Supporting Information
'Roads to Rule, Roads to Rebel: Relational State Capacity and Conflict in Africa'

Carl Müller-Crepon, Philipp Hunziker, and Lars-Erik Cederman, Journal for Conflict Resolution, 2020.

Table of Contents

A1 Digitizing road maps . . . A2
A1.1 The Michelin map corpus . . . A2
A1.2 Map digitization as a computer vision task . . . A2
A1.3 Our FCNN: Architecture and training . . . A4
A1.4 Vectorization and results . . . A5
A2 Constructing and validating road networks . . . A7
A2.1 Network construction . . . A7
A2.2 Validation . . . A7
A3 Summary statistics . . . A9
A3.1 Correlation between RSC measures and control variables . . . A10
A4 Robustness checks . . . A11
A4.1 Alternative dependent variables . . . A11
A4.2 Alternative functional forms . . . A13
A4.3 Measuring RSC on 1990 road networks . . . A14
A4.4 Adding and dropping control variables . . . A14
A4.5 Restricting the sample . . . A15
A4.6 Alternative units of analysis . . . A16
A4.7 Country-level jackknife . . . A16
A4.8 Cross-sectional analysis . . . A17
A4.9 Alternative mechanisms . . . A18
A5 Instrumental variable approach: Simulating road networks . . . A20
A6 Instrumental variable approach: Additional robustness checks . . . A22
A6.1 First stage by size of country and ethnic group . . . A22
A6.2 Assessing bias from contemporary census data in HYDE . . . A26
A6.3 IV results by resolution of simulated rod networks . . . A26
A6.4 Disaggregated relational state capacity . . . A28

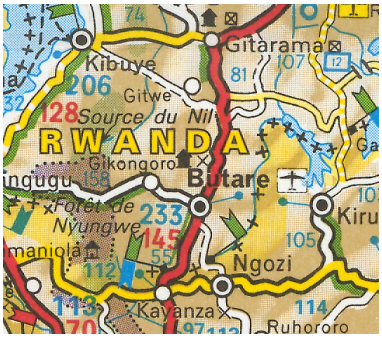


Figure A1: Small excerpt from the 2003 Michelin map for Southern Africa, showing southern Rwanda.

A1 Digitizing road maps

A1.1 The Michelin map corpus

Our source for road network data is the African Michelin map corpus, a collection of large topographical maps at a resolution of 1:4,000,000, showing detailed information on transport infrastructure with a consistent cartographic symbology. While coverage before the 1960s is sporadic, Michelin has covered the entire African continent at intervals of at most 5 years beginning in 1966. This makes the Michelin corpus an unparalleled source for time-variant road-network information. In the present paper, we make use of the 6 maps published in 1966 and 1990.

We digitize this map 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 negligible.

A1.2 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 A1). Therefore, heuristic algorithms that distinguish only colors or lines fail to classify roads correctly. Instead, we implement 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, we borrow 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 outputting a scalar that is passed on to the next layer.[1] 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*.[2] 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 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

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

[2] Here, "depth" refers to the third dimension of an image. An RGB image has depth 3.

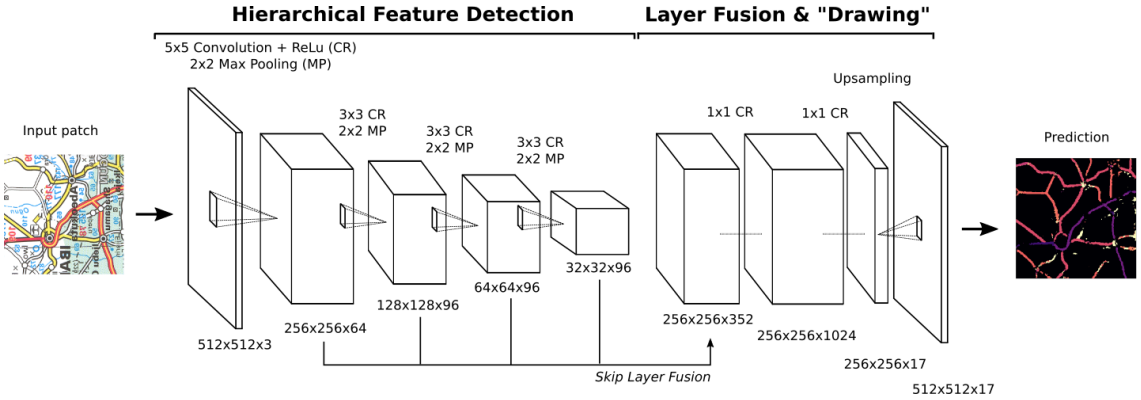


Figure A2: Architecture of our custom fully convolutional neural network.

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.

A1.3 Our FCNN: Architecture and training

To solve the semantic segmentation problem on the Michelin map material we implement a version of Shelhamer et al.'s 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.[3] The precise architecture of our model is shown in Figure A2. The model is described in canonical notation, see Bengio, Goodfellow 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.[4] These are color-images of dimension $512 \times 512 \times 3$ that superficially look like real road maps, but which we create programatically 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. 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.[5]

Interpreting trained artificial neural networks is notoriously difficult, as the learned

[3] An additional reference category identifies background pixels.

[4] For 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.

[5] All 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 (2014) with a batch-size of 2.

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 A3 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.

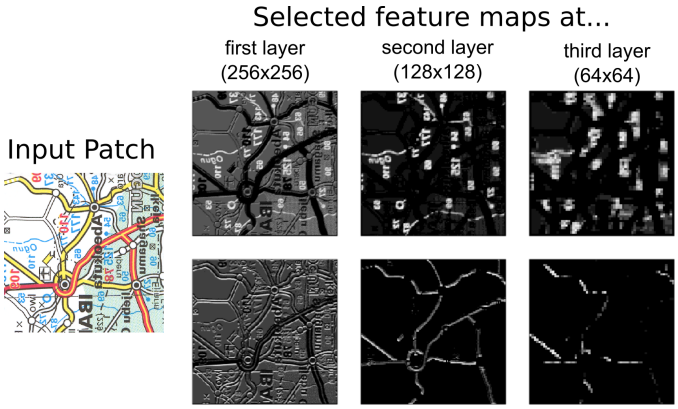


Figure A3: Selected feature maps from the trained FCNN.

Finally, it is instructive to demonstrate the trained FCNNs predictive performance visually. Figure A4 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.

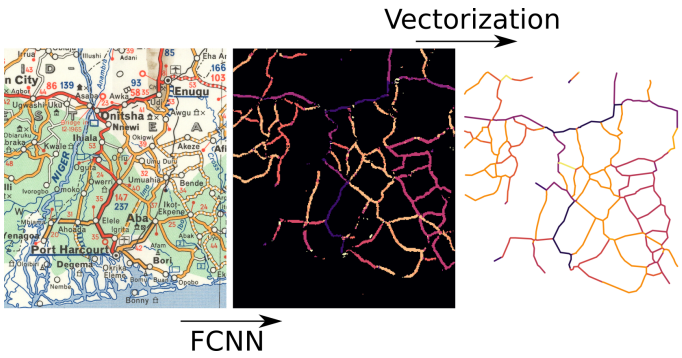


Figure A4: Predictive performance of the digitization procedure.

A1.4 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:

	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.

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

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 A4 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. *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 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.[6]

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 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.

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

A2 Constructing and validating road networks

A2.1 Network construction

We transform the Michelin data into a planar graph that uniformly covers each African country. We do so in a step-wise manner:

1. **Foot-path network:** The basis of our planar graph consist of a network of 8-connected 'foot-paths', shown for the case of Uganda in Figure A5a. The graph's nodes are the centroids of a raster of population estimates from the HYDE 3.1 data (Klein Goldewijk, Beusen and Janssen 2010) for 1960 at a resolution of $.1667 \times .1667$ decimal degrees (or ca. 20 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.

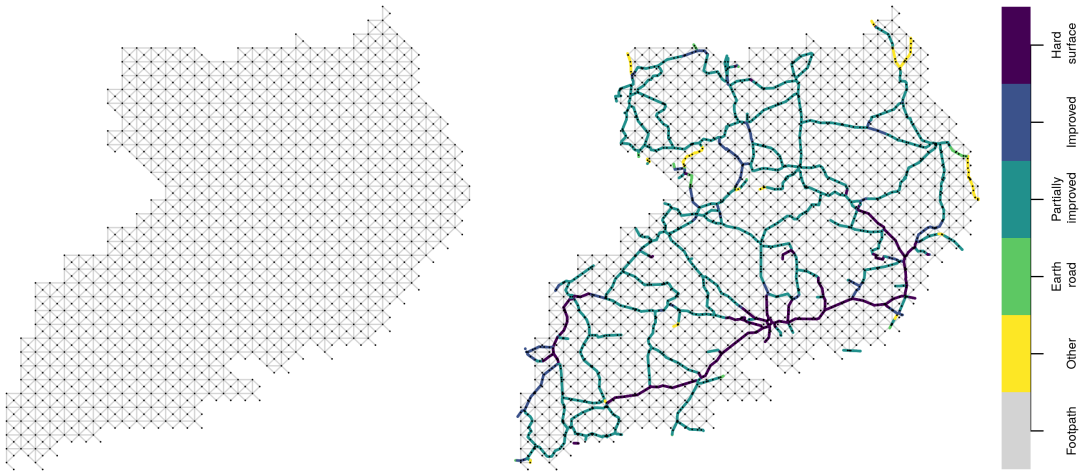
2. **Adding roads:** We overlay the basic foot-path network with the spatial lines extracted from the Michelin maps (see Figure A5b). We create additional nodes wherever two roads or foot-paths cross, thus retaining the planar graph property. These additional nodes' purpose is to serve as intersections. 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.

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. Before assigning these edge weights, we first collapse the 16 road types on the Michelin maps to 6 main categories. We obtain estimated travel speeds for each of these categories 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 A6a).[7] 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 (2016)).

A2.2 Validation

We validate the travel times computed on our Michelin-based network using the Google Maps API. Since Google only offers contemporary data, we base our comparison on road

[7]Note 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.



(a) Foot-path network (b) ... adding Michelin roads observed in 1966

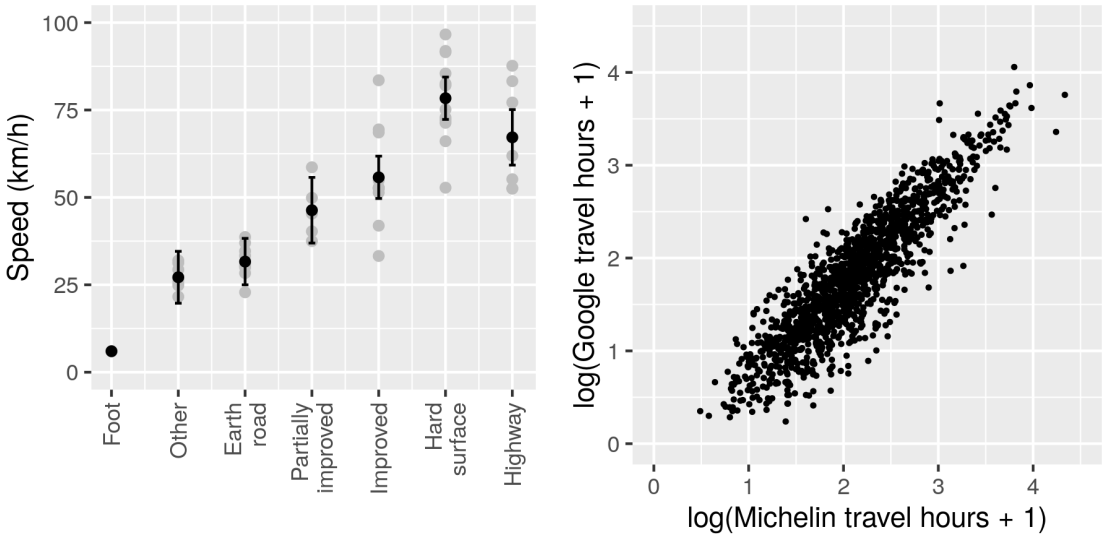
Figure A5: Constructing road networks that regularly cover geographical space. Additional vertices are added to the graph where foot-paths and roads intersect.

networks constructed with Michelin data from 2003, the most contemporary source at our disposal. For each country, we draw 50 source- and 50 destination nodes, each with a probability relative to a node's population size.[8] 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.

We compare the results from the 1,416 successful queries with our Michelin-based computations. Figure A6b 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'.

This comparison does not only highlight the quality of our measurement of travel times. It also sheds light on one of the key shortcomings of Google Maps as an alternative resource for measuring travel times. Since Google Maps does not allow for querying paths between arbitrary coordinates, but makes such searches contingent on the presence of roads, it is impossible to use their services for our purposes.

[8]Fewer if the country in question does not have 50 populated nodes.



(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 A6: Construction and validation of edge-weights.

A3 Summary statistics

Table A2: Summary statistics

Statistic	N	Mean	St. Dev.	Min	Max
Challengers	31,780	0.51	2.41	0	128
Challenger Events	31,780	1.02	13.33	0	739
State Events	31,780	0.99	12.86	0	881
RSC 1966 (log)	42,878	−1.21	0.79	−4.09	0.49
State access 1966; road (log)	42,878	−2.76	0.72	−5.06	−0.06
Internal connect. 1966; road (log)	42,878	−1.55	0.59	−4.01	0.00
State access 1966; foot (log)	42,878	−4.22	0.81	−5.94	−0.23
Internal connect. 1966; foot (log)	42,878	−1.98	0.80	−4.88	0.00
State access 1880; road (sim; log)	42,196	−2.45	0.64	−4.40	−0.11
Internal connect. 1880; road (sim; log)	42,196	−1.35	0.67	−4.04	0.00
Distance to border	42,851	601.64	452.68	3.80	2,427.10
Capital dummy	42,851	3.91	1.21	1.00	9.00
Median altitude	42,851	1,617.84	274.68	1,073.75	2,509.15
Median slope	42,851	1,166.26	554.99	3.24	3,217.60
Evapotranspiration	42,851	4.42	1.53	1.00	8.00
Precipitation	42,851	24.85	2.89	11.32	29.94
Evapotranspiration / Precipitation	42,851	0.36	0.15	0.00	0.80
Temperature	42,851	0.35	0.25	0.00	0.99
Cash crop suitability	42,878	5.83	4.08	0.0001	16.10
Agricultural suitability	42,878	1.74	2.29	0.001	20.07
Distance to coast	42,878	0.12	0.32	0	1
Distance to nav. river	42,878	17.35	64.23	0.11	1,311.14
Mineral deposit	42,878	463.54	1,846.02	0.003	37,581.51
Area (1000 km^2)	42,878	178.03	1,159.41	0.00	30,564.21
Population (1000s)	42,878	103.45	106.13	0.03	590.94
Urban population (1000s)	42,878	0.03	0.16	0	1

A3.1 Correlation between RSC measures and control variables

Table A3: Effect of covariates on RSC and its components

	Dependent variable (1966; logged)		
	RSC	State access	Internal connectedness
	(1)	(2)	(3)
State access 1966, foot (log)	0.683***	0.702***	0.019*
	(0.015)	(0.015)	(0.011)
Internal connectedness 1966, foot (log)	−0.624***	0.014	0.638***
	(0.020)	(0.014)	(0.020)
Population (log)	0.017**	0.058***	0.042***
	(0.007)	(0.006)	(0.007)
Urban population (log)	−0.009***	0.004***	0.013***
	(0.001)	(0.001)	(0.001)
Area (log)	0.028**	−0.056***	−0.084***
	(0.013)	(0.010)	(0.013)
Median altitude	−0.00002	−0.0001**	−0.0001*
	(0.0001)	(0.0001)	(0.0001)
Median slope	0.009	0.021***	0.013**
	(0.006)	(0.006)	(0.006)
Precipitation	−0.00004	−0.0002***	−0.0002***
	(0.0001)	(0.00005)	(0.0001)
Evapotranspiration	−0.00004	−0.00005	−0.00001
	(0.0001)	(0.0001)	(0.0001)
Evapotranspiration / Precipitation	0.010	0.055**	0.045*
	(0.023)	(0.022)	(0.024)
Temperature	−0.007	−0.021*	−0.014
	(0.011)	(0.011)	(0.011)
Cash crop suitability	0.053	0.112***	0.059
	(0.046)	(0.042)	(0.047)
Agricultural suitability	0.066*	0.075**	0.009
	(0.034)	(0.030)	(0.034)
Mineral deposit (0/1)	−0.063***	0.026	0.089***
	(0.019)	(0.018)	(0.020)
Distance to coast	−0.005	−0.007	−0.002
	(0.008)	(0.008)	(0.007)
Distance to nav. river	−0.003	0.001	0.004
	(0.005)	(0.005)	(0.005)
Capital dummy	−0.224***	−0.044	0.180***
	(0.036)	(0.043)	(0.038)
Distance to border (log)	0.012*	0.025***	0.013**
	(0.006)	(0.006)	(0.006)
Country-year FE:	yes	yes	yes
Mean DV	-1.21	-2.76	-1.55
Observations	42,851	42,851	42,851
Adjusted R^2	0.931	0.932	0.879

Notes: OLS models. Two-way clustered standard errors in parentheses (ethnic group and country-year clusters). Significance codes: *p<0.1; **p<0.05; ***p<0.01.

Table A3 presents the results of a set of simple models that regress our main independent variable, RSC, as well as its components, state access and internal connectedness, on the full vector of control variables.

Unsurprisingly, the two 'foot-travel' measures of state access and internal connectedness that capture only geodesic distances are highly correlated with the respective measures

computed as travel times on the road network. Furthermore, we see a range of control variables that are correlated with our measure of RSC and its components. In particular, we see that highly populated and smaller ethnic groups have higher levels of state access and internal connectedness. Related to the colonial origins of road networks as tools for resource extraction, we see that ethnic groups with high levels of cash crop suitability and mineral deposits feature denser local road networks that increase their internal connectedness. Lastly, groups closer to national borders profit from less roads than groups farther away from borders. These correlational patterns coincide with Herbst's (2000) descriptive evidence on road networks and state reach in Africa in general.

A4 Robustness checks

We apply an extensive set of robustness checks to our main analysis. The following pages give an overview over the motivations, implementation, and results of each additional test. To facilitate interpretation, most robustness checks are presented as coefficient plots, in particular and unless otherwise noted, in Figure A7 below. Detailed reports will be available with the replication data.

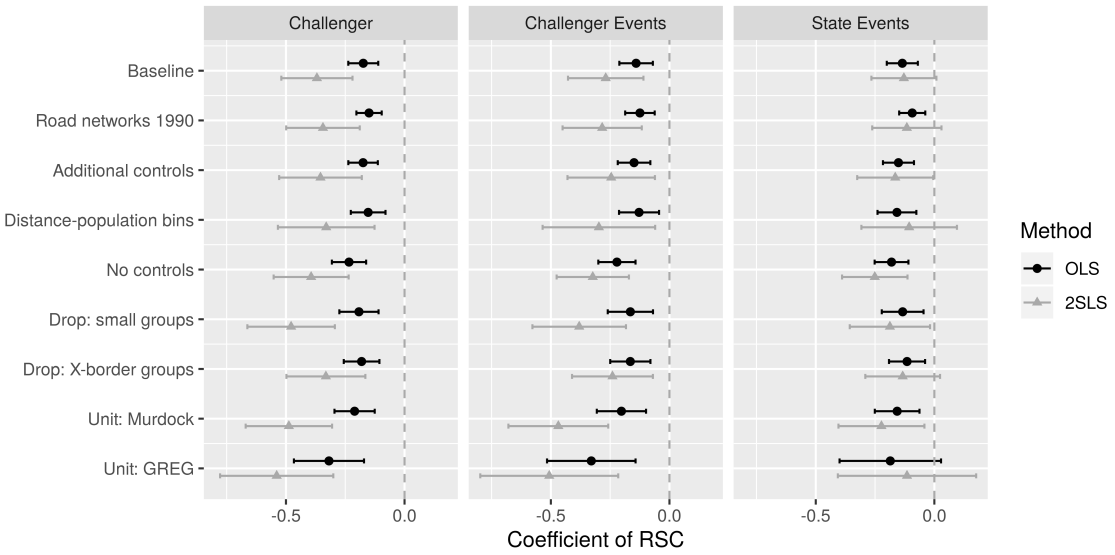


Figure A7: Coefficients of RSC in robustness checks. Bars indicate 95% confidence intervals.

A4.1 Alternative dependent variables

We first gauge whether our baseline effects on the number of and violent events associated with challengers to state power are driven by rebel groups or militias, defined as as ACLED's class of political and ethnic militias. In its three panels, Figure A8 plots the estimated effect of RSC on (1) the number of rebel groups and militias active in an ethnic settlement area (first and second row), (2) the number of battles between either

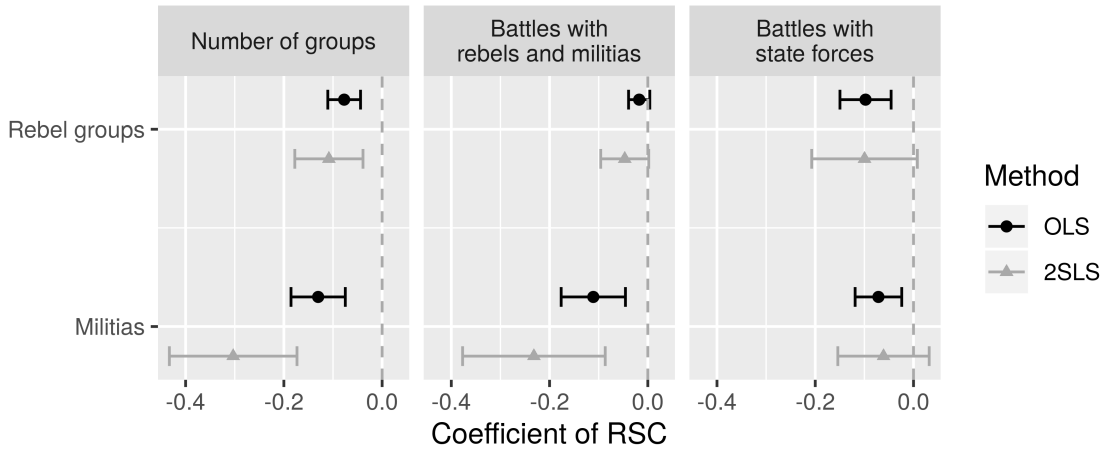


Figure A8: Coefficients of RSC when when disaggregating challengers to local state rule into rebel groups and militias.

rebel groups or militias with both, rebel groups and militias, as well as (3) the number of battles between state fores and rebel groups and militias. The results show that the baseline patterns found for the aggregate composite of 'challengers' are not solely driven by either rebel groups or militias.

We heed the advice of Hegre and Sambanis (2006) and subject our main model to (1) alternative specifications of our outcomes of interest and (2) additional outcomes from various datasets on political violence. Figure A9 plots the results for using (1) linear models of the probability that our main outcomes take a value > 0 and (2) the logged number of fatalities of the event types as the dependent variable. The respective results are consistent with those that are based on pure event counts. In addition, the Figure shows results when conducting the repeating our analyis with alternative outcome data from the (3) ACLED (Raleigh et al. 2010), (4) UCDP GED (Sundberg and Melander 2013), and (5) SCAD datasets (Salehyan et al. 2012). The results show that low levels of relative state capacity are robustly associated with higher event counts across almost all categories of political violence.

Two classes of exceptions exist. First, the results show no evidence of an impact of RSC on remote violence, i.e. aerial bombings, as measured by ACLED. This is not too surprising, since these events are seldom and have little to do with physical accessibility. Also, We do not find effects of RSC on the number of violent incidents committed by pro-government militias taken from the SCAD data. These oftentimes happen in capitals, where RSC is by definition high. The second class of exceptions consist in that, in our instrumental variable approach, we find no evidence of an effect of RSC on (1) fatalities of state-challenger battles and (2) UCDP GED civil war events. Both exceptions relate to the generally more mixed results of the effect of RSC on fighting between challengers and state forces discussed in the respective section of the main text.

Notwithstanding these deviating patterns, the results from analyzing the effect of RSC

on alternative measures of violent events in ethnic groups highlight that challenger-related violence in the periphery of a state also brings along other forms of political violence, most importantly violence against civilians.

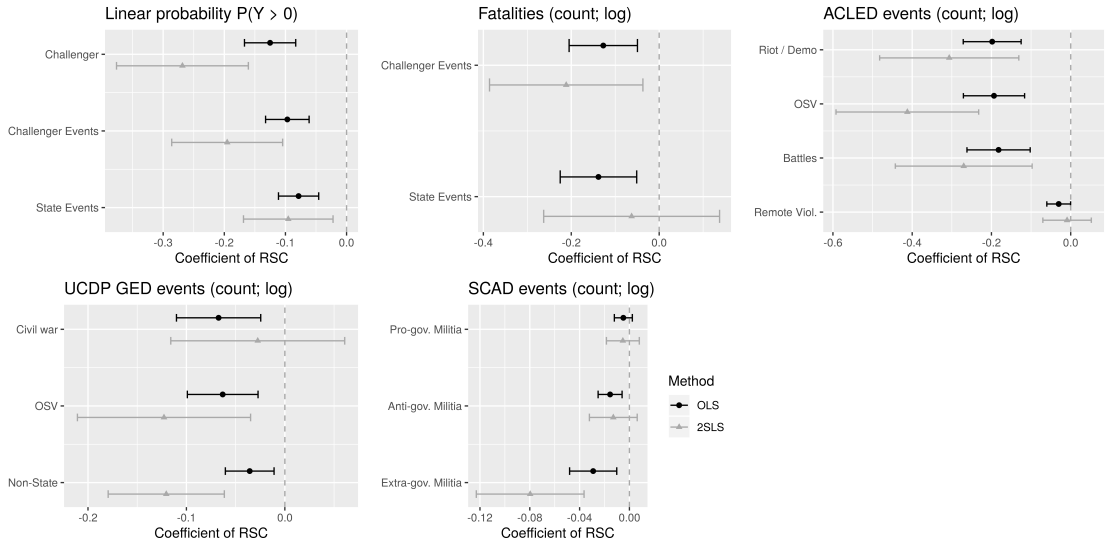


Figure A9: Effect of RSC on alternative outcome variables. Bars indicate 95% confidence intervals.

A4.2 Alternative functional forms

Although linear count models have the advantage of allowing for a very flexible specification of fixed effects, the bias introduced by the miss-specification of the distribution of the dependent variables might drive our results. Figure A10 therefore presents results based on logistic and negative binomial models that use the same vector of explanatory variables as our main model.[9] Due to computational complications in estimating such models with a large number of country-year dummies, we do control for country- and year-fixed effects rather than country-year fixed effects. Because our independent variables are

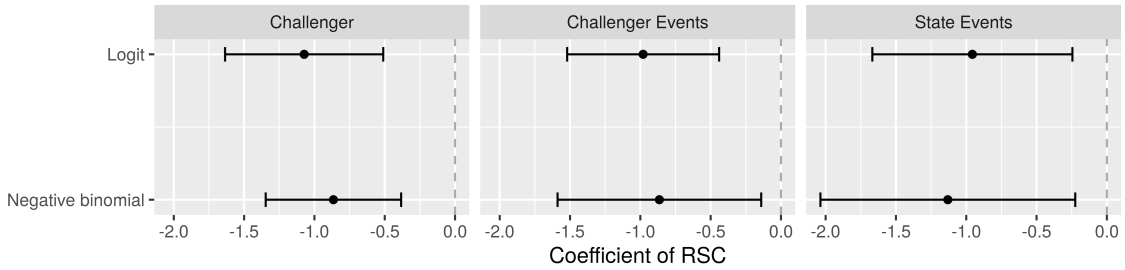


Figure A10: Coefficients of RSC when choosing different functional forms. Bars indicate 95% confidence intervals.

[9]We do not estimate Poisson models here because of the significant overdispersion of the dependent variables.

cross-sectional in nature, this limitation has negligable effects. The results mirror the ones presented in the main paper, indicating that the latter are not due to the choice of linear count models.

A4.3 Measuring RSC on 1990 road networks

In order to gauge whether the reliance on 1966 road network data significantly affects the results, we re-estimate the baseline specifications of the naive OLS and the instrumental variable analysis using data on the Michelin road network observed in 1990. For the IV-specification, this implies that we simulate a new set of road networks for each country using the road budget observed within a country's borders in 1990, and then using the simulation to instrument for RSC as measured on the 1990 road networks. The results of the resulting models are plotted in the second row of Figure A7 and show no significant deviation from the baseline estimates, except for the IV-estimate of the effect of RSC on the number of battles between the state and its challengers which slightly decreases in precision ($p = .11$) but comes with a point estimate that is equivalent to the baseline estimate.

A4.4 Adding and dropping control variables

To account for potential omitted variable bias, we include a set of additional control variables. These are:

- *Precolonial characteristics:* Data from Murdock (Murdock 1959, 1967) to account for precolonial characteristics of ethnic groups that might affect the extent and structure of colonial road building as well as contemporary conflict risk. In particular, we control for precolonial (1) economic practices – ethnic groups' dependence of hunting, fishing, animal husbandry, and agriculture –, (2) intensity of agriculture, and (3) political centralization.

- *Landcover:* Since landcover might affect conflict risk and the extend of transport infrastructure, we control for the percentage of ethnic groups' settlement areas that covered by (1) pasture land that is used for grazing, (2) savanna, (3) tropical woodlands, and (4) tropical forests (FAO 2015). With average values over 10 percent, these are the most prevalent land-cover types in Africa.

- *Characteristics of the area between an ethnic group and the capital:* Lastly, the quality of the road connection between an ethnic group and the capital is likely correlated with characteristics of the area between an ethnic group and the capital. Such characteristics might also conflict risk. For this reason, we compute the geographical area that lies between ethnic groups and their countries' capitals, recompute all baseline controls for these areas,[10] and add these new variables as controls.

[10] I.e, these areas' distance to the coast, their local climate (mean temperature, precipitation, evapo-

The third row of Figure A7 shows that including these three vectors of control variables does not affect the baseline results.

In addition to the effect of potential omitted variables, the results reported above might be driven by non-linear effects of (1) pure geographic distances, in particular $\text{state access}_g^{foot}$ and $\text{internal connectedness}_g^{foot}$, and/or (2) population counts. To control for this caveat, we create yearly bins of ethnic groups that have similar values on these three dimensions. Within each country-year, we divide each of the three variables into bins of approximately 25 observations, which when all combined and interacted with make a total of 10'279 unique bins populated by observed ethnic groups. We then add one fixed effect per distance-population-country-year-bin to our baseline specification. Econometrically, this method is akin to matching observations from the same country-year on these three variables. The results of the re-estimated models are presented in the fourth row of Figure A7. The coefficients for RSC are similar those at the baseline and statistically significant. This further suggests that the negative effect of relational state capacity on conflict risk is not due to country-specific, nonlinear effects of (the combination of) pure geographic distances and population sizes.

Lastly, the fifth row of Figure A7 reports the results of estimating the main specifications without any control variables, except for foot-travel times to capitals and within ethnic groups. The results are not driven by the inclusion of the various controls. If at all, the specification shows that the inclusion of controls decreases the estimated effect of RSC, in particular its effect on the number of battles between state forces and challengers, which is now highly significant, both in the naive OLS and the IV-specification.

A4.5 Restricting the sample

Ethnologue lists (1) many very small ethnic groups, and (2) groups that are not strictly nested in (changing) country-borders. The first caveat may lead to internal connectedness measures that are hardly influenced by any roads. The second caveat produces small 'rump' groups where ethnic settlement patterns are cut by country borders, leaving potentially insignificant parts of an ethnic group on one side of a border. Such ethnic groups might inherently feature different types of conflict risks (Michalopoulos and Papaioannou 2011) which might bias our results.

To address both caveats, we restrict the sample of ethnic groups in two ways. The first is to drop ethnic groups with an area that is below the 25th percentile of the distribution of ethnic groups' areas (1295 km^2). The second robustness check drops ethnic groups that are divided by a state border. The results (Figure A7) of both tests are consistent with our main results. All coefficients are statistically indistinguishable from the baseline.

transpiration, and the ratio of the latter two), and their average altitude and roughness. We also include their contemporaneous (urban) population (both logged), as well as the logged size of their area.

A4.6 Alternative units of analysis

Next, we reevaluate our hypotheses using two alternative data sets of ethnic settlement patterns, namely GREG (Weidmann, Rød and Cederman 2010) and Murdock's (Murdock 1959) ethnic map. GREG is based on the Soviet Atlas Narodov Mira, and Murdock's Atlas is coded on the basis of ethnographic evidence available in the 1950s and has been digitized by Nunn and Wantchekon (2011). Ethnic groups encoded in both data sets are on average bigger than those of the Ethnoglogue data. This might reduce potential bias introduced by very small ethnic groups (see also above, Subsection A4.5). The ensuing results, plotted in Figure A7, are mostly equivalent to the baseline results. As noted in the main text, this is with the exception of the IV-estimate of the effect of RSC on the number of state-challenger battles which does not decrease in magnitude but is estimated to be insignificantly different from zero. The respective naive OLS estimate is statistically significant at $p < .1$.

A4.7 Country-level jackknife

To assess the effect single countries have on our estimates, we implement a country-by-country jackknife approach, where we iteratively delete observations from each country from our sample. The estimates for the coefficient of relational state capacity RSC are plotted in Figure A11. No single country affects the statistical significance of the effect of RSC on the number of challengers and fighting between them. Not surprisingly however, we do see lower associations of RSC with our conflict-related outcomes when we drop countries such as the DR Congo, Ethiopia, Algeria, or Kenya. In particular the first two countries host peripheral ethnic groups which are prime examples of relatively weak relational state capacity leading to violent competition over local power. With regard to the third main outcome, battles between state forces and armed groups, Figure A11 shows that the baseline estimate of the IV specification are mainly driven by patterns of relational state capacity and conflict in the DR Congo. Although the country, in particular its Eastern part, hosts many proto-typical ethnic groups with low levels of relational state capacity and high levels of conflict, its influence on the results casts doubt on its representatives of conflict dynamics elsewhere on the continent. We discuss the implications of this finding in the main text.

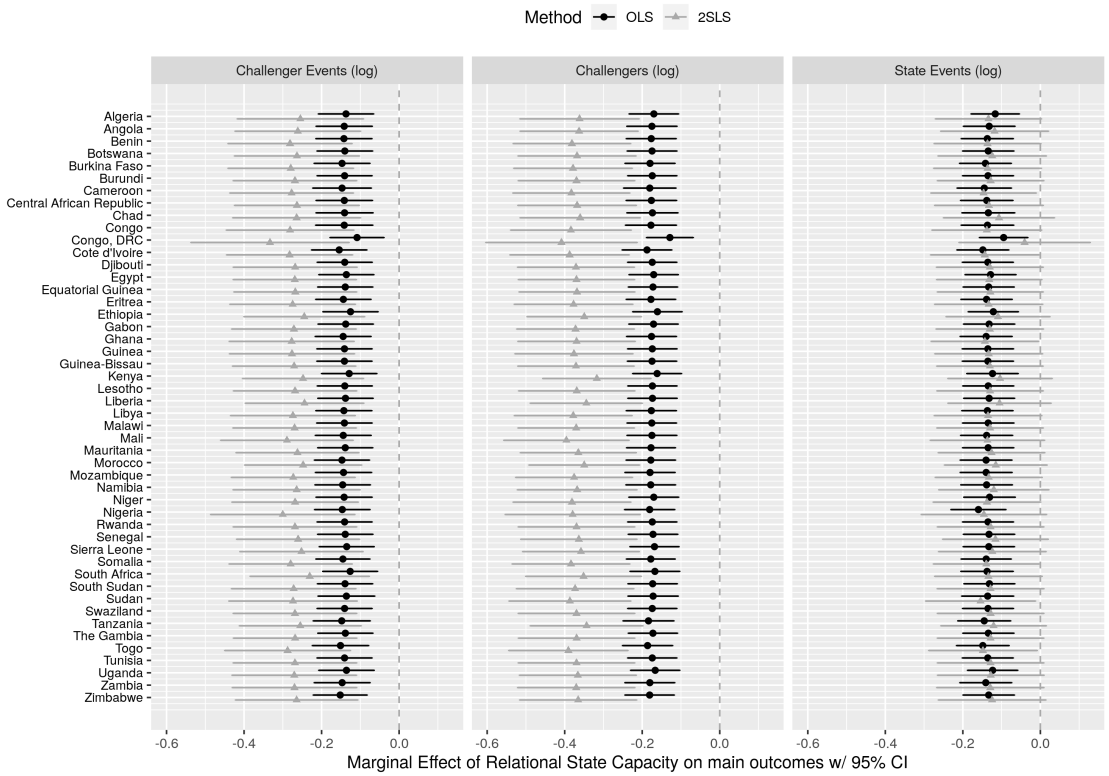


Figure A11: Country-by-country jackknife applied to main models (Table 2, main paper). Bars indicate 95% confidence intervals.

A4.8 Cross-sectional analysis

To reflect border changes in Africa after 1997,[11] we have so far analyzed panel data. Because such border changes might bias our results, we turn towards a cross-sectional analysis. The first cross-section is based on state borders and capitals observed in 1997 (the start of ACLED). The second cross-section chooses countries' borders and capitals locations at the time of their independence.[12] The dependent variables consist of the logged sums of our baseline outcomes between 1997 and 2016. Using the cross-sectional data, we estimate our main models with country-fixed effects, and base data for the population controls on the year 1966 (1990) for the first (second) cross-section.

The results of these analyses are presented in Figure A12. They show that the insights gained from our baseline models hold irrespective of the choice of a panel- or cross-sectional design. Only the IV-estimate of the effect of RSC on violence between the state and challengers turns statistically insignificant, but is associated with the a point-estimate

[11] Mainly the secession of South Sudan in 2011, but, where the SCAD and UCDP GED data is used, also the independence of Eritrea and Namibia.

[12] Dropping ethnic groups from Eritrea, South Sudan, and Namibia from the sample, and replacing them with the groups we would have observed had these countries not become independent from Ethiopia, Sudan, and South Africa.

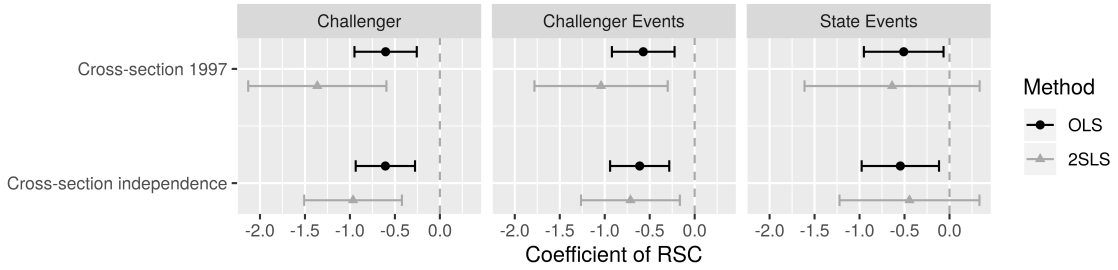


Figure A12: Coefficients of RSC in the cross-sectional analysis. Bars indicate 95% confidence intervals.

that is indistinguishable from the one derived from the respective naive OLS-estimate.[13] In sum, these patterns suggest that it is unlikely that the main results are due to endogenous, postcolonial changes of borders and capital locations.

A4.9 Alternative mechanisms

Our theoretical argument is that *RSC* affects conflict risk through social control. However, our empirical measure of RSC, being based on travel times, might affect conflict risks through different mechanisms. In the following, we test two such alternatives.

First, roads will increase market access and thus foster local development (e.g. Donaldson and Hornbeck 2016), which in turn increases the odds of peace. To control for this alternative causal mechanism, we control for (1) ethnic groups' nightlight emissions as a proxy for local development (Henderson, Storeygard and Weil 2012),[14] as well as (2) the total, quality-weighted road mileage present in an ethnic groups settlement area in 1966.[15] The latter controls for the historical level of economic development. After controlling for the length of the local road network, the RSC-measure picks up effects only due to the structure but not the local length of the road network. It therefore is a conservative test.

Second, ethnic groups that are well connected to the capital might also be well connected to entirety of the country's population. Such physical integration might promote socio-economic integration and increased trans-ethnic interactions, both of which may foster peace. To avoid picking up effects that are due to the general level of connectedness between an ethnic group and its country, we include a measure for the extent of 'external connectedness' of an ethnic group.[16]

Using the additional covariates to control for these two alternative causal mechanisms,

[13]Note however, that the size of coefficients is smaller than at the baseline, once they are standardized by the means of the dependent variables.

[14]We include the logged value of average nightlight emissions in a settlement area and a dummy for whether any light is emitted in a certain group-year. Data on nightlights is available from 1992 to 2015 from National Geophysical Data Center (2014). We substitute the missing 2016 values for data from 2015.

[15]This measure is simply the sum of all road kilometers multiplied by the average speed attainable on them.

[16]Specifically and similar to the computation of the measure *internal.connectedness* (Equation 2, main text), we compute

Table A4: Alternative Mechanisms, OLS: Alternative mechanisms

	Dependent variable (logged)		
	Challengers	Challenger Events	State Events
	(1)	(2)	(3)
RSC 1966 (log)	−0.149***	−0.115***	−0.125***
	(0.032)	(0.036)	(0.033)
State access 1966; foot (log)	0.020	−0.012	−0.016
	(0.027)	(0.033)	(0.028)
Internal connectedness 1966; foot (log)	−0.112***	−0.095***	−0.100***
	(0.029)	(0.032)	(0.029)
External connectedness (log)	0.009	0.026	−0.011
	(0.056)	(0.061)	(0.063)
External connectedness; foot (log)	0.065	0.048	0.121
	(0.068)	(0.079)	(0.081)
Roads (km x quality)	−0.002	−0.004**	−0.002
	(0.002)	(0.002)	(0.002)
Nightlights (log)	0.043***	0.047***	0.025***
	(0.006)	(0.007)	(0.006)
Nightlights >0	−0.176***	−0.208***	−0.087***
	(0.031)	(0.040)	(0.031)
Country-year FE:	yes	yes	yes
Controls:	yes	yes	yes
Mean DV	0.21	0.17	0.15
F-Stat:	24.1	21.42	16.49
Observations	31,740	31,740	31,740
Adjusted R^2	0.408	0.379	0.316

Notes: OLS models. Control variables consist of the total and urban population (log), groups' area (log), the mean annual temperature, precipitation, evaporation, the ratio of precipitation and evaporation, the mean altitude and slope of a group's settlement area, its cash crop and agricultural suitability, a mineral deposit dummy, as well as groups' logged distance to the closest coast, navigable river, and border. Two-way clustered standard errors in parentheses (ethnic group and country-year clusters). Significance codes: *p<0.1; **p<0.05; ***p<0.01.

we re-estimate the baseline models (Table A4).[17] Of the additional variables, only our proxy for local development is consistently associated with challengers to state power and

$$external.connectedness_g = \left(\frac{1}{I_g * K_g} * \sum_{k=0}^{K_g} \sum_{i=0}^{I_g} time_{k,i} \right)^{-1},$$

where $i \in I_g$ denotes the inhabitants of the settlement area of group g that live at a distance of travel time $time_{k,i}$ from their compatriots $k \in K_g$ that live in all other ethnic settlement areas. To differentiate effects of pure geography from the effects of the road network, we again compute the measure $\text{external connectedness}_g^{foot}$ on the foot-path network and include it as a covariate.

[17] We do not implement this robustness check for the IV-analysis because the contemporary data (population and nightlights) would introduce post-treatment bias.

conflict. In particular, local nightlight emissions have a negative association with conflict at the extensive, and a positive one at the intensive margin. External connectedness has no discernible effect on conflict patterns. Importantly however, the estimate of the effect of relational state capacity remains almost unchanged from the baseline.

A5 Instrumental variable approach: Simulating road networks

To simulate realistic, yet simplified road networks, we assume that the road-builder aims to minimize the following objective function:

$$LOSS = \frac{1}{I^2} * \sum_{j=0}^{I} \sum_{i=0}^{I} time_{j,i}, \tag{1}$$

where $i, j \in I$ are the inhabitants of the territory on which roads are built. They are separated by a distance of travel time $time_{i,j}$. The road builder seeks to minimize the average travel time between any two inhabitants of the territory.[18]

To build an instrument as clean as possible, we would need data on population distributions that are unaffected by modern transport infrastructure. Unfortunately, the first comprehensive censuses on the continent were taken during colonial rule, after the introduction of roads and railroads. Furthermore, more spatially disaggregated population data only exists for the very contemporary period since 1990. Confronting this dearth of spatially disaggregated population data needed to simulate the structure of road networks, we turn to *estimates* of the African population distribution in 1880 from the HYDE 3.1 data (Klein Goldewijk, Beusen and Janssen 2010). This estimate is based on ingredients at varying levels of analysis: broad, macro level population estimates of the evolution of the African population and urbanization rates by (backprojected) country (e.g., Maddison 2001), subnational census data with temporal coverage varying by country, and geographic data on soil productivity, distance to water, and the landcover and population distribution in the year 2000. Described in more detail in Klein Goldewijk, Beusen and Janssen (2010), the estimation procedure leverages and combines these data into historical population estimate at a resolution of .0833 decimal degrees (ca. 10km). We aggregate these data for most countries in our sample (see Table A5). While constituting the most comprehensive and detailed population data for our purpose, the use of contemporary data (partly subnational census data, contemporary geographic land cover and population estimates) can bias our results. We discuss such biases and our respective empirical strategy in the main text.

[18]Note that Burgess et al. (2015) add a distance penalty to each pair of citizens to the loss function for their simulation of 'optimal' road investment. In our IV-setup, this approach would introduce an additional parameter (distance) and further complicate the control for omitted variables.

Road building is constrained by the road budget:

$$B_q = \sum_{k=q}^{Q} length_k^{observed}, \tag{2}$$

which consists of different qualities $q \in Q$ of roads, each of which corresponds to the observed Michelin road types (see A2). For each road type, the road-builder receives the road mileage of this type of road and all superior types of roads observed in the Michelin map of 1966.

Roads are 'built' on a pre-determined network of footpaths.[19] Given computational constraints in the repeated computation of the loss function (Eq. 1), we adjust the resolution of our baseline network to countries' size (see Table A5). As detailed by the following description of our road-building algorithm, the road-builder builds one type of road after the other, by upgrading existing roads sequentially. The budget constraint ensures that the total road mileage per type of road on the simulated network corresponds to the one in the observed network.

Table A5: Mean of input values to road network simulation

Resolution	Countries	Example	Population	Vertices	B_1	B_2	B_3	B_4	B_5
0.083	11	Togo	352,680.9	450	1.5	1.5	1.0	0.8	0.2
0.167	33	Nigeria	3,306,608.0	1,547.2	11.6	11.6	8.2	7.4	2.9
0.25	22	DR Congo	3,827,680.0	1,984.9	21.5	21.5	15.9	14.6	7.7

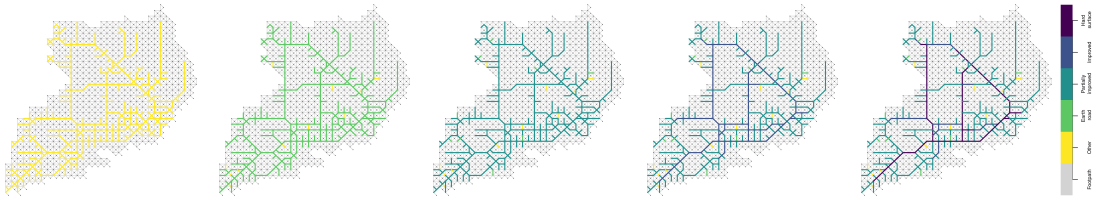
Networks' resolution is measured in decimal degrees and road budgets B_q in 1000 kilometers.

Algorithm:

1. Round; $q = 1$
   (a) Draw 10 seed edges, upgrade to q
   (b) Select neighboring edges $q_e < q$ of all edges with quality $q_e = q$, evaluate and keep 10 most promising edges as E_p.
   (c) Upgrade edge $e \in E_p$ that minimizes $LOSS$. Select neighboring edges $q_e < q$ of e and add to E_p. Update $B_q = B_q - length_e$.
   (d) Repeat step (c), and, in every 10th round, step (b), until budget B_q is spent.

2.-5. Round; $q \in [2, 3, 4, 5]$
   (a) Select all edges with quality $q_e = q - 1$, evaluate and keep 10 most promising edges as E_p.

[19] These footpath-networks are of the same kind as the one used to transport the Michelin maps into networks. In particular, the networks consist of (1) vertices distributed in a grid-like manner in space, and (2) 8-connected 'foot-path'-edges.



(a) $q = 1$ (b) $q = 2$ (c) $q = 3$ (d) $q = 4$ (e) $q = 5$

Figure A13: Simulation of the Ugandan road network in 1966, using population data from 1880.

(b) Upgrade edge $e \in E_p$ that minimizes $LOSS$. Select neighboring edges $q_e = q-1$ of e and add to E_p.

(c) Repeat step (b), and, in every 10th round, step (a), until budget B_q is spent.

(d) Proceed to the next higher road quality and start again at (a).

The algorithm generates realistic road networks (Figure A13 and A14) for all country-periods in Africa since independence. To speed up computation, we make use of a 90 CPU high performance computing cluster. While small countries can be simulated in minutes, bigger countries require up to 48 hours of run time, simply because each time we recalculate the loss value we have to compute up to 3415^2 shortest paths.[20]

Over subsequent iterations of edge-by-edge road building, the within-country connectedness increases and the LOSS improves (Figure A15a), but with decreasing marginal returns due to substantive scale effects of roads. Figure A15b compares the LOSS values achieved by the simulated networks and observed networks. The simulated networks are consistently better in connecting countries' populations.

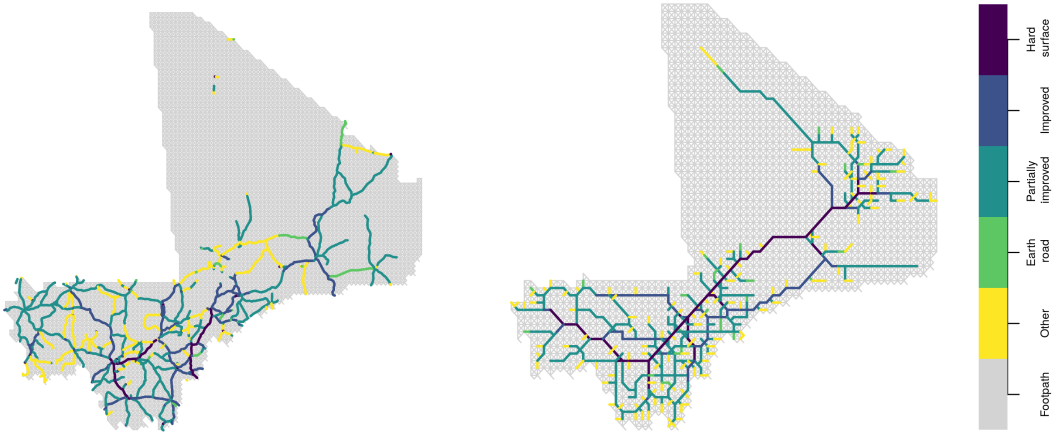
A6 Instrumental variable approach: Additional robustness checks

In addition to the sensitivity analyses presented above, this section discusses robustness checks tailored specifically to the IV-approach.

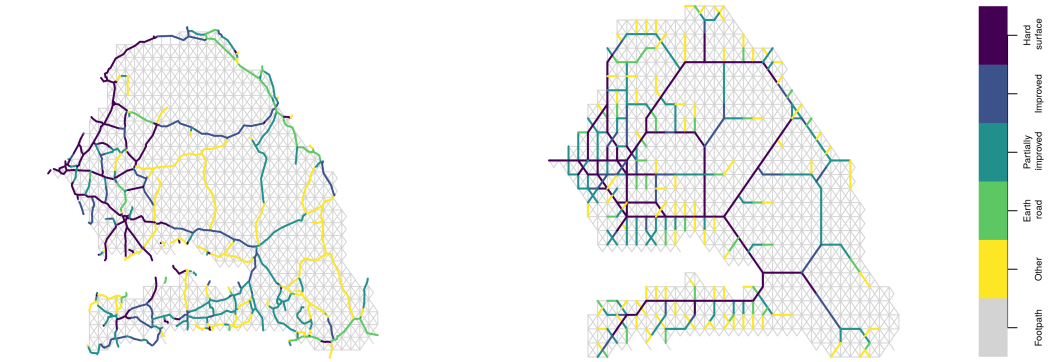
A6.1 First stage by size of country and ethnic group

Because we vary the resolution of the road simulations depending on countries' size (see Table A5), our first stage results are driven by intermediate and large countries. In contrast, small countries do not feature enough variation in our instruments. Splitting up the sample along the three levels of resolution used for simulating road networks and estimating the first stage separately illustrates this pattern (Table A6). Our two instruments possess no significant explanatory power in the 11 very small countries (think of Burundi

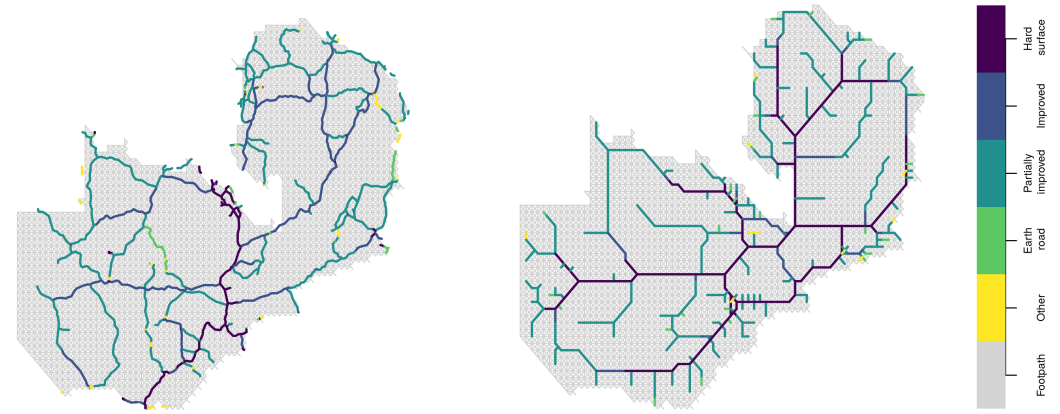
[20] We make use of a path updating algorithm with $\mathcal{O}(|V^2|)$ efficiency.



(a) Mali: Observed and... (b) ... simulated network

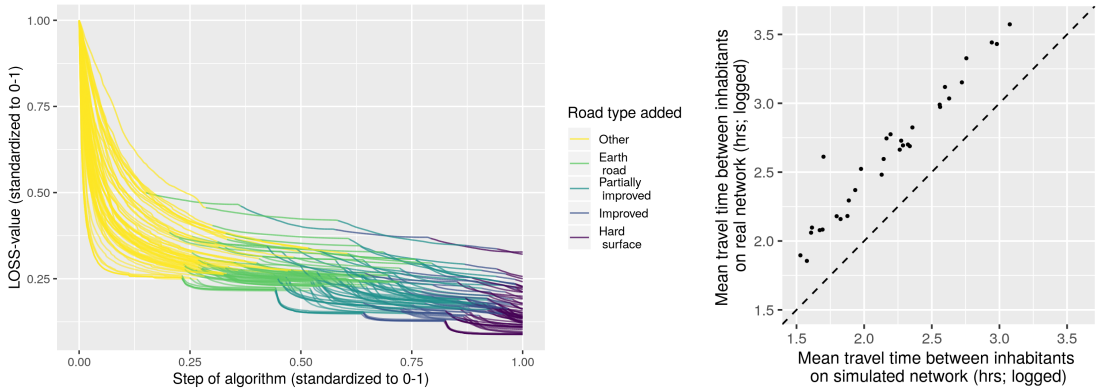


(c) Senegal: Observed and... (d) ... simulated network



(e) Zambia: Observed and... (f) ... simulated network

Figure A14: Observed and simulated road networks in Mali, Senegal, and Zambia, 1966.



(a) Improvement of loss during application of road-building algorithm

(b) Comparison of loss on observed and optimized networks (both with population data for 1960).

Figure A15

Table A6: First stage estimation across resolutions of optimized networks

	Dependent variable: RSC 1990 (log)		
	(1)	(2)	(3)
State access 1880 (sim; log)	−0.039	0.479***	0.914***
	(0.365)	(0.070)	(0.109)
Internal connectedness 1880 (sim; log)	−0.241	−0.276***	−0.258***
	(0.179)	(0.040)	(0.047)
Resolution (dec. degrees):	0.083	0.167	0.25
Country-year FE:	yes	yes	yes
Controls:	yes	yes	yes
Mean DV	-0.68	-1	-1.38
F-Stat:	84.23	329.3	375.02
Observations	1,740	19,045	10,495
Adjusted R^2	0.919	0.897	0.902

Notes: OLS models. Control variables consist of the total and urban population (log), groups' area (log), the mean annual temperature, precipitation, evaporation, the ratio of precipitation and evaporation, the mean altitude and slope of a group's settlement area, its cash crop and agricultural suitability, a mineral deposit dummy, as well as groups' logged distance to the closest coast, navigable river, and border. Two-way clustered standard errors in parentheses (ethnic group and country-year clusters). Significance codes: *p<0.1; **p<0.05; ***p<0.01.

Table A7: First stage estimation across geographic size of ethnic groups

	Dependent variable: RSC 1990 (log)			
	(1)	(2)	(3)	(4)
State access 1880 (sim; log)	0.539***	0.530***	0.771***	0.925***
	(0.107)	(0.097)	(0.116)	(0.121)
Internal connectedness 1880 (sim; log)	0.019	−0.186***	−0.463***	−0.367***
	(0.072)	(0.051)	(0.066)	(0.074)
Group area quartile:	1	2	3	4
Country-year FE:	yes	yes	yes	yes
Controls:	yes	yes	yes	yes
Mean DV	-1.58	-1.29	-0.98	-0.59
F-Stat:	48.82	79.05	98.48	95.11
Observations	7,485	7,910	7,950	7,935
Adjusted R^2	0.814	0.885	0.901	0.916

Notes: OLS models. Control variables consist of the total and urban population (log), groups' area (log), the mean annual temperature, precipitation, evaporation, the ratio of precipitation and evaporation, the mean altitude and slope of a group's settlement area, its cash crop and agricultural suitability, a mineral deposit dummy, as well as groups' logged distance to the closest coast, navigable river, and border. Two-way clustered standard errors in parentheses (ethnic group and country-year clusters). Significance codes: *p<0.1; **p<0.05; ***p<0.01.

or Lesotho) in our sample. However, in the remaining bulk of countries, our instruments behave as expected.

In a similar vein, it might be argued that only very large ethnic groups drive the first stage. We therefore split the sample along the quartiles of the total geographic area of ethnic groups and re-estimate the first stage regression. Table A7 shows that the first stage results are not driven only by large ethnic groups. Indeed, although the coefficient size of our first instrument, $\mathsf{state\ access}^{sim}$ increases with group size, its precision does not. For the smallest quartile of ethnic groups, where we expect least variation in the internal connectedness, our second instrument, $\mathsf{internal\ connectedness}^{sim}$ does not have explanatory leverage over RSC. However, it does have such power for the three remaining quartiles. These results show that our first stage estimation is not driven by the biggest ethnic groups although, as one would expect, the first stage is strongest there.

A finial finding worth discussing is that across country and group sizes, the coefficient for the state access instrument is larger than that of the internal connectedness instrument. This difference is due to the manner in which we construct our road graph, in which the observed road networks are superimposed on the 8-connected 'foot-path' network. This superposition implies that to travel on the road network, each traveler has to first 'walk' the next road network vertex. This walk inflates within-group travel times proportionally more than travel towards capitals. Because this extra 'walk' is not necessary on the simulated networks, the coefficient for the internal connectedness instrument is pushed towards zero.

A6.2 Assessing bias from contemporary census data in HYDE

As discussed in the main text, one important caveat of the HYDE population data used for simulating road networks is that it is projected back on the basis of (post)colonial census data. The subnational census counts used in the HYDE projections comprise all the data for all censuses available to Klein Goldewijk, Beusen and Janssen (2010). These were mostly retrieved from the PopulStat database (Lahmeyer 2004).[21] The temporal coverage of the subnational census data used by Klein Goldewijk, Beusen and Janssen (2010) varies for each country, ranging from, for example, the mid 19th century for South Africa, to 1921 for Tanzania, 1956 for Sierra Leone, 1971 for Botswana and as late as 1996 for Burkina Faso.[22] For slightly less than half of our sample (42%), the HYDE estimates are based on subnational population data collected before 1960, before we observe the first Michelin Map.

Absent any precolonial population data on the African continent to compare our results against, one strategy to judge whether the use of the backprojected data introduces bias into the analysis follows the assumption that the HYDE data that is based on earlier subnational census counts is less biased that the data that is only based on contemporaneous counts. Making this assumption, we can compare results from cases with and without data available for early years with our baseline estimate. Figure A16 plots the results of this exercise, splitting our sample into cases with and without subnational census data from before 1960 (the median value) used by HYDE. The results are remarkable similar across the two subsamples, with the "early" sample exhibiting, if at all, stronger results. This is evidence against the suspicion that our results are caused by bias produced by HYDE reliance on contemporary census data.

A6.3 IV results by resolution of simulated rod networks

A second test of reverse causality introduced be the HDY population estimates gauges the extent of bias arising from the geographic data on population counts and landcover from the year 2000 that are "baked" into the HYDE data. The HYDE estimation process uses these data along with the other geographic data on land productivity and distance to water bodies to derive the fine-grained spatial distribution of higher-level population estimates derived from the subnational census data. Because the HYDE data may thereby pick up information on (contemporary) road networks and other post-treatment outcomes, our road network simulation may be biased. These effects are likely quite localized and nested within administrative regions, because the overall HYDE population count of regions depends on the (extrapolated) census data discussed above. The related reverse causality

[21] Note that this database is no longer openly available, but can be accessed through the internet archive: http://web.archive.org/web/20170712102100/http://www.populstat.info/.

[22] These years are listed in Appendix A of Klein Goldewijk et al. (2017) who update HYDE to version 3.2, using the same PopulStat data. Because this update post-dates the simulation of road networks, we have used HYDE version 3.1. Deviations of the cited years from the earliest PopulStat data are explained by divergent administrative unit boundaries that cannot be harmonized over time.

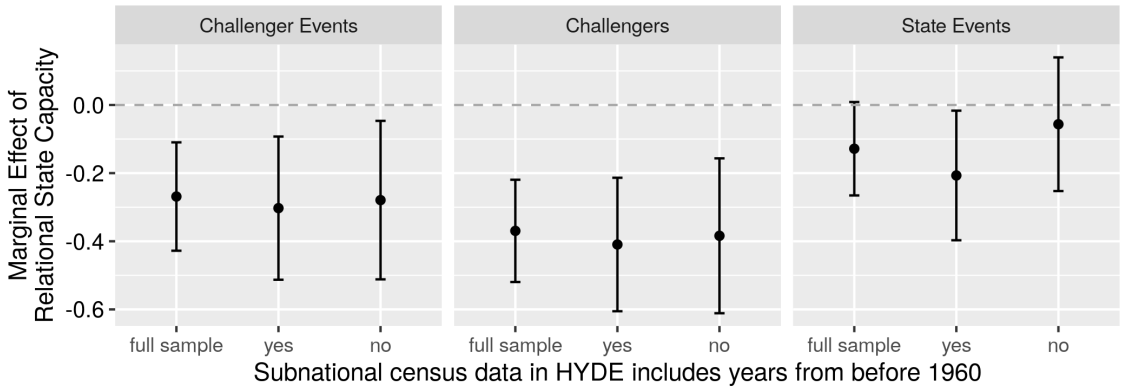


Figure A16: Distinguishing cases with and without early population data used by Klein Goldewijk, Beusen and Janssen (2010) to estimate the HYDE data.

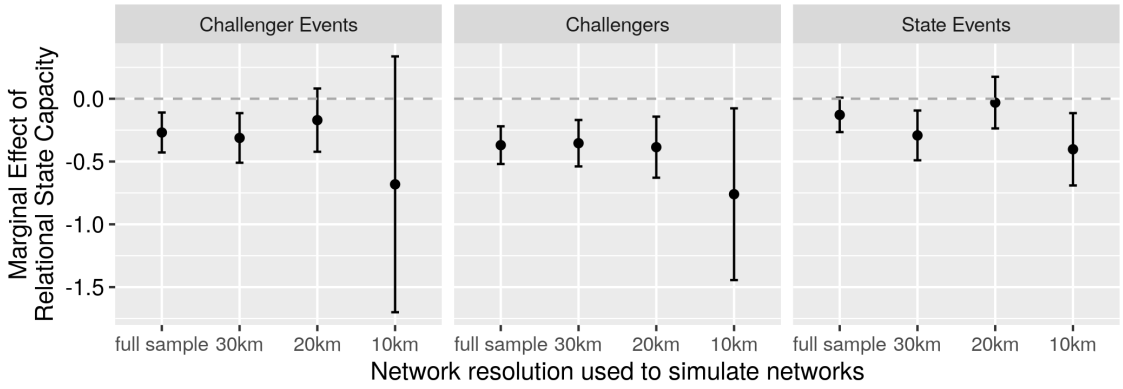


Figure A17: IV results by resolution of network used for simualting road networks.

bias should therefore disappear as we aggregate the HYDE population data to coarser raster data, thereby dissipating spatial detail. While the original HYDE data comes as a raster with a 10km resolution, we simulated roads on networks with a 20 and 30km resolution for larger countries (see Table A5 above) to avoid simulation costs that rise exponentially with networks' size. We use this variation between countries to examine whether our results are driven by road simulations based on the most detailed population data. These are the cases where we would expect most bias.

To do so, we split the sample into the three levels of spatial resolution used to simulate road networks. The results in Figure A17 show that the results in the sample with the coarsest network resolution (30km, 22 countries) almost exactly reflects the baseline results for the first two outcomes and exhibits stronger patterns for the number of violent events between challengers and state forces. The results for the sample with networks simulated at a resolution of 20km (33 countries) are similar for the first two outcomes and null for the last outcome. The sample of a very detailed resolution comprises only cases from 11 relatively small countries. Notably, the first stage within this sample amount to only 2.7, statistically well below the threshold value of at least 7 (Table A6, page A24). This explains the noisy and relatively large estimates in Figure A17. Note however, that these

estimates are not significantly different from the baseline results or those among the sets of observations with road networks simulated at a coarser resolution. We interpret the finding that the results are similar among observations for which we simulate road networks at different spatial resolutions as suggestive evidence that potential reverse causality bias introduced through HYDE's reliance on spatial data from 2000 does not cause the patterns we observe in the data.

A6.4 Disaggregated relational state capacity

To test whether it is indeed the combination of ethnic groups' internal connectedness and state access, coined RSC, that drives the results, we conduct a robustness check where we instrument for both constituents of RSC separately. Table A8 shows that both instruments are valid for the respective endogenous variable. The instrumented measures of state access and internal connectedness are consistently related to the number of challengers and violence among them. In these two models, we can also not reject the null that the coefficients of state access and internal connectedness are of equal absolute size. This supports the use of our aggregated measure of RSC. Mirroring the results of the reduced form estimates reported in the main text in Table 3, the results show that the number of battles between state forces and armed group decreases with state access to ethnic groups ($p = .057$), but does not increase with their internal connectedness. The coefficient of internal connectedness is half the size of that of state access, and associated with a large standard error. Please refer to the main text for further discussion of the implications of this pattern.

Table A8: Effect of the components of RSC, 2SLS

	Dependent variable (logged)				
	Stage 1		Stage 2		
	State access	Internal connectedness	Challengers	Challenger events	State
	(1)	(2)	(3)	(4)	(5)
State access 1880 (sim; log)	0.665***	0.020			
	(0.051)	(0.060)			
Int. connect. 1880 (sim; log)	−0.077***	0.185***			
	(0.024)	(0.033)			
β_3: State access 1990; (log)			−0.348***	−0.246***	−0.137*
			(0.084)	(0.088)	(0.072)
β_4: Int. connect. 1990; (log)			0.517***	0.428***	0.065
			(0.154)	(0.159)	(0.143)
State access 1880; foot (log)	0.273***	0.012	0.151***	0.074	0.007
	(0.032)	(0.038)	(0.056)	(0.061)	(0.050)
Int. connect. 1880; foot (log)	0.053***	0.262***	−0.171***	−0.132**	−0.012
	(0.017)	(0.031)	(0.064)	(0.066)	(0.058)
$\beta_3 + \beta_4$			0.17	0.18	-0.07
			(0.16)	(0.16)	(0.14)
Country-year FE:	yes	yes	yes	yes	yes
Controls:	yes	yes	yes	yes	yes
Mean DV	-2.6	-1.49	0.21	0.17	0.15
F-Stat:	445.97	120.14	20.56	19	15.46
F-Stat Stage 1 (state):			185.93	185.93	185.93
F-Stat Stage 1 (internal):			39.81	39.81	39.81
Observations	31,280	31,280	31,280	31,280	31,280
Adjusted R^2	0.932	0.786	0.352	0.343	0.307

Notes: 2SLS-IV models. Control variables consist of the total population in 1880 (log), groups' area (log), the mean annual temperature, precipitation, evaporation, the ratio of precipitation and evaporation, and the mean altitude and slope of a group's settlement area, its cash crop and agricultural suitability, a mineral deposit dummy, as well as groups' logged distance to the closest coast, navigable river, and border. Two-way clustered standard errors in parentheses (ethnic group and country-year clusters). Significance codes: *p<0.1; **p<0.05; ***p<0.01.

References

Bengio, Yoshua, Ian J Goodfellow and Aaron Courville. 2016. *Deep Learning*. MIT Press.

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 and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A "Market Access" Approach." *Quarterly Journal of Economics* 131(2):799–858.

FAO. 2015. *Global Agro-Ecological Zones: Crop Suitability Index. [Computer File]*. Rome: Food and Agricultural Organization (FAO). Retrieved June 1, 2017 (http://gaez.fao.org/Main.html).

Hegre, Håvard and Nicholas Sambanis. 2006. "Sensitivity analysis of empirical results on civil war onset." *Journal of conflict resolution* 50(4):508–535.

Henderson, J .Vernon, Adam Storeygard and David N Weil. 2012. "Measuring Economic Growth From Outer Space." *American Economic Review* 102(2):994–1028.

Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton: Princeton University Press.

Jedwab, Remi and Adam Storeygard. 2016. "The Heterogeneous Effects of Transportation Investments: Evidence from sub-Saharan Africa 1960-2010." *Unpublished Working Paper. Retrieved February 28, 2018 (https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=NEUDC2015&paper_id=503)* .

Kingma, Diederik and Jimmy Ba. 2014. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* .

Klein Goldewijk, Kees, Arthur Beusen, Jonathan Doelman and Elke Stehfest. 2017. "New anthropogenic land use estimates for the Holocene: HYDE 3.2." *Earth System Science Data* 9(2):927–953.

Klein Goldewijk, Kees, 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.

Lahmeyer, J. 2004. "Populstat database. Growth of the population per country in a historical perspective, including their administrative divisions and principal towns." *Available at: http://web.archive.org/web/20170712102100/http://www.populstat.info/* .

LeCun, Yann, Yoshua Bengio and Geoffrey Hinton. 2015. "Deep learning." *Nature* 521(7553):436–444.

Maddison, Angus. 2001. *The world economy: A millennial perspective*. Paris: OECD.

Michalopoulos, Stelios and Elias Papaioannou. 2011. "The Long-Run Effects of the Scramble for Africa." *American Economic Review* 106(7):1802–1848.

Murdock, George Peter. 1959. *Africa. Its Peoples and Their Culture History.* New York: McGraw-Hill Book Company.

Murdock, George Peter. 1967. *Ethnographic Atlas.* Pittsburgh: University of Pittsburgh 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.* .

Nunn, Nathan and Leonard Wantchekon. 2011. "The slave trade and the origins of Mistrust in Africa." *American Economic Review* 101(7):3221–3252.

Raleigh, Clionadh, Andrew Linke, Havard Hegre and Joakim Karlsen. 2010. "Introducing ACLED: An Armed Conflict Location and Event Dataset." *Journal of Peace Research* 47(5):651–660.

Salehyan, Idean, Cullen S. Hendrix, Jesse Hamner, Christina Case, Christopher Linebarger, Emily Stull and Jennifer Williams. 2012. "Social Conflict in Africa: A New Database." *International Interactions* 38(4):503–511.

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.

Sundberg, Ralph and Erik Melander. 2013. "Introducing the UCDP georeferenced event dataset." *Journal of Peace Research* 50(4):523–532.

Weidmann, Nils B., Jan Ketil Rød and Lars-Erik Cederman. 2010. "Representing Ethnic Groups in Space: A new dataset." *Journal of Peace Research* 47(4):491–499.

Zeiler, Matthew D and Rob Fergus. 2014. Visualizing and understanding convolutional networks. In *European Conference on Computer Vision.* Springer pp. 818–833.