The Persistent Effect of the African Slave Trades on Development: Difference-in-difference evidence

PRELIMINARY AND INCOMPLETE: COMMENTS WELCOME

PLEASE DO NOT CITE OR CIRCULATE WITHOUT AUTHOR PERMISSION *

Grady Killeen

University of California, Berkeley

April 5, 2021

Abstract

A small economics literature examines the role of the slave trades in shaping differences in African productivity. Extant studies often use national productivity data and rely on instrumental variable approaches that instrument participation in the slave trade with geographic variables to identify the causal effect of slavery. This paper builds on the existing literature by examining long-run productivity differences at the level of historic ethnic boundaries using nightlights data. I estimate the effect of the slave trade using a difference-in-difference approach in which I compare areas in which the slave trade occurred to those where it did not and productivity among persistent ethnic groups to the productivity of areas in which the dominant ethnic group has changed. I find a large and negative long-run effect of the slave trade: point estimates indicate that the slave trade decreased long-run GDP per capita by about 33 log points ($p < .01$) and GDP by over 20 log points ($p < .1$). These results indicate that social capital may significantly contribute to economic development and suggest that shocks to social cohesion can be extremely persistent.

*gkilleen@berkeley.edu

1 Introduction

Despite a recent increase in growth rates in Africa, the continent remains economically underdeveloped relative to other regions of the world. Africa suffers from high rates of poverty – a significant majority of the world's 25 poorest countries were located in Africa in 2020 – and substantial inequality both across and within nations (United Nations Development Programme, 2017). A growing body of evidence indicates that the economic history of Africa, particularly the slave trades and European colonization, may help explain the continent's economic struggles. This paper contributes to a recent literature that aims to quantify the long-run effects of the slave trades on African development.

Nunn (2008) was the first quantitative study to examine the economic effect of the slave trades on Africa. Nunn compiles data from shipping records and combines it with a variety of sources documenting the ethnicity of slaves sold into the trans-Atlantic, Indian, Red Sea, and trans-Saharan trades. This data is then used to estimate the number of slaves exported from each (modern day) country. Cross-sectional OLS estimates show that contemporary economic output is negative correlated with the number of slaves that originated in a country. Examination of historical evidence indicates that areas that selected into the slave trade were wealthier, suggesting that this relationship is causal. Nunn further estimates an instrumental variable model in which the number of slaves captured from a country is instrumented with distance to areas in which demand for slaves was high. The instrumental variable results are consistent with the interpretation that historical participation in the slave trades reduced long-run productivity.

Nunn and Puga (2012) presents further evidence that the slave trades had a negative impact on African development. Nunn and Puga construct measurements of terrain ruggedness and find that output today is higher in areas of Africa that are more difficult to navigate. In contrast, ruggedness is negatively related to productivity in other regions of the globe. Nunn and Puga present evidence that rugged areas had better long-run economic outcomes because they were less susceptible to slave raids, providing further evidence that the slave trades depressed economic development in Africa.

Several studies have also attempted to isolate the channels through which the slave trades affected development. Nunn and Wantchekon (2011) merges slave ethnicity data from the trans-Atlantic and Indian slave trades compiled in Nunn (2008) to individual-level survey data from the 2005 Afrobarometer survey. Nunn and Wantchekon find that trust is lower among ethnic groups heavily exposed to slavery, indicating that mistrust resulting from the slave trade may impede economic development. Whatley and Gillezeau (2011) examine the relationship between ethnic stratification and the slave trades using a tribal map from Murdock (1959) which estimates the boundaries of ethnic groups prior to colonization. Whatley and Gillezeau find that ethnic groups are more dense in areas where more people were enslaved. An instrumental variable approach similar to that employed in Nunn (2008) suggests that the slave trades made ethnic stratification more likely.

This paper builds on the existing literature on the African slave trades by employing a novel approach to measuring the long-run effect of the trades on economic development. Nunn (2008) and Nunn and Puga (2012) leverage variation

in geography in an effort to identify the causal effect of the slave trades. In contrast, I exploit the evolution of ethno-linguistic groups from pre-colonial to contemporary times to estimate the effect. We would expect the effect of the slave trade to be the largest in areas in which an ethno-linguistic group that historically governed a region remains the predominant ethno-linguistic group in modern times. I leverage this variation by estimating a difference-in-difference model in which I compare the difference in productivity between persistent and new ethno-linguistic groups in regions that participated in the slave trade to the difference in regions that did not participate in the slave trade.

I find that participation in the slave trades had a negative effect on long-run development, measured using nighttime luminosity data, that is both statistically and economically significant. Point estimates indicate that the slave trades reduced GDP by about 20 log points ($p < .1$) and GDP/capita by over 32 log points ($p < .01$). These results confirm previous results that slavery negatively affected long-term economic development in Africa. Furthermore, the findings suggest that the international slave trades, which ended over 100 years ago, remain significant contributors to African poverty and inequality.

There are four reasons why this contribution to the slave trade literature is valuable. First, the estimates presented in this paper are robust to potential identification issues associated with the use of geographical variables to estimate the effect of slavery. Nunn (2008) reports first-stage F statistics below 5, well below the Staiger-Stock rule of thumb cutoff of 10 for weak instruments (Staiger and Stock, 1997). Hence, these geographic variables may not be adequately strong instruments to produce unbiased estimates of the effect of the slave trades. In addition, the set of instruments presented in Nunn (2008) may affect development through other channels. For instance, countries closer to shorelines may have developed more quickly in the absence of slavery because of easier access to international trade routes. This paper is able to qualitatively reproduce the results of other papers using a distinct identification strategy, suggesting that prior results were not driven by identification issues.

Second, this study improves out understanding of the causal channels through which the legacy of the slave trades affects development. Since this paper leverages variation in ethnic persistence to estimate the persistent effect of the slaves trades, the effects were transmitted through cultural channels. Hence, the findings of this paper support the argument of Nunn and Wantchekon that mistrust is important to the slave trades' long-run impact, and this study quantifies the economic costs of the degradation of social capital caused by slavery.

Third, the estimates in this paper may more accurately capture the effect of the slave trade than estimates that examine cross-country income data. Using nightlights data, I examine productivity at the level of the ethno-linguistic boundaries that existed prior to European colonialism. These boundaries reflect the political divisions that governed Africa during the slave trades, whereas the colonial boundaries that dictate modern national divisions were imposed with little consideration to existing divisions. National boundaries may also include both ethno-linguistic groups from which slaves were exported and those that did not participate in the slave trade.

Fourth, the difference-in-difference approach employed in this paper is less susceptible to data quality issues. Slave

trade data is used only to determine whether an ethnicity participated in the slave trade, and so results are not sensitive to estimates of the number of slaves taken from a particular area. And output is measured using satellite luminosity data that is consistently measured across the units of observation, so measurement error is orthogonal to variables such as institution strength that may be affected by slavery.

This paper also contributes to the literature on social capital. Studies such as Putnam (2000) and Temple and Johnson (1998) suggest that social capital may be important to growth, and a dense literature demonstrates that social capital affects variables that are likely to affect growth such as financial development (Guiso et al., 2004; Karlan, 2005) and political institutions (Satyanath et al., 2017). But despite an abundance of research illustrating the importance of social capital, relatively few studies have quantified the causal impact of social capital or trust on output. One exception is Algan and Cahuc (2010) which concludes that trust is important to cross-country income differences. This paper extends the evidence that social capital is important to economic development by estimating the long-run effect of a negative shock to social cohesion. The results provide evidence that social capital is an important determinant of regional income differences in developing economies. Moreover, this paper suggests that shocks to social capital can be extremely persistent.

2   Historical context

Slavery has a long history in Africa, and the existence of markets for exporting slaves predates the establishment of the trans-Atlantic slave trade. For instance, it is estimated that about 5 million slaves were exported through the trans-Saharan trade between 650 and 1600, and another 1.6 million were exported through the Red Sea trade. Although commercial slave trades existed, historical evidence suggests that early slave trades were driven by politics. Many slaves were captured in wars and raids, and over time it became common to enslave prisoners (Lovejoy, 2011).

The expansion of the trans-Atlantic trade caused a structural change in African slavery. The trade was established by the Portuguese in the 16th century. Beginning in around 1600, massive demand for slaves to labor on plantations in the Americas dramatically increased the volume of slaves exported from Africa and the price of slaves. Lovejoy (2011) estimates that about 1 million slaves were exported from Africa in the 16th century, close to 3 million were exported in the 17th century, and almost 8 million were exported in the 18th century. The expansion of slavery was driven almost entirely by the trans-Atlantic trade: about 75% of the 11,659,000 slaves exported from Africa between 1500-1800 were sold into the trans-Atlantic trade, and by the 18th century the trans-Atlantic trade accounted for over 80% of exports.

Historical evidence indicates that – unlike previous slave trades primarily driven by politics which created a supply of slaves – the growth of the trans-Atlantic slave trade was primarily motivated by demand. Lovejoy (2011) reproduces estimates from Bean (1975) that the real price of slaves sold in West Africa increased by about 4.8 times from 1663 to

1775. The supply of slaves is estimated to have been relatively elastic with respect to changes in price, supporting the interpretation that the growth of the trans-Atlantic trade transformed slavery into an institution driven by commercial interests. Bean (1975) estimates the price elasticity of supply was between 0.75 and 0.83, and LeVeen (1977) calculates a similar estimate of 0.81.

The expansion of the slave trades had significant effects on African political development. Lovejoy (2011) contends that the increased extraction of slaves prevented the establishment of large and unified states as groups raided each other for slaves, eroding ties and thus making it more difficult to unite groups into larger political entities. Whatley and Gillezeau (2011) finds strong econometric evidence in support of this hypothesis.

The export of slaves from Africa fell sharply during the 19th century as abolition movements outlawed participation in the international slave trade. Countries began outlawing the import of slaves at the end of the 18th century, and by 1807 both the United States and Britain had passed laws outlawing the trans-Atlantic slave trade. In 1808, the British navy imposed a blockade aimed at eliminating the international trade Falola and Warnock (2007). The illegal trade of slaves persisted for some time, but by the end of the 19th century the international slave trade was largely eliminated. After the abolition of the international slave trade, the institution persisted in Africa as slaves were used to produce goods such as palm oil. However, domestic slavery was at odds with abolitionist pressures from European powers. As a result, domestic slavery gradually receded at the of the 19th and early 20th century as Europe colonized the continent (Lovejoy, 2011).

European colonialism also affected the distribution of ethnic groups. National boundaries imposed by colonial powers were constructed with little regard for existing institutions or ethnic boundaries. As a result, ethnic groups were frequently partitioned by national boundaries. Michalopoulos and Papaioannou (2016) show that partitioned ethnic groups are more prone to ethnic violence and have worse economic outcomes. Blanton et al. (2001) similarly present evidence that characteristics of the power that colonized an area affected ethnic fragmentation. Robinson (2014) shows that, despite the tension between colonial and ethnic boundaries, colonialism did in fact lead to national over ethnic identification. Hence, colonialism represented a significant shock to existing ethnic identities, and this shock was plausibly exogenous to the dynamics of the slave trades.

3 Data

The unit of observation for this paper is given by pre-colonial ethnic boundaries. As a result, geospatial data sources are used so that statistics can be calculated based on non-standard boundaries. All calculations were performed using Python or the Google Earth Engine Python API. All tables were constructed in Python, and figures were constructed using Python and ArcGIS Pro.

This section describes the data source and processing of key variables. A detailed list of each data source used in this

paper is available is Appendix Table 1.

3.1 Ethnic boundaries

Pre-colonial ethnic boundaries are from Murdock (1959). The data was digitized by Nunn (2008), and the version used in this paper was downloaded from the replication files for Nunn and Wantchekon (2011). Murdock details the boundaries of ethno-linguistic groups in Africa in the late 19th century prior to the onset of European colonial rule. Hence, the data reflect ethnic boundaries at roughly the time that international slave trades ended, and near the end of slavery as an institution. The shapefile downloaded contains entries on 843 entries. However, eight regions in the data are identified as historically uninhabited, so they are omitted from analysis, leaving 835 observations. These 835 polygons are the primary unit of observation for this study, and variables are calculated at the level of these boundaries.

I also examine contemporary ethno-linguistic boundary data. I use data from Felix and Meur (2001) which was digitized by the Harvard Center for Geographic Analysis' AfricaMap project. Of the 835 ethno-linguistic groups identified in the data from Murdock (1959), 736 (88%) have a close match in the Felix and Meur (2001) data, defined as a token set ratio similarity score above 0.8 between the Murdock name and either the primary of variant name in Felix and Meur, and 423 observations (51%) perfectly match between the two data sets. Hence, there is strong consistency between the two data sources, and differences between the data thus seem to be reflective of evolving ethnic boundaries and not differences in cartography.

I construct an indicator variable called "Persistent" that takes on a value of 1 if the Murdock ethnic group is still the dominant ethnic group in the geographic area in the Felix and Meur (2001) data, defined as a close or perfect match. In particular, for each ethnic boundary in the pre-colonial data, I calculate the percent overlap with each of the contemporary ethnic boundaries. If the contemporary ethnic group with the highest percent overlap is the same as that which historically occupied the land, I code the variable to 1. If the group differs, I code "Persistent" to 0.

The variable "Persistent" should be interpreted as measuring highly persistent as opposed to low persistence ethnic groups. That is, area in which "Persistent" is coded to 0 likely did not cease to exist in the area. Rather, the prominence of the ethno-linguistic group is likely lower than areas in which "Persistent" is equal to 1.

1 plots the pre-colonial and contemporary ethnic boundaries over a map of Africa. 2 plots pre-colonial boundaries where shaded polygons indicate a persistent ethnic group and solid polygons indicate that a new ethnic group now occupies the area.

3.2 Slave trade data

Data on the slave trades was initially compiled and digitized in Nunn (2008), and the version used in this paper was obtained from the replication files for Nunn and Wantchekon (2011). Nunn and Wantchekon matched data on slave exports from the trans-Atlantic and Indian slave trades to the ethno-linguistic groups identified in Murdock (1959). Data from the trans-Saharan and Red Sea trades is not included because adequate ethnicity data is not available. The omission of trans-Saharan and Red Sea data is unlikely to affect results given that the trans-Atlantic trade was much larger than the other three trades combined. In addition, Nunn and Wantchekon (2011) show that the estimates used in this paper are consistent with historical sources describing from where slaves were taken.

The way that the data is used in this paper also reduces the likelihood that the omission of trans-Saharan and Red Sea data affects estimates. In regression estimating the effect of the slave trade, an indicator variable called "Slave trade" is used as opposed to the raw estimates. This variable is coded to 1 if any slaves from the ethno-linguistic group were sold into the trans-Atlantic or Indian slave trades and is otherwise equal to zero. Given the prevalence of the trans-Atlantic slave trade, it is unlikely that an ethno-linguistic group had significant involvement in international slave trades but no documented participation in the trans-Atlantic slave trade. Hence, the indicator for participation in the slave trade is likely an accurate measurement of which ethnic groups had significant involvement in the slave trade even if the data is imperfect.

1 presents a cross-tabulation of "Persistent" against known participation in the slave trades. Of the 835 ethnic groups, 299 are known participants in the slave trades. About half (153) of these ethnicities were persistent. Of the 536 ethnic groups from which no slaves were sold, 203 persisted. 2 plots pre-colonial ethnic groups in which green polygons indicate areas that did not participate in the slave trades and red polygons indicate areas from which slaves were sold into the international slave trades.

3.3 Long-run economic development

The primary outcome examined in this paper is long-run economic development. Development is measured using night light data VIIRS Stray Light Corrected Nighttime Day/Night Band Composites processed on the Google Earth Engine. Productivity is measured using luminosity data because the units of observation are non-standard, so official estimates of GDP are not available. In addition, luminosity data is consistently measured across geographies, so errors are orthogonal to the effectiveness of government institutions which may be correlated with participation in the slave trade. A large number of studies such as Henderson et al. (2012) and Chen and Nordhaus (2011) demonstrate that luminosity data is an effective proxy for GDP.

Night light data is obtained from the VIIRS mission because the satellite constellation is better able to detect low levels of light than the DMSP-OLS mission that predated it. This is important for measuring productivity in Africa

where poor areas often have little to no detectable light. I begin with monthly composites that have been processed to remove stray light, and then take the median pixel value across each satellite image from a year. Due to processing (e.g. moonlight corrections), areas with no stable light can sometimes have negative pixel values. These pixels are coded to zero, indicating there is no stable light, to avoid issues when logarithms are taken. I use night lights data from 2014-2018 to limit the influence of stray lights and fluctuations in GDP. For each year and each observation in Murdock (1959), the sum of night lights values contained in each polygon is calculated.

The log of GDP is estimated from the log of 1 plus night lights readings. To estimate the log of GDP per capita, the night lights readings are normalized by population, and then the log is taken after adding 1. Population data comes from LandScan gridded population data produced by Oak Ridge National Laboratory.

3 plots pre-colonial ethnic boundaries over night lights data from 2014.

3.4   Control variables

Several regressions include controls for the number of diamond and gold deposits contained within the ethnic group's boundary, an indicator for whether the area includes known oil deposits, temperature suitability for malaria (P. vivax and P. falciparum), and annual rainfall. The source of each of these data sources is described in Appendix Table 1. The number of neighbors – in the pre-colonial boundaries – is also calculated from the Murdock (1959) data. In analysis examining factors affecting the persistence of ethnicities, indicators are created to capture whether each European colonial power had a presence in the area, and the total number of powers which occupied an area.[1]

4   Empirical methodology

I estimate the average effect of the slave trades on long-run economic development using the difference-in-difference specification

$$Y_{it} = \alpha_t + \beta_0 ST_i + \beta_1 PERS_i + \gamma ST_i \times PERS_i + \lambda' X_i + \epsilon_{it} \tag{1}$$

where Y_{it} is the log of GDP (or GDP/capita) measured in pre-colonial ethnic boundary i in year t, α_t are year fixed-effects, ST_i is an indicator denoting participation in the slave trade, $PERS_i$ is an indicator denoting ethnic group persistence, and X_i is a vector of control variables. The coefficient γ is the difference-in-difference estimate of the long-run effect of slavery on economic development.

Productivity likely exhibits limited spatial dependence since historical ethnic groups often fall within the same national boundaries and unobserved variables that affect output may vary geographically. OLS produces consistent coefficient

---

[1]This variable is designed to capture whether an area was partitioned by national boundaries, and so "Independent" is treated as a colonial power in these calculations.

estimates in the presence of limited spatial dependence, but standard errors must account for the lack of independence. Hence, I report spatial HAC standard errors that are robust to spatial correlation across ethnic boundaries and serial correlation from repeated nightlight measurements in each entity (Conley, 1999, 2016). In particular, let S_{it} denote an observation and define

$$
K_N\left(S_{it}, S_{jk}\right) = \begin{cases} 0 & i \neq j \wedge t \neq k \\[2mm] 1 & i = j \wedge t \neq k \\[2mm] 0 & d(S_{it}, S_{jk}) > 1,500 \wedge t = k \\[2mm] 1 & d(S_{it}, S_{jk}) \leq 1,500 \wedge t = k \end{cases} \tag{2}
$$

where $d(S_{it}, S_{jk})$ denotes the distance, in kilometers, between the centroids of the two observations. Under the assumption that observations further than $1,500$ km from each other are not spatially correlated, Conley (1999) and Conley (2016) show that the asymptotic distribution of $\hat{\beta}$, where β is the population vector of regression coefficients and $\mathbf{x_{it}}$ is a vector of covariates, is given by

$$
\hat{\beta} \stackrel{A}{\sim} \mathcal{N}\left(\beta, \frac{1}{TN}\left(\frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{x_{it}x_{it}}'\right)^{-1}\hat{V}_N\left(\frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{x_{it}x_{it}}'\right)^{-1}\right) \tag{3}
$$

We estimate

$$
\hat{V}_N = \frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{k=1}^{T}K_N\left(S_{it}, S_{jk}\right)\hat{u_{it}}\hat{u_{jk}}\mathbf{x_{it}x_{jk}}' \tag{4}
$$

where $\hat{u_{it}}$ are the OLS residuals. Note that these standard errors allow for arbitrary serial correlation within entities across time, and if we used a distance cutoff near 0 km instead of $1,500$ km we would be estimating standard errors clustered by historical ethnic group.

The identification strategy is based on the intuition that the long-run effects of slavery should be larger in areas where ethnic groups from which slaves were sold persisted into the present than in areas where these groups were supplanted by new ethnicities. This is because, based on historical and economic history evidence, many of the channels through which slavery may affect long-run development operate through cultural channels, such as trust and institution strength. The primary identifying assumption in this paper is that absent slavery, the difference in long-run productivity between areas where the pre-colonial ethnic group persisted versus those where it did not persist would not be larger in areas that participated in the slave trade relative to those that did not. If this assumption is satisfied, then γ offers an unbiased estimate of the lower bound of the effect of slavery on long-run development.

I interpret γ as a lower bound on the effect of slavery on development because the persistence of ethnic groups may be endogenous to participation in the slave trade, but an endogenous process would likely lead to understating the negative effect of the slave trades.

Based on historical evidence, it is likely that a large source of variation in ethnic group persistence is plausibly exogenous to slavery, e.g. due to European colonialism. This is particularly true since the boundaries used in this paper capture the distribution of ethnic groups at the end of the slave trades, so the effect of slavery on ethnic fragmentation demonstrated in Whatley and Gillezeau (2011) were already realized. However, the areas most badly affected by slavery may be less likely to persist if social ties and institutions are weaker. But these areas are also likely would have had low productivity if the ethnicity had persisted, and so output in persistent ethnic groups relative to non-persistent groups would be higher absent the attrition of these ethnicities. If this is the case, γ is larger than the true value, and so the difference-in-difference specification underestimates the negative effects of slavery. Hence, this bias does not pose a threat to this paper's conclusion that slavery had a large and negative effect on development, but is rather a reason that the estimates should be viewed as lower bounds.

The potential endogeneity of ethnic persistence with respect to participation in the slave trade is investigated at the end of Section 5, and results are consistent with the interpretation of γ as a lower-bound. However, it is not possible to definitively rule out bias due to a relationship between persistence and the slave trade.

5 Results

5.1 Summary statistics

I begin by reporting summary statistics of key variables, along with t-tests for equality between areas that did and did not participate in the slave trade and were and were not persistent, in 2. Column (1) reports the mean and standard deviation across the full sample, column (2) across observations that did not participate in the slave trade, and column (4) across non-persistent ethnicities. Column (3) reports a difference in means and standard error of the difference between areas that did participate in the slave trade relative to those that did not. Column (5) is similar, but examines persistence.

Areas that participated in the slave trade are more prone to malaria (based on geographic conditions), had more neighbors in pre-colonial times (reflecting greater ethnic fragmentation), have higher annual rainfall, and have a higher contemporary population. We can strongly reject the hypothesis that areas that did and did not participate in the slave trades were the same: the p-value on an f-test of joint orthogonality across each variable is well below 0.01. Hence, simple cross-sectional comparisons are likely to return biased results, and the difference-in-difference approach is necessary to obtain an unbiased estimate of the effect of slavery on development.

We similarly see that areas in which ethnic groups persisted vary systematically from those in which a different ethnic group now occupies the area. Persistent areas are, on average, less susceptible to malaria, have lower rainfall, and had more neighbors in pre-colonial times. These relationships are quite sensible and support the interpretation that the

measure of persistence is accurate and that factors outside of the slave trade affect the persistence of ethnic groups.

5.2   GDP estimation

I next construct estimates of the log of GDP and the log of GDP per capita using night lights data. Since GDP data is not available for the ethnic groups identified in Murdock (1959), I calibrate a model using national level GDP data obtained from the World Development Indicators. Tables 3 and 4 demonstrate that the luminosity data is an effective proxy for GDP, and the relationship between the variables appears to be linear. The adjusted R^2 is 0.66 in the case of the log of GDP and 0.396 in the case of GDP/capita. I then use the calibrated parameters to construct GDP estimates at the level of the Murdock ethnic boundaries. The slope and intercept are allowed to vary from year to year to account for changing spacecrafts in the VIIRS constellation.

5.3   Cross-sectional comparisons of productivity

Table 4 presents OLS regressions examining the difference in long-run productivity in areas that participated in the slave trade relative to areas that did not. Without including any control variables, point estimates indicate that GDP/-capita is over 36 log points lower in areas from which slaves were sold ($p < .01$). There is no statistically significant difference in GDP. Once including control variables, the estimated effect of the slave trades on GDP is statistically 0, and GDP is 39 log points higher in areas from which slaves were captured ($p < .01$). No inferences can be drawn from these results about the causal effect of the slave trades on long-run development since areas that participated in the slave trade varied systematically and the coefficient on participation in the slave trade varies dramatically based on the control variables that are included. Hence, we must turn to difference-in-difference estimation.

5.4   The effect of the slave trades on long-run development

Difference-in-difference estimates indicate that the slave trades had an economically large, statistically significant, and negative effect on long-run economic development in Africa. Table 5 presents estimates of the effect on the log of GDP/capita (columns (1) and (3)) and GDP (columns (2) and (4)). The point estimate of the effect of the slave trades on GDP/capita in column (3) indicates that the slave trades reduced output by about 33 log points ($p < .01$), and the 95% confidence interval ranges from a 17 log point to a 50 log point reduction in GDP. These values are particularly striking given that the estimate presented may be a lower bound on the effect of slavery. Estimates are robust to the inclusion of control variables. The results also indicate that the slave trades reduced GDP, but the result in column (4) is only marginally significant. Overall, results indicate that the slave trade harmed long-run development, and that slavery is important in explaining poverty and inequality in Africa. Appendix Table 2 demonstrates that results are similar if we use raw night lights values as the dependent variable.

10

In Appendix Table 3 we report difference-in-difference estimates of the effect of the slaves trades on GDP/capita and GDP values winsorized at the 5th and 95th percentiles. The results are qualitatively similar, although the estimates are about $\frac{1}{3}$ smaller in magnitude. We may thus conclude that the reported effect of the slave trade is not entirely attributable to outlying growth in a very small number of areas that did not participate in the slave trades (e.g. South Africa).

5.5   Predicting persistence

I next turn to examining the relationship between the persistence of ethnic groups and the slave trades. Columns (1) and (2) of Table 6 demonstrate that the persistence of ethnic groups is not orthogonal to involvement in the slave trade at the extensive margin. Ethnicities from which slaves were sold are more likely than those that did not engage in the trades to occupy pre-colonial lands today. This is perhaps unsurprising since ethnic groups engaged in the slave trade varied systematically from those that did not. However, this effect vanishes when variation at the intensive margin is included. In columns (3) and (4), the number of slaves sold from each ethnicity is included in the regression instead of an indicator, and the coefficient on the number of slaves sold is statistically zero. In columns (5) and (6), the coefficient on the number of slaves captured from the ethnic group is negative if we restrict the sample to groups that were involved in the slave trades, although the coefficient is not statistically significant. Hence, it appears that the intensity with which an area was engaged in slavery either had no effect on the persistence of the ethnic group, or that areas the most affected by slavery were perhaps more likely to see a change in ethnic demographics than those less intensely engaged.

5.6   Difference-in-difference estimation using the exogenous component of persistence

A final approach to rule out bias from endogeneity in ethnic group persistence is to estimate counterfactual persistence based on factors exogenous to the slave trade. To do so, I use a random forest classifier to predict persistence trained on a random subset of 80% of the sample of ethnic groups that did not engage in the slave trade. I then evaluate the model out of sample on the other 20% of this sample. Figure 5 plots the performance of the model against a random baseline. The model modestly outperforms the random baseline, although it is imperfect – either because omitted variables are important or because the sample size is relatively small and limits the ability to extract complex relationships from the data. The precision of the model, the ratio of true positives to true positives plus false positives, is about 20 points higher in the model than the random baseline.

Using the model for predicting persistence trained on the sample of observations not engaged in the slave trade, I predict persistence across the entire sample. I then recalculate the difference-in-difference specification using predicted persistence in place of actual persistence in Table 7. The estimated effect of the slave trade on long-run economic development remains negative and statistically significant, supporting the validity of the initial results.

6 Conclusion

Using variation in the persistence of ethnic groups from pre-colonial to contemporary times, I find that the slave trades had a large and statistically significant negative effect on long-run African development. I estimate that the slave trade reduced long-run per capita GDP by about 33 log points and can reject effect sizes below 17 with 95% confidence. The estimates are robust to the inclusion of a rich set of controls, and tests indicate that endogeneity in the relationship between ethnic group persistence and the slave trades is not likely a source of downward bias in our estimate in the effect of the slave trades. These results confirm the findings of Nunn (2008) and show that these findings were not the result of weak instruments or a failure of exclusion restrictions.

The magnitude of the effect of slavery is striking. This suggests that slavery remains an important contributor to poverty and inequality in Africa over 100 years after the abolition of international slave trades. In addition, the findings demonstrate that differences in social capital can drive significant income differences, and suggest that shocks to social capital can be extremely long-lasting.

The methodology employed in this paper relies on modern technology to estimate the effect of slavery, and so it is poorly suited to test whether the effects of the slave trades are diminishing over time. This suggests the need for future research aimed at answering this question. However, the magnitude of the persistent effect size after over 100 years indicates that if the effect of the slave trades is diminishing, it is occurring at a relatively slow rate. Hence, development research should also examine whether interventions can reduce the persistent effects of slavery.

Similarly, research into the dynamics of shocks to social capital may be valuable. This paper presents causal evidence that a shock to social capital over 100 years ago remains an important driver of productivity. This suggests that shocks to social capital may follow an auto-regressive process. A better understanding of these dynamics would thus contribute to our understanding of how disparities in social capital emerge and yield evidence for policy makers about the value of responding proactively to shocks to social capital.

References

Yann Algan and Pierre Cahuc. Inherited trust and growth. *American Economic Review*, 100(5):2060–2092, 12 2010. ISSN 00028282. doi: 10.1257/aer.100.5.2060. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.100.5.2060.

Richard Nelson Bean. *The British trans-Atlantic slave trade, 1650-1775*. Arno Press, New York, 1975. ISBN 0405072562.

Robert Blanton, T. David Mason, and Brian Athow. Colonial Style and Post-Colonial Ethnic Conflict in Africa. *Journal of Peace Research*, 38(4):473–491, 7 2001. ISSN 0022-3433. doi: 10.1177/0022343301038004005. URL http://journals.sagepub.com/doi/10.1177/0022343301038004005.

Xi Chen and William D Nordhaus. Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences of the United States of America*, 108(21):8589–94, 5 2011. ISSN 1091-6490. doi: 10.1073/pnas.1017031108. URL http://www.ncbi.nlm.nih.gov/pubmed/21576474http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3102367.

T. G. Conley. GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45, 9 1999. ISSN 03044076. doi: 10.1016/S0304-4076(98)00084-0.

Timothy G Conley. Spatial Econometrics. In *The New Palgrave Dictionary of Economics*, pages 1–9. Palgrave Macmillan UK, London, 2016. ISBN 978-1-349-95121-5. doi: 10.1057/978-1-349-95121-5{_}2023-1. URL https://doi.org/10.1057/978-1-349-95121-5_2023-1.

Toyin Falola and Amanda Warnock. *Encyclopedia of the middle passage*. Greenwood Press, Westport, Conn., 2007. ISBN 9780313334801.

Marc Felix and Charles Meur. *Peoples of Africa Atlas: An ethnolinguistic atlas of Africa*. 2001. URL https://worldmap.harvard.edu/data/geonode:etnicity_felix.

Elisabeth Gilmore, Nils Petter Gleditsch, Päivi Lujala, and Jan Ketil Rod. Conflict Diamonds: A New Dataset. *Conflict Management and Peace Science*, 22(3):257–272, 7 2005. ISSN 0738-8942. doi: 10.1080/07388940500201003. URL http://journals.sagepub.com/doi/10.1080/07388940500201003.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. The role of social capital in financial development, 6 2004. ISSN 00028282.

J Vernon Henderson, Adam Storeygard, and David N Weil. Measuring Economic Growth from Outer Space. *American Economic Review*, (2):994–1028, 2012. doi: 10.1257/aer.102.2.994.

Dean S. Karlan. Using experimental economics to measure social capital and predict financial decisions, 12 2005. ISSN 00028282.

E Phillip LeVeen. *British slave trade suppression policies, 1821-1865*. Arno Press, New York, 1977. ISBN 0405107927.

Paul E. Lovejoy. *Transformations in Slavery*. Cambridge University Press, 2011. doi: 10.1017/cbo9781139014946. URL https://www-cambridge-org.libproxy.berkeley.edu/core/books/transformations-in-slavery/97658D86435B0A9D2D2AB9126404CADA.

Päivi Lujala, Jan Ketil Rod, and Nadja Thieme. Fighting over Oil: Introducing a New Dataset. *Conflict Management and Peace Science*, 24(3):239–256, 7 2007. ISSN 0738-8942. doi: 10.1080/07388940701468526. URL http://journals.sagepub.com/doi/10.1080/07388940701468526.

Stelios Michalopoulos and Elias Papaioannou. The long-run effects of the scramble for Africa. *American Economic Review*, 106(7):1802–1848, 7 2016. ISSN 00028282. doi: 10.1257/aer.20131311. URL http://dx.doi.org/10.1257/aer.20131311.

George Peter Murdock. *Africa, its peoples and their culture history*. McGraw-Hill, New York, 1959.

Nathan Nunn. The Long-Term Effects of Africa's Slave Trades. *Quarterly Journal of Economics*, 123(1):139–176, 2 2008. ISSN 0033-5533. doi: 10.1162/qjec.2008.123.1.139. URL https://academic.oup.com/qje/article-lookup/doi/10.1162/qjec.2008.123.1.139.

Nathan Nunn and Diego Puga. Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics*, 94(1):20–36, 1 2012. ISSN 00346535. doi: 10.1162/REST{_}a{_}00161. URL https://www.mitpressjournals.org/doix/abs/10.1162/REST_a_00161.

Nathan Nunn and Leonard Wantchekon. The slave trade and the origins of Mistrust in Africa. *American Economic Review*, 101(7):3221–3252, 12 2011. ISSN 00028282. doi: 10.1257/aer.101.7.3221.

Robert D Putnam. *Bowling alone: the collapse and revival of American community*. Simon & Schuster, New York, 2000. URL https://libproxy.berkeley.edu/login?qurl=https%3a%2f%2fsearch.ebscohost.com%2flogin.aspx%3fdirect%3dtrue%26db%3dreh%26AN%3dATLA0000014109%26site%3deds-live.

Amanda Lea Robinson. National versus ethnic identification in Africa modernization, colonial legacy, and the origins of territorial nationalism. *World Politics*, 66(4):709–746, 9 2014. ISSN 10863338. doi: 10.1017/S0043887114000239. URL https://www.cambridge.org/core/terms.

Shanker Satyanath, Nico Voigtländer, and Hans-Joachim Voth. Bowling for Fascism: Social Capital and the Rise of the Nazi Party. *Journal of Political Economy*, 125(2):478–526, 4 2017. ISSN 0022-3808. doi: 10.1086/690949. URL https://www.journals.uchicago.edu/doi/10.1086/690949.

Douglas Staiger and James H. Stock. Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3):557, 5 1997. ISSN 00129682. doi: 10.2307/2171753.

J. Temple and P. A. Johnson. Social Capability and Economic Growth. *The Quarterly Journal of Economics*, 113(3): 965–990, 8 1998. ISSN 0033-5533. doi: 10.1162/003355398555711. URL https://academic.oup.com/qje/article-lookup/doi/10.1162/003355398555711.

United Nations Development Programme. Income Inequality Trends in Sub-Saharan Africa: Divergence, Determinants and Consequences. Technical report, 2017. URL https://www.africa.undp.org/content/rba/en/home/library/reports/income-inequality-trends-in-sub-saharan-africa--divergence--dete.html.

Warren Whatley and Rob Gillezeau. The impact of the transatlantic slave trade on ethnic stratification in Africa. In *American Economic Review*, volume 101, pages 571–576, 5 2011. doi: 10.1257/aer.101.3.571.

7   Figures

Figure 1: Pre-colonial and contemporary ethnic boundaries

(a) A. Pre-colonial

(b) B. Contemporary

Panel A reports the pre-colonial boundaries of ethnic groups as presented in Murdock (1959) and digitized in Nunn (2008). Panel B presents contemporary ethnic boundaries as presented in Felix and Meur (2001) and digitized by the Center of Geographic Analysis at Harvard University.

Figure 2: Pre-colonial ethnic boundaries:
Slave trade participation and persistence
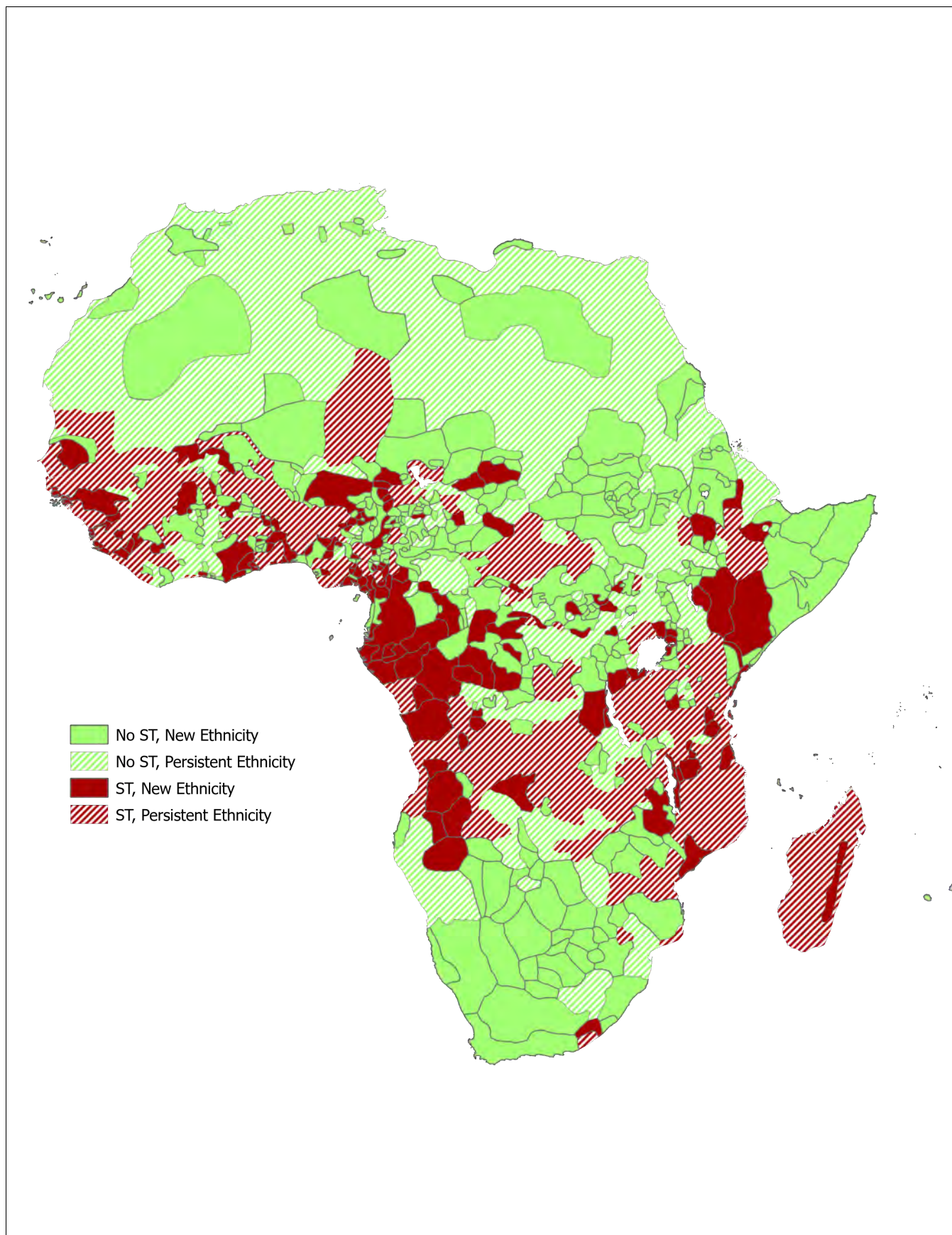
Figure 2 plots pre-colonial ethnic boundaries as presented in Murdock (1959). The territories are colored green if no members of the ethnicity were documented as being sold into the trans-Atlantic or Indian slave trades and red if at least one member of the ethnic group was sold into slavery. The polygons are solid if the pre-colonial ethnic group is not the current dominant ethnic group in the area. They are hashed if the pre-colonial ethnic group is the dominant contemporary ethnic group.

Figure 3: Night lights by pre-colonial ethnic boundary
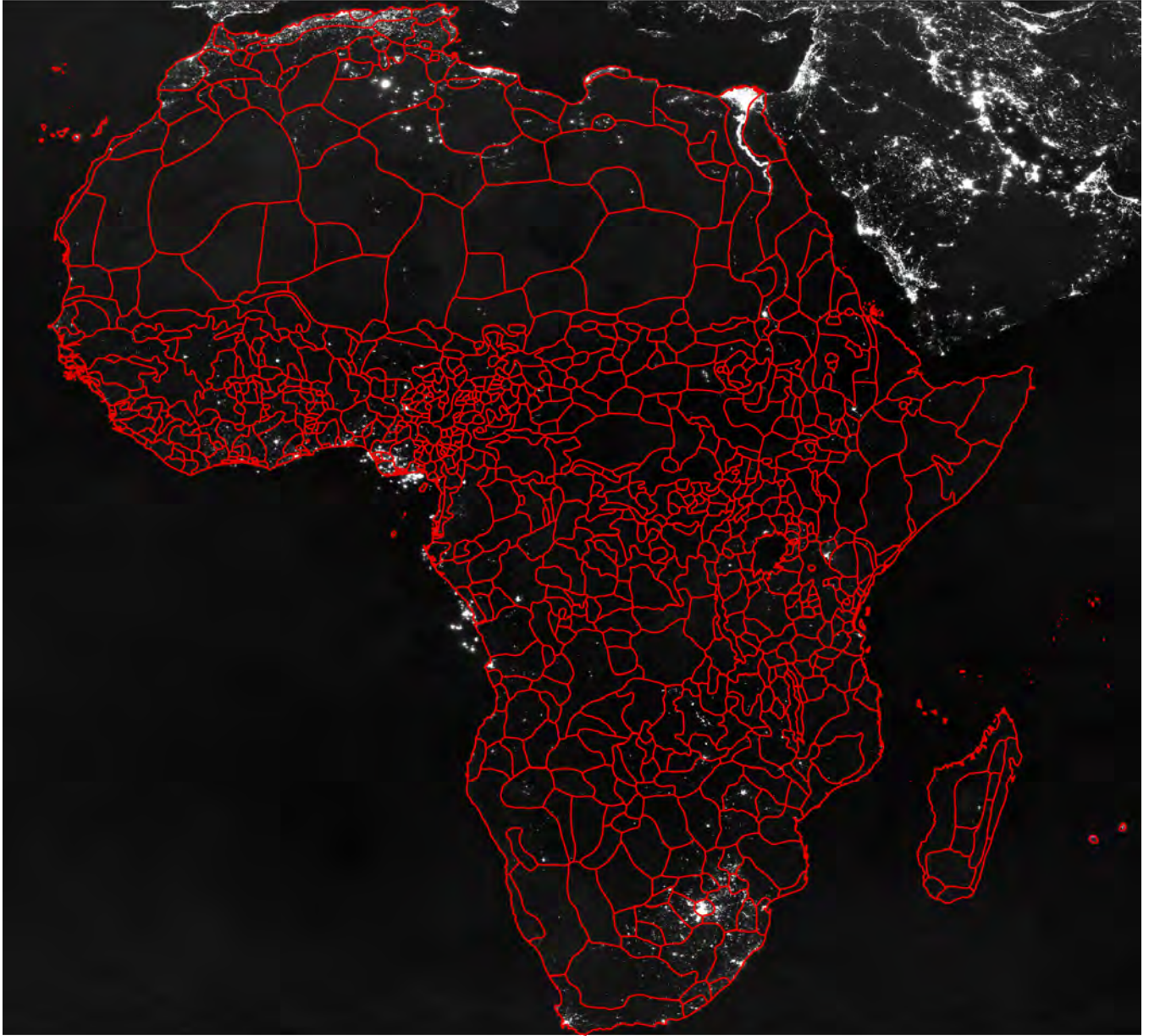
Figure 3 plots the log of 2014 luminosity readings on pre-colonial ethnic boundaries. Night light data is from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB). Night lights values were constructed using the Google Earth Engine beginning with stray light corrected monthly composites. The median pixel value across each reading from 2014 was then taken. I then plotted the natural log of 1 plus each pixel value.

Figure 4: African national output vs. night lights
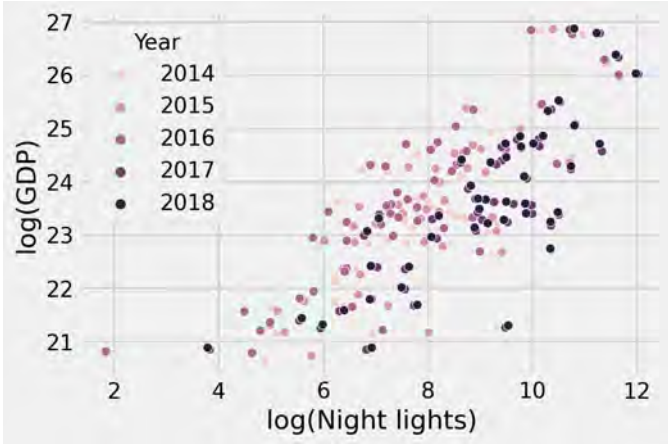
(a) A. $\log(GDP)$ vs $\log(NL)$ (b) B. $\log(GDP/capita)$ vs $\log(NL/capita)$
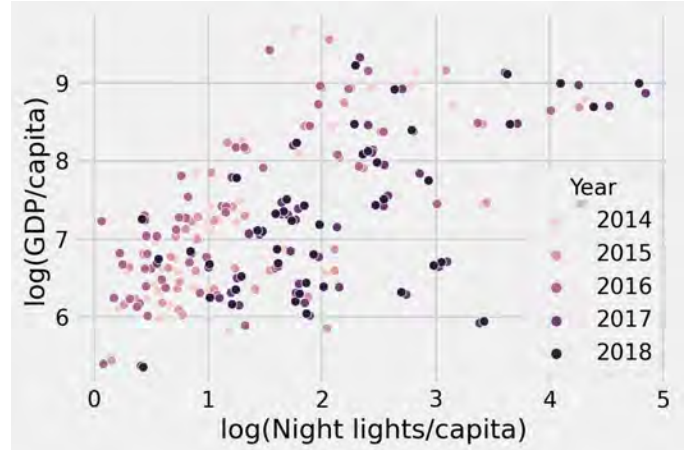
Panel A presents a scatter plot of the natural log of GDP vs the natural log of luminosity readings for 48 countries in Africa using annual observations from 2014-2018. Panel B plots (the log of) GDP/capita and night lights per capita. Night light data is from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB). Night lights values were constructed using the Google Earth Engine beginning with stray light corrected monthly composites. Values were then clipped below at 0 and the sum of all luminosity values observed in the boundaries of a country was calculated for each composite, then the median was taken across each image obtained in a year.

Figure 5: ROC plots of persistence model performance
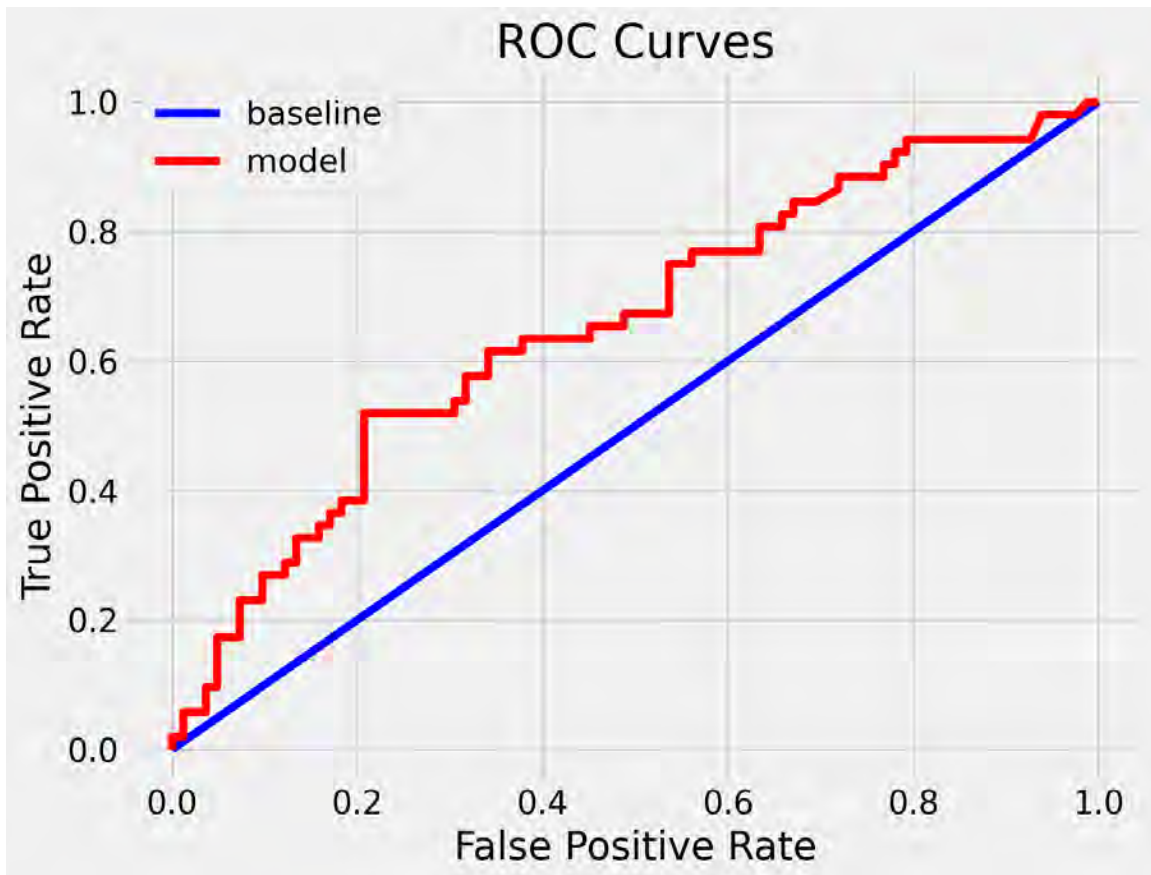
Figure 5 plots ROC curves from the random forest classifier used to predict ethnic group persistence. The random forest model was trained using the subset of ethnic groups that did not participate in the slave trade. Persistence was predicted using the number of diamond deposits, an indicator for oil deposits, temperature suitability for malaria transmission, the number of neighbors, area, the number of powers that colonized the area, and indicators for whether each colonial power colonized the area in question. The ROC curve plots the true vs false positive rates with different cutoff probabilities. The performance of the model is plotted against a baseline of random assignment which is the 45 degree line.

Tables

Table 1: The persistence of ethnic groups by participation in the slave trade

Persistent ethnic group Participated in slave trade	No	Yes	Total
No	333	203	536
Yes	146	153	299
Total	479	356	835

Rows denote whether the ethnic groups recorded in Murdock (1959) participated in the slave trade. An ethnic group is considered to have participated in the slave trade if one or more slaves from the trans-Atlantic or Indian slave trade records from Nunn (2008) listed the ethnicity. Columns record whether the ethnic group retained control of their historical lands to modern time. We define the variable persistence as equal to 1 if the dominant ethnic group identified in the area in Felix and Meur (2001) is equal to that in Murdock (1959).

Table 2: Summary statistics

	(1) Full sample	(2) ST = 0	(3) ST: 1 - 0	(4) Persistent = 0	(5) Persistent: 1 - 0
Diamond deposits	0.516 [2.448]	0.397 [2.128]	0.332* (0.192)	0.468 [2.302]	0.114 (0.175)
Oil	0.105 [0.307]	0.108 [0.311]	-0.008 (0.022)	0.088 [0.283]	0.042* (0.022)
Gold deposits	0.631 [4.886]	0.793 [5.967]	-0.452 (0.275)	0.835 [6.349]	-0.478 (0.298)
Malaria (Pf)	0.440 [0.174]	0.409 [0.175]	0.085*** (0.012)	0.457 [0.173]	-0.040*** (0.012)
Malaria (Pv)	0.588 [0.205]	0.547 [0.209]	0.113*** (0.014)	0.607 [0.201]	-0.045*** (0.014)
Annual rainfall (mm/day)	2.925 [1.587]	2.539 [1.546]	1.079*** (0.106)	3.048 [1.575]	-0.287*** (0.111)
Neighbors	5.468 [2.503]	5.278 [2.297]	0.531*** (0.190)	5.194 [2.387]	0.643*** (0.176)
Colonizers (number)	1.388 [0.571]	1.364 [0.547]	0.068 (0.042)	1.370 [0.563]	0.043 (0.040)
Population (2014, mil)	1.427 [3.976]	1.182 [4.306]	0.657*** (0.252)	1.139 [2.452]	0.628** (0.286)
Belgium (colonized by)	0.132 [0.338]	0.125 [0.331]	0.019 (0.025)	0.129 [0.336]	0.005 (0.024)
Britain	0.491 [0.500]	0.519 [0.500]	-0.077** (0.036)	0.520 [0.500]	-0.068* (0.035)
France	0.321 [0.467]	0.319 [0.466]	0.005 (0.034)	0.282 [0.450]	0.092*** (0.033)
Germany	0.228 [0.419]	0.188 [0.391]	0.109*** (0.031)	0.246 [0.431]	-0.044 (0.029)
Italy	0.038 [0.192]	0.058 [0.233]	-0.054*** (0.011)	0.038 [0.190]	0.002 (0.013)
Portugal	0.078 [0.268]	0.037 [0.190]	0.113*** (0.022)	0.061 [0.239]	0.041** (0.019)
Spain	0.022 [0.145]	0.030 [0.170]	-0.023*** (0.009)	0.015 [0.120]	0.016 (0.011)
Observations	835	536	835	479	835
p-val joint orthogonality			0.000		0.000

Standard deviations in brackets. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column (1) reports the mean and standard deviation of the indicated variable across the full sample. Column (2) reports the mean and standard deviation among the restricted sample of ethnic groups that did not participate in international slave trades. Column (4) is the similar but the sample is instead restricted to non-persistent ethnic groups. Column (3) reports the difference in mean of each variable among groups that did participate in the slave trades relative to those that did not, and standard errors of this pairwise t-test. Column (5) reports similar results, but based on ethnicity persistence instead of slave trade participation.

Table 3: Predicting GDP with luminosity data

	(1) $\log(GDP)$	(2) $\log(GDP/capita)$
Constant	18.291***	6.292***
	(0.564)	(0.144)
Night lights	0.666***	0.719***
	(0.073)	(0.102)
2015	-0.039	-0.005
	(0.189)	(0.065)
NL x 2015	-0.007	-0.040
	(0.022)	(0.045)
2016	0.147	0.095**
	(0.124)	(0.041)
NL x 2016	0.012	0.049
	(0.016)	(0.040)
2017	-0.896***	-0.156
	(0.316)	(0.158)
NL x 2017	0.018	-0.191**
	(0.037)	(0.087)
2018	-0.935***	-0.142
	(0.326)	(0.170)
NL x 2018	0.027	-0.185*
	(0.038)	(0.094)
Observations	237	237
Adj R^2	0.660	0.396
Entities	48	48
Time periods	5	5

Clustered standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column (1) reports the results of a regression of the natural log of GDP on the log of the sum of night lights values recorded in the boundary of each country. Column (2) reports the results of a regression of the log of GDP/capita on the log of night lights per capita. Night light data is from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB). Night lights values were constructed using the Google Earth Engine beginning with stray light corrected monthly composites. Values were then clipped below at 0 and the sum of all luminosity values observed in the boundaries of a country was calculated for each composite, then the median was taken across each image obtained in a year. The log VIIRS value used in column 1 is defined as the log of 1 plus the raw luminosity calculation, and in column 2 the log of 1 plus 10,000 times the raw nightlights per capita value is considered.

Table 4: Correlations between output and the slave trade

	(1) log(GDP/capita)	(2) log(GDP)	(3) log(GDP/capita)	(4) log(GDP)
Slave trade	-0.364***	0.108	-0.063	0.390***
	(0.058)	(0.081)	(0.063)	(0.071)
Diamond deposits			0.046***	0.042***
			(0.010)	(0.012)
Oil			0.760***	1.207***
			(0.130)	(0.120)
Gold deposits			0.006*	0.033***
			(0.003)	(0.006)
Malaria (Pf)			1.680*	0.775
			(0.978)	(1.100)
Malaria (Pv)			-1.558*	-1.684*
			(0.908)	(0.960)
Annual rainfall (mm/day)			-0.248***	-0.201***
			(0.032)	(0.028)
Neighbors			-0.015	0.129***
			(0.011)	(0.011)
Time FE	Yes	Yes	Yes	Yes
Observations	4,175	4,175	4,175	4,175

Spatial HAC standard errors in parenthesis (Conley, 1999, 2016). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable in columns (1) and (3) is the log of GDP/capita estimated using the log of night lights per capita measured using VIIRS luminosity data. In columns (2) and (4), the log of GDP is the dependent variable and is estimated using the log of night lights. All variables are calculated over the historical ethnic boundaries presented in Murdock (1959). The variable "Slave trade" is an indicator recording whether the ethnic group historically participated in the slave trade.

Table 5: Difference-in-difference estimate of the effect of the slave trade on output

	(1) log(GDP/capita)	(2) log(GDP)	(3) log(GDP/capita)	(4) log(GDP)
Slave trade	-0.255***	0.097	0.054	0.424***
	(0.074)	(0.105)	(0.086)	(0.088)
Persistent ethnicity	0.467***	0.653***	0.337***	0.426***
	(0.103)	(0.106)	(0.078)	(0.067)
Slave trade x Persistent	-0.334***	-0.149	-0.329***	-0.202*
	(0.122)	(0.148)	(0.115)	(0.106)
Diamond deposits			0.046***	0.042***
			(0.011)	(0.012)
Oil			0.726***	1.167***
			(0.128)	(0.113)
Gold deposits			0.008**	0.035***
			(0.003)	(0.006)
Malaria (Pf)			1.879*	1.116
			(0.983)	(1.114)
Malaria (Pv)			-1.678*	-1.894**
			(0.906)	(0.965)
Annual rainfall (mm/day)			-0.242***	-0.190***
			(0.031)	(0.028)
Neighbors			-0.018	0.122***
			(0.012)	(0.011)
Time FE	Yes	Yes	Yes	Yes
Observations	4,175	4,175	4,175	4,175

Spatial HAC standard errors in parenthesis (Conley, 1999, 2016). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable in columns (1) and (3) is the log of GDP/capita estimated using the log of night lights per capita measured using VIIRS luminosity data. In columns (2) and (4), the log of GDP is the dependent variable and is estimated using the log of night lights. All variables are calculated over the historical ethnic boundaries presented in Murdock (1959). The variable "Slave trade" is an indicator recording whether the ethnic group historically participated in the slave trade. The variable "Persistent" indicates whether the ethnic group that historically occupied the area is still the dominant ethnicity in the area. The coefficient on the interaction between "Slave trade" and "Persistent" is the difference-in-difference estimate of the effect of the slave trade on output.

Table 6: The relationship between ethnic group persistence and participation in the slave trade

	Full sample				Participants in the slave trade	
	(1) Indicator	(2) Indicator	(3) # Enslaved	(4) # Enslaved	(5) # Enslaved	(6) # Enslaved
Slave trade	0.133***	0.162***				
	(0.036)	(0.039)				
Total enslaved (100,000s)			0.005	-0.002	-0.001	-0.008
			(0.015)	(0.015)	(0.012)	(0.010)
Diamond deposits		0.000		0.001		0.003
		(0.006)		(0.006)		(0.009)
Oil		0.062		0.068		0.077
		(0.056)		(0.057)		(0.090)
Gold deposits		-0.006***		-0.006***		0.008
		(0.002)		(0.002)		(0.021)
Malaria (Pf)		-1.376*		-1.625**		-0.862
		(0.713)		(0.721)		(1.142)
Malaria (Pv)		0.974		1.231*		0.420
		(0.643)		(0.651)		(1.059)
Annual rainfall (mm/day)		-0.034**		-0.021		-0.033
		(0.016)		(0.016)		(0.026)
Neighbors		0.022***		0.026***		0.021**
		(0.007)		(0.007)		(0.010)
Colonizers (number)		0.035		0.021		0.103
		(0.070)		(0.069)		(0.145)
Belgium (colonized by)		0.000		-0.022		-0.141
		(0.083)		(0.082)		(0.153)
Britain		-0.059		-0.053		-0.009
		(0.073)		(0.073)		(0.148)
France		0.045		0.073		-0.003
		(0.077)		(0.075)		(0.154)
Germany		-0.098		-0.081		-0.046
		(0.073)		(0.072)		(0.147)
Italy		-0.061		-0.059		-0.823***
		(0.121)		(0.121)		(0.299)
Portugal		0.029		0.097		0.039
		(0.092)		(0.091)		(0.156)
Spain		0.078		0.052		-0.602***
		(0.122)		(0.122)		(0.232)
Observations	835	835	835	835	299	299
Adjusted R^2	0.015	0.060	-0.001	0.040	-0.003	0.041

Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable in each regression is an indicator that takes on a value of 1 if the dominant ethnic group in an area is the same ethnicity that historically occupied the land. In columns (1) and (2), the dependent variable is regressed on an indicator for whether members of the historical ethnic group were sold into the slave trade. In columns (3) - (6) the independent variable is the total number of slaves from the ethnic group sold into the trans-Atlantic and Indian slave trades. Columns (1) - (4) examine the entire sample. Columns (5) - (6) are restricted to ethnicities that participated in the slave trade.

Table 7: Difference-in-difference estimate of the effect of the slave trade on output
Predicted persistence

	(1) log(GDP/capita)	(2) log(GDP)	(3) log(GDP/capita)	(4) log(GDP)
Slave trade	-0.186*** (0.062)	0.324*** (0.098)	0.047 (0.075)	0.497*** (0.083)
Predicted persistence	0.679*** (0.123)	0.876*** (0.101)	0.591*** (0.095)	0.613*** (0.069)
Slave trade x Predicted persistence	-0.542*** (0.155)	-0.645*** (0.146)	-0.268** (0.133)	-0.249** (0.112)
Diamond deposits			0.048*** (0.012)	0.044*** (0.012)
Oil			0.679*** (0.123)	1.125*** (0.112)
Gold deposits			0.009*** (0.003)	0.037*** (0.006)
Malaria (Pf)			2.224** (0.977)	1.340 (1.098)
Malaria (Pv)			-1.868** (0.895)	-2.003** (0.950)
Annual rainfall (mm/day)			-0.255*** (0.031)	-0.209*** (0.028)
Neighbors			-0.026** (0.011)	0.117*** (0.010)
Time FE	Yes	Yes	Yes	Yes
Observations	4,175	4,175	4,175	4,175

Spatial HAC standard errors in parenthesis (Conley, 1999, 2016). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable in columns (1) and (3) is the log of GDP/capita estimated using the log of night lights per capita measured using VIIRS luminosity data. In columns (2) and (4), the log of GDP is the dependent variable and is estimated using the log of night lights. All variables are calculated over the historical ethnic boundaries presented in Murdock (1959). The variable "Slave trade" is an indicator recording whether the ethnic group historically participated in the slave trade. The variable "Predicted persistence" is a random forest prediction of whether the ethnic group that historically occupied the area is still the dominant ethnicity in the area. The random forest model was trained using the subset of ethnic groups that did not participate in the slave trade. Persistence was predicted using the number of diamond deposits, an indicator for oil deposits, temperature suitability for malaria transmission, the number of neighbors, area, the number of powers that colonized the area, and indicators for whether each colonial power colonized the area in question. The random forest produced a precision score (true positives over true positives plus false positives) of 0.56 relative to a baseline value of 0.39 from random assignment. The coefficient on the interaction between "Slave trade" and "Predicted persistence" is the difference-in-difference estimate of the effect of the slave trade on output.

Appendix A: Tables

Appendix Table 1: Data sources

Description	Source	Processing
Pre-colonial ethnic boundaries	The pre-colonial ethnic boundaries were originally constructed by Murdock (1959). The version used in this paper was downloaded from the replication data for Nunn and Wantchekon (2011).	Data from areas that were unpopulated in pre-colonial times are omitted from analysis.
Contemporary ethnic boundaries	Data on contemporary ethnic boundaries is from Felix and Meur (2001). The data was digitized by Harvard University's AfricaMap project.	The Python package "Geopandas" was used to identify persistent ethnic groups as discussed in Section 3.
Slave trade data	Data on the trans-Atlantic and Indian slave trades was originally compiled on Nunn (2008). The data used in this paper was downloaded from the replication files for Nunn and Wantchekon (2011).	An indicator variable "Slave trade" is coded to 1 if at least one slave was exported from an ethnic group in the raw data. The variable "Total slaves" records the sum of all slaves exported from an ethnic group across all years and trades.
Nighttime lights data	VIIRS Stray Light Corrected Nighttime Day/Night Band Composites Version 1 accessed from the Google Earth Engine	Luminosity data was processed on the Google Earth Engine. The median value was calculated for each pixel across all images from a year for the years 2014-2018. Pixels with a value below 0 were recoded to 0, then the sum of all pixels contained in each polygon from the Murdock (1959) was calculated.
Population data	LandScan Global rasters from Oak Ridge National Laboratory	The total number of people contained in each polygon from Murdock (1959) was calculated by summing each pixel contained in each polygon. Calculations were performed in Python using the "rasterstats" package.
Petroleum	Lujala et al. (2007)	An indicator was constructed in Python using the package "Geopandas" that is coded to 1 if an oil reserve intersects with the boundaries of a polygon from Murdock (1959).
Diamonds	Gilmore et al. (2005)	The number of diamond deposits contained in each polygon from Murdock (1959) was calculated in Python using the "Geopandas" package.
Gold	United States Geological Survey Mineral Resource Data System	The number of gold deposits contained in each polygon was calculated using Python.
Malaria	The temperature suitability indices for P. falciparum and P. vivax transmission were downloaded from the Malaria Atlas Project	The average index value was calculated separately for each index across each Murdock (1959) observation using Python.
GIS data on colonizers of Africa	The data was uploaded by Bucknell University on ArcGIS and is available under the name "Colonial Africa" for download.	Python was used to determine which colonial powers occupied each polygon from Murdock (1959). An indicator was constructed to indicate if each power occupied the area, and the number of colonizers was recorded.
National GDP/capita and population data	World Development Indicators	Used for evaluating night lights data and estimating GDP from luminosity. GDP was calculated by multiplying GDP by population.

Appendix Table 2: Difference-in-difference estimate of the effect of the slave trade on luminosity

	(1) log($NL/capita$)	(2) log(Night lights)	(3) log($NL/capita$)	(4) log(Night lights)
Slave trade	-0.415*** (0.117)	0.144 (0.156)	0.051 (0.132)	0.629*** (0.131)
Persistent ethnicity	0.703*** (0.152)	0.968*** (0.157)	0.507*** (0.118)	0.631*** (0.100)
Slave trade x Persistent	-0.505*** (0.187)	-0.221 (0.219)	-0.500*** (0.176)	-0.300* (0.158)
Diamond deposits			0.070*** (0.016)	0.063*** (0.017)
Oil			1.065*** (0.185)	1.732*** (0.169)
Gold deposits			0.010* (0.005)	0.053*** (0.008)
Malaria (Pf)			2.560* (1.498)	1.706 (1.655)
Malaria (Pv)			-2.234 (1.387)	-2.851** (1.434)
Annual rainfall (mm/day)			-0.376*** (0.046)	-0.282*** (0.042)
Neighbors			-0.026 (0.018)	0.181*** (0.016)
Time FE	Yes	Yes	Yes	Yes
Observations	4,175	4,175	4,175	4,175

Spatial HAC standard errors in parenthesis (Conley, 1999, 2016). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable in columns (1) and (3) is the log of night lights per capita measured using VIIRS luminosity data. The dependent variable is columns (2) and (4) is the log of night lights. All variables are calculated over the historical ethnic boundaries presented in Murdock (1959). The variable "Slave trade" is an indicator recording whether the ethnic group historically participated in the slave trade. The variable "Persistent" indicates whether the ethnic group that historically occupied the area is still the dominant ethnicity in the area. The coefficient on the interaction between "Slave trade" and "Persistent" is the difference-in-difference estimate of the effect of the slave trade on luminosity.

Appendix Table 3: Difference-in-difference estimate of the effect of the slave trade on output
Winsorized data

	(1) log(GDP/capita)	(2) log(GDP)	(3) log(GDP/capita)	(4) log(GDP)
Slave trade	-0.234***	0.081	0.017	0.379***
	(0.067)	(0.098)	(0.076)	(0.079)
Persistent ethnicity	0.347***	0.553***	0.248***	0.349***
	(0.077)	(0.095)	(0.060)	(0.062)
Slave trade x Persistent	-0.225**	-0.067	-0.224**	-0.122
	(0.102)	(0.140)	(0.098)	(0.101)
Diamond deposits			0.044***	0.042***
			(0.009)	(0.011)
Oil			0.601***	0.991***
			(0.095)	(0.092)
Gold deposits			0.009***	0.029***
			(0.003)	(0.005)
Malaria (Pf)			1.707*	1.147
			(0.895)	(1.057)
Malaria (Pv)			-1.446*	-1.791*
			(0.818)	(0.924)
Annual rainfall (mm/day)			-0.200***	-0.179***
			(0.026)	(0.026)
Neighbors			-0.019*	0.115***
			(0.010)	(0.010)
Time FE	Yes	Yes	Yes	Yes
Observations	4,175	4,175	4,175	4,175

Spatial HAC standard errors in parenthesis (Conley, 1999, 2016). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependent variable in columns (1) and (3) is the log of GDP/capita estimated using the log of night lights per capita measured using VIIRS luminosity data. In columns (2) and (4), the log of GDP is the dependent variable and is estimated using the log of night lights. Both variables were winsorized at the 5th and 95th percentiles. All variables are calculated over the historical ethnic boundaries presented in Murdock (1959). The variable "Slave trade" is an indicator recording whether the ethnic group historically participated in the slave trade. The variable "Persistent" indicates whether the ethnic group that historically occupied the area is still the dominant ethnicity in the area. The coefficient on the interaction between "Slave trade" and "Persistent" is the difference-in-difference estimate of the effect of the slave trade on output.