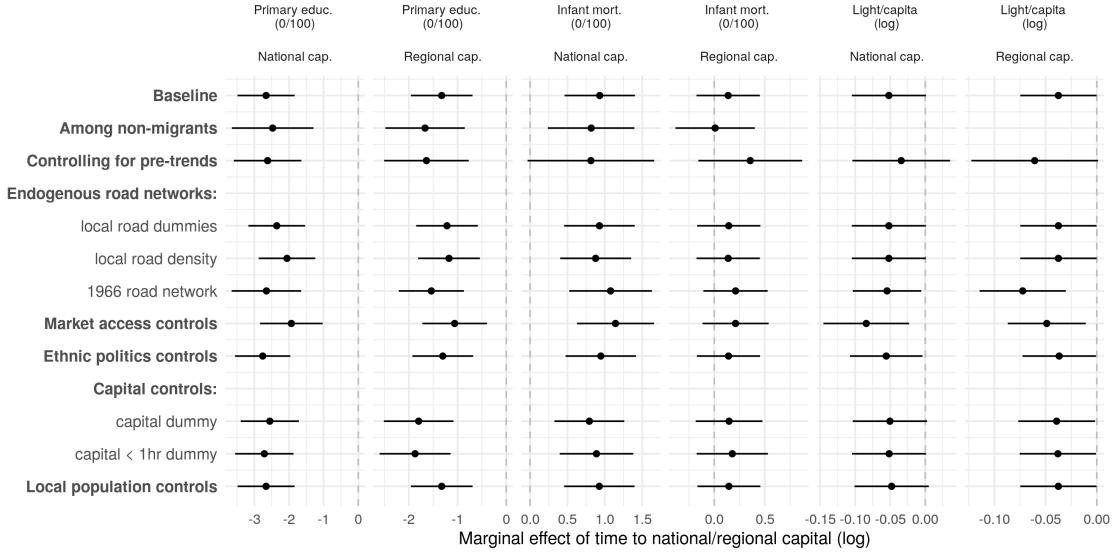


Figure 5: Robustness checks. Coefficient estimates from varying specifications with 95% CIs.

of a Voronoi cell decreases from two to one 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.2). The estimate's 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, more gradual changes in road networks that only have second order impacts on local development (Appendix E.4) drive the estimate.

Figure 5 presents the results of robustness checks that are discussed in Appendix E and briefly summarized here. First, I show that the effects of travel times to capitals hold among non-migrant DHS respondents. Second, there is no significant effect of differential pre-trends, except for those that affect the nightlight model. Third, the results are robust to controlling for potentially endogenous changes in local road networks accounted for with road-type dummies, density-measures, or by using the first available, time-invariant road data from 1966. Fourth, results are also not due to a spurious correlation of travel times to capitals with access to the economic markets in the 1530 biggest African towns and cities. Fifth, they are also robust to accounting for ethno-regional political inclusion that might come with favoritism, observations in regional or national capitals, and local population growth. Lastly, Appendix E.5 demonstrates robustness across alternative outcome specifications.

In general, these results show positive developmental effects of increases in state capacity as proxied by travel times to administrative 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. They also do so when an area becomes better connected to its national capital, but this effect is in part driven by pre-trends, which raises concerns about reverse causality.

Conclusion

One important determinant of states' capacity and their ability to foster development is their physical access to the population. [Herbst \(2000\)](#) argues that 'difficult geographies' with populated regions far away from national capitals curb African states' capacity and development. Measuring African states' varying success to shrink the distance to their citizenry, this letter has introduced new data on post-colonial state capacity proxied by travel times between administrative capitals and citizens. Employing the new data, I have documented generally positive effects of increases in local state capacity on local measures of wealth and well-being of citizens. This supports the long-standing argument that difficult geographies with disperse and hard-to-access populations are an impediment to African development. However, the results at the same time reject geographic determinism in showing that states can improve their reach and thereby foster development by closing the spatial gap between their administrations and their citizens.

With new data and empirical results, this letter opens up avenues for future research. A first series of questions concerns the origins of the uncovered variation in the effects of state capacity, which might relate to the nature of state-provided goods or heterogeneous effects under different institutional settings. Second, there might be important heterogeneity in *when* local state capacity matters most for citizens' welfare. Citizens may suffer most from weak statehood in times of crisis inflicted by droughts, floods, or violent conflicts. Lastly, the new data on local state capacity can be used to study state-society interactions more generally, touching on topics that range from resource extraction, over accountability and corruption, to conflict processes and peacekeeping.

References

- Acemoglu, Daron, Camilo García-Jimeno and James A. Robinson. 2015. "State Capacity and Economic Development: A Network Approach." *American Economic Review* 105(8):2364–2409.
- Ades, Alberto F. and Edward L. Glaeser. 1995. "Trade and Circuses: Explaining Urban Giants." *Quarterly Journal of Economics* 110(1):195–227.
- Afrobarometer. 2018. "Afrobarometer Data." Available at <http://www.afrobarometer.org>.
- A'Hearn, Brian, Jörg Baten and Dorothee Crayen. 2009. "Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital." *The Journal of Economic History* 69(03):783.
- Banerjee, Abhijit, Rema Hanna, Jordan Kyle, Benjamin A. Olken and Sudarno Sumarto. 2018. "Tangible Information and Citizen Empowerment: Identification Cards and Food Subsidy Programs in Indonesia." *Journal of Political Economy* 126(2):451–491.
- Boone, Catherine. 2003. *Political Topographies of the African State*. Cambridge: Cambridge University Press.
- Boulding, Kenneth. 1962. *Conflict and Defense*. Lanham: University Press of America.
- Campante, Filipe R. and Quoc Anh Do. 2014. "Isolated capital cities, accountability, and corruption: Evidence from US States." *American Economic Review* 104(8):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, Kees Klein, Arthur Beusen and Peter Janssen. 2010. "Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1." *The Holocene* 2010(1):1–9.
- Grossman, Guy, Jan H. Pierskalla and Emma Boswell Dean. 2017. "Government Fragmentation and Public Goods Provision." *Journal Of Politics* 79(3):823–839.
- Grossman, Guy and Janet I. Lewis. 2014. "Administrative Unit Proliferation." *American Political Science Review* 108(01):196–217.
- Hendrix, Cullen S. 2010. "Measuring state capacity: Theoretical and empirical implications for the study of civil conflict." *Journal of Peace Research* 47(3):273–285.
- Henn, Sören. 2018. "Complements or Substitutes: State Presence and the Power of Traditional Leaders." *Unpublished Working Paper*. http://soerenhenn.com/files/Henn_Chefs.pdf
- Herbst, Jeffrey. 2000. *States and Power in Africa*. Princeton: Princeton University Press.

- Huntington, Samuel P. 1968. *Political Order in Changing Societies*. New Haven: Yale University Press.
- Krishna, Anirudh and Gregory Schober. 2014. "The Gradient of Governance: Distance and Disengagement in Indian Villages." *Journal of Development Studies* 50(6):820–838.
- Lee, Melissa M. and Nan Zhang. 2017. "Legibility and the Informational Foundations of State Capacity." *The Journal of Politics* 79(1):118–132.
- Lujala, Päivi, Jan Ketil Rød and Nadja Thieme. 2007. "Fighting over Oil: Introducing a New Dataset." *Conflict Management and Peace Science* 24(3):239–256.
- Mann, Michael. 1984. "The autonomous power of the state: its origins, mechanisms and results." *European Journal of Sociology* 25(2):185–213.
- Migdal, Joel S. 1988. *Strong societies and weak states: state-society relations and state capabilities in the Third World*. Princeton: Princeton University Press.
- 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.*
- Rogowski, Jon C, John Gerring, Matthew Maguire and Lee Cojocaru. 2019. "Public Infrastructure and Economic Development: Evidence from Postal Systems." *American Journal of Political Science, forthcoming*.
- Scott, James C. 2017. *Against the Grain. A Deep History of the Earliest States*. New Haven: Yale University Press.
- Timmons, Jeffrey F. 2005. "The Fiscal Contract: States, Taxes, and Public Services." *World Politics* 57(4):530–567.
- Tollefson, Andreas Forø and Halvard Buhaug. 2015. "Insurgency and inaccessibility." *International Studies Review* 17(1):6–25.
- Weidmann, Nils B and Sebastian Schutte. 2017. "Using night light emissions for the prediction of local wealth." *Journal of Peace Research* 54(2):125–140.
- Weidmann, Nils and Kristian Skrede Gleditsch. 2010. "Mapping and Measuring Country Shapes: The cshapes Package." *R Journal* 2(1):18–24.
- Wig, Tore and Andreas Forø Tollefson. 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>.

Supporting Information

State Reach and Development in Africa since the 1960s: New data and analysis

Table of Contents

A	Regions in post-colonial Africa	A2
B	Road data from the Michelin map corpus	A4
B.1	The Michelin map corpus	A4
B.2	Map digitization as a computer vision task.	A5
B.3	FCNN: Architecture and training	A7
B.4	Vectorization and results	A8
B.5	Retrieving travel speeds	A10
B.6	Network construction	A10
B.7	Network Validation.	A11
C	Travel times to capitals as proxy for state reach: Data description and validation	A12
C.1	Validation	A12
C.2	Description	A13
D	Data and summary statistics	A19
D.1	Units of the nightlight analysis: Voronoi cells.	A19
D.2	Summary statistics	A22
E	Robustness checks	A25
E.1	Re-locations of national capitals vs. road-network development:	A25
E.2	Testing for pre-trends:	A26
E.3	Migration	A29
E.4	Omitted variables:	A30
E.5	Varying education- and health-related outcomes:	A35
E.6	Additional robustness checks:	A36
F	References	A39

A Regions in post-colonial Africa

To collect a full panel on the borders and capitals of first-level administrative regions in Africa since countries' independence, I draw on the qualitative accounts of unit changes from the [statoids.com](#) database and encode each administrative unit change geographically. I therefore rely primarily on maps from the GAUL database, ([FAO 2014](#)). Although this data is contemporary (reaching back until 1990), it allows me to trace back all unit-splits – this type of unit-change constitutes the vast majority of cases – by simply merging the units observed after the split which results in the original unit. Where units have been merged or the administrative map of a country has been redrawn completely, I make use of more than 100 digitized maps, mostly from the CIA Base Map series as well as other GIS data, such as the GADM database. Each region-period is associated with its capital, as listed in most cases by [statoids.com](#). Missing capitals are searched on the maps and in secondary sources. The capitals are then geocoded using the [geonames.org](#) gazetteer. Changes in the location of capitals within the same boundaries of a region naturally result in new region-periods. Each region-period is associated with a start and end-year. To ensure consistent and temporally non-overlapping coding, region-periods that start after January 1st are coded as starting in the next year. The final data set covers 1763 unique region-periods, covering each African country from independence to 2016. The evolution of the number of regions in the data set is traced in Figure A1. Figure A2 plots the data for the year 2016.

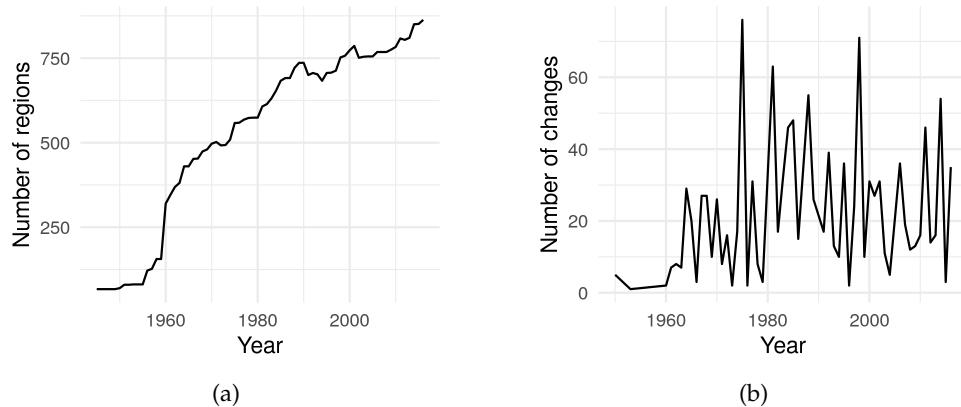
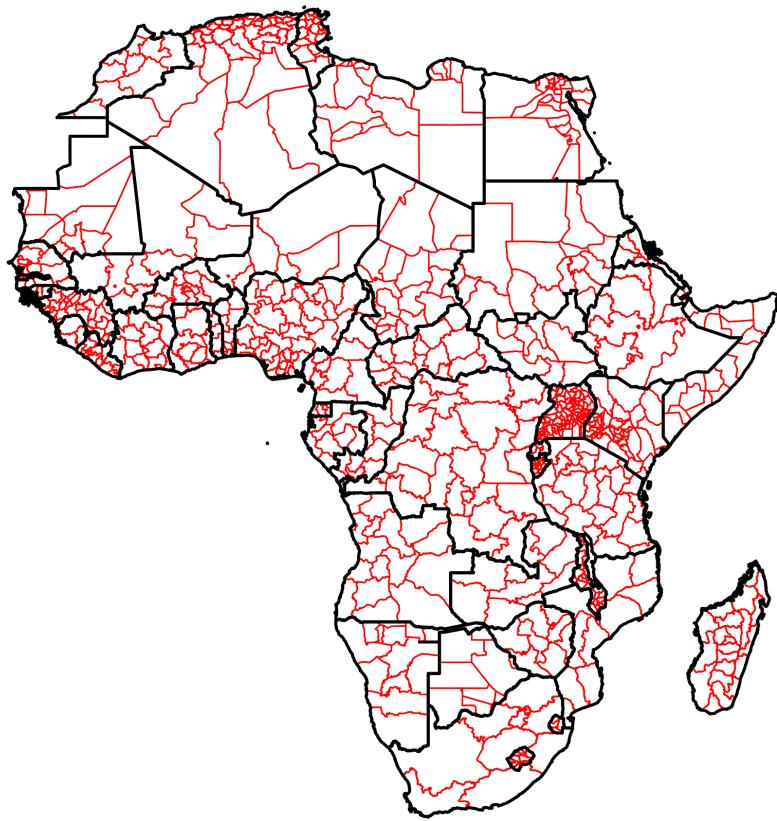
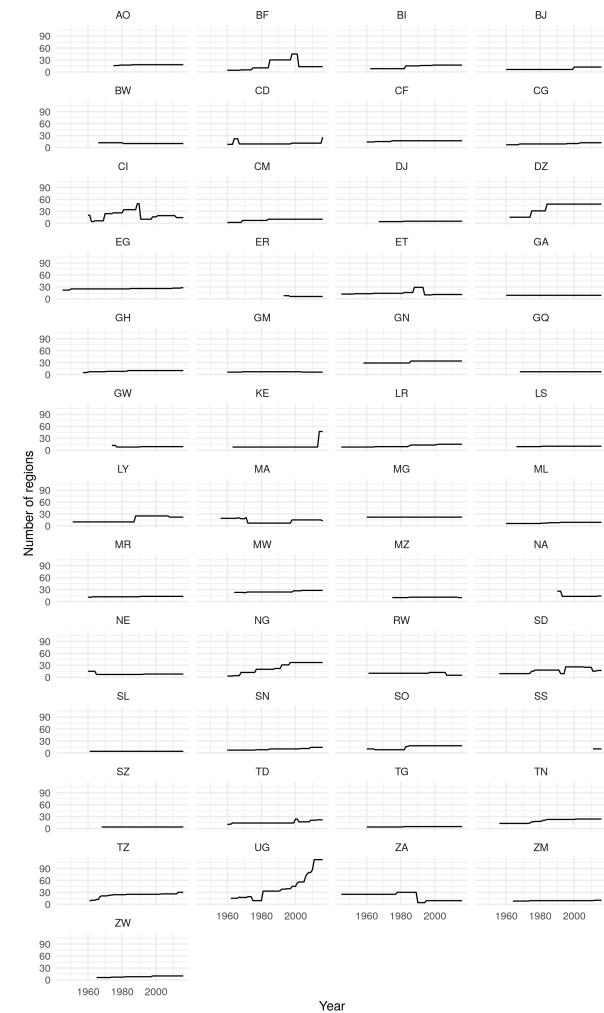


Figure A1: Description of newly collected data on first-level administrative units in Africa.

A3



(a) First-level administrative units in Africa 2016



(b) Number of admin-1 units by country and year

Figure A2: Overview over first-level administrative unit data.

B Road data from the Michelin map corpus

This section provides the details on how I digitize the Michelin map corpus. Note that this procedure was developed in previous, yet unpublished work with co-authors.¹² Data collection has been extended for the purpose of this article from the initial 6 to 23 road map cross-sections. Subsection B.1 describes the map corpus, Subsection B.2 introduces the extraction of information from maps as a computer vision task, and Subsection B.3 discusses the architecture of the Fully Convolutional Neural network. The two last subsections present the results of the digitization procedure as well as the coding of travel speeds.

B.1 The Michelin map corpus

The source for road network data for post-colonial Africa is the African Michelin map corpus, a collection of large topographical maps at a resolution of 1:4,000,000. Each map shows detailed information on road infrastructure with a consistent cartographic symbology for about a third of the continent (see Figure A3). While coverage before the 1960s is sporadic, Michelin has covered the entire African continent at intervals of approximately 5 years beginning in 1964 (see Figure A4). This makes the Michelin corpus an unparalleled source for time-variant road-network information. I digitize 34 map sheets published between 1964 and 2017, which combine into 23 maps of the entire continent.

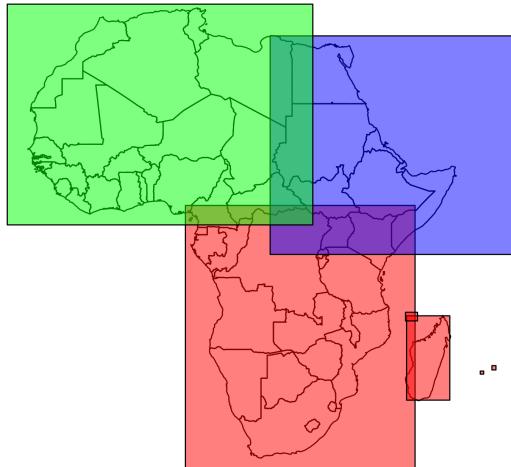


Figure A3: Spatial coverage of Michelin Map types.

Note that Madagascar is not covered in 1966. Green: North-West. Blue: North-East. Red: Center-South.

¹²Full reference to this work will be given in the final version of this manuscript. Because this section is adapted from joint work with co-authors, I use the pronoun ‘we’ when describing this work.

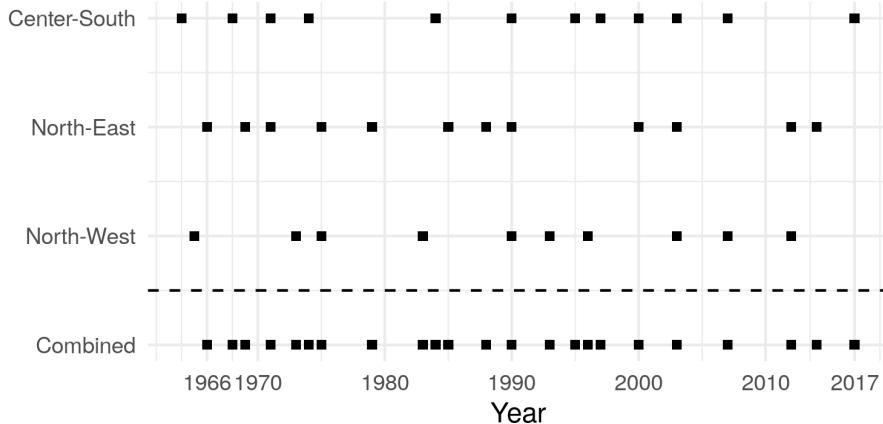


Figure A4: Temporal coverage of Michelin Maps

I digitize the maps collection automatically. Apart from being relatively cheap, the automatic digitization approach features a number of additional benefits:

- *Consistency:* The cartographic information is extracted in a highly consistent manner, avoiding errors due to human fatigue and less-than-perfect inter-coder reliability.
- *Replicability:* The entire data set can be reproduced at will.
- *Extendability:* After the initial system is set up, the marginal costs of adding new cartographic material (including from other sources) are low.

B.2 Map digitization as a computer vision task

A critical first step for extracting information from geospatial imagery is to distinguish between areas representing objects of interest and background. Roads in the Michelin maps are drawn as complex features with multiple color and line-patterns, and often interrupted by other objects (see Figure A5). Therefore, heuristic algorithms that distinguish only colors or lines fail to classify roads correctly. Instead, the maps are digitized using a method that “looks at” entire map segments at once, and is able to distinguish between lines and other object types using contextual visual information.

To do so, the method borrows from recent advancements in the machine learning literature and implement a *Convolutional Neural Network*-based system for road network extraction. Convolutional Neural Networks (CNNs) have recently emerged as a powerful method for computer vision applications, outperforming other approaches across a variety of classification problems ([LeCun, Bengio and Hinton 2015](#)). Fundamentally, CNNs are feedforward artificial neural networks (ANNs). They consist of multiple layers of neurons, each neuron representing a non-linear function associated with a trainable weights vector, accepting a linear combination of inputs from the previous layer, and out-

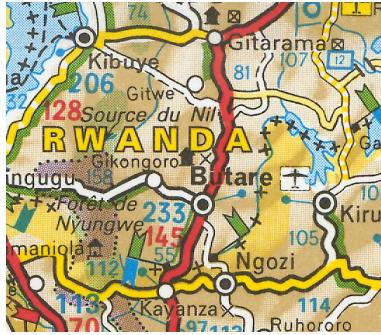


Figure A5: Small excerpt from the 2003 Michelin map for Southern Africa, showing southern Rwanda.

putting a scalar that is passed on to the next layer.¹³ In the language of ANNs, the vector of predictors associated with a single observation is then called the “input layer,” whereas the prediction produced by the ANN is called the “output layer.” Note that in computer vision problems, the input layer typically consists of raw image data, structured as a pixel-image with multiple color bands.

While the most basic variants of feedforward ANNs feature fully connected architectures where each neuron accepts inputs from *all* neurons of the previous layer, *convolutional* neural networks restrict the visual receptive field of each neuron to a small, spatially contiguous patch of input data, thus retaining the spatial structure of the inputs. Moreover, CNNs feature a shared-weights architecture, whereas neurons reuse the same set of parameters to “look” at all locations of the input image. This ensures that CNNs are shift-invariant: They are able to detect objects regardless of their spatial location in the input image.

Neurons in CNNs typically implement two types of operations. A *convolution* operation, computing the dot product between a patch of input data and the neuron’s weights, and a pooling operation, which downsamples the input image to a lower resolution by some given factor. Productive CNNs typically feature multiple convolution- and pooling-layers in succession, giving rise to a complex non-linear function that transforms a given input image into a series of images with decreasing resolution, but higher depth, called *feature maps*.¹⁴ This architecture gives rise to the key advantage of CNNs: their ability to learn features relevant for classification from raw, unprocessed input imagery ([Zeiler and Fergus 2014](#)). Hence, instead of the researcher having to pre-process the input data and extract variables that are useful for classification (e.g. whether particular shapes or color patterns are present), CNNs are capable of learning important features by themselves. The layers close to the input image recognize low-level features such as edges or blobs of a particular color, which are then fed to the higher-level layers that capture more complex features at lower resolutions, like specific line patterns or shapes with particular

¹³The following discussion of ANNs and CNNs draws from [Goodfellow, Bengio and Courville \(2016, ch. 6 & 9\)](#).

¹⁴Here, “depth” refers to the third dimension of an image. An RGB image has depth 3.

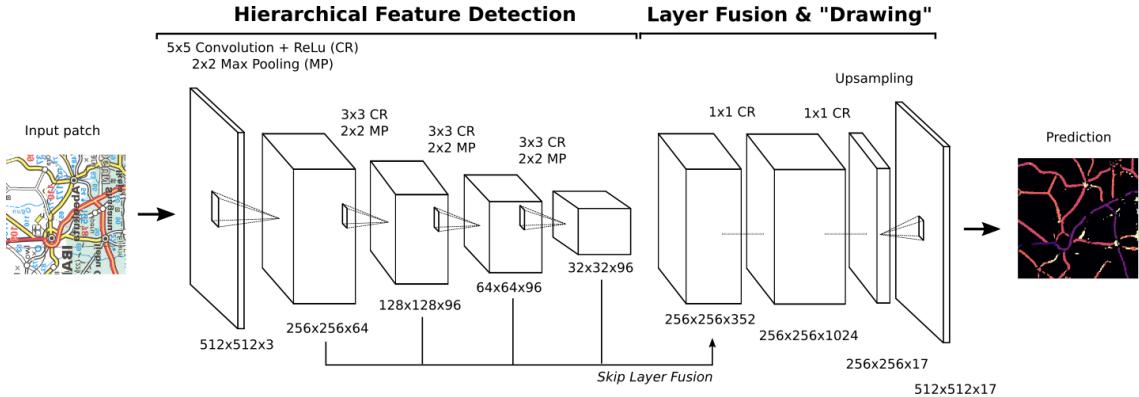


Figure A6: Architecture of our custom fully convolutional neural network.

textures.

For image segmentation, [Shelhamer, Long and Darrell \(2017\)](#) have recently proposed what they call a *fully convolutional* approach. Here, the feature maps produced by the regular convolutional and pooling layers are used as inputs for a set of upsampling, or *deconvolutional* layers. These consist of neurons that implement a reverse convolutional operation, mapping lower resolution feature maps onto higher resolution outputs via a trained interpolation function. Hence, fully convolutional neural networks (FCNNs) have an “abstraction stage,” where convolutional and pooling layers learn to recognize complex image features, and a “drawing stage.” Here, the information from the lower-resolution feature maps is mapped back onto the scale of the original input image, yielding a full semantic segmentation.

B.3 FCNN: Architecture and training

To solve the semantic segmentation problem on the Michelin map material we implement a version of [Shelhamer et al.’s \(2017\)](#) FCNN model that takes RGB image patches of dimension $512 \times 512 \times 3$ pixels as input, and maps them onto output segments of size $512 \times 512 \times 17$. The output image depth arises from the fact that Michelin identifies 16 road categories.¹⁵ The precise architecture of our model is shown in Figure A6. The model is described in canonical notation, see [Goodfellow, Bengio and Courville \(2016\)](#) and [Shelhamer, Long and Darrell \(2017\)](#) for more information.

We pursue a transfer-learning approach and pre-train the FCNN on 2000 artificial map images.¹⁶ These are color-images of dimension $512 \times 512 \times 3$ that superficially look like real road maps, but which we create programmatically by drawing arbitrary planar networks together with other map-like shapes and text labels of arbitrary color, size, orientation, etc. Each simulated map image is paired with a “ground truth” label of dimension $512 \times 512 \times 2$ that highlights the location of the road-network to be detected.

¹⁵An additional reference category identifies background pixels.

¹⁶For pre-training, the outcome layer is only of depth two (instead of 17) because construct the artificial training labels such that they only identify the *presence* of roads, but not their type.

With the pre-trained model, we then proceed to the training of the main model using actual, hand-annotated training data from the Michelin maps.¹⁷

Interpreting trained artificial neural networks is notoriously difficult, as the learned parameters have little intuitive meaning by themselves. However, one commonly employed strategy is to show the neural activations of the network’s feature maps for some input image. Six such feature maps shown in Figure A7 demonstrate how different neurons capture different types of information. The feature maps in the top row appear to recognize numbers, those in the bottom row identify road-related features. They also show that feature maps at lower resolutions tend to capture more abstract, higher-level objects, reflecting the hierarchical logic of feature detection in CNNs.

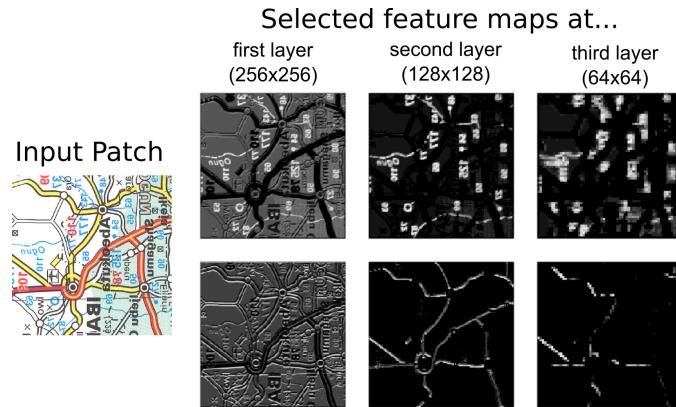


Figure A7: Selected feature maps from the trained FCNN.

Finally, it is instructive to demonstrate the trained FCNNs predictive performance visually. Figure A8 shows an excerpt from a 1966 map segment for Southern Nigeria (left panel), together with the corresponding road predictions obtained from the trained FCNN model (middle panel). The different colors in the predicted image correspond to different road types. We highlight that the FCNN is able to distinguish between even subtle differences in line types, e.g., lines of the same color, but with different border thickness.

B.4 Vectorization and results

Given the FCNN predictions, we implement and apply a four-step algorithm to convert the pixel-based FCNN output into vectorized road-network data:

1. The Zhang-Suen *topological thinning* algorithm is applied to the input images, leading to single-pixel-width road representations.

¹⁷All layers up to the second-to-last one (exclusive) are initialized with pre-trained weights, whereas the weights of last two layers are initialized randomly. For training both the initial “artificial” model as well as the final model, we use the stochastic gradient descent (SGD) based optimizer introduced by Kingma and Ba (2015) with a batch-size of 2.

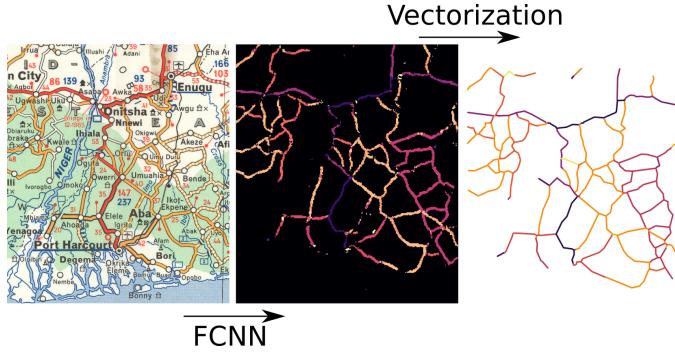


Figure A8: Predictive performance of the digitization procedure.

	Binary	Categorical
Precision	0.988	0.888
Recall	0.986	0.964

Table A1: Evaluation statistics for the full digitization pipeline based on a hold-out sample.

2. The thinned images are fed to a *line-tracing* algorithm, transforming the road network information to a vector-based representation.
3. A *line-splicing* algorithm is then applied to fill small, unlikely gaps in the vectorized road network.
4. A sequential, *hidden-Markov* style model is used to smooth the road type classification, leading to the removal of short segments with misclassified road types.

The right-most panel of Figure A8 illustrates the result of this vectorization procedure. To assess the accuracy of the digitization pipeline, we generate vectorized predictions for two hand-coded hold-out maps, each covering about 1000 square kilometers. We split the ground-truth and predicted networks into 5 km long road segments and calculate two evaluation metrics.¹⁸ *Precision* measures the proportion of predicted road segments that are proximate to a ground-truth road segment. *Recall* measures the proportion of ground-truth road segments that are proximate to a predicted road segment. We use 5 km error bands ($\approx 1\text{mm}$ on the original maps) to establish whether two road segments are proximate. We also calculate variants of these metrics that take road types into account. Here, predicted and a ground-truth segments are only coded as proximate if they are also of the same road type.¹⁹

The result of this evaluation exercise is summarized in Table A1. We find that our digitization procedure is highly accurate. Over 98.8% of all extracted roads are present in the Michelin maps, and 98.6% of all Michelin roads are extracted. The corresponding

¹⁸All metrics are based on length-weighted averages.

¹⁹Note that for this evaluation, we employ the 6-category road type coding used in the paper, not the 16-category coding used during digitization.

figures are somewhat lower if we take road categories into account, but still 88.8 and 96.4, respectively. We note, however, that in those cases where the model misclassifies the road type, the error is typically small. Across all cases where roads are correctly extracted but assigned the wrong category, the mean absolute error on the ordinal road-type scale is 1.38. In other words, misclassifications typically take the form of a partially improved road erroneously being classified as an improved road, rather than an earth road being mislabeled as a highway. The lower category-precision is due to very small stretches of missclassified roads that should only marginally affect the estimates of travel times. In addition, we see no reason to believe that the FCNN introduces non-random errors.

B.5 Retrieving travel speeds

We obtain estimated travel speeds for each of the six main road categories in the Michelin data²⁰ by querying the mapping tool on the Michelin website (www.viamichelin.com). For each road category, we identify a random selection of trips on roads of that category, and record the travel speed returned by the Michelin querying tool (see Figure A10a).²¹ We set the traveling speed on foot-paths to 6 km (about 4 miles) per hour. This corresponds to walking-time estimates on www.maps.google.com (see also [Jedwab and Storeygard 2018](#)).

B.6 Network construction

I transform the road data from the Michelin map corpus into planar graphs that uniformly cover geographic space. I do so in a step-wise manner:

1. **Foot-path network:** The basis of the planar graphs consist of network of 8-connected ‘foot-paths’, shown for the case of Uganda in Figure A9a. The graph’s nodes are the centroids of a raster of population estimates from the HYDE 3.1 data ([Goldewijk, Beusen and Janssen 2010](#)) for 1960 at a resolution of $.04167 \times .04167$ decimal degrees (or ca. 5 km at the equator). Each node is connected with a foot-path to its 8 nearest neighbors using queen moves. This setup allows for much more flexible applications than travel-query APIs such as Google Maps which do not process queries from/to points that are too distant from the next road.
2. **Adding roads:** I then overlay the basic foot-path network with the spatial lines extracted from each map corpus (see Figure A9b) after aligning them all to the last

²⁰The 16 types of roads in the Michelin data are collapsed into the 6 main categories.

²¹Note that the average speed returned for ‘highways’ is somewhat lower than that returned for ‘hard surface’ roads. Highways are almost non-existent in Africa. They constitute only .06 percent of the total road mileage observed in 1966 and cluster in the immediate neighborhood of large cities where speed is slowed by congestion. To preserve the rank-ordering of roads (which is important for our road simulation), we recode all highways as hard surface roads.

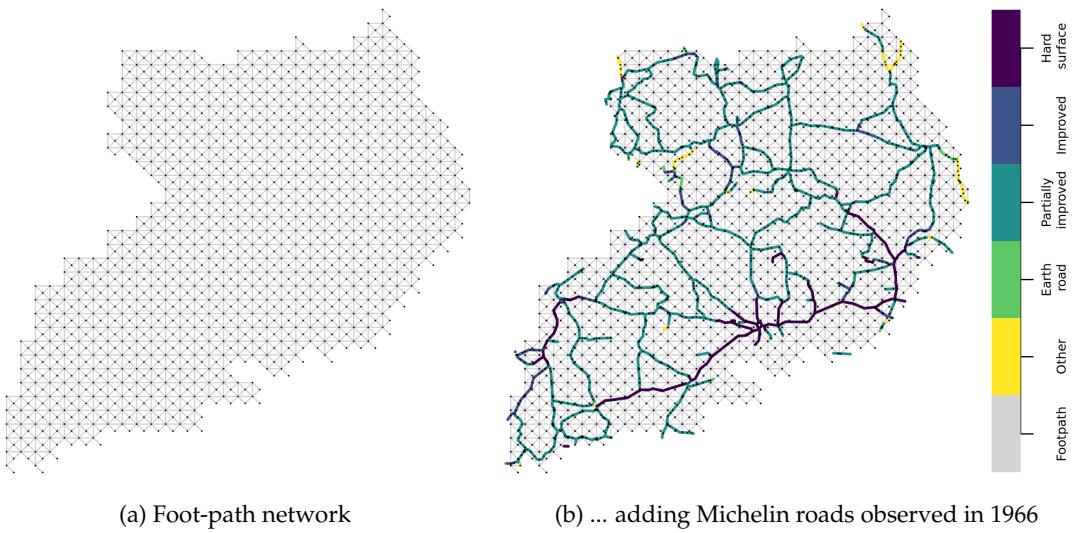


Figure A9: Constructing road networks that regularly cover geographical space. Additional vertices are added to the graph where foot-paths and roads intersect.

and most extensive network observed in 2017.²² I create additional nodes wherever two roads or foot-paths cross, thus retaining the planar graph property. These additional nodes' purpose is to serve as intersections. They are not associated with any population data. Hence, travel between two populated nodes will typically start by taking a foot-path to a road, and end by traveling from a road to the target node on another foot-path. Note that the occasional imperfect spatial alignment of road networks observed in consecutive Michelin maps causes variation in the length of these first or last foot-paths of some trips. The resulting variation in travel times is however deemed negligible and, importantly, random.

3. **Calculating edge weights:** Each edge on the network is associated with an edge weight which is equivalent to the estimated time it takes to traverse the edge (see Subsection B.5).

B.7 Network Validation

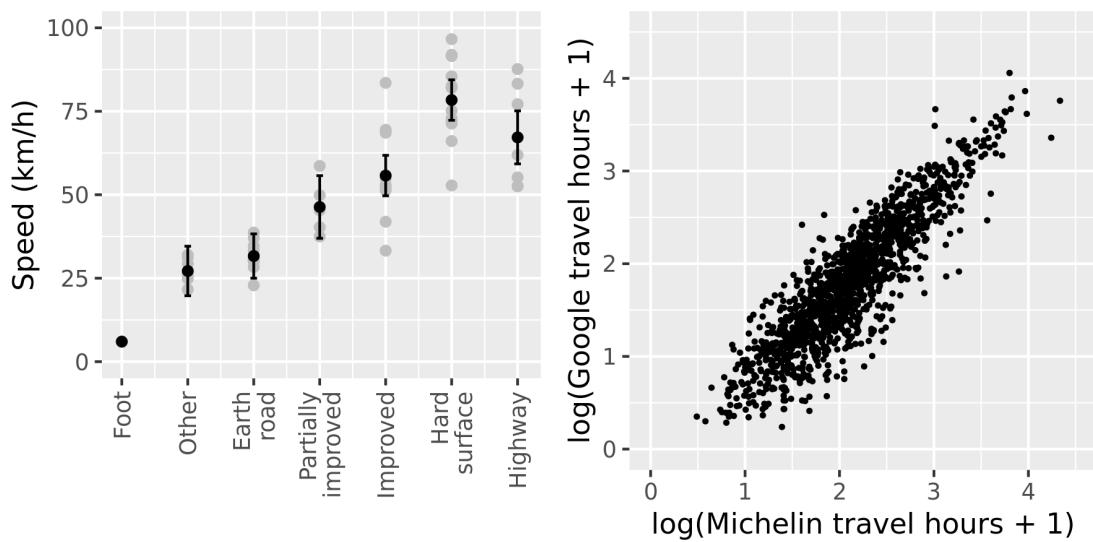
We validate the travel times computed on the Michelin-based network using the Google Maps API. Since Google only offers contemporary data, we base our comparison on road networks constructed with Michelin data from 2003.²³ For each country, we draw 50 source- and 50 destination nodes, each with a probability relative to a node's population

²²This alignment is necessary to preclude small ‘jumps’ in the location of roads, caused by the digitization procedure, to introduce random noise into the measure and downwards bias the estimation results. The alignment of roads is computed via ArcGIS’s align_feature function.

²³This data validation was conducted with the original set of 6 digitized road maps, where the map from 2003 was the youngest.

size.²⁴ These nodes make up the start- and end-points of 50 paths, for which we compute both foot-travel and road-travel times on our network. We query the travel time between the two coordinates on the Google Maps API. Since the API allows only the search of geographical paths which start and end in close proximity to a road, only 60% of our queries are successful.

I compare the results from the 1,416 successful queries with our Michelin-based computations. Figure A10b plots the two data sources against each other. The figure shows a high correlation of ≈ 1 which is least precise at low travel times. This imprecision likely results from the fact that, in certain areas, Google Maps uses data on very small roads whereas our Michelin-based networks approximate such roads as ‘foot-paths’.



(a) Estimate of travel speed on different road types (b) Comparison of travel times on the Michelin-based road network (roads from 2003) and travel times queried from the Google Maps API.

Figure A10: Construction and validation of edge-weights.

C Travel times to capitals as proxy for state reach: Data description and validation

C.1 Validation

I use data from the [Afrobarometer \(2018\)](#) surveys to validate travel times to national and regional capitals as proxies for subnational state capacity. In particular, the surveys contain information provided by the enumerators about the presence of state organs and services in each enumeration area (EA). The respective items range from the local provision of electricity, water and sewage, over the presence of a school, clinic, or post office, to

²⁴Fewer if the country in question does not have 50 populated nodes.

the presence of police and military forces. All variables are coded as dummies. I combine them into a joint index of local state capacity by taking their first principal component, which explains 36.2 percent of the variation in its constitutive parts. Furthermore, I compute for each EA the time to its regional and national capital at the time of the survey.²⁵

Figure A11 plots the association between EA's travel times to their regional and national capital and each indicator, demeaned by country. All indicators correlate with travel times to capitals, which I take as a first indication of their quality as proxies for local state capacity.

Table A2 goes a step further and compares the association between travel times and the local state capacity index with the correlation between mere geodesic distances to capitals. The results show that both distance measures correlate with the index. However, once both are included in Models 3 and 6, the coefficient of geodesic distances becomes much smaller and loses significance in Model 3. I take this as evidence that travel times are superior to simply taking geodesic distances. After all, state agents typically rely on earth bound vehicles and do not fly as crows.

Table A2: Logged distances to national and regional capitals correlate with state capacity index

	State capacity index					
	(1)	(2)	(3)	(4)	(5)	(6)
Nat. capital: geodesic	-0.295*** (0.037)		-0.026 (0.056)			
Nat. capital: time		-0.635*** (0.080)	-0.588*** (0.117)			
Reg. capital: geodesic			-0.414*** (0.029)			-0.199*** (0.043)
Reg. capital: time				-1.006*** (0.059)	-0.608*** (0.080)	
Unit:	EA	EA	EA	EA	EA	EA
Survey FE:	yes	yes	yes	yes	yes	yes
Mean DV:	-0.0042	-0.0042	-0.0042	-0.0042	-0.0042	-0.0042
Std.-dev. DV:	2	2	2	2	2	2
Observations	11,302	11,302	11,302	11,302	11,302	11,302
Adjusted R ²	0.311	0.317	0.317	0.348	0.351	0.356

Note:

*p<0.1; **p<0.05; ***p<0.01

C.2 Description

Three main factors influence the difficulties of a state to reach out to its population: the location of administrative borders and capitals, and the structure of the transportation network that links the state to its subjects, and the geographic distribution of its population. By changing their geography along each of these three dimensions, states can increase their reach and improve their capacity to govern. First, states can optimize the location

²⁵To do so, I use the geocodes provided by [Ben Yishay, Ariel Rotberg et al. \(2017\)](#).

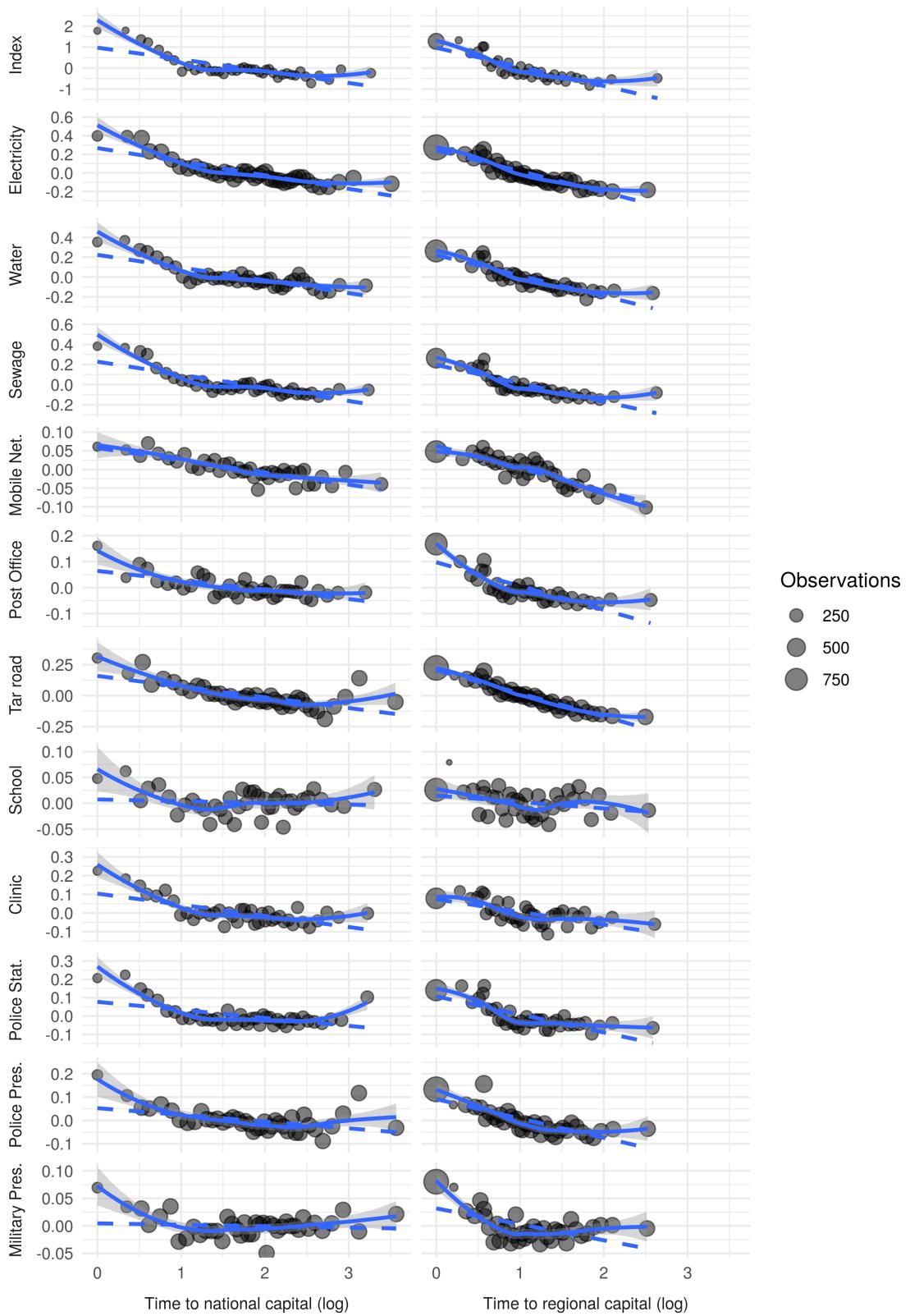


Figure A11: Correlation between travel times to national and regional capitals and state capacity index

The state capacity index is constructed from enumeration area level data in the Afrobarometer survey. The plot shows values of the index that are demeaned by country \times survey round and averaged within 40 quantiles of the travel time to regional and national capitals.

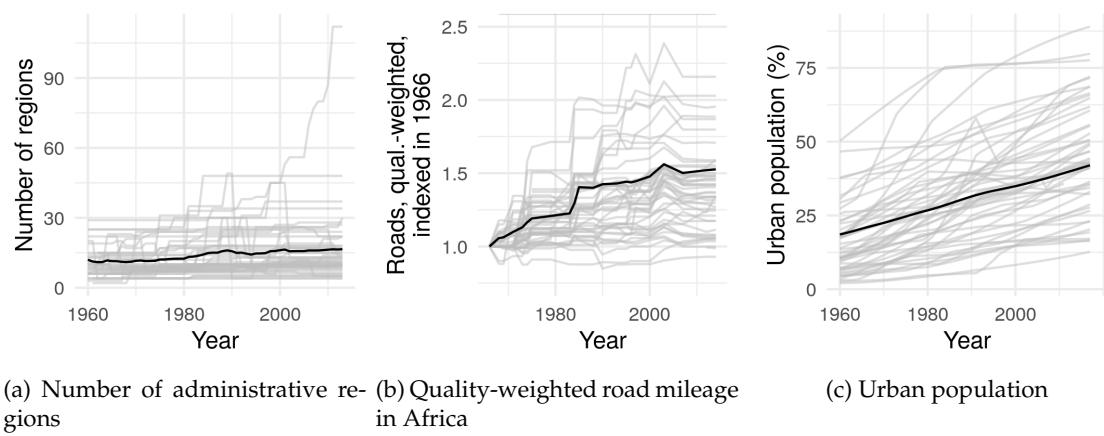


Figure A12: Drivers of expanding state reach in Africa 1960–2015.

Sources: (a) own data; (b) Michelin road map corpus; (c) World Development Indicators ([World Bank 2018](#)).

of its headquarters, open new branches of state agencies, and shift the boundaries of administrative units ([Fesler 1949](#)). Since independence, the Côte d'Ivoire, Nigeria, and Tanzania have relocated their national capitals²⁶ and most have increased the number of administrative units (Figure A12a, see also [Grossman and Lewis 2014](#); [Grossman, Pier-skalla and Dean 2017](#)).²⁷ Similarly, the independence of Eritrea and South Sudan has reduced the distance between their population and the capital.²⁸

Second, states can improve transport networks to access certain areas. Although the main backbones of African road networks are of colonial origin, the networks' extent has increased by about 50 percent since independence (Figure A12b). While this figure is not necessarily impressive ([Herbst 2000](#)), it has shortened the distance between states and their citizens.

Lastly, states can incentivize their population to concentrate and urbanize, thereby increasing governments' economies of scale of reaching out to a particular populated place (e.g. [Scott 2017](#)). While in 1960 only about 20% of Africans have lived in cities, the proportion of urban residents in 2016 has risen above 40% (Figure A12c). Among other social changes brought about by this development, rural-urban migrants experience a steep increase in state reach since administrations and state institutions are typically based in cities. Equivalent state-led population concentration also occurred in the countryside. In particular villagization programs, such as the resettlement of millions of Tanzanians into so-called 'Ujamaa-Villages' in the 1970s (e.g. [Miguel 2004](#)), have made rural populations more accessible to the state.

Together, administrative unit changes, road building, and population concentration

²⁶Côte d'Ivoire (Abidjan to Yamoussoukro in 1983), Nigeria (Lagos to Abuja in 1991), and Tanzania (Dar es Salaam to Dodoma in 1974). The change in Tanzania was less de facto than de jure. Until today, all ministries are located in Dar es Salaam.

²⁷Note the case of Uganda being the outlier with the steepest increase in Figure A12a.

²⁸Their independence of course also affected other dimensions of the distance between the state and citizens, in particular the ethnic distance.

since the 1960s have decreased the distance between the state's headquarters and citizens, thereby extending states' reach over the continent. As seen in Figure A13, the average travel time between African national capitals and citizens has decreased from 11.7 hours in 1966 to 9.3 hours in 2016, a change of about 20.8 percent. Moving to the level of regional administrations where changes in the design of units are more common, we observe a steeper trend. While citizens in 1966 had an average travel time of 5.1 hours to their regional capital, they had to travel 'only' 3.5 hours in 2016 – a decrease of 32.2 percent.

As expected from the large diversity of countries and their geographies, these continent-wide aggregates mask substantial heterogeneity across and within states. Figure A15 visualizes this variation and plots population-weighted densities of travel times to national capitals for five countries in 2016 and their change since 1966. From a cross-sectional perspective, Subfigure A15a shows how states differ in their reach towards their population. Capitals of countries with "difficult geographies" ([Herbst 2000](#)) and poor infrastructure such as the DR Congo are farthest away from their median inhabitant (34.3 hours). In the mid-range, we find Mali where one travels 7.2 hours from Bamako to the median citizen. Lastly, capitals of small countries such as Rwanda naturally are closest to their median citizen (2.6 hours). Similar variation marks changes in the accessibility of the population since 1966 (Subfigure A15b). Here, the populations of states that seceded (Eritrea, Namibia, and South Sudan) and of those that relocated their capital (Nigeria, Côte d'Ivoire, Tanzania) profited the most. Other states, such as the DR Congo, significantly improved their reach in absolute term. However, in relative terms, these improvements look less impressive.

Even more striking than the variation across countries is the variation observed within countries. The density plots in Subfigure A15a visualize high levels of inequality in state reach in some countries. In particular states that [Herbst \(2000\)](#) associates with 'difficult geographies'²⁹ exhibit large variation in travel times – to the point where the distribution of travel times in the DR Congo is heavier in its right than left tail. Similarly, new roads, borders, and capitals do not have a geographically uniform effect. For example, relocating the Nigerian capital from Lagos to Abuja in 1991 increased state reach towards the northern areas of the country, while decreasing it around Lagos in the southwest (Figure A14).

²⁹Angola, DR Congo, Ethiopia, Mozambique, Namibia, Nigeria, Senegal, Somalia, Sudan (borders of 2000), Tanzania. Cf. Herbst ([2000](#), p. 161).

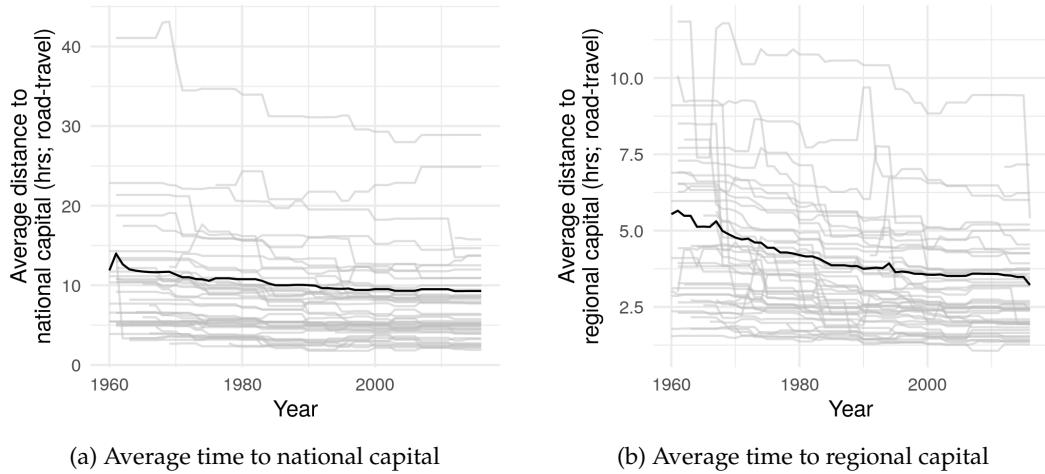


Figure A13: Decreasing travel times to regional and national capitals in Africa 1960–2015.

All averages are population weighted. Sources: Own calculations based on HYDE population estimates ([Goldewijk, Beusen and Janssen 2010](#)), Michelin-based road networks, Cshapes ([Weidmann and Gleditsch 2010](#)), and own data on administrative regions and their capitals.

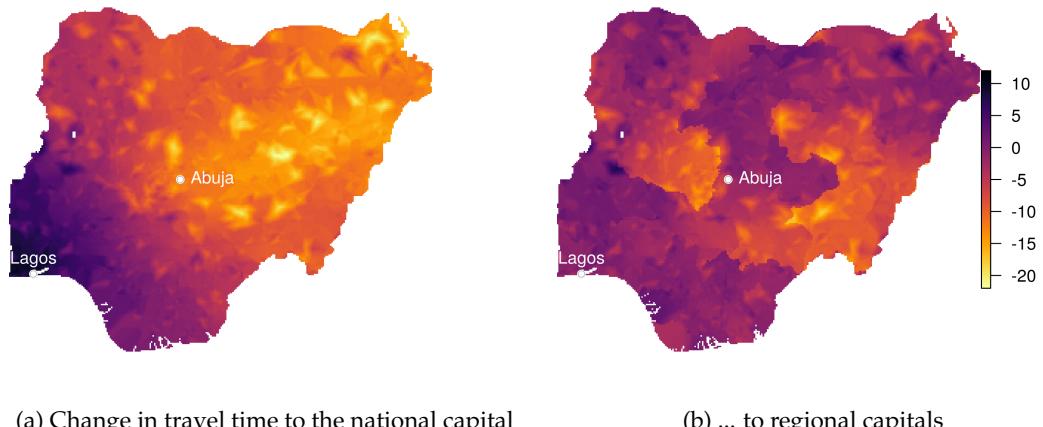


Figure A14: Change of travel times to capitals in Nigeria 1966–2016.

Note: Brighter colors indicate a decrease of travel times, thus an improvement in state reach. Sources: See Figure A13.



Figure A15: Population-weighted distribution of state reach in Africa in 2016 and development since 1966.

Note: Brighter colors identify better outcomes. To calculate population-weighted distributions for single countries, I fix the spatial distribution of the population and the borders of countries to their status in 2016. For areas that became independent after 1966 no change can be calculated. Sources: Own calculations based on 2016 WorldPop population estimates, Michelin-based road networks, and country-borders from 2016 from Cshapes ([Weidmann and Gleditsch 2010](#)).

D Data and summary statistics

D.1 Units of the nightlight analysis: Voronoi cells

In order to divide an arbitrary geographical space into units of roughly equal size and high levels of compactness (i.e. similarity to a circle),³⁰ this section introduces a spatial clustering algorithm that is combines the advantages of the k-means clustering algorithm ([Lloyd 1982](#)) that requires a finite sample of points to cluster, and the Voronoi tessellation that is used to transform the centers of the k-means clusters into continuous areas. The algorithm proceeds as follows:

1. Draw a large number of points P from the area of polygon t in set T . T is defined in the present application as the constitutive parts of all administrative regions observed between 1992 and 2013. These polygons are computed as cutting the territory of each country with all regional borders existing between 1992 and 2013. The resulting polygons are strictly nested within all regional boundaries. Points are sampled from these polygons on the basis of a raster with a resolution of $\approx 1\text{km}$ (.01 decimal degrees).
2. Conduct a k -means clustering ([Lloyd 1982](#)) of points P into N clusters, with $N = \text{round}(A_T/A_{target})$, thus N being the number of units to create so that the average area of each unit comes closest to the target size of units. For the main analysis, the target size is 400 km^2 , further variations are conducted in the robustness check presented below in Subsection E.6. For best results, I initialize the k-means algorithm with a random spatial sample of N points from P .
3. Take the centroids of the clusters thus computed and conduct a Voronoi tessellation around them.
4. Crop the resulting Voronoi polygons with the target polygon T .

We can now compare the Voronoi cells with the more commonly used quadratic grid cells. Across various target sizes, the Voronoi cells resulting from the relatively simple (and fast) algorithm significantly improve over the quadratic grid cells, both in terms of their distribution around the target size and in terms of their level of compactness. First, Figure A17 shows that many regular quadratic grid cells are smaller than the target size – this occurs wherever a cell is cut by a regional border or the coast line. Thus, the heterogeneity in units' sizes is correlated with their closeness to the coast and border, a feature which might introduce slight bias into an analysis. Second, Figure A18 proves that Voronoi cells are much more compact than grid cells, which – as quadratic shapes and in particular where cut by borders – are shaped in a more irregular and less “circle-like” manner.

³⁰The shape that can continuously cover an area with the highest level of compactness is the hexagon. However, where the honeycomb reaches a border, hexagonal cells must be cut or reshaped, thus deviating from the requirements of uniform size and compactness.

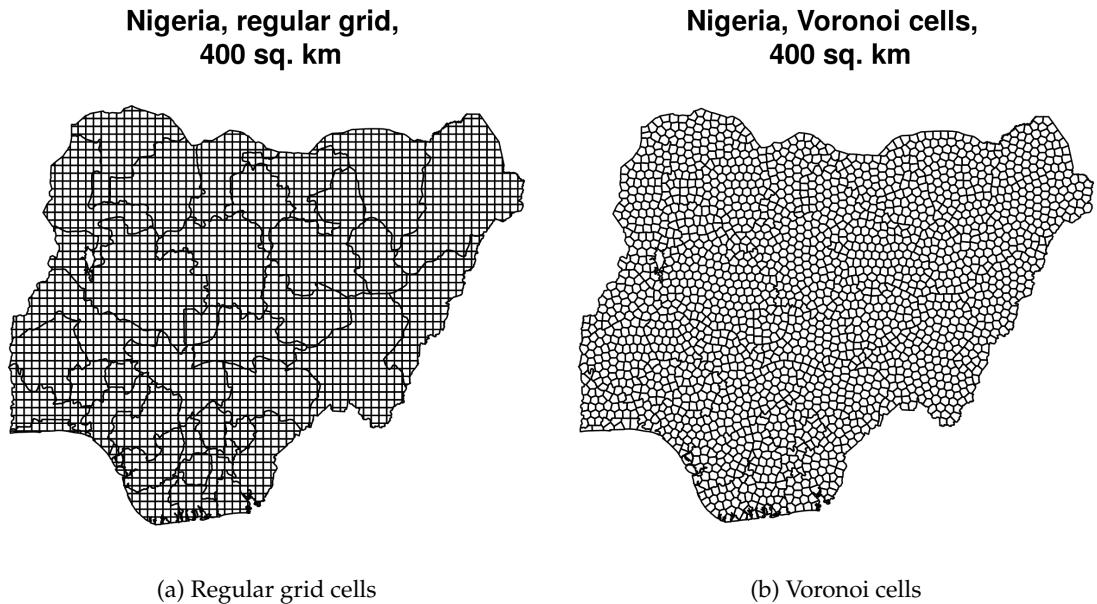


Figure A16: Regular grid and Voronoi cells for Nigeria with regional borders observed in 1992–2013.

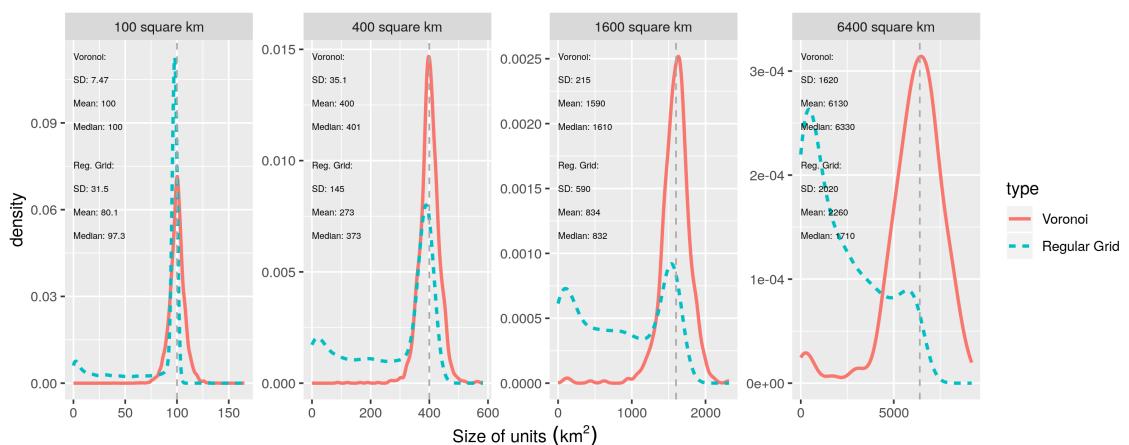


Figure A17: Size of Voronoi and regular grid cells for varying target sizes.

Vertical dashed line indicates the target size of units. Cells are constructed for all countries in mainland Africa, using borders from the year 2000.

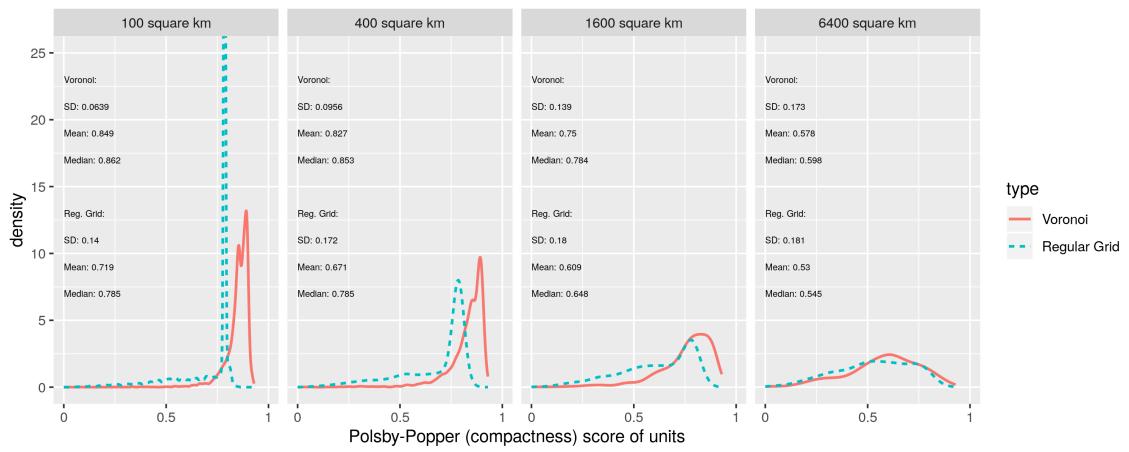


Figure A18: Compactness of Voronoi and regular grid cells for varying target sizes.

The compactness of each unit is calculated according to [Polsby and Popper \(1991, p. 349\)](#) as: $(4\pi A_i)/P_i^2$, where A_i is the size of unit i and P_i its perimeter. Cells are constructed for Nigeria, using regional borders observed between 1992 and 2013. Note that densities above 25 are censored to improve the readability of the graphs.

D.2 Summary statistics

Table A3: Summary statistics: DHS education data (Personal Recode)

Statistic	N	Mean	St. Dev.	Min	Max
Primary educ. (0/100)	1893067	70.22	45.73	0	100
Female	1893067	0.53	0.50	0	1
Age	1893067	27.92	9.66	15	57
Time to nat. capital (log)	1893067	1.92	0.81	0.00	4.32
Time to reg. capital (log)	1893067	1.17	0.65	0.00	3.93

Table A4: Summary statistics: DHS infant mortality data (Children Recode)

Statistic	N	Mean	St. Dev.	Min	Max
Infant mort. (0/100)	2644591	9.89	29.86	0	100
Female	2644591	0.49	0.50	0	1
Birth-order	2644591	3.30	2.25	1	18
Twin	2644591	0.03	0.18	0	1
Mother's age at birth	2634666	24.64	6.42	10	49
Time to nat. capital (log)	2644591	1.94	0.72	0.00	4.26
Time to reg. capital (log)	2644591	1.21	0.62	0.00	3.93

Table A5: Summary statistics: Nightlight data (Voronoi cells, 400 km²)

Statistic	N	Mean	St. Dev.	Min	Max
Light/capita (log)	1506991	-6.51	1.28	-6.91	10.12
Time to nat. capital (log)	1506991	2.92	0.77	0.26	4.97
Time to reg. capital (log)	1506991	2.30	0.83	0.26	4.94

Table A6: Samples across data sources, DHS rounds and nightlight observations

Country	Adults	Children	Households	Nightlight-cells
Algeria				1992–2013
Angola	7.1	6.1, 7.1	5.1, 6.1, 7.1	1992–2013
Benin	4.1, 6.1	4.1, 6.1	6.1	1992–2013
Botswana				1992–2013
Burkina Faso	2.1, 3.1, 4.1, 6.1	2.1, 3.1, 4.1, 6.1	4.1, 6.1, 7.1	1992–2013
Burundi	6.1, 6.2, 7.1	6.1, 6.2, 7.1	6.1, 6.2, 7.1	1992–2013
Cameroon	2.1, 4.1, 6.1	2.1, 4.1, 6.1	4.1, 6.1	1992–2013
Central African Republic	3.1	3.1		1992–2013
Chad				1992–2013
Congo				1992–2013
Côte D'Ivoire	3.1, 3.2, 6.1	3.1, 3.2, 6.1	6.1	1992–2013
Djibouti				1992–2013
DR Congo	5.1, 6.1	5.1, 6.1	5.1, 6.1	1992–2013
Egypt	2.1, 3.1, 4.1, 4.2, 5.1, 5.2	2.1, 3.1, 4.1, 4.2, 5.1, 5.2	5.1, 5.2	1992–2013
Equatorial Guinea				1992–2013
Eritrea				1993–2013
Ethiopia	4.1, 5.1, 6.1, 7.1	4.1, 5.1, 6.1, 7.1	5.1, 6.1, 7.1	1992–2013
Gabon	6.1	6.1	6.1	1992–2013
Gambia				1992–2013
Ghana	3.1, 4.1, 4.2, 5.2, 7.1	3.1, 4.1, 4.2, 5.2, 7.1	4.2, 5.2, 7.1, 7.2	1992–2013
Guinea	4.1, 5.1, 6.1	4.1, 5.1, 6.1	5.1, 6.1	1992–2013
Guinea Bissau				1992–2013
Kenya	4.1, 5.1, 7.1	4.1, 5.1, 7.1	4.1, 5.1, 7.1, 7.2	1992–2013
Lesotho	4.1, 6.1, 7.1	4.1, 6.1, 7.1	4.1, 6.1, 7.1	1992–2013
Liberia	5.1, 6.2	0.1, 5.1, 5.2, 6.1, 6.2	5.1, 5.2, 6.1, 6.2, 7.1	1992–2013
Libya				1992–2013
Malawi	4.1, 4.2, 6.1, 7.2	4.1, 4.2, 6.1, 6.2, 7.2	4.2, 6.1, 6.2, 7.1, 7.2	1992–2013
Mali	3.1, 4.1, 5.1, 6.2	3.1, 4.1, 5.1, 6.2	5.1, 6.2, 7.1	1992–2013
Mauritania				1992–2013
Morocco	4.1	4.1	4.1	1992–2013
Mozambique	5.1, 6.1, 7.1	6.1, 7.1	5.1, 6.1, 7.1	1992–2013
Namibia	4.1, 5.1, 6.1	4.1, 5.1, 6.1	5.1, 6.1	1992–2013
Niger	2.1, 3.1	2.1, 3.1		1992–2013
Nigeria	2.1, 4.2, 5.1, 6.1, 6.2, 7.1	2.1, 4.2, 5.1, 6.1, 6.2	4.2, 5.1, 6.1, 6.2, 7.1	1992–2013
Rwanda	5.1, 6.1, 7.1	5.1, 5.2, 6.1, 7.1	5.1, 5.2, 6.1, 7.1	1992–2013
Senegal	2.1, 4.2, 6.1, 6.2	2.1, 3.1, 4.2, 5.2, 6.1, 6.2	4.2, 5.2, 6.1, 6.2	1992–2013
Sierra Leone	5.1, 6.1	5.1, 6.1	5.1, 6.1, 7.1	1992–2013
Somalia				1992–2013
South Africa				1992–2013
South Sudan				2011–2013
Sudan				1992–2013
Swaziland	5.1	5.1	5.1	1992–2013
Tanzania	4.1, 4.2, 5.1, 6.1, 6.2	4.1, 5.1, 6.1, 6.2	4.2, 5.1, 6.1, 6.2	1992–2013
Togo	3.1, 6.1	0.1, 3.1, 6.1	6.1	1992–2013
Tunisia				1992–2013
Uganda	4.1, 5.1, 6.1, 6.2	4.1, 5.1, 5.2, 6.1	5.1, 5.2, 6.1, 6.2	1992–2013
Zambia	5.1	5.1	5.1	1992–2013
Zimbabwe	4.1, 5.1, 6.1	4.1, 5.1, 6.1	5.1, 6.1	1992–2013

Note that the divergence between the samples used from the DHS stems from the fact that not all surveys enlist the level of education of household members, come with the Child Recode file needed to derive infant mortality rates, or include the DHS wealth index.

A24

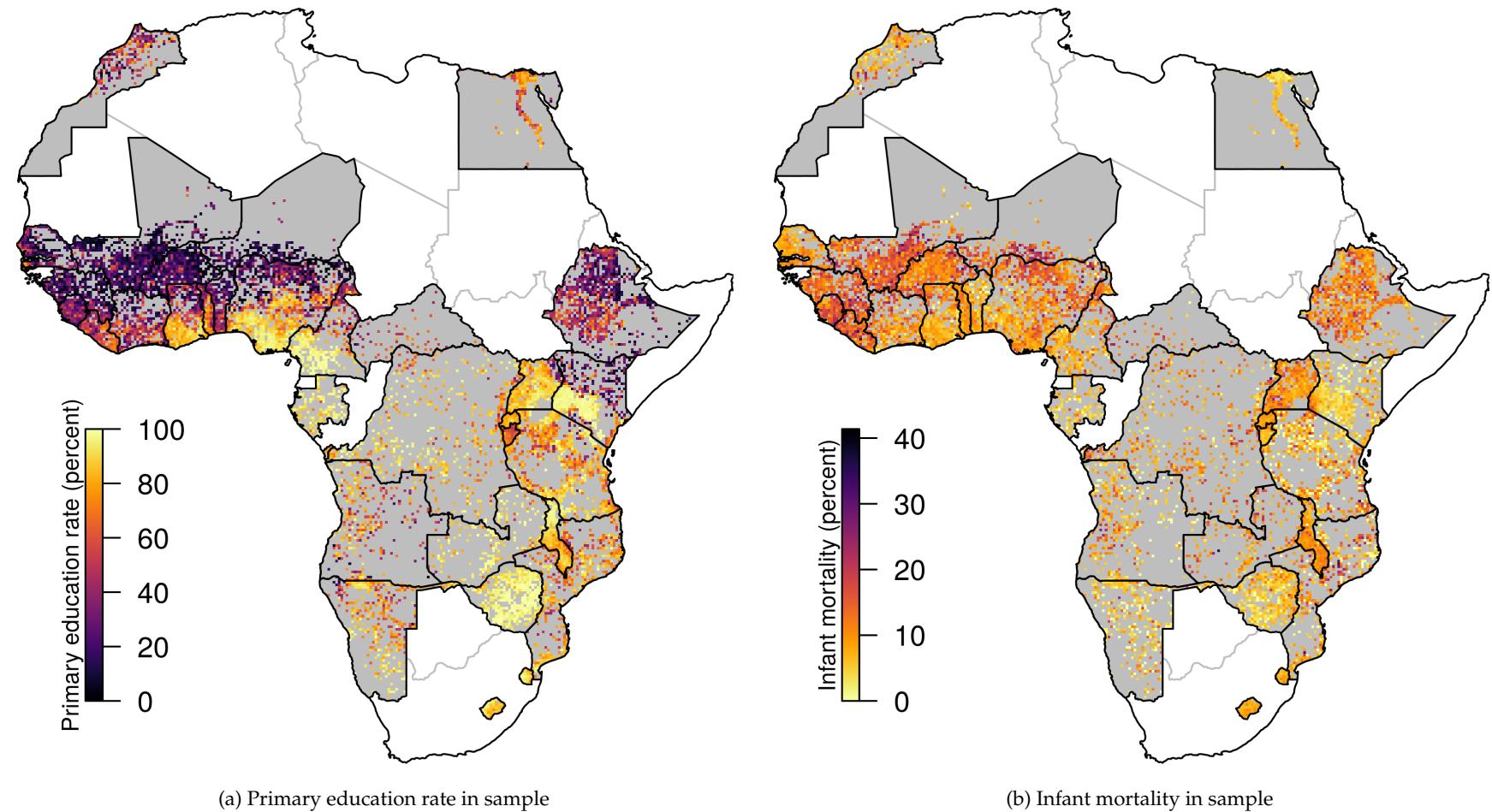


Figure A19: Primary education and infant mortality in the DHS Data, aggregated to .25 raster cells.

E Robustness checks

This section describes additional analyses that probe the robustness of the analysis of the effects of changes in the distance to national and regional capitals on local development. Figure A20 provides an overview over the various robustness checks. All additional models are described in detail below. Where Figure A20 captures the main insights from an additional analysis, I do not report detailed results as a table. However, the reader may note that all results will be made available as tables with the replication data.

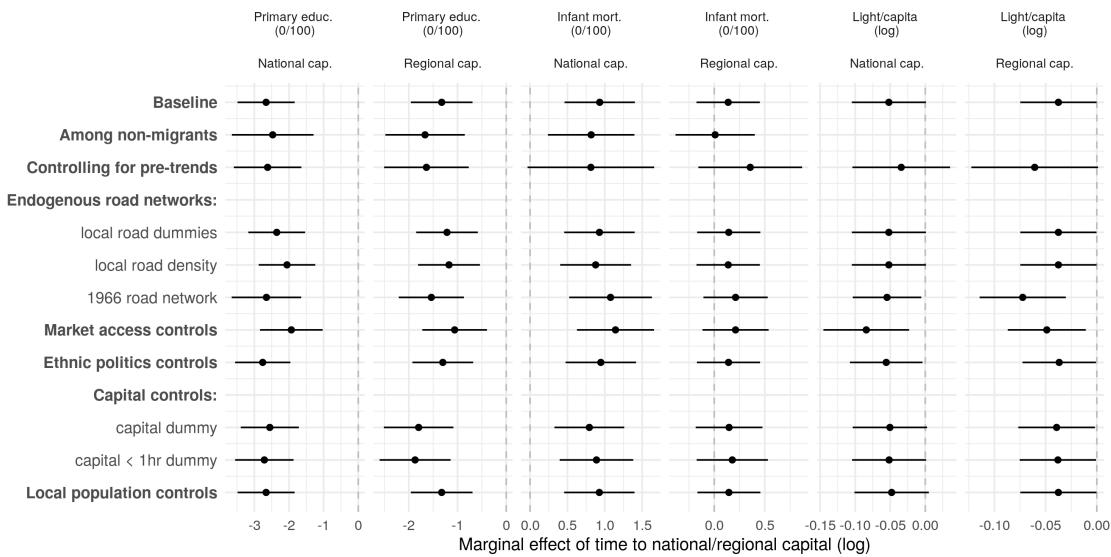


Figure A20: Panel analysis: Summary of robustness checks.

E.1 Re-locations of national capitals vs. road-network development:

One main concern of the analysis of changes in travel times to national capitals is that we observe only very few relocations of capitals that drive the results. In the sample of DHS respondents, capital re-locations have occurred in Côte d'Ivoire (Abidjan to Yamassoukro), Nigeria (Lagos to Abuja), and Tanzania (the de jure move from Dar es Salaam to Dodoma). The data on nightlights include the secessionist cases of Eritrea and South Sudan. In order to gauge whether the baseline estimates are due to these relocations and secessions or whether they are driven by changes in the road networks that link national capitals to their citizens, Table A7 presents the results of estimating the baseline panel specification on the split samples. For primary education rates we observe a larger effect in the sample of DHS respondents from countries with capital relocations, than from those without. However, also in the latter case the estimate is substantive and statistically significant, meaning that improved road connections to national capitals come with increases in local primary education rates. In the case of infant mortality rates, the results show them to be driven by capital relocations rather than road network improvements. Lastly and consistent with the baseline results, the results show only insignificant effects

of either better road connections or new capitals on nightlight emissions. Note however, that the secessionist cases of Eritrea and South Sudan are each only observed one year pre-/post-treatment. We would hardly expect the new capitals to have such an immediate and sudden effect on nightlight emissions, in particular since both cases were riven by civil war before (and after, in the case of South Sudan) their secession.

Table A7: Changes in time to national/regional capital and local development: Cases where national capitals changed

	Primary educ. (0/100)		Infant mort. (0/100)		Light/capita (log)	
	(1)	(2)	(3)	(4)	(5)	(6)
Time to nat. capital (log)	-1.409** (0.645)	-3.392*** (0.523)	0.105 (0.306)	1.484*** (0.318)	-0.044 (0.034)	-0.054 (0.037)
Time to reg. capital (log)	-1.617*** (0.379)	-0.784 (0.614)	0.051 (0.169)	1.065** (0.462)	-0.039* (0.021)	-0.039* (0.023)
$\beta_1 + \beta_2:$	-3.026*** (0.674)	-4.176*** (0.778)	0.156 (0.31)	2.549*** (0.574)	-0.083** (0.034)	-0.094** (0.039)
Cap.-re-loc. cases:	drop	only	drop	only	drop	only
Point FE:	yes	yes	yes	yes	yes	yes
Country-year FE:	yes	yes	yes	yes	yes	yes
Survey FE:	yes	yes	yes	yes	yes	yes
Controls:	yes	yes	yes	yes	-	-
Mean DV:	70	71	9.7	11	-6.5	-6.8
Observations	1,583,837	309,230	2,246,915	387,294	1,467,083	39,908
Adjusted R ²	0.445	0.446	0.051	0.049	0.838	0.395

Notes: OLS linear models. Control variables for models with primary education as the dependent variable consist of respondents' 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

E.2 Testing for pre-trends:

As highlighted in the main text, one important threat to inference in the panel analysis is that the baseline estimates may be biased by differential pre-trends in local development that reversely cause the extension of state reach. I here present the full details of the empirical test that accounts for such trends. In particular, I re-estimate the baseline specification, adding measures of future travel times to capitals in $t + x$. These leads of the main treatments capture differential changes in local development that occur before changes in travel times affect localities. More specifically, for the analysis of local primary education rates, which are affected by state reach during age 6 to 11 of respondents, I estimate the effect of the average travel time to capitals during age 6-11, and add five separate controls for the travel times to capitals at age 12 to 16. For infant mortality rates, which are affected only in the year of an infants' birth, I add the time to capitals at age 2 to 6. Similarly, for local nightlight emissions, I add the respective variables for $t + 1 \dots 5$.

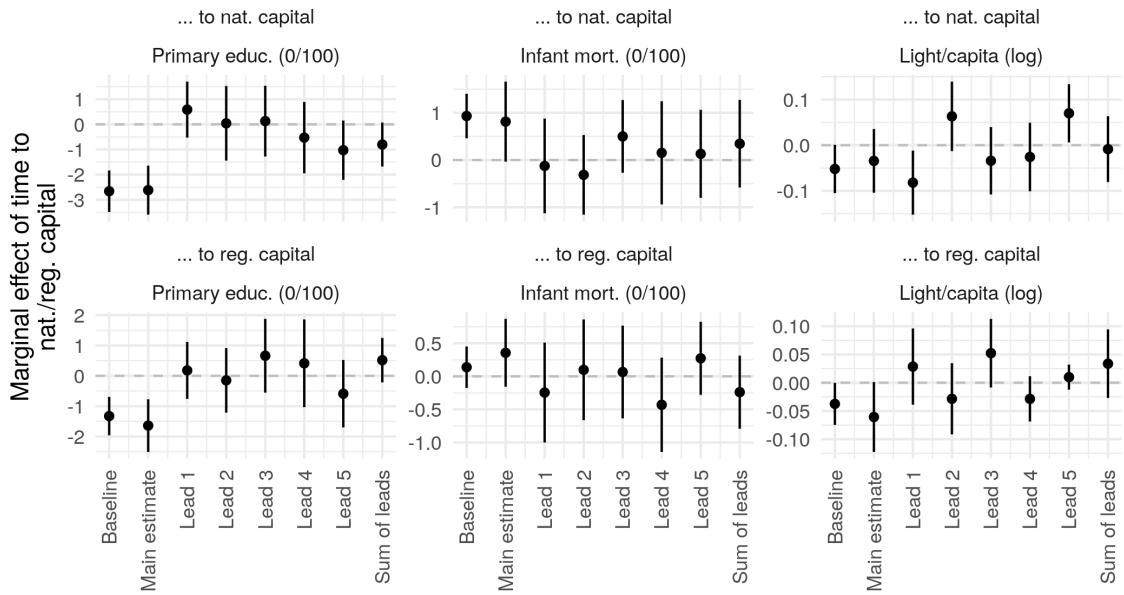


Figure A21: All leads, based on Table A8.

Table A8 and Figure A21 show the results of these demanding³¹ specifications. The first thing to note is that the main estimates of the effects of travel times to capitals are only marginally and insignificantly different from those obtained in the baseline specification. This is a first sign that these are not affected by differential pre-trends. Second however, and as Table A8 shows, the sum of leads is negative and marginally significant in the estimation of the effect of travel times to national capitals on education rates. This suggests that education rates increase before capitals move closer to a location through roads or relocation. The leads in the infant mortality analysis are in sum close to zero and show few signs of divergent pre-trends. The leads in the nightlight analysis have very heterogeneous estimates, but are, for the most part statistically insignificant and in sum not different from zero. Both patterns suggest that the main estimate is not systematically affected by differential pre-trends.

One reason for the negative and marginally significant lead effect of travel times to national capitals on primary education rates consists in biased migration patterns by which individuals select in and out of treatment after having gone to school. These patterns can be picked up by the lead effects because I attribute travel times on the basis of the current location of DHS respondents. In order to test for this possibility, I re-estimate the respective specification, now interacting the treatment variables and their leads with a dummy for migrants.³² The respective variable is only available for the reduced sample of DHS respondents that have gone through the entire interview and is based on whether they

³¹The specifications are demanding because of the high correlations between the actual treatments and its leads.

³²The baseline results for this migrant \times travel times interaction without the leads is reported below in Subsection E.3.

Table A8: Changes in time to national/regional capital and local development: Controlling for leads

	Primary educ. (0/100)	Infant mort. (0/100)	Light/capita (log)
	(1)	(2)	(3)
Time to nat. capital (log)	-2.618*** (0.498)	0.813* (0.432)	-0.034 (0.036)
Time to reg. capital (log)	-1.638*** (0.443)	0.356 (0.262)	-0.061* (0.032)
$\beta_1 + \beta_2$:	-4.256*** (0.629)	1.169*** (0.444)	-0.095*** (0.036)
Sum of leads (nat. cap.):	-0.804* (0.447)	0.346 (0.473)	-0.009 (0.037)
Sum of leads (reg cap.):	0.516 (0.373)	-0.24 (0.282)	0.034 (0.031)
Time to cap. $t+1, \dots, t+5$:	yes	yes	yes
Point FE:	yes	yes	yes
Country-year FE:	yes	yes	yes
Survey FE:	yes	yes	yes
Controls:	yes	yes	-
Mean DV:	70	10	-6.5
Observations	1,887,591	2,561,533	1,369,902
Adjusted R ²	0.445	0.050	0.840

Notes: OLS linear models. Control variables for models with primary education as the dependent variable consist of respondents' 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

have 'always' lived in their current place of residence.³³

Reassuringly, the results in Table A9 show that the lead effects are only negative for the migrants in the sample, but positive for the non-migrants. Both sums of leads are statistically insignificant, presumably due to the smaller sample size. Furthermore, the main effect associated with travel times to national and regional capitals is much larger in the non-migrant sample than in the migrant sample, which is consistent with the fact that non-migrants' primary education is affected by local state reach to greater extent than that of migrants. The results for the respective analysis of the mortality of infants of non-migrant mothers mirror those described above, albeit with the caveat that the children of migrant mothers are slightly – but only weakly significantly – more likely to survive close to capitals than those of non-migrant mothers.

³³Note that the respective question does not allow to distinguish individuals who have moved within the same neighborhood from those who have migrated from one place to another. The migrant dummy therefore overestimates migration.

Table A9: Changes in time to national/regional capital and local development: Controlling for leads, migrants and non-migrants

	Primary educ. (0/100)	Infant mort. (0/100)
	(1)	(2)
Migrant	-3.483*** (0.810)	0.720 (0.535)
Non-migrants: Time to nat. capital (log)	-1.730*** (0.622)	0.144 (0.363)
Non-migrants: Time to reg. capital (log)	-3.971*** (0.795)	0.641*** (0.247)
Migrants: Time to nat. capital (log)	-1.988** (0.825)	-0.023 (0.871)
Migrants: Time to reg. capital (log)	0.050 (0.809)	-0.037 (0.552)
Non-migrants: $\beta_1 + \beta_2$:	-5.213*** (0.996)	0.864 (0.578)
<i>Non-migrant leads:</i>		
Sum of leads (nat. cap.):	0.137 (0.697)	0.36 (0.586)
Sum of leads (reg. cap.):	-0.236 (0.533)	-0.133 (0.388)
<i>Migrant leads:</i>		
Sum of leads (nat. cap.):	-0.655 (0.705)	1.365 (0.897)
Sum of leads (reg. cap.):	-0.473 (0.761)	-0.146 (0.555)
Time to cap. $t+1, \dots, t+5$:	yes	yes
Point FE:	yes	yes
Country-year FE:	yes	yes
Survey FE:	yes	yes
Controls:	yes	yes
Mean DV:	67	11
Observations	671,851	1,465,462
Adjusted R ²	0.479	0.052

Notes: OLS linear models. Control variables for models with primary education as the dependent variable consist of respondents' 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

E.3 Migration

Over their lifetime, DHS respondents might have moved towards or away from changing regional (and national) capitals in a manner correlated with their level of education and wealth. Such migration patterns might bias the results. If that was the case, we should see differential effects of travel times among migrants and non-migrants. In particular, if migrants were driving the results, no effect of changes in the travel time towards capitals should be visible among non-migrants. Table A10 demonstrates that this is not the case. The effect of travel times on education rates is significantly larger for migrants than for non-migrants. In the case of infant mortality rates, the difference between the two sample

is mostly insignificant, except for Model (4), which suggest that longer travel times to regional capitals have a more *negative* effect on infant mortality rates among migrant- than non-migrant mothers. Their absolute effect is however insignificantly different from zero in both cases (see also the lead-analysis above, Table A9).

Table A10: Changes in time to national/regional capital and local development: Migrants and non-migrants

	Primary educ. (0/100)		Infant mort. (0/100)	
	(1)	(2)	(3)	(4)
Time to nat. capital (log)	-2.472*** (0.602)	-2.362*** (0.591)	0.818*** (0.295)	0.643** (0.288)
Time to reg. capital (log)	-1.666*** (0.416)	-1.582*** (0.413)	0.009 (0.200)	-0.004 (0.207)
Migrant	-3.903*** (0.783)	-13.391*** (2.290)	0.654*** (0.243)	1.965*** (0.363)
Migrant×Time to nat. capital (log)	0.652** (0.312)	0.567* (0.308)	0.224* (0.132)	0.207 (0.136)
Migrant×Time to reg. capital (log)	1.542*** (0.264)	1.308*** (0.266)	-0.151 (0.154)	-0.154 (0.154)
Non-migrants: $\beta_1 + \beta_2$:	-4.138*** (0.704)	-3.944*** (0.694)	0.826** (0.325)	0.639** (0.322)
Migrant × controls	no	yes	no	yes
Point FE:	yes	yes	yes	yes
Country-year FE:	yes	yes	yes	yes
Survey FE:	yes	yes	yes	yes
Controls:	yes	yes	yes	yes
Mean DV:	67	67	11	11
Observations	673,335	673,335	1,485,168	1,495,445
Adjusted R ²	0.479	0.480	0.052	0.034

Notes: OLS linear models. Control variables for models with primary education as the dependent variable consist of respondents' 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

E.4 Omitted variables:

Controlling for roads: The main results might also be driven by roads that are built at the local level around specific towns and villages. Because such road building inherently lowers the distance to all capitals, it might lead to spurious results if it was caused by increasing levels of development in these areas. To exclude such omitted variable bias, I employ two strategies. First, I control for the mileage of roads in the geographic neighborhood of respondents (20km) / in a Voronoi cell (see Figure A20). This is similar to the strategy of [Donaldson and Hornbeck \(2016\)](#) in that the variation left stems from changes in the road network outside a particular point's neighborhood. The second and more stringent test is to re-estimate all models with non-time variant road networks. In such a setting, all variation within points stems solely from changes in the administrative design

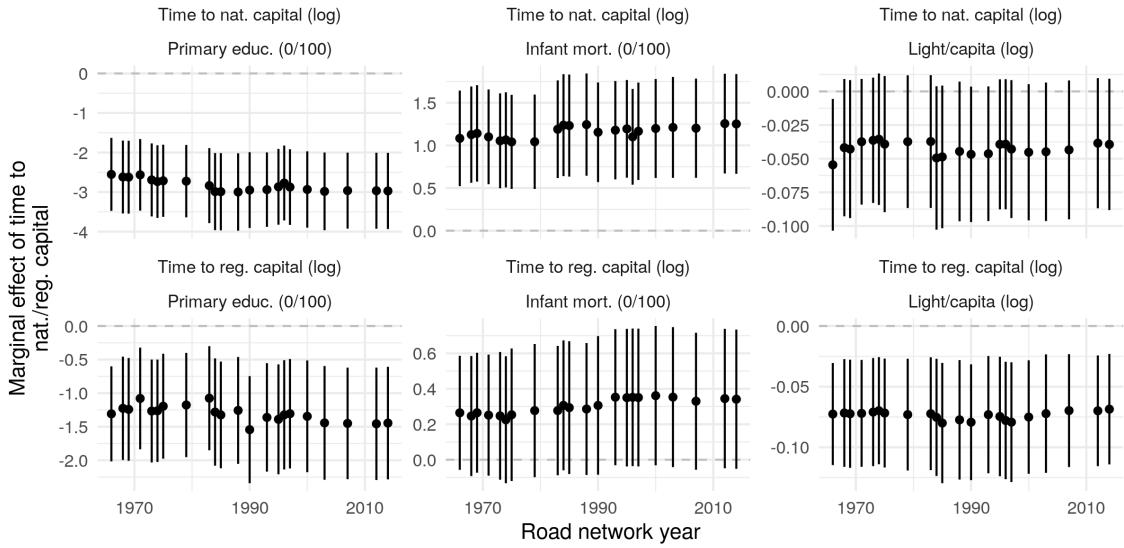


Figure A22: Estimating the baseline models with constant road networks from varying years

of a country – that is the location of boundaries and capitals. I do so with each Michelin network that I observe. The results, plotted in Figure A22, are insensitive towards these changes – it appears that the magnitude of baseline results are unaffected by purging the model of temporal variance induced through changes in road networks. If at all, the absolute size of the estimates slightly increases.

Controlling for market access: Travel times to regional and national capitals might not only proxy for transaction costs between governments and citizens, but more broadly transaction costs between the participants of economic markets. Variation in economic market access might lead to spurious results, since it results in higher levels of economic activity and development (e.g. [Donaldson 2018](#); [Donaldson and Hornbeck 2016](#); [Eaton and Kortum 2002](#); [Jedwab and Moradi 2016](#); [Jedwab and Storeygard 2018](#)). I thus follow the economic literature on the effect of market access on economic growth and calculate, for each year, the road-network based access to national and international markets. Following [Donaldson \(2018\)](#) and [Eaton and Kortum \(2002\)](#), I define the measure as:

$$MA_{p,t} = \sum_{m=1}^M c_{p,m,t}^{-\theta} * P_{m,t},$$

where the market access of point p in year t is the sum of the market potential P of a market m in year t multiplied by the travel time between p and m calculated on the road network and discounted by a trade elasticity θ . Because [Donaldson \(2018\)](#) and [Eaton and Kortum \(2002\)](#) estimate different trade elasticity measures ($\theta = 8.28$ and 3.2 respectively), I construct the market access measure for both parameters. Because the effects of access to national and international markets might differ, I calculate MA separately for markets inside and outside of p 's country. I define markets as the 1530 biggest cities and towns

in Africa. These are all cities that ever reached more than 50'000 inhabitants since 1950³⁴. Each city's market potential P is approximated by its population as measured in each decade. Controlling for the four resulting variables in Table A11 indicates does not affect the magnitude of the effects associated with the measures of state reach. This indicates that a low distance towards capital cities increases local development above and beyond the effect of the economic markets they harbor. The effect of market access on education rates is slightly positive, once we take the sum of the respective coefficients. They have a mixed effect of infant mortality and nightlight emissions.

Table A11: Changes in time to national/regional capital and local development: Controlling for market access

	Primary educ. (0/100) (1)	Infant mort. (0/100) (2)	Light/capita (log) (3)
Time to nat. capital (log)	-1.933*** (0.461)	1.143*** (0.263)	-0.084*** (0.031)
Time to reg. capital (log)	-1.059*** (0.339)	0.210 (0.167)	-0.049** (0.019)
MA, internat. (log; $\theta = 3.8$)	2.204*** (0.511)	0.465* (0.257)	-0.146*** (0.018)
MA, nat. (log; $\theta = 3.8$)	-0.840*** (0.179)	-0.049 (0.084)	0.082*** (0.008)
MA, internat. (log; $\theta = 8.28$)	0.489 (0.343)	0.248 (0.190)	-0.009 (0.012)
MA, nat. (log; $\theta = 8.28$)	-0.150 (0.123)	-0.102 (0.078)	0.034*** (0.006)
$\beta_1 + \beta_2$:	-2.992*** (0.561)	1.353*** (0.313)	-0.133*** (0.035)
Point FE:	yes	yes	yes
Country-year FE:	yes	yes	yes
Survey FE:	yes	yes	yes
Controls:	yes	yes	-
Mean DV:	70	9.9	-6.5
Observations	1,889,905	2,634,150	1,506,991
Adjusted R ²	0.446	0.051	0.836

Notes: OLS linear models. Control variables for models with primary education as the dependent variable consist of respondents' 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: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Controlling for ethnic politics and war: Ethnic politics are an important driver of both, development and state reach. Research on administrative unit reforms has found that they can reward government allies (Green 2010; Hassan 2016; Gottlieb et al. 2019) or harm opponents (Resnick 2017), strategies which may well be used in processes of ethnic accommodation or exclusion. At the same time, ethnic and regional favoritism of governments has been found to affect investments into road infrastructure (Burgess et al. 2015)

³⁴Data comes from Africapolis.org.

and local development in general (Franck and Rainer 2012; Hodler and Raschky 2014). To capture such dynamics, I draw on the most comprehensive and geocoded data on ethnic power dynamics in Africa since independence, the Ethnic Power Relations data set (Vogt et al. 2015).³⁵ I use the geodata of ethnic groups to map respondents and Voronoi cells to the database's coding of ethnic inclusion and exclusion as well as the occurrence and history of ethnic civil wars. Table A12 shows the results from the baseline analysis with the resulting variables as additional controls. The additional controls do not affect the results. While in particular the time since the last ethnic civil war positively affects education rates, ethnic inclusion has a positive effect on infant survival rates and nightlight emissions.

Table A12: Changes in time to national/regional capital and local development: Controlling for ethnic representation and civil war

	Primary educ. (0/100)	Infant mort. (0/100)	Light/capita (log)
	(1)	(2)	(3)
Time to nat. capital (log)	-2.760*** (0.407)	0.946*** (0.241)	-0.056** (0.026)
Time to reg. capital (log)	-1.302*** (0.318)	0.141 (0.159)	-0.037** (0.018)
Ethnic inclusion (0/1)	0.861 (0.708)	-0.589** (0.231)	0.084*** (0.025)
Ethnic exclusion (0/1)	0.837 (0.702)	-0.348 (0.219)	0.036* (0.020)
Ethnic civil war (0/1)	0.358 (0.786)	-0.043 (0.222)	0.037*** (0.010)
Eth. war since indep. (0/1)	0.613 (0.900)	-0.357 (0.311)	-0.048*** (0.017)
Time since eth. war	0.431*** (0.151)	0.013 (0.037)	-0.001 (0.001)
Time since eth. war ²	-0.014*** (0.004)	0.001 (0.001)	0.00000 (0.00002)
$\beta_1 + \beta_2:$	-4.062*** (0.477)	1.087*** (0.274)	-0.092*** (0.029)
Point FE:	yes	yes	yes
Country-year FE:	yes	yes	yes
Survey FE:	yes	yes	yes
Controls:	yes	yes	-
Mean DV:	70	9.9	-6.5
Observations	1,893,067	2,626,280	1,506,991
Adjusted R ²	0.445	0.051	0.836

Notes: OLS linear models. Control variables for models with primary education as the dependent variable consist of respondents' 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

³⁵All data can be freely downloaded from growup.ethz.ch.

Controlling for capitals: Another potential danger is that as some cities may become, for a variety of reasons, richer with time, and then benefit from a political upgrade and get their own administrative unit. In such cases, changes in state reach in that city would be endogenous to local development. In an additional robustness check I therefore include dummies for whether an interview-location was (1) in and (2) closer than 1 hour to a regional and national capital in time t . For the Voronoi units, I create analogous measures when they either contain a capital or have an average distance to a capital of below 1 hour. The results in Table A13 highlight that changes in distances towards capitals in locations which are not capitals themselves drive the baseline patterns. Above and beyond the effect associated with a reduction in travel times, becoming a regional capital is associated with *lower* education rates, and more mixed patterns in the other outcomes. Relocations of national capitals are associated with reductions of infant mortality rates in the new capitals.

Table A13: Changes in time to national/regional capital and local development: Controlling for capitals

	Primary educ. (0/100)		Infant mort. (0/100)		Light/capita (log)	
	(1)	(2)	(3)	(4)	(5)	(6)
Time to nat. capital (log)	-2.555*** (0.428)	-2.712*** (0.430)	0.793*** (0.238)	0.888*** (0.251)	-0.050* (0.027)	-0.052* (0.027)
Time to reg. capital (log)	-1.798*** (0.364)	-1.871*** (0.372)	0.147 (0.168)	0.178 (0.180)	-0.039** (0.019)	-0.038** (0.019)
National capital (0/1)	-0.364 (3.426)		-5.790** (2.331)		0.200 (0.271)	
Regional capital (0/1)	-3.194*** (1.135)		-0.003 (0.457)		-0.049 (0.053)	
Time to nat. cap. < 1hr		-0.594 (0.631)		-0.217 (0.396)		0.099** (0.044)
Time to reg. cap < 1hr		-1.283*** (0.390)		0.093 (0.180)		-0.028 (0.033)
$\beta_1 + \beta_2$:	-4.353*** (0.501)	-4.584*** (0.512)	0.94*** (0.275)	1.067*** (0.29)	-0.09*** (0.029)	-0.09*** (0.029)
Point FE:	yes	yes	yes	yes	yes	yes
Country-year FE:	yes	yes	yes	yes	yes	yes
Survey FE:	yes	yes	yes	yes	yes	yes
Controls:	yes	yes	yes	yes	-	-
Mean DV:	70	70	9.9	9.9	-6.5	-6.5
Observations	1,893,057	1,893,067	2,634,195	2,634,209	1,506,991	1,506,991
Adjusted R ²	0.445	0.445	0.051	0.051	0.836	0.836

Notes: OLS linear models. Control variables for models with primary education as the dependent variable consist of respondents' 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: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

E.5 Varying education- and health-related outcomes:

Lastly, two sets of additional analyses gauge whether the results are applicable to alternative education and health care outcomes. First, Table A14 shows very similar effects of the travel time to regional and national capitals on (1) whether a respondent has spent any time in school, (2) on her years of schooling – logged and linear, and (3) on a simple secondary education dummy. The main deviation from the baseline model is that the distance to the national capital does not seem to impact secondary education levels. With regard to infant mortality, Table A15 indicates that increased infant mortality in regions of low state reach can indeed be related to a lower availability of professional prenatal assistants. Similarly, changes in travel times to national capitals come with an increased chance that a child is born in a public clinic and positively relate to the receipt of professional assistance during delivery. Conversely, if children are born under low levels of state reach from the national capital, assistance is given more often in traditional manner. Consistent with the earlier results, the distance to regional capitals does not have any significant effect on the receipt of professional prenatal or birth assistance. Regarding local development measured through nightlight emissions, Table A16 shows that the choice of outcomes – whether nightlights per capita (log), absolute nightlights (log), or a dummy for whether a Voronoi cell exhibits any nightlight emissions does not produce different conclusions. In all three cases, changes in travel times to capitals are in sum associated with more nightlights. The estimated effect of changes in travel times towards national capitals are statistically significant, with the exception of their effect on the logged amount of absolute nightlight emissions.

Table A14: Changes in time to national/regional capital and various education outcomes

	Any educ. (0/100) (1)	Educ. years (linear) (2)	Educ. years (log) (3)	Sec. educ. (0/100) (4)
Time to nat. capital (log)	-2.655*** (0.421)	-0.227*** (0.048)	-0.054*** (0.009)	-0.070 (0.441)
Time to reg. capital (log)	-1.301*** (0.327)	-0.075** (0.034)	-0.026*** (0.007)	-0.955*** (0.334)
$\beta_1 + \beta_2$:	-3.956*** (0.487)	-0.303*** (0.054)	-0.081*** (0.011)	-1.025** (0.511)
Point FE:	yes	yes	yes	yes
Country-year FE:	yes	yes	yes	yes
Survey FE:	yes	yes	yes	yes
Controls:	yes	yes	yes	yes
Mean DV:	70	5.4	1.5	35
Observations	1,890,391	1,890,391	1,890,391	1,615,848
Adjusted R ²	0.443	0.459	0.483	0.356

Notes: OLS linear models. Control variables include respondents' age and age squared, as well as a female dummy. Standard errors clustered on the point and country-year levels. Significance codes: *p<0.1; **p<0.05; ***p<0.01

Table A15: Changes in time to national/regional capital and quality of prenatal care and birth assistance (in percent)

	Prof. prenatal care (1)	Birth in public inst. (2)	Prof. birth assist. (3)	Trad. birth assist. (4)
Time to nat. capital (log)	-9.605*** (2.457)	-3.138** (1.284)	-7.310*** (2.024)	4.057** (1.586)
Time to reg. capital (log)	0.268 (1.056)	0.482 (1.084)	0.512 (1.019)	0.921 (0.742)
$\beta_1 + \beta_2$:	-9.338*** (2.661)	-2.656* (1.435)	-6.798*** (2.179)	4.977*** (1.663)
Point FE:	yes	yes	yes	yes
Country-year FE:	yes	yes	yes	yes
Survey FE:	yes	yes	yes	yes
Controls:	yes	yes	yes	yes
Mean DV:	81	21	54	18
Observations	419,768	579,787	578,573	573,685
Adjusted R ²	0.450	0.543	0.444	0.307

Notes: OLS linear models. Control variables include an infant's mother's age and age squared, 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

Table A16: Changes in time to national/regional capital and various nightlight measures

	Light/capita (log) (1)	Light (log) (2)	Any Light (0/100) (3)
Time to nat. capital (log)	-0.052* (0.027)	-0.113 (0.073)	-1.215** (0.601)
Time to reg. capital (log)	-0.037** (0.019)	-0.112* (0.063)	-0.870* (0.510)
$\beta_1 + \beta_2$:	-0.089*** (0.029)	-0.225*** (0.072)	-2.085*** (0.583)
Unit FE:	yes	yes	yes
Country-year FE:	yes	yes	yes
Survey FE:	yes	yes	yes
Controls:	-	-	-
Mean DV:	-6.5	-5.4	12
Observations	1,506,991	1,506,991	1,506,991
Adjusted R ²	0.836	0.838	0.784

Notes: OLS linear models. Standard errors clustered on the point and country-year levels. Significance codes: *p<0.1; **p<0.05; ***p<0.01

E.6 Additional robustness checks:

Country-weights: Because the DHS has not regularly sampled all African countries, the weights each country receives in the baseline specifications vary considerably. If the effects of state capacity on development vary systematically with the number and size of survey the DHS by countries, the results would be biased towards the most-sampled set of countries. Table A17 addresses this caveat by weighting each observation by the inverse of the number of observations from its country (Models 1 and 3) and from the re-

spective cohort in the same country (Models 2 and 4). The latter serves to prevent that the biggest cohorts drive the results at the expense of dynamics in smaller cohorts observed in the data.³⁶ Though coefficients slightly change, the results remain generally consistent with those reported at the baseline. Lastly, Model 5 addresses the potential problem that until now I have treated all Voronoi cells in the same manner, thus giving equal weight to areas with many and few inhabitants. Weighting the Voronoi units by their population roughly doubles the estimated effect of the travel time to regional capitals on nightlight emissions. This is reassuring, since it shows that the results above are not driven by areas with very low population densities that are prone to produce outliers in the per-capita nightlight measure. As in some previous results, the effect of the travel time to national capitals, is statistically insignificant.

Table A17: Changes in time to national/regional capital and local development: Alternative weights

	Primary educ. (0/100)		Infant mort. (0/100)		Light/capita (log)
	(1)	(2)	(3)	(4)	(5)
Time to nat. capital (log)	-1.908*** (0.440)	-2.538*** (0.489)	0.593** (0.249)	0.768* (0.407)	-0.031 (0.040)
Time to reg. capital (log)	-1.508*** (0.322)	-1.609*** (0.339)	0.143 (0.170)	0.194 (0.366)	-0.089** (0.037)
$\beta_1 + \beta_2$:	-3.417*** (0.5)	-4.147*** (0.546)	0.735*** (0.279)	0.962* (0.525)	-0.12** (0.049)
Weights	country	country cohort	country	country cohort	population
Point FE:	yes	yes	yes	yes	yes
Country-year FE:	yes	yes	yes	yes	yes
Survey FE:	yes	yes	yes	yes	yes
Mean DV:	70	70	9.9	9.9	-6.5
Observations	1,893,067	1,893,067	2,634,209	2,634,209	1,506,991
Adjusted R ²	0.427	0.422	0.054	0.155	0.949

Notes: OLS linear models. Control variables for models with primary education as the dependent variable consist of respondents' 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

Varying the size of Voronoi units: This last robustness check assesses the impact of the choice of the size of Voronoi units used in the baseline analysis (400km^2). Figure A23 plots the results of the baseline specification estimated with units with exponential increases in their size from 100 to 6400km^2 . Reassuringly, the Figure shows that small units generally give rise to smaller estimated and standardized effects. The increases in the respective effects that comes with larger units might be due to the unit constant added to the nightlight measure, which has a larger effect in small units in which more units have no observed nightlight emissions. Throughout, decreases in travel times to regional capitals are associated with more nightlights. Decreases in travel times to national capitals

³⁶Because of the sampling scheme of the DHS and attrition-by-death, cohorts close to the date of the survey are bigger than those in the past.

are only associated with more nightlights in large units. This result adds to the uncertainty discussed in the main text on whether travel times to national capitals are indeed related to local nightlight emissions or not.

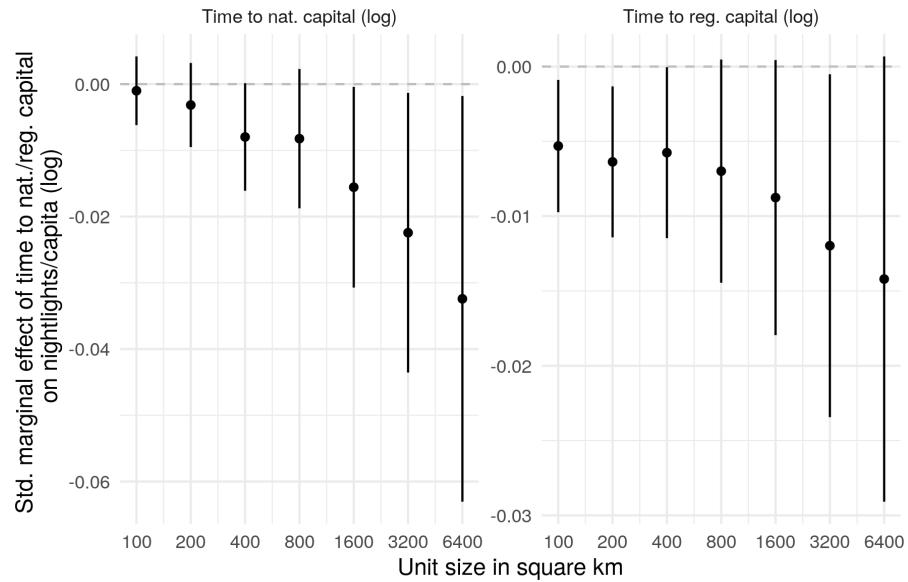


Figure A23: Effect of travel times to national and regional capitals on nightlight emissions in Voronoi units of increasing size.

Estimated coefficients and their 95% CI are standardized by dividing them by the mean dependent variable ($\log(.001 + \text{nightlight p.c.})$).

F References

- Afrobarometer. 2018. "Afrobarometer Data." Available at <http://www.afrobarometer.org>.
- Ben Yishay, Ariel Rotberg, Renee, Jessica Wells, Zhonghui Lv, Seth Goodman, Lidia Kovacevic and Dan Runfola. 2017. *Geocoding Afrobarometer Rounds 1-6: Methodology and Data Quality*. Williamsburg, VA: AidData.
- Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria and Gerard Padró i Miquel. 2015. "The value of democracy: Evidence from road building in Kenya." *American Economic Review* 105(6):1817–1851.
- Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* 108(4-5):899–934.
- Donaldson, Dave and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A "Market Access" Approach." *Quarterly Journal of Economics* 131(2):799–858.
- Eaton, Jonathan and Samuel Kortum. 2002. "Technology, Gravity, and Trade." *Econometrica* 70(5):1741–1779.
- FAO. 2014. "Global Administrative Unit Layers." Available from: <http://data.fao.org/map?entryId=f7e7adb0-88fd-11da-a88f-000d939bc5d8&tab=metadata>.
- Fesler, James. 1949. *Area and Administration*. Birmingham: University of Alabama Press.
- Franck, Raphaël and Ilia Rainer. 2012. "Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education, and Health in Sub-Saharan Africa." *American Political Science Review* 106(2):294–325.
- Goldewijk, Kees Klein, Arthur Beusen and Peter Janssen. 2010. "Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1." *The Holocene* 2010(1):1–9.
- Goodfellow, Ian, Yoshua Bengio and Aaron Courville. 2016. *Deep Learning*. Cambridge, MA: MIT Press.
- Gottlieb, Jessica, Guy Grossman, Horacio Larreguy and Benjamin Marx. 2019. "A Signaling Theory of Distributive Policy Choice: Evidence from Senegal." *Journal of Politics* 81(2):631–647.
- Green, Elliott. 2010. "Patronage, district creation, and reform in Uganda." *Studies in Comparative International Development* 45(1):83–103.
- Grossman, Guy, Jan H. Pierskalla and Emma Boswell Dean. 2017. "Government Fragmentation and Public Goods Provision." *Journal Of Politics* 79(3):823–839.

- Grossman, Guy and Janet I. Lewis. 2014. "Administrative Unit Proliferation." *American Political Science Review* 108(01):196–217.
- Hassan, Mai. 2016. "A State of Change: District Creation in Kenya after the Beginning of Multi-party Elections." *Political Research Quarterly* 69(3):510–521.
- Herbst, Jeffrey. 2000. *States and Power in Africa*. Princeton: Princeton University Press.
- Hodler, Roland and Paul A. Raschky. 2014. "Regional Favoritism." *The Quarterly Journal of Economics* 129(2):995–1033.
- Jedwab, Remi and Adam Storeygard. 2018. "The Average and Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa 1960-2010." *Unpublished Working Paper*.
- Jedwab, Remi and Alexander Moradi. 2016. "The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa." *Review of Economics and Statistics* 98(2):268–284.
- Kingma, Diederik P. and Jimmy Ba. 2015. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980*.
- LeCun, Yann, Yoshua Bengio and Geoffrey Hinton. 2015. "Deep learning." *Nature* 521(7553):436–444.
- Lloyd, Stuart P. 1982. "Least Squares Quantization in PCM." *IEEE Transactions on Information Theory* 28(2):129–137.
- Miguel, Edward. 2004. "Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania." *World Politics* 56(April 2004):328–362.
- Polsby, Daniel D. and Robert Popper. 1991. "The Third Criterion: Compactness as a Procedural Safeguard Against Partisan Gerrymandering." *Yale Law & Policy Review* 9(2):301–353.
- Resnick, Danielle. 2017. "Democracy, decentralization, and district proliferation: The case of Ghana." *Political Geography* 59:47–60.
- Scott, James C. 2017. *Against the Grain. A Deep History of the Earliest States*. New Haven: Yale University Press.
- Shelhamer, Evan, Jonathan Long and Trevor Darrell. 2017. "Fully Convolutional Networks for Semantic Segmentation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39(4):640–651.
- Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp M Hunziker and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and

Conflict: The Ethnic Power Relations Dataset Family." *Journal of Conflict Resolution* 59(7):1327–1342.

Weidmann, Nils and Kristian Skrede Gleditsch. 2010. "Mapping and Measuring Country Shapes: The cshapes Package." *R Journal* 2(1):18–24.

World Bank. 2018. "World Development Indicators." Accessed on 2019/03/20 from <https://databank.worldbank.org/data>.

Zeiler, Matthew D. and Rob Fergus. 2014. Visualizing and Understanding Convolutional Networks. In *Computer Vision – ECCV 2014*, ed. David Fleet, Tomas Pajdla, Bernt Schiele and Tinne Tuytelaars. Cham: Springer pp. 818–833.