Transborder Ethnic Kin and Local Prosperity : Evidence from Night-Time Light Intensity in Africa*

Christophe Muller and Pierre Pecher

Aix-Marseille Univ., CNRS, EHESS, Centrale Marseille, AMSE

26 May 2018

Abstract

Ethnicity often occupies a core role in integrated social, economic, and political development processes, which have mostly been studied within specific countries. Across countries, social and economic development may be supported by political capabilities achieved by ethnic kin abroad, although there is little hard evidence on politico-economic interactions through ethnic networks. We fill this gap by providing the first robust empirical evidence of the substantial effects of political predominance of transborder ethnic kin on local economic development in Africa. This is achieved by specifying and estimating dynamic spatial models of geolocalised luminosity and matching these data with other geolocalised information on geographic, political, and ethnic characteristics. Spatial and ethnic network effects are separately identified and jointly analysed. Not only distinct spatial effects and transborder ethnic effects are exhibited, but also are their complex dynamics and spatial distribution features in terms of local development. The results draw attention to the relevance of a broader international perspective on policies affecting ethnic politics within countries.

JEL Classification: D72, R11, O43.

Keywords: Local Development, Ethnic Networks, Institutions.

*This research was carried out thanks to the support of the A*MIDEX project (No. ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French government programme, managed by the French National Research Agency (ANR), and the ANR project (No.ANR-17-CE39-0009-01) funded by the French National Research Agency (ANR), within the framework of the TMENA and TMENA2 projects (http://christophemuller.net/tmena/). We are grateful to participants at several conferences and seminar presentations for their comments.

1 Introduction

1.1 The Issue

Ethnic dimensions play prominent roles in many spheres of social and economic life in less developed countries, particularly in Africa. The political interactions of ethnic groups have been regarded as potentially core determinants of development outcomes.[1]

Within countries, power struggles at the state level often track ethnic demarcations that also shape the allocation of public resources.[2] Across countries, the partitioning of historical ethnicities into modern nations in the aftermath of colonisation, as well as migrations, created situations with strong transborder solidarities. These social and economic links through ethnic networks may be supported by the political capabilities achieved by ethnic kin abroad.[3] If that is the case, development policies that have political consequences on any given government, and which are typically evaluated from a national perspective, should also be analysed in terms of their economic repercussions that may operate through ethnic networks on neighbouring countries. For example, the conquest of the central power in Zaire by Tutsi militaries supporting Laurent-Desire Kabila economically benefited Rwanda, including through exclusive access to adjacent Congolese mineral resources.

However, the measurement of the economic consequences of transborder ethnic links is challenging because they may be confused with other spatio-economic effects as most ethnic groups are spatially connected. This may explain why so little hard evidence is available on politico-economic interactions in ethnic networks across borders. We fill this gap.

In this paper, we investigate these issues, and we provide the first empirical evidence of the substantial economic effect of connections to politically dominant ethnic kin across borders on local development. To be able to exhibit this so far neglected transborder politico-ethnic phenomenon, we proceed in several stages:

i. We construct fine-scale spatial satellite measures of local light emissions. We then match these measures with information on the geographical location of ethnic groups, their own political participation in the central government and that of their ethnically-related kin in neighbouring countries, besides other variables on violence and climate.

ii. For the first time in this literature, we specify and estimate spatial panel data models that account for these ethnic channels, on top of other spatial interactions. We mitigate endogeneity concerns by using an identification strategy that relies on variations of emissions of light and political variables within the same country and in the same year, while controlling for fixed local factors, country$\times$year fixed-effects, and other controls.

[1] See Collier and Hoeffler (2004), Alesina and La Ferrara (2005), Cederman et al. (2010) and Francois et al. (2015).

[2] For instance, see Esteban and Ray (1994), Kramon and Posner (2013), Hodler and Raschky (2014), Burgess et al. (2015), and De Luca et al. (2018).

[3] See Gleditsch (2007) and Gurses (2015).

iii. We exploit the estimated model to investigate the spatial and ethnic externalities using simulation methods. In particular, we exhibit distribution effects that cannot be summarised by the sole knowledge of the coefficients of the model.

iv. We apply a broad range of robustness checks. Notably, we estimate models that account for dynamic interactions, including with spatial effects; network spatial effects, and heterogeneity of political effects. Finally, we confirm our main result with a causal analysis conducted with an adjusted difference-in-differences method.

1.2 Spatial and Ethnic Kin Interactions

Several non-exclusive channels may generate ethnic and spatial externalities for economic development, such as technology diffusion, insurance and other solidarity mechanisms, human capital accumulation, migration and marriage networks, trade, violence, market competition, or institutions.[4] Let us examine some of these briefly.

Economic activities tend to cluster in space for many reasons that operate irrespective of ethnic dimensions. For instance, theoretical trade models often imply that bilateral flows of goods and services are the largest between nearby places because of lower transaction costs. Technology diffusion and investment also entail strong spatial interactions. Similarly, human capital accumulation may be stimulated locally by enhanced job opportunities, and facilitated by former migrations and their associated remittances flows.[5] Furthermore, local infrastructure often determine the geographical patterns of economic activities. For example, Jedwab and Moradi (2016) and Storeygard (2016) find that railroad and road connections in Africa have lasting effects on local development, notably by allowing for the emergence of new profitable activities thanks to reduced transport costs. Finally, conflicts and violence also tend to diffuse across space with their adverse consequences (Guidolin and La Ferrara, 2007; Berman et al., 2017).

In conjunction with these general effects, some distinct causes — mainly political — work predominantly through ethnic channels. For instance, the transborder ethnic kin bonds may facilitate the collective action required to take up arms in civil conflicts. Such support may take the form of financial and military assistance or a safe haven for the retreat of rebels.[6] Nonetheless, as Gleditsch (2007) and Gurses (2015) argue, these ethnic bonds should not necessarily foster conflict but might instead give rise to superior general outcomes for the kin factions by shifting the balance of political power in their favour. For example, in the context of weakly institutionalised states, the share of public expenditure allocated to a faction may be proportional to its expected utility in case of opposition

[4] See Krugman (1991), Acemoglu et al. (2001) and Klenow and Rodriguez-Clare (2005).

[5] On the effects of trade, see Tinbergen (1962) and Fujita and Thisse (2013). On technology diffusion, see Klenow and Rodriguez-Clare (2005) and Ertur and Koch (2007). On labour market opportunities, see Lucas (1990), and on migration, see Giuliano and Ruiz-Arranz (2009), Yang (2011) and Giovanni et al. (2015).

[6] See Cederman et al. (2009), Aidt and Leon (2016) and Haynes (2016).

(outside option), as predicted by theoretical models of ethnic bargaining.[7]

Another important channel where ethnicity plays a major role is the formation of risk sharing networks (De Weerdt and Dercon, 2006; Dubois et al., 2008; Ambrus et al., 2014). In the developing world where rainfall or illness shocks generate significant income fluctuations and where people lack formal insurance opportunities, these informal arrangements, which entail gifts and transfers, provide a remedy against adverse consequences of these misfortunes. They tend to follow kinship lines to resolve information and enforcement issues.[8] Marriage, migration and remittances form distant solidarity bonds to benefit from a spatially diversified network that may therefore be resilient to local economic shocks (Ioannides and Datcher Loury, 2004; Munshi and Rosenzweig, 2016). Recently, the possibility to carry out cash transfers through mobile phone instant messaging has greatly improved the efficiency in Africa, by reducing theft and delay (Yang, 2011; Jack and Suri, 2014).

One may suspect that these ethnic externalities operate in parallel with those of the general type, reinforcing them in some cases. For instance, a group in power could decide to construct a road to improve the connection with their kin abroad, who would benefit economically from easier trade or other newly available opportunities.

A concurrent explanation of the effect of dominant transborder kin groups could be that it is, in fact, merely a consequence of spatial spillovers because kin ethnicities are often located on both sides of a country's border. In particular, a positive effect could result from interactions distinct from a shift of power in internal politics and instead might stem from the enrichment of the neighbouring dominant group even if nothing operates specifically through the ethnic bond. This generates an identification issue that we will address in this paper by distinguishing and modelling those ethnic and spatial interactions, for the first time in this literature.

In short, we will investigate whether transborder ethnic relations generate externalities through increased central government participation; and whether these externalities have measurable consequences for local prosperity. An example of transition on which our identification strategy relies is Mali in 1993, when the Blacks coalition uniting the Peul, Mande and Voltaic people gained a monopoly over power to the detriment of the northern Tuareg and Moors groups. We look at the consequences of this power shift in neighbouring Guinea for the powerless Malinke and Peul groups, in a country where the Susu group was dominant at that time.

To investigate these issues, we estimate spatial panel models that relate transborder politico-ethnic variables to local luminosity measured by satellite imaging. The units of observation in this study are *spatial units*, i.e., small geographic areas. In our baseline specification with 2,929 spatial units and 21 consecutive years from 1992 to 2012, we find that having more politically dominant transborder ethnic kin significantly increases economic activity as measured by luminosity in the corresponding ethnic homeland. Our Spatial Autoregressive (SAR) estimates of the coefficient of the variable Ethnic Kin Dominant (EKD) imply that, on average, a group gains around 2.7 per cent in light intensity whenever a transition to dominance occurs among its neighbouring and related ethnic

[7] See Driscoll (2008, 2012), Bidner et al. (2014) and Francois et al. (2014, 2015).

[8] See Grimard (1997), La Ferrara (2003), Fafchamps and Gubert (2007), and Leider et al. (2009)

network. This corresponds to an approximate 0.9 per cent rise in income in the long run. Such an effect remains positive and statistically significant over a broad range of alternative specifications and robustness checks.

1.3 Related Literature

Our work is related to the literature that uses night-time light as a measure of development. In a celebrated seminal paper, Henderson et al. (2012) introduce these data in economics as a measure of growth because it is less prone to errors in countries that notoriously report highly contaminated GDP figures, like in Africa. The global coverage and fine-grain of the data have further advantages. For instance, using these data, Alesina et al. (2016) construct measures of ethnic inequality and show that they provide better predictors of sluggish development than traditional ethnic diversity indices.

Papers exploring partitioned ethnicities on the African continent, such as Michalopoulos and Papaioannou (2013a), on long-run comparative development, and Michalopoulos and Papaioannou (2016), on social unrest, emphasise the empirical relevance of ethnic ties. The former show that precolonial centralisation positively affects modern wealth. The latter argue that the drawing of boundaries by colonial powers in the nineteenth century was unrelated to the location of historical homelands of African ethnicities and that the resulting separations can thus be regarded as random. The authors use this accident as a quasi-natural experimental setting and show that the prevalence of conflicts is higher in the long run in regions that are inhabited by split groups.

Our work is akin in spirit to that of Hodler and Raschky (2014) and De Luca et al. (2018). The former paper assesses the effect of the leader's birthplace on regional favouritism as measured by changes in night-time light intensity in administrative regions, while the latter evaluate the effect of the leader's ethnicity on ethnic favouritism in ethnic homelands. Both these papers find positive correlations. Our study thus extends the work by Hodler and Raschky (2014), who were the first to assemble a panel dataset of night-time light intensity for subnational units to study the effect of the political leader's birthplace. However, we much broaden the scope of the investigated effects by (1) considering political participation by ethnic constituencies instead of the sole position of the president, (2) moving away from the apportionment in administrative regions to consider ethnic homelands combined with fixed spatial grids and (3) dealing with ethnic and spatial externalities.

Our study is also related to the huge theoretical and empirical literature on economic growth, such as Barro and Sala-i-Martin (2004), and the institutional economics literature that examines the root causes of development, such as Acemoglu et al. (2001) and Glaeser et al. (2004). Many authors concentrate on the consequences of autocracy, democracy or other types of power-sharing arrangements and the transitions between these political states.[9] However, *no evidence of transborder kin solidarity related to power participation is discussed in the literature.*

[9] Meyersson (2016) examines the economic consequences of coups, while Papaioannou and Siourounis (2008) challenge the view that democratisations do not have a favourable impact on growth considering consolidated transitions only. Mitton (2016) considers the role played by geography and institutions in a panel of sub-nations in 101 countries.

We focus our attention on Africa for two main reasons: first, because of the importance of ethnic ties that cross national borders (Englebert et al., 2002); and second, because of the prominent role played by ethnicity in African politics (Posner, 2004). Finally, the similar geographic and cultural features of many African regions make the parameter homogeneity hypothesis in our estimations much more likely than in a global sample (Durlauf and Johnson, 1995).

The next section presents our empirical models and estimation strategies. In Section 3, we discuss the econometric issues that are related to our specification choices. Section 4 presents the data construction. The baseline results and robustness checks are reported and analysed in Section 5. Finally, Section 6 concludes.

2 Empirical Strategy

Our cautious empirical strategy is based on standard panel data methods, in combination with the the estimation of sophisticated spatial econometric models. This ensures that our result of interest is not caused by the peculiarities of spatial estimation procedure. Moreover, it facilitates comparability with the existing literature that is restricted to classical fixed-effects models.

2.1 Panel Data Model

Our first empirical specification is based on a standard panel data model with the measure of development as the dependent variable, as in De Luca et al. (2018). Expression (1) describes the equation that we estimate:

$$\begin{aligned} \text{Log(Light)}_{i,t} &= \gamma \text{EKD}_{i,t-1} + X_{i,t-1}\beta + \alpha_i + \delta_{c(i),t} + \epsilon_{i,t}\,, \\ &\quad \text{for } i = 1, \dots, N \text{ and } t = 1, \dots, T\,, \end{aligned} \tag{1}$$

where $\text{Log(Light)}_{i,t}$ is our dependent variable: the logarithm of total light intensity emitted by spatial unit i in year t.[10] The variable $\text{EKD}_{i,t-1}$ (*Ethnic Kin Dominant*) is our independent variable of interest. It is the number of politically dominant ethnic friends of the ethnic group located in spatial unit i at time $t-1$ (described in detail in Subsection 4.3). We choose to lag all explanatory variables by one year instead of using the contemporary values to allow for delays and to mitigate for the occurrence of simultaneity issues in the estimations.[11] The variable EKD is constructed using the information contained in the Transborder Ethnic Kin supplementary material of the Ethnic Power Relations (EPR,

[10] We follow the literature and use the Sum of Lights (SoL) statistic, which is the sum of the Digital Numbers (DNs) of the pixels of the Spatial Unit.

[11] We specify a yearly panel for two reasons: first, to exploit all the information contained in the 21 years of light data; and second, to match the timing of the observed changes in the political status of the ethnic groups.

Vogt et al., 2015) database. For each ethnic group, we can identify their transborder ethnic connections that generally result from ethnic partitioning and which can, therefore, be considered as exogenous according to the same arguments proposed by Michalopoulos and Papaioannou (2016). The EPR-TEK database contains the political status of these groups at each point in time in their respective countries. Consequently, we can compute the time-varying EKD index that reflects the political transitions in the sample.

Vector $X_{i,t-1}$ contains $k-1$ control variables, encompassing an indicator variable for political dominance in domestic politics. We also control for population, meteorological factors (precipitations and temperatures) and the occurrence of civil conflicts.

The α_i's are the spatial unit fixed-effects that control for all the covariates that do not fluctuate over time.[12] This includes many geographical factors, such as the presence of mountains or rivers and the distance to the coastline or the capital city. The notation $\delta_{c(i),t}$ stands for country×year fixed-effects.[13] In other words, these unobserved variables account for any factors fixed over each combination of country and year. It is important to integrate these controls as Pinkovskiy (2017) demonstrates the presence of growth discontinuities at borders. These variables control for a broad range of observed and unobserved factors that may vary over years but still characterise a country's context, such as major institutional changes or shifts in its international relations with aid donors or commercial partners. These shocks or events would, therefore, not generate omitted variable biases in our estimates. Finally, $\epsilon_{i,t}$ is an error term such that

$$E[\epsilon_{i,t} \mid \mathrm{EKD}_i, X_i, \alpha_i, \delta_{c(i)}] = 0, \qquad \text{for } i = 1, \ldots, N \text{ and } t = 1, \ldots, T\,,$$

which is the usual strict exogeneity assumption with EKD_i, X_i and $\delta_{c(i)}$ being the respective notations for stacking all periods of $EKD_{i,t}$, $X_{i,t}$ and $\delta_{c(i),t}$ and

$$V[\epsilon_i \mid \mathrm{EKD}_i, X_i, \alpha_i, \delta_{c(i)}] = \sigma^2 \times I_T, \qquad \text{for } i = 1, \ldots, N\,,$$

where ϵ_i is the vector of the stacked $\epsilon_{i,t}$ and σ^2 is a scalar parameter. The latter assumption can be relaxed by considering more general forms for the variance-covariance matrices of the idiosyncratic error vector, while still preserving consistency. Moreover, the hypothesis of strict exogeneity can be relaxed into a hypothesis of pre-determinedness if wished. In that case, robust estimators of the standard errors must be used, as we did. By piling all individual observations numbered by $i = 1, \ldots, N$ of each cross-section in a certain year t in vectors and matrices, we obtain the following matrix expression of the model:

$$Y_t \;=\; \gamma \mathrm{EKD}_{t-1} + X_{t-1}\beta + \alpha_0 + \delta_{c,t} + \epsilon_t\,, \qquad \text{for } t = 1, \ldots, T\,, \tag{2}$$

[12] Random-Effects specifications are rejected by the results of Hausman tests in favour of Fixed-Effects specifications (with p-value below 0.0001).

[13] $c(i)$ denotes a mapping of the index of a spatial unit to the index of the country that contains it.

where we define Y_t to be equal to $\text{Log(Light)}_t = (y_{1,t}, y_{2,t}, \ldots, y_{N,t})'$, the $N \times 1$ vector of the logarithm of luminosity for all spatial units in year t. Our main independent vector of interest is EKD_{t-1}, the $N \times 1$ stacked vector of the numbers of neighbouring kin groups that are politically dominant at $t-1$. Similarly, X_{t-1} is an $N \times (k-1)$ matrix of time-varying controls, again lagged one year, whereas α_0 is an $N \times 1$ vector of spatial unit fixed-effects. The notation $\delta_{c,t}$ stands for the $N \times 1$ vector of country×year fixed-effects common to all spatial units in country c at year t.[14] Finally, $(\gamma', \beta')'$ is a $k \times 1$ parameter vector to be estimated, along with the means of the fixed-effects α_0 and $\delta_{c,t}$, $t = 1, \ldots, T$.

2.2 Spatial Panel Data Models

So far, two of the lingering concerns with our model is the possible occurrence of spatially correlated omitted variables, on the one hand, and spatial heterogeneity and spatial spillover effects, on the other. To deal with these issues, we specify spatial versions of our model. Methods for the maximum likelihood estimation, testing, and inference of this kind of panel data spatial econometric models have recently been developed.[15] To take spatial interactions into account, we extend equation (1) by adding a spatially-lagged dependent variable. Our baseline specification is the Spatial Autoregressive (SAR) model, which is a special case of the family of models analysed in Yu et al. (2008) and Lee and Yu (2010).[16] As in the previous subsection, we proceed stepwise by first presenting the model in observation-by-observation form, which is now:

$$\begin{aligned} Y_{i,t} \;&=\; \rho \sum_{j=1}^{N} w_{i,j} Y_{j,t} + \gamma EKD_{i,t-1} + X_{i,t-1}\beta + \alpha_i + \delta_{c(i),t} + \epsilon_{i,t}, \\ &\quad \text{for } i = 1, \ldots, N \text{ and } t = 1, \ldots, T. \end{aligned} \tag{3}$$

The new term in (3), as compared with (1), involves a single synthetic spatial regressor that is coupled with the coefficient ρ, which is denominated as the 'spatial autoregressive parameter' (Kelejian and Prucha, 2010). The associated variable is a linear combination of the dependent variable of the neighbouring units, with spatial weights described by $w_{i,j}$. We discuss their formulation in greater detail in the next section. There are alternative ways of setting them, but in general, a positive value of the weight $w_{i,j}$ indicates that units i and j are neighbours in some sense, in this case geographic contiguity, and $w_{i,j}$ is equal to zero otherwise. A unit is never a neighbour to itself, that is: $w_{i,i} = 0$,

[14] In some specifications, we replace the country×year fixed-effects by year fixed-effects noted δ_t, which are constant for each year in the sample.

[15] See Baltagi et al. (2003), Kapoor et al. (2007), Yu et al. (2008), LeSage and Pace (2009), Lee and Yu (2010) and Elhorst (2013).

[16] Using various robustness checks, we show that our results stand for the broader range of specifications proposed in these papers, which includes, for instance, adding a temporarily lagged dependent variable or a temporally and spatially lagged dependent variable.

for all i. These weights are constant over time and they are non-stochastic. Finally, the $\epsilon_{i,t}$ represent the error terms. Although the maximum likelihood estimation rests on the assumption that their conditional distribution is Gaussian with mean zero, this is not necessary for obtaining consistency and consistent standard errors.

By piling up equation (3) as above, using the $N \times N$ spatial weights matrix W whose $(i, j)^{th}$ element is $w_{i,j}$, the general form of our spatial model reads:

$$Y_t = \rho W Y_t + \gamma EKD_{t-1} + X_{t-1}\beta + \alpha_0 + \delta_{c,t} + \epsilon_t, \qquad \text{for } t = 1, \dots, T, \tag{4}$$

where $(\gamma', \beta', \rho)'$ is a $(k + 1) \times 1$ parameter vector to be estimated, again along with the means of the fixed-effects α_0 and $\delta_{c,t}$, $t = 1, \dots, T$. We refer to WY_t as the spatial lag of Y_t. In our baseline estimations, W is an $N \times N$ spectral-normalised matrix of spatial weights. While our findings are robust to a large spectrum of alternatives concerning the construction of the spatial weights matrix, its normalisation, and model specifications, the present specification is convenient as it allows for spatial dependence between geographically close units through the dependent variable. In sub-section 5.3, we shall further extend this model by introducing dynamics and diverse interaction terms.

3 Econometric Issues

In this section, we discuss the main econometric issues and the strategies that we use to deal with them.

3.1 Omitted Variable Bias

The endogeneity of our political variables may be of concern if there are local unobserved characteristics at the ethnic group level, such as social norms and local institutions, which would be correlated to its economic activity, on the one hand, and political domination by an ethnic group or its neighbouring kin abroad, on the other hand. In such a case, our estimates may be inconsistent because they would reflect the impact of both the political edge variables and the unobserved omitted variable. However, this would only happen whenever these omitted factors change over time differently across regions, since spatial unit fixed-effects and country×year fixed-effects are included. In other words, all the factors that correspond to common events, shocks, or characteristics at the country level, or that are constant in time, are perfectly dealt with by our estimation strategy. Therefore, our country-by-year fixed-effects control all factors that we can think of which could cause this type of endogeneity issue. Only the remaining co-movements between our explanatory variables and unobservables must be acknowledged. Moreover, we incorporated a few additional time-varying controls, namely meteorological factors and conflict incidence. Then, we believe that we have surveyed and included all the possibly available data with geographic and temporal variations and sufficient coverage. As a result, our estimates should entail the lowest possible amount of omitted variable bias.

3.2 Reverse Causality

Reverse causality may also cause concern. The chain of causation claimed in this paper runs in the following way. In one country, say country A, an ethnic group, say A1, gains monopoly over central power at the expense of the other groups in country A : A2, A3 and so on. The effect that we purport happens in country B, which is A's neighbour. We investigate whether the change described above affects, on average, the local level of prosperity as measured by light, perhaps through diverse political or economic mechanisms in favour of group B1, which has ethnic ties with A1. One channel through which this effect may operate is the change in the balance of power between the ethnic groups in country B, as depicted by Francois et al. (2015). However, as discussed in the Introduction, one cannot exclude other possible channels triggered by the ethnic connection. In these conditions, a typical chain of events that would produce reverse causality would be a situation in which group B1 gains economic importance, thereby causing a political transition in country A. Likewise, a poorer ethnic group may generate larger external effects in the form of grievances on neighbouring kin groups, inducing them to enter the political arena.

However, this source of endogeneity is benign for several reasons. A first reason is that we lag all our explanatory variables by one year, which rules out short-term reverse effects. Indeed, for such reverse effects to arise, there should be rational expectations about ethnically heterogeneous local future income shocks affecting a political transition for some ethnic kin abroad. This kind of hypothesis sounds like a long shot. First, these expectations are difficult to form. Second, they would have to be devised by less informed foreign actors. Therefore, such prospect looks very unlikely here.

An additional reason relates to the nature of the variation profiles of these two elements and the scale at which they are measured. Light is measured for small geographic areas, while ethnic power participation happens at the national level. Moreover, the political variables are persistent with rare transitions, unlike the emitted lights that follow a trend captured at the country level by the fixed-effects. These contrasts between scales and dynamic features mitigate the concern for reverse causality. As a matter of fact, spatial unit fixed-effects, year fixed-effects and country$\times$year fixed-effects should control for most reverse causality patterns, if they exist. Finally, the spatial autoregressive specification is also likely to absorb many of the reverse causality effects through the spatial interaction term. Even with this tight requirement, dynamical effects not present in the model could still generate problems. To deal with this concern, we estimate extensions of the model with a time-lagged dependent variable or spatial- and time-lagged dependent variable, or a combination of the two to capture these potential sources.

3.3 Measurement Error

The econometric strategy is based on a heavy introduction of controls in the specification — including climatic variables and country-by-year fixed-effects — and can, to some extent, be seen as a substitute for unsafe instrumentation. However, the issue of measurement errors is vital for the estimation strategy. Indeed, the simultaneous inclusion of spatial unit fixed-effects, year fixed-effects or country-by-year fixed-effects, and an autoregressive spatial term may exacerbate the impact of data contamination on the identification and the accuracy of the estimates. The success of the estimation greatly relies on the exceptionally intensive work in measuring, constructing and cleaning the main data used, especially for the light, ethnic, and political variables.

For instance, the location of the ethnic groups contained in the Geographic Ethnic Power Relations (GeoEPR, Wucherpfennig et al., 2011) database could be mistaken by a non-negligible margin. We address this issue by referring to Wucherpfennig et al. (2011), who describe the meticulous methodology that was carried out in the digitization of these data, updated from the earlier GREG (Geographic Representation of Ethnic Groups, Weidmann et al., 2010), and based on the work by numerous country experts. GeoEPR improves upon GREG by considering the ethnic cleavages that are politically salient, in coordination with the EPR database. By focusing on the politically relevant groups, EPR reduces the number of listed groups and, therefore, attenuates the errors that may be associated with smaller and more obscure factions. The strict geographic digitization procedure required experts specialists of each country to provide a map of the ethnic group's settlements. Evidently, our measures rely on the settlements that can be spatially identified; that is, the regionally-based groups that constitute most of the entries, at least in Africa. These maps were subsequently passed on to Geographic Information System (GIS) experts, who then transformed them into machine-readable information. A final step involved cross-validation by fellow experts.

Errors in the coding of the political status of the ethnic groups might arise, although they are little likely for the type of transition that we use, namely, between political domination and other forms of control, because this is the clearest cut. Our 'Dominant' political status corresponds to the levels of power 'Monopoly' and 'Dominant' in the EPR conventions.[17] The measured Transborder Ethnic Kin reciprocal links are most convincing because they derive from inherited associations between great political players and are, therefore, manifest and easy to identify with precision.

Our geographically-based variables — Light, Population, Meteorological Factors and Violence — are also probably measured with some error, but, in the absence of systematic errors, this should not generate estimation biases. We correct the lights' DNs (Digital Numbers) with the procedure of Elvidge et al. (2014) and adjust for gas flares using masks.

[17] Compared with 'Monopoly', 'Dominant' allows for token representation by a minority ethnic group; that is, puppet representatives without any effective power.

4 The Data

This section describes the information that we use. We start with the construction of the spatial units.

4.1 The Spatial Units

We construct our dataset by combining georeferenced information using a Geographic Information System, which is a computer program specialised in the treatment of geographic information. These data come in the form of shapefiles for areas, lines, and points or in 'rasters', which are chequered images of pixels that each have a numerical value. The light data comes in the form of a raster and it is possible to calculate the sum of the values associated with the pixels falling in a given region defined by a shapefile. There are 194 relevant ethnic regions from the GeoEPR dataset in the whole period under consideration, 1992–2012.

Our geographical grid is constructed using the International Conflict Research's GeoEPR dataset, which is a geocoded version of the EPR dataset that contains the regional homelands of the politically relevant ethnic groups across space and time (Wucherpfennig et al., 2011). This methodology allows us to identify the ethnicity of the population living in each spatial unit for each year. GeoEPR derives from the Geographic Representation of Ethnic Groups (GREG) dataset (Weidmann et al., 2010) that digitizes the global information of the Atlas Narodov Mira in GIS format. Unlike GREG, which is based on the outdated data collected by Soviet ethnographers in the 1960s, GeoEPR reflects the changes over time in settlement patterns resulting from migration or ethnic cleansing, for instance.

Furthermore, GeoEPR codes the actual spatial distribution of ethnic group members rather than their ancient kingdoms and the most significant changes in location are captured by an expert assessment procedure. Finally, thanks to its association with the EPR list, GeoEPR aggregates the groups based on their political relevance instead of language criteria, such as in the Atlas and GREG. This distinction is of importance for our analysis because the hypothesised transmission mechanism may be political.[18]

Figure I represents the ethnic regions for the whole continent. No politically relevant ethnic groups are based in the unshaded regions. Our sample spans over 48 countries. Some of the countries of the sample are not found in the EPR data: Tunisia, Somalia, Burkina Faso, Equatorial Guinea, Lesotho and Swaziland. However, we still include them in the analysis because a simply connected space is necessary for correct spatial estimations, and because they do not hamper the estimation thanks to the presence of fixed-effects. We checked that the results of the panel estimation stay valid when we exclude these countries.[19]

To construct our subdivision, we combine information on the geographical location of the ethnopolitical groups with the borders of the countries of the African continent and a grid similar to

[18]For example, the Hutus and Tutsis in Rwanda constitute a single linguistic group in GREG but they are listed separately in GeoEPR.

[19]We replace the values of the ethnic variables by zeros for these observations.

that of Alesina et al. (2016): that is, a mesh fitted together with meridians and parallels of the globe. Our units of analysis are the ensuing 2,929 disjoint areas forming a partition of the territory, in the baseline case.[20] In our baseline estimations, the grid has a comparable size of 2 by 2 degrees, in place of Alesina et al. (2016)'s 2.5 by 2.5 grid.[21] A too thin grid would excessively increase the number of observations and prevent the estimation of the spatial models that involve maximisations of likelihoods that are cumbersome. However, we want our spatial unit to be small enough to have a clear interpretation, distinguishing, for instance, coastal areas from the hinterland. This allows for a more precise specification of the geographical unobserved heterogeneity through fixed effects. Fixed geographical characteristics, such as the distance to the border, which may inform on the distance to the transborder ethnic kin, or the distance to the capital city of different places of the same ethnic homeland are encompassed in these fixed effects. We also want to identify the effect in places where the homelands of more than one group overlap, and the fixed effects help us to account for this situation. Besides, creating a distinct spatial unit at such intersection allows us to relate luminosity to the combination of the foreign political links of the two overlapping groups.

We illustrate this procedure for Guinea, which is Mali's neighbour, in West Africa. Guinea is a former French colony that gained independence in 1958. The largest three ethnic groups, which make up 86 per cent of the population, are: the Fula (also called Peul), the Mandinka and the Susu. The Fula are Muslim nomadic pastoral people with their own language. Nigeria, Guinea and Mali are the countries inhabited by the largest Fula communities, while this group also has connections in Guinea-Bissau, the Gambia, Senegal, Mauritania, Niger and Cameroon, with Hausa, Toucouleur and other Fula people. The Mandinka people are also Muslim and heirs of the Mali Empire. They are part of the Mande family and they enjoy an extensive network of connections involving the Mandinka in Gambia, the Mande in Mali, the Mandingue in Senegal, the Mande and Gur in Cote d'Ivoire, and the Indigenous Peoples and Mandingo in Liberia. The Susu people form a patrilineal society and they speak their own language. They are descendants of the animist local Mande people that historically converted to Islam following the Fula domination. Figure II shows the location of the politically relevant ethnic groups in Mali and Guinea. In Mali, the map shows the Arabs and Moors in the northern deserts, the Tuaregs in the east, and the Blacks coalition in the south. In Guinea, the Malinke reside in the east, the Peul reside in the centre, and the Susu reside in the coastal region.

This example is typical of our identification strategy in that it relies on political transitions. At the time of the settlement of the Tuareg rebellion in 1993, Bréhima Siré Traoré — a representative of the central government — and Zahabi Sidi Mohamed — a spokesperson of the rebel armed forces — signed a national pact that granted a particular status to the northern parts of Mali and de facto control of

[20] There are 43 countries with information on politically relevant ethnic groups. The mean number of Politically Relevant Ethnic Groups by country is 5.48 with a standard deviation of 3.36. The mean number of spatial units by country is two (in Swaziland and Rwanda) and the maximum is 302 (in Sudan). The average number of spatial units is 59.77 with a standard deviation of 59.71. Concerning the area of the spatial units, the average is 8,665.5 square kilometres with a standard deviation of 11,033.39 square kilometres.

[21] We reconstructed the data with alternative grid spacings, e.g., 3 by 3 and 4 by 4, and found that the results are robust with our baseline specifications.

the central state to the Blacks coalition. This coalition has ethnic connections to the Malinke and Fula groups in Guinea, as shown in Figure III, along with the Susu group, who was politically dominant there. Figure III illustrates how we use a grid to construct our subdivision. The horizontal and vertical dashed lines are two degrees apart and they partition the territory *in combination* with the ethnic zones and country borders. In places where some groups overlap, which are depicted by the two crossed regions, a separate unit is created to distinguish the effect of the political transitions in these areas from those where a single group dwells.

4.2 The Dependent Variable

Log Light. The dependent variable in (1) and (3), 'Log Light', is the logarithm of the pixels' total intensity of luminosity in the considered area. To form our proxies of local development, we use satellite images of luminosity at night, which are available from the National Oceanic and Atmospheric Administration. The initial datasources are satellite measures from the United Air Force Defense Meteorological Satellite Program. The National Oceanic and Atmospheric Administration processes the raw daily data by removing all non human-made lights, which for instance, may be due to the bright half of the lunar cycle, daytime lights in summer months, forest fires, and the northern and southern aurora lights. An average over all valid days, with measures taken from 8:30 pm to 10:00 pm in the absence of cloud coverage, is converted to satellite-year raster data that are subsequently distributed to the public. When data exist from two satellites for the same year, we take the average to construct our variables. The data are recorded from 1992 onwards and are currently available for the years up to 2013.[22]

Henderson et al. (2012) demonstrate that night light data are a useful proxy for economic growth that tracks its short-term fluctuations. They find a statistically significant elasticity of GDP with respect to light intensity of 0.3. The availability of these data for almost all inhabited parts of the world and the possibility to adjust the spatial unit of observation to the research question are of crucial importance for our purpose.

We use the method proposed by Elvidge et al. (2014) to attenuate measurement errors in light intensities. This correction is also necessary because the internal setting of the measurement device changes how lights are amplified between years. We notably apply a mask to suppress gas flares. Extensive discussions of these issues can be found in the remote sensing literature, for instance in Bennett and Smith (2017).

4.3 The Main Explanatory Variable of Interest

Ethnic Kin Dominant (EKD). This variable is a count taking integer values between zero and three in our sample. It counts the number of kin groups that are politically dominant in neighbouring countries.

[22] We use the years up to 2012 because of the limitation imposed by Elvidge et al. (2014), that provide coefficients to adjust the light data up to that year.

Countries are neighbours here in the sense that their borders touch each other, in the baseline case.[23] Specifically, we follow the definitions of political statuses proposed by Vogt et al. (2015), i.e., a group in power is dominant when it does not share power with other ethnic groups.

The last version of the EPR data is used to construct the network of transborder ethnic bonds. For each group, this takes the form of the list of ethnic groups that are their related kin, including ethnic groups spread over several countries. To construct our variable 'Ethnic Kin Dominant', we first count, for each ethnic group, the number of their transborder ethnic kin, which are coded 'politically dominant', while restricting the scope to neighbouring countries to exclude the link between far apart Arabs in Morocco and Arabs in Egypt, for instance. When more than one ethnic group occupy a spatial unit, our aggregation rule in the baseline case uses the maximum number of ethnically related dominant neighbouring groups to construct EKD, to avoid double counting.[24]

4.4 The Controls

To limit omitted variable bias, we control for many factors that may affect the dependent variable at the local level. Aidt and Leon (2016) stresses the importance of the meteorological factors in agriculture-based African economies. We also control for violence hampering local development (Guidolin and La Ferrara, 2007; Berman et al., 2017). Here is the list of the other (non-fixed) controls that we use in our estimations:

Dominant Ethnic Group (DOM). This is a binary variable taking the value one in year t if the ethnic group of the spatial unit is the only one represented in the central government of the country in this year, and zero otherwise. We adopt the definition of political representation in the sense of the EPR data source (Vogt et al., 2015); that is, whenever the group has a monopoly over central power or is sole in charge with the other groups enjoying only token representation.

GeoEPR enable us to measure the access of the groups to central state power, as delineated by executive public positions.[25] An ethnic group is coded as participating in the central state decision instances if some of its members occupy the presidency, cabinet minister positions, or senior posts in the administration or army. Francois et al. (2015) and Cederman et al. (2010) show that this is likely to affect distributive policies. To construct our political variable 'Dominant' in spatial units with several ethnic groups, we use the maximum value among these groups. Using the total instead preserves the result.

[23] Or within a 100 or 1,000 km radius of each other in alternative specifications for which the results are qualitatively similar.

[24] This is the most conservative practice. Using the total value instead preserves the result. In our baseline sample, the proportion of spatial units with single ethnic group is 53 per cent (1,549 out of 2,929). There are 888 spatial units with no politically relevant ethnic group (30 per cent), 408 with two overlapping groups, 81 with three, and three spatial units with four overlapping groups.

[25] EPR uses a definition of ethnicity as a "subjectively experienced sense of commonality based on the belief of a common ancestry and shared culture". Unlike other lists of ethnic groups, its classification is based on how the groups act politically, see Vogt et al. (2015) for further details.

Population. To construct this variable, we use data from the Gridded Population of the World, which provides georeferenced population estimates between 1990 and 2010 from the Socioeconomic Data and Applications Center.[26] These data are available at five-year intervals, and we use a log-linear interpolation for the missing years. The variable included in the model is the logarithm of the total number of people in the spatial unit, lagged by one year. By controlling for population with a free parameter, we deal with Cogneau and Dupraz (2014)'s criticism on the methodology of Michalopoulos and Papaioannou (2013b), according to which the results may hinge on the combination of the sparsely populated areas and the functional form of the dependent variable. Besides, Michalopoulos and Papaioannou (2017) also recommend to insert population in the right-hand side.

Meteorological Conditions. We retrieve monthly rainfall and temperature data from the National Oceanic and Atmospheric Administration.[27] We average these monthly raster data to construct measures of yearly precipitations and temperatures everywhere in Africa, up to 2012. In each spatial unit, we then calculate the temperature mean value and the rainfall total value of the raster pixels for each year to construct our 'Mean Temperature' and 'Total Precipitations' variables.

Violence. For our measures of violent instability, we use the Georeferenced Event Dataset from the Uppsala Conflict Data Program, which provides the most accurate information on violent strife since 1975.[28] Our variable 'Violence' is a binary variable that takes a value one if civil violence incidents have occurred within an area-year.

Table I shows the basic descriptive statistics of the pooled data, for the 2,929 spatial units and 21 years in the baseline case. The mean of our dependent variable (Log Light) is 8.776 with a standard deviation of 1.833. The number of dominant ethnic kin connections ranges from zero to three and is on average 0.194. Likewise, 9.1 per cent of the observations have a politically dominant ethnic group. Table II shows the descriptive statistics of the between- and within-transformed panel data. For most variables, the between variability is larger than the within variability, which suggests that fixed effects should be an essential component of models. Moreover, the time-variability of the independent variables of interest, measured by the standard deviation of the within transformed data, is of comparable magnitude to its between variability (0.117 against 0.565 for 'Ethnic Kin Dominant', and 0.107 against 0.267 for 'Dominant'). This implies that panel data analyses should be appropriate to capture their effect. Table III shows the number of transitions of the variable 'Ethnic Kin Dominant'. The total number of changes across distinct states is 357 (104 from state 0 to 1, 174 from 1 to 0, 36 from 2 to 1, and 43 from 3 to 2). This means that, on average, close to 12 per cent of the spatial units of our sample experience a change during the period under consideration. Table IV shows the number of transitions between the values 0 and 1 of the variable 'Dominant' as our fixed-effects estimation relies on this type of transition. The total number of changes across distinct states is 310 (242 to dominance and 68 from dominance). On average, around 10 per cent of the spatial units of our sample experience

[26]See http://sedac.ciesin.columbia.edu/data/collection/gpw-v3

[27]UDel-AirT-Precip data provided by the NOAA/OAR/ESRLPSD, Boulder, Colorado, United States, from their website at http://www.esrl.noaa.gov/psd/.

[28]Downloaded from http://www.ucdp.uu.se/ged/.

this kind of change during the period under consideration.

5 Results

5.1 Baseline Estimations

Table V reports the baseline estimation results of (2); that is: the panel Fixed Effects model linking log-luminosity to our transborder kin index (the number of politically dominant neighbouring kin) and other controls; and of (4): the Spatial Panel Fixed Effects SAR model linking these variables plus a spatially lagged dependent variable. All estimates reported in Table V have robust estimators of standard errors clustered at the spatial unit level, and we use the dataset constructed with a two-by-two degrees grid, where the transborder ethnic kin relationships are limited to direct neighbour countries. All spatial models are estimated using a spectrally normalised queen-neighbours weights matrix. Columns 1 and 3 contain the Fixed Effects estimates, and Columns 2 and 4 contain the SAR estimates. All estimations control for spatial unit fixed-effects. In Columns 1 and 2, we only control for year fixed-effects, whereas in Columns 3 and 4, we control for country$\times$year fixed-effects instead. Finally, all columns in this table are based on the whole sample of African countries. As mentioned before, our most important independent variable of interest in this investigation is 'Ethnic Kin Dominant$_{t-1}$' (EKD). This variable receives significantly positive estimated coefficients at the one per cent level in all models. This is the main point of the paper.

In our preferred estimation in Column 4, the coefficient estimate for EKD is equal to 0.0275, which is slightly lower than the FE estimate of 0.0340. In the presence of spatial interactions, the FE estimates may be inconsistent. Furthermore, as mentioned in LeSage and Pace (2009), omitted variables affect spatial regressions less than least squares methods. This justifies our preference, which is also supported by the significance of the spatial coefficient ρ. Even if the coefficients of spatial models are not directly interpretable in terms of marginal effects (Elhorst, 2013), a useful rule of thumb to compare changes in light emissions and economic activity is to multiply the elasticity by the factor 0.3 that was computed by Henderson et al. (2012). For certain specifications, we explicitly computed these marginal effects by using the method of LeSage and Pace (2009) to break down the total effect into a direct and an indirect component along with the simulated standard errors. These are available upon request. We found that the direct effects have an amplitude and significance that are in line with the estimates reported in Table V. However, because of the complexities brought by the spatial patterns and the ethnic networks, the effects of EKD can hardly be captured solely by considering the value of its coefficient in the model. A more complete understanding of these effects can be achieved through simulations, as in the next sub-section.

Regarding the other time-varying controls, their significance depends upon the type of fixed-effects included in the equation. For instance, 'Log Population' has a positive coefficient, significant at the one per cent level with year fixed-effects, and a small and insignificant coefficient with country×year fixed-effects, because the population trends at the country level are captured by this more precise specification. 'Violence' has a small coefficient, insignificant in Columns 1 and 2, and significant in Columns 3 and 4, with country×year fixed-effects, at the five and one per cent level, respectively. Like 'Log Population', the variable 'Rainfall' is very significant in the year fixed-effects formulation, but becomes insignificant or mildly significant (10%) in the country×year fixed-effects specifications of Columns 3 and 4. The coefficient is positive, which is line with the fact that African economies rely heavily on agriculture, in which rainfall is of utmost importance. Among these variables, only 'Temperature' has a coefficient that retains a stable negative sign and significance over all columns. The coefficient stays in the interval [-0.0036,-0.0020] and is significant at the one per cent level in the FE specifications and at the five per cent level in the SAR specifications. This suggests, as expected, that too high temperature shocks are bad for local activity.

5.2 Simulations

5.2.1 Setup

In this subsection, we investigate the impact of power changes using two simulation experiments. In the first experiment, we proceed by artificially changing the power status of the demographically largest group in every country to 'Dominant'. This thought experiment corresponds to a situation where the political power stems from the population, in place of entrenched leaders and vested interests. In the second experiment, the power status of the second largest group is changed to 'Dominant' in every country. We compare the values predicted by the model in these alternative situations to the initially predicted values. It is important to proceed with simulations instead of direct calculations, because with our model that integrates transborder aspects, a change in the political situation in one country has repercussions in neighbouring countries that are not accounted for in standard econometric predictions. Note first that when the political status of a group is changed to dominant, the status of the other groups of the country is accordingly adjusted to non-dominant, and the EKD index must be recomputed with all these changes.

An additional difficulty arises when running these experiments because the spatial unit fixed-effects also capture local effects of political dominance in our baseline estimations because of the inertia of the political situations. Therefore, the within estimates do not necessarily include all the effects of the dominance variable. To circumvent this issue, we extract the estimated fixed effects of equations (2) and (4) and estimate their cross-sectional multivariate correlations with dominance, transborder connections and other controls. The equation estimated here over the cross-section of the 2,929 extracted spatial units is:

$$\hat{\alpha}_i = \overline{\mathrm{EKD}}_i \gamma^* + \bar{X}_i \beta^* + \epsilon_i, \tag{5}$$

where $\hat{\alpha}_i$ is the estimated fixed-effect for spatial unit i, from either the non-spatial panel model or the SAR model. Equation (5) is a standard linear model of regression on spatial-unit means over the

21 years of data for a few characteristics. $\overline{EKD}_i$ is the mean of 'Ethnic Kin Dominant', and $\bar{X}_i$ includes the means of 'Dominant', 'Log Population', 'Temperature', and 'Rainfall' and 'Log Surface'. γ^* and β^* are parameters that can be associated with parameters γ and β in the baseline model.

Table VI shows the estimates of an OLS regression where the dependent variable is the spatial unit fixed-effect from the model of equation (2), in Column 1. In this estimation, 'Mean Dominant' has a coefficient of 0.060, significant at the five per cent level and 'Mean Ethnic Kin Dominant' has a coefficient of 0.048, significant at the one per cent level, using the robust standard errors displayed in parentheses. Because 21 years already provide quite a few degrees of freedom for estimating each of the means and fixed-effects with reasonable accuracy, we can neglect to correct for their first-stage estimations, while estimating (5).

Nevertheless, assuming that our sample is representative of the population, it is possible to obtain bootstrap estimates of the standard errors. We perform 1,000 bootstrap iterations, estimating (2) from each randomly drawn sample of spatial units from the original data.[29] Each time, we extract the fixed-effects from this estimation and build a cross-section to estimate (5). We store the estimates of γ^* and β^* from the 1,000 replications and then estimate the standard errors using these distributions. The bootstrap standard errors are displayed in square brackets. They have comparable magnitudes to the robust asymptotic standard errors. The significance levels of all independent variables are similar to those for the previous statistical tests, except for the variable 'Dominant', for which it is now 10 per cent.

In Column 2 of Table VI, we display the estimates of an OLS regression where the dependent variable is the estimated spatial unit fixed-effect from the model of equation (4). In this estimation, 'Mean Dominant' has a coefficient of 0.078, significant at the five per cent level and 'Mean Ethnic Kin Dominant' has a coefficient of 0.038, significant at the one per cent level, using robust asymptotic standard errors.[30] These estimates, which capture long-term impacts, indicate that the effect of local dominance on luminosity is about twice as large as that of dominance by transborder kin groups. Even if we are not necessarily making strong causality claim from these estimates, we use them to construct the predicted values in our simulations, by adjusting the fixed effects according to the correlations stemming from Table VI.

5.2.2 Simulated Values

Experiment 1

In Table VII, we report the mean and the standard deviation of the distribution of the outcome variable and their changes for these experiments. These statistics are disaggregated in four sub-samples according to two criteria: (1) 'Ethnic Network' that refers to the spatial units of the groups

[29]This sample is clustered by spatial unit, i.e., when some spatial units are selected at random, all years of those spatial units are also selected, and stratified at the country level, i.e., the bootstrap sample contains precisely the same number of spatial units per country as the original sample.

[30]It is not technically feasible to obtain bootstrap standard errors as previously, because each iteration would last about three days to estimate the SAR model.

that have indirectly benefited from changes in political status, i.e., they have transborder ethnic connections to the groups that have benefited from the experimented change in power status; 'Ethnic Outsiders' refers to the spatial units of the groups that have not indirectly benefited; (2) 'Favoured Groups' that refers to the spatial units of the groups that have directly benefited from changes in political status from the experimented change, i.e., their political status was changed to 'Dominant'; and 'Non-Favoured' refers to the spatial units of the groups that have not directly benefited from the change.

Table VII reports the summary statistics for the predicted distribution and the change in the predicted distribution for the four possible sub-samples. For comparison, we provide the descriptive statistics for the corresponding sub-samples for the predicted values before the experiment. We construct these four sub-samples to distinguish the effect of local dominance from the transborder politico-ethnic effect, controlling for spatio-economic interactions. The observations of the sub-sample 'Ethnic Network/ Favoured Groups' benefit from the two effects. The intermediate sub-samples 'Ethnic Network/ Non-Favoured Groups' and 'Ethnic Outsiders/ Favoured Groups' benefit from only one effect, either the transborder effect or the dominance effect, respectively. The last sub-sample, 'Ethnic Outsiders/ Non-Favoured Groups' does not benefit from any effect, but might instead suffer from them because they are affected by unfavourable power changes to their own political status or that of their ethnic connections, in addition to spatio-economic interactions.

For instance, we observe that, taking into account the adjustment of the fixed effects, the change in the predicted values of 'Log Light' in the spatial units of the category 'Ethnic Network'/ 'Favoured Groups' is almost equal to the sum of the estimated coefficients of 'Dominant' and 'EKD' in the baseline equation (4) and the fixed-effects equation (5).

The average variation in the dependent variable can be compared across the 'Favoured'/'Non-Favoured' categories, conditional on 'Ethnic Network'/'Ethnic Outsiders', to find out if there is a systematic difference. The increase in 'Log Light' for the observations which belong at the same time to the favoured ethnic groups and the ethnic network of other favoured groups is 0.1400, indeed. For comparison, the corresponding growth in the spatial units that correspond to the 'Ethnic Outsiders/ Favoured Groups' sub-sample is only 0.06. The corresponding statistics for the 'Non-Favoured Groups' category are 0.0779 and -0.0002, respectively. These comparisons suggest that local dominance and transborder ethnic effects have similar magnitude, and that the transborder ethnic effect is positive, controlling for local political status and spatio-economic interactions.

It is possible to inspect the whole distribution of the predicted values. In Figures IV and V, we can examine whether the effect varies, for example, at the top, middle or bottom of the distribution, or if irregular features arise from the slope of the density. Figure IV plots kernel density estimates of the distribution of predicted 'Log Light' after the simulated change, compared with the initially predicted values, by sub-sample. They are displayed for the 'Ethnic Network' and 'Favoured Group' sub-sample in panel 1, the 'Ethnic Network' and 'Non-Favoured Group' sub-sample in panel 2 and the 'Ethnic Outsiders' and 'Favoured Group' sub-sample in panel 3. A first unusual feature to note is that, in general, these distributions have two or more modes.[31] Because of the very large dispersion of the 'Light' variable with its logarithm ranging from around 4 to 13, we have to display the density of the logarithm of 'Light'. Note that changes that are economically significant will appear small. For instance, according to the numbers in Table VII, because the dependent variable is in logarithm, the apparently small shift to the right of around 0.14 in panel 1 corresponds to a luminosity change of 14 per cent, which is substantial. In panels 2 and 3, only one among the two effects is at work, the transborder effect in panel 2, and the local dominance effect in panel 3. While small, these effects are still visible on these graphs. Moreover, comparing panels 1 and 3, we observe that among the groups that benefit from the transition, the distribution of those that have some transborder ethnic kin that also benefits from this transition bears a substantial shift to the right. This corresponds to first order stochastic dominance, consistent with an increase of 0.1400, instead of 0.0620 on average in the absence of transborder connection. This result supports again the finding that transborder groups benefit from the political dominance of their related kin.

Figure IV suggests that all spatial units of the respective sub-samples are affected by about the same amount resulting in a shift to the right of the distribution, at least in panel 1. This characteristic is not surprising because of the assumed additive homogeneity of the effect of 'Ethnic Kin Dominant' in the model and the construction of the experiment. However, it may have been the case that the spatial interactions affect different types of spatial units to varying degrees. In that case, the shape of the distribution could have changed, as we shall see in the next experiment. However, this is not the case here, at least perceptibly in the figures.

Experiment 2

Let us now move to Experiment 2. We reproduce the previous analysis, but instead of changing the political status of the largest ethnic group of the country, we now change that of the second largest ethnic group in the country to dominant, which corresponds to a case of state capture by a politically powerful ethnic minority.[32] Therefore, the observations belonging to the sub-samples are modified accordingly. In the second panel of Table VII, we report the corresponding numbers for Experiment 2. There, the increase in 'Log Light' for the observations that belong at the same time to the favoured ethnic groups and the ethnic network of other favoured groups is on average 0.1557, compared with 0.0482 in the absence of connection to a favoured group, a difference even larger than in the previous case. The corresponding difference for 'Non-Favoured' groups is 0.0296, of comparable magnitude, though smaller than before.

Figure V plots the corresponding kernel density estimates for Experiment 2. The general pattern is similar to the previous figure, with an increase in the spatial units that benefit from the two discussed effects. The increase is smaller, but still visible in panels 2 and 3, where there is only one effect that operates. In contrast to the previous figure, the spatial interaction affects the shape of the simulated distribution, which cannot be merely described by a right shift of the distribution be-

[31] This is perfectly justifiable by the nature of the data, characterised by high variability, and the construction of the sub-samples, based on discrete politico-ethnic definitions.

[32] If this second largest group was not already dominant, obviously.

fore the experiment. This is discernible in panel 1, where beyond the general shift to the right, the slope of the distribution gains amplitude for values of 'Log Light' between around 7 and 9. Moreover, panel 1 shows that the observations between 9.5 and 11 lose, while those at the right of the distribution benefit the most. This heterogeneous-effect phenomenon might be linked to the location of the concerned ethnic groups in relation to their ethnic friends abroad in Africa and their spatial concentration. Indeed, whenever an ethnic group that benefits directly or indirectly from the transition, is geographically concentrated or close to its transborder ethnic kin group, the politico-ethnic and spatial effects reinforce each other. At the opposite, the effects of local dominance and transborder ethnic kin connections could be thwarted for remote groups.

5.2.3 Parameter Restrictions in the Simulations

We now investigate the relative importance of the various transmission channels of the political shocks by imposing nullity restrictions on some of the parameters of equation (4) in our simulations. We consider three restrictions. The first one cancels the spatial externality channel by imposing $\rho = 0$ (Case 1) for the term $\rho W Y_t$. This may be interpreted as removing economic side-effects like those resulting from technology diffusion, trade, and solidarity mechanisms.

Table VIII presents the results. The top panel corresponds to Experiment 1, and the bottom panel to Experiment 2. As in Table VII, the numbers display changes in 'Log Light' in the simulated experiment as compared with the baseline. The numbers in parentheses below the changes show the standard deviations in the corresponding sub-samples. By comparing Tables VIII-Case 1, where we impose $\rho = 0$, to Table VII, it can be seen that the effects of the political variables on luminosity are reduced compared with the case in which the spatial externalities are operative. This discrepancy may be the consequence of the spatial proximity of the groups that have politico-ethnic connections. Consequently, neglecting these spatial externalities in some policy experiments may obscure the relationship between politico-ethnic variables and local development across countries.

The second restriction cancels the effect of our variable of interest, 'Ethnic Kin Dominant' ($\gamma = 0$, Case 2). It corresponds to severing the channel of transmission between political power by an ethnic group in one country to the local economic activity of ethnic groups in the political network in neighbouring countries. Third, we remove the direct local effect of the variable 'Dominant' ($\beta_1 = 0$, Case 3).[33] When we impose these restrictions one by one, the other effect is still operative, in conjunction with the spatial externalities. In cases 2 and 3, we observe that the change in the sub-samples that benefit from certain types of political channels fade away when we exclude the corresponding channel. For instance, in Case 2, when we impose the constraint $\gamma = 0$, i.e., we cancel the effect of 'Ethnic Kin Dominant', the change for the observations of the category 'Ethnic Network/ Non-Favoured' becomes 0.0004 on average, instead of 0.0779 in Experiment 1, and -0.0121 instead of 0.0286 in Experiment 2.

[33] To obstruct the channels fully, when we impose the restrictions $\gamma = 0$ or $\beta_1 = 0$, we also impose the restriction on the corresponding parameter of equation (5), $\gamma^* = 0$ or $\beta_1^* = 0$, respectively.

Finally, we consider the effect of imposing two restrictions at the same time, each time ruling out one political channel and the spatial externality with $\gamma = 0$ and $\rho = 0$ (Case 4) and $\beta_1 = 0$ and $\rho = 0$ (Case 5).[34] In cases 4 and 5, again imposing $\rho = 0$ reduces the simulated effects of the politico-ethnic variables as compared with the case with spatial externalities (Cases 2 and 3). As in Case 2, imposing the constraint $\gamma = 0$ in Case 4 reduces the change in the sub-sample that benefit from 'EKD'. The analogous reduction occurs for the sub-sample that benefits from dominance in Case 5, after imposing $\beta_1 = 0$ as in Case 3.

The general lesson from this analysis is that severing a channel affects mainly the category of observations that was benefiting from it. The size of the effects of EKD and DOM, which are much larger than that of the spatial term, depend upon the experiment. For instance, in Experiment 1, the effect of EKD is stronger than that of DOM, while the opposite happens in Experiment 2. Lastly, cancelling the spatial term slightly reduces the effects of the political variables. This coincides with the results of Table V, where the estimated coefficients of EKD and DOM decline when the spatial term is included, because it captures a fraction of the effect that was formerly incorrectly attributed to EKD and DOM. In the next sub-section, we discuss diverse robustness checks and extended models.

5.3 Robustness Checks

5.3.1 Alternative Spatial Specifications

We show that the results hold for a variety of alternative specifications, not all shown for the sake of brevity. For instance, we use a spectral-normalised contiguity matrix as the weight matrix in (4), while several other matrices have also been used and deliver comparable results.[35] Table IX reports those estimates and we comment mainly on our variable of interest, which is Ethnic Kin Dominant.

The Spatial Error Model (SEM), characterised by a spatially auto-correlated error term, is described in the following equation, where the stochastic process ϵ_t is identically and normally distributed with mean zero:

$$\begin{aligned} Y_t &= \gamma EKD_{t-1} + X_{t-1}\beta + \alpha_0 + \delta_{c,t} + u_t \\ u_t &= \epsilon_t + \rho W u_t \qquad \text{for } t = 1, \dots, T. \end{aligned} \tag{6}$$

We also estimated an extension of this model by including spatially lagged explanatory variables, which is called the Spatial Durbin Model (SDM). This notably allows us to investigate whether the political effect of interest is specific to ethnic kin, by including spatially lagged explanatory variables

[34] We skip the presentation of the result with $\gamma = 0$ and $\beta_1 = 0$ because this combination of constraints absorbs all the effects of the simulated political change and is therefore little interesting.

[35] Spectral normalization corresponds to the transformation of the symmetric contiguity matrix, which is initially populated by zeros and ones. It consists of the division of all entries by the modulus of the largest eigenvalue of the original matrix. An alternative is row normalization, in which each row is divided by the corresponding row total. Both ways result in a spatial weights matrix that has eigenvalues below one, restricting the autoregressive parameter to be within the $[-1, 1]$ interval, to facilitate the interpretation.

of the political status 'Dominant' of the ethnic groups in the nearby spatial units. The equation of the Spatial Durbin Model is:

$$Y_t = \rho W Y_t + \gamma EKD_{t-1} + X_{t-1}\beta + \nu W Z_{t-1} + \alpha_0 + \delta_{c,t} + \epsilon_t, \qquad \text{for } t = 1, \dots, T. \tag{7}$$

The term $\nu W Z_{t-1}$ represents a spatially lagged exogenous independent variable: that is, Z_{t-1} is a subset of X_{t-1}, here the variable 'Dominant'. The SEM and SDM (Columns 1 and 2) are estimated using a spectrally normalised queen-neighbours weights matrix. The estimated coefficients in these two columns are very close to the corresponding estimates in Column 4 of Table V and the significance is again strong. The estimated value for the spatial parameter ν in the SDM of Column 4 is 0.0173, and not significant at conventional levels with an estimated standard error of 0.0202 (not shown in the table).

To check that our results do not depend on our choices in the specifications of the spatial weights matrix, we estimate a SAR with a rook-contiguity weights matrix in Column 3. With a rook-contiguity matrix, spatial units are neighbour only if they share an edge, not merely a corner. Given that this change is minor, it is not surprising that the results change neither regarding size nor significance. In Column 4, we change the normalisation method and estimate a SAR with a row-normalised weights matrix instead. Even though the size of the estimated coefficient drops, 'EKD' comes in again significantly.

In our baseline specifications, we limited the transborder ethnic relations to direct neighbour countries. In Columns 5 and 6, we extend these transborder ethnic kin relationships to pairs of countries that are within a distance of 100 kilometres (in Column 5) or 1,000 kilometres (in Column 6). We obtain similar and highly significant estimates of the EKD coefficient. The coefficient is slightly lower than in the baseline when we specify transborder ethnic relationships up to 100 km and the estimated value returns closer to the baseline with the 1,000 km range.

To show that our results do not hinge on the size the spatial units, we report in Columns 7 and 8 respectively, the estimates of the FE and SAR models with a four-by-four degrees grid instead. This enlarged grid reduces the number of observations to 1,762 units, and the estimated coefficient become slightly smaller than in the baseline, at 0.0242 in the FE model and 0.0227 in the SAR, still significant at the five per cent level. All the estimations in this table are controlled for spatial unit fixed-effects and country×year fixed-effects.

Finally, the results are robust to the addition of controls for the other possible political status of the EPR. For example, we used a count variable of transborder ethnic friends with the political status 'in power' instead of 'dominant' and an indicator of 'inclusion in the local government', in addition to political dominance. The results concerning EKD did not change much and the significance was unchanged.

5.3.2 Heterogeneous Effects: Interactions with Institutions and Legal Origin

To explore heterogeneity in the effect of connections to kin ethnic groups and political dominance across the continent, we interact the variables 'Ethnic Kin Dominant' and 'Dominant' with the following two indicators: (i) Institutions (Autocracy or Democracy) and (ii) Legal Origin (French or British). The equation that we estimate for the first interaction is the following:

$$\begin{aligned} Y_t &= (\gamma_0 EKD_{t-1} + \beta_{10} DOM_{t-1}) \circ (\text{Autocracy}_t) + (\gamma_1 EKD_{t-1} + \beta_{11} DOM_{t-1}) \circ (\text{Democracy}_t) \\ &\quad + \rho W Y_t + X_{-1,t-1}\beta_{-1} + \alpha_0 + \delta_{c,t} + \epsilon_t, \qquad \text{for } t = 1, \dots, T, \end{aligned} \tag{8}$$

where the element-by-element matrix multiplication (Hadamard product) is denoted by a circle '$\circ$', Autocracy_t is a N-vector of indicator variables, that takes the value one to indicate 'Autocracy', that is, whenever the Polity IV Index of Marshall and Cole (2011), is below -5, at time t, in each of the N spatial units. The indicator vector Democracy_t is one minus Autocracy_t. This model is identical to the baseline SAR model, equation (4), except for the interactions of the political variables with these indicators. The coefficients γ_0 and β_{10} correspond to the transborder effect and the local dominance effects, respectively, in an 'Autocracy' country. The coefficients γ_1 and β_{11} are the corresponding coefficients in a 'Democracy' country. The reason to make this distinction is that the absence of democracy and of the accompanying efficient institutions is associated with weak checks and balances that allow leaders to engage in favouritism (Glaeser et al., 2004; Acemoglu et al., 2005). The term $X_{-1,t-1}\beta_{-1}$ stands for all the other controls previously present in model (4), except 'Dominant', which is now written out explicitly.

The results are displayed in Table X, where we concentrate on the most demanding version of the model, that with country×year fixed-effects. We report the FE equation estimates in Column 1 and the SAR estimates in Column 2. When we interact our political variables with an Autocracy/ Democracy indicator, the variable 'Ethnic Kin Dominant' remains significant at one per cent for both categories. The size of the coefficient is twice as large for the 'Autocracy' category compared with the 'Democracy' category. This result is consistent with the idea that autocratic rulers can more easily reallocate resources to their co-ethnics than in democracies when checks and balances are strong (Francois et al., 2015).

Furthermore, in the spirit of La Porta et al. (1997, 1999), we follow the idea that differences in legal origin induce norms and rules in the judicial system, which may influence partiality in public spending. To distinguish the effect of the political variables by legal origin, we estimate equation (9).

$$\begin{aligned} Y_t &= (\gamma_0 EKD_{t-1} + \beta_{10} DOM_{t-1}) \circ (\text{French LO}) + (\gamma_1 EKD_{t-1} + \beta_{11} DOM_{t-1}) \circ (\text{U.K. LO}) \\ &\quad + \rho W Y_t + X_{-1,t-1}\beta_{-1} + \alpha_0 + \delta_{c,t} + \epsilon_t, \qquad \text{for } t = 1, \dots, T, \end{aligned} \tag{9}$$

where 'Ethnic Kin Dominant' and 'Dominant' are interacted with two $N \times 1$ dummy vectors for French and British legal origins, constructed using the data of La Porta et al. (1999). The variable 'French LO' equals one in the countries that were colonised by France and zero otherwise. Similarly, 'U.K. LO' equals one in the countries that were colonised by the United Kingdom and zero otherwise. We focus on these two categories because they are the most frequent. Regarding the legal origin interaction, Column 3 contains the FE equation estimates, and Column 4, the SAR estimates. The variable 'Ethnic Kin Dominant' is strongly significant in both these columns, with a coefficient greater than in the baseline in countries of French legal origin. However, the effect loses significance in countries of British legal origin. The estimates of the spatial autoregressive coefficient remain stable in the two spatial estimations, at values close to 0.09, comparable to the baseline.

5.3.3 Extended Dynamics: Lagged Dependent Variable and Spatially- and Temporally-Lagged Dependent Variable

In this subsection, we investigate robustness to extended specifications with lags of the dependent variable and/or the spatial lag of the dependent variable. Alternative specifications include a time-lagged dependent variable, a spatially- and temporally-lagged dependent variable or the combination of the two, in addition to the spatially lagged dependent variable that is already present. We do not choose these specifications as our benchmark because they may raise additional challenges related to the endogeneity of all these lagged dependent variables. Nonetheless, we now discuss these extensions of the model (4), whose statistical foundations are discussed in Yu et al. (2008) and Lee and Yu (2010). Note in particular that the data, with many spatial units and many years, fit well the conditions for consistency of the estimation in these papers: N and T tend to infinity, and T is not too small relative to N. The equation is:

$$\begin{aligned} Y_t &= \rho W Y_t + \omega W Y_{t-1} + \lambda Y_{t-1} + \gamma \mathrm{EKD}_{t-1} + X_{t-1}\beta + \alpha_0 + \delta_{c,t} + \epsilon_t, \\ &\quad \text{for } t = 1, \ldots, T. \end{aligned} \tag{10}$$

The terms $\omega W Y_{t-1}$ — a spatially- and temporally-lagged dependent variable — and λY_{t-1} — a time-lagged dependent variable — are incorporated into equation (4). We estimate all the possible combinations with each of the two terms or with both terms at the same time. We include either year fixed effects or country$\times$year fixed effects in the equation.

Table XI displays the results, where Columns 1 to 4 contain the estimation results with only year fixed-effects, while Columns 5 to 8 have the full set of country$\times$year fixed-effects. In Columns 1 and 5, we estimate the fixed-effect panel model with a lagged dependent variable. In Columns 2 and 6, we estimate the SAR model with a lagged dependent variable. Columns 3, 4, 7 and 8 have in addition a spatially- and time-lagged dependent variable — $\omega W Y_{t-1}$ — with (in 4 and 8) or without (in 3 and 7) a simply time-lagged dependent variable — λY_{t-1}. Because of the lagged dependent variable, the initial year of the data is lost for the estimation.

Adding this battery of lagged variables that control for many unobserved and unmeasurable elements, which may vary across time in the spatial unit and its surroundings, reduces the estimated coefficient of the variable 'Ethnic Kin Dominant' to 0.0166 in Column 8, instead of 0.0295 in the baseline. The significance of this variable, however, is still strong, with a p-value below one per cent. By comparing the coefficients of 'EKD' by type of fixed effects across models, we observe that the coefficient is always larger in the country×year fixed-effects specification than in the corresponding year fixed-effects specification. For instance, the coefficient in Column 7, 0.0347, is larger than the coefficient in Column 3, 0.0310. However, we obtain values smaller than in the baseline, except in Columns 3 and 7. The fact that the variable 'Rainfall' loses significance in the richest versions of Columns 7 and 8 may indicate that including such a wide range of controls produces a stringent test on the statistical relationship that may exist between the explanatory and dependent variables. However, it is somewhat comforting that such a rich dynamic specification, including even country×year fixed-effects in addition to spatial unit fixed-effects, turns out to be estimable, a primer in the empirical spatial econometric literature, to the best of our knowledge. The coefficient of the lagged dependent variable, λ, which is always precisely estimated, stays in a bracket from 0.58 to 0.71. This coefficient corresponds to persistence or inertia of the dependent variable, and the estimated values seem plausible compared to the autocorrelations usually estimated in GDP series. The coefficient of the spatially- and temporally-lagged dependent variable is also very precisely estimated, at values around 0.045–0.049. The estimated values of the coefficient ρ increase when we move from Column 2 to 4 and 6 to 8, that is, when we strengthen the controls.

Fundamentally, the inclusion of all these additional controls undermines neither the sign nor the significance of the coefficient of the variable 'Ethnic Kin Dominant', even if the coefficient estimates fall to levels that are slightly higher than 50 per cent of their initial value in the baseline case.

5.3.4 Ethnic Network Transmission

As mentioned earlier, the spatio-economic interactions may constitute a competing explanation to the political effect of transborder ethnic kin groups, which is why we control for them. In this subsection, we investigate whether there might exist contiguous and ethnically homogeneous regions where these spatio-economic interactions are stronger. As in the previous case, such phenomenon could cast doubt on our main result, if for instance, the measured effect of transborder connections would stem from especially intensive spatial effects through ethnic networks. Furthermore, this is another way to account for some heterogeneity in spatial effects. To investigate the influence of this alternative transmission mechanism on the significance of 'Ethnic Kin Dominant', we estimate Spatial Durbin Models of the form

$$Y_t = \rho W Y_t + \delta W_{\text{Ethnic}} Z_t + \gamma \text{EKD}_{t-1} + X_{t-1}\beta + \alpha_0 + \delta_{c,t} + \epsilon_t, \qquad \text{for } t = 1, \ldots, T, \tag{11}$$

where W_{Ethnic} is a spatial weights contiguity matrix, structured according to the ethnicity of the groups. That is, $W_{\text{Ethnic},i,j} = 1$ if spatial units i and j are contiguous, and contain the same ethnic group or transborder ethnic kin groups, and zero otherwise.[36] Apart from this addition, equation (11) is identical to (4). We estimate three different specifications, by removing or not the term $\rho W Y_t$, and by changing the definition of the variable Z_t, which interacts with W_{Ethnic}, and is either Y_t or the product of Y_t and 'Dominant$_{t-1}$'.

The results are presented in Table XII where the new explanatory variable Z_t is constructed using either log-luminosity (in Columns 1 and 2) or the interaction of log-luminosity and lagged local political dominance (in Column 3). That is: the new spatial term $W_{\text{Ethnic}} Z_t$ is a linear combination of the dependent variable in the neighbouring spatial units that belong to the ethnic network. This variable replaces the spatial term in Column 1, and comes in addition to it in Column 2. In Column 3, we alter the definition of Z_t so that $W_{\text{Ethnic}} Z_t$ corresponds to the linear combination of the dependent variable in the neighbouring spatial units of the ethnic network that contains dominant groups.

The SAR model of Column 1, with the constraint $\rho = 0$, where the ethnic spatial weights matrix interacts with log-luminosity, controls for an alternative transmission channel where the economic externalities go only through the ethnic network. In that case, the estimated coefficient for the variable 'Ethnic Kin Dominant' is of 0.0406 , again significant at the one per cent level. This value is higher than 0.0295, obtained in the corresponding baseline result. Although the new control may capture only a fraction of the spatial interaction, the link remains strong and significant, which supports the main claim of this paper. In the model of Column 2, we estimate an SDM model with the two types of spatially-lagged dependent variables. The estimated value of ρ, at 0.0859, remains comparable to the baseline, and highly significant. The spatio-ethnic coefficient δ, is also significant, with a value of 0.0242. Now, the estimated coefficient for EKD is 0.0361, again significant at the one per cent level. In the model of Column 3, Z_t denotes the interaction of log-luminosity and lagged local political dominance, to capture the possibility that ethnic groups can benefit from the economic activity of their transborder kin groups conditional on political dominance, unlike previously where it occurred irrespective of the political status of the groups. However, we still obtain similar results concerning the estimated coefficients and their significance. On the whole, although these additional controls could in principle capture the effect of connections to politically dominant transborder ethnic groups, we do not observe this in our estimations. In all cases, the significance and magnitude of the coefficient of EKD, and of ρ, remain stable.

5.3.5 Difference-In-Differences

To further support a strictly causal interpretation of the mechanism studied in this paper, we finally propose a difference-in-differences approach (Card and Krueger, 1994; Di Tella and Schargrodsky, 2004). This may be useful if some determinants of the changes in political dominance are complex economic causes that may extend over borders and be relatively persistent.

[36]For the estimation, we use row-normalisation for W_{Ethnic} because of the sparsity of the matrix.

Our focus is the average treatment effect on the treated (ATT). The treatment considered is a reduction in the number of dominant transborder ethnic kin groups. We choose to define the treatment as a reduction in the number of dominant transborder ethnic kin because, over the studied period, these are the only types of changes that have significantly taken place, even though, in the short-run, there are also a few occurrences of positive transitions. This pattern results from the wave of democratisations in Africa and the increasing prevalence of power sharing arrangements. Since the treatment may occur at different dates for different units, we take advantage of a generalisation of the typical difference-in-differences estimation method to this case, which is exposed in Muller (2018). This is made possible by an assumption on the stability of the treatment effect across the considered periods. Another important implication of the method is to define the treatment group and the control group within the same common geographical area — here the country — so as to allow for many unobserved common factors, including geographic, climatic, institutional, cultural and historical characteristics.

This methodology solves a great deal of the potential endogeneity problems possibly arising in the models of equations (2) and (4). Indeed, even if the underlying data generating process involves idiosyncrasies in the spatial units, or the specific circumstances of the countries at certain times, this strategy will take them into account, while avoiding the necessity to estimate all the corresponding fixed-effects, thanks to differencing.

The estimation is based on group averages over the spatial units. In each country, we define the treatment group as including the spatial units inhabited by an ethnic group that, at some point during the period 1992–2012, loses a connection to a politically dominant ethnic group in the neighbouring country. That is: one of their politico-ethnic connections ceases to be dominant in its own country. To be able to define a control group with a corresponding year of transition, we restrict the analysis to the countries containing ethnic groups that have experienced this particular change over the period, with this transition year varying across countries.[37] To circumvent the problem resulting from the non-simultaneity of the treatment across countries, we construct a new two-period dataset from our original data. For each spatial unit of a country, we define an ex-ante period and an ex-post period. The first period lasts until the year just before the occurrence of the treatment for at least one ethnic group in the country, and the second period starts after that year. The transition year is therefore country-specific. The dependent variable is computed as the mean of the dependent variable, 'Log Light' at the spatial unit level, for four sub-samples of observations of the original panel. These sub-samples are identified by the fact that their spatial unit belong, or do not belong, to the treatment group, and jointly whether these observations occur before or after the transition. Precisely, we aggregate luminosity for the sub-samples: 'Treatment Group/After Treatment', 'Treatment Group/Before Treatment', 'Control Group/After Treatment' and 'Control Group/Before Treatment'.

If we follow the logic of our previous results, we can expect an adverse effect of the treatment on emitted luminosity. In Figure VI, we investigate the presence of this pattern by subdividing the

[37] Among the 49 countries composing our sample, 17 did undergo such treatment. However, because the countries involved are bigger than the average, this concerns 1,346 out of the 2,929 spatial units, about 46 per cent.

two periods (before- and after-treatment) into four sub-periods. Because the timing of the treatment is location-dependent, we define the time thresholds accordingly. That is: the time line is country-specific since we never observe treatments for two distinct ethnic groups in the same country in our data. Specifically, we divide the period before the treatment into two sub-periods of equal length and the period after the treatment into two other equally long sub-periods. Then, we aggregate the data by taking the average of the dependent variable by sub-period and 'Treatment/ Control' group. Because of the wide dispersion in the dependent variable due to the variety of surface area, wealth level, and population density of the spatial units, the values shown on the Y-axis are mean deviations from the time mean at the spatial unit level. To take deviations from the mean at the spatial unit level, we regress 'Log Light' on a set of spatial unit fixed-effects and extract the residuals.

The obtained profile in Figure VI is compelling. The dashed line corresponds to the control group, i.e., the spatial units of the ethnic groups that did not lose any connections to dominant groups. The solid line corresponds to the treatment group, i.e., the spatial units of the ethnic groups that experienced such change. Before the treatment, which occurs between sub-periods two and three, the two groups follow almost the same upward trend. In sub-periods three and four, the upward trend is still present for both groups, but the treatment group experiences a noticeable drop, as compared with the control group. We depict confidence intervals at the five per cent level on the figure, which indicate that the means were distinct before the treatment. In sub-period three and four, the positions of the means are inverted compared with the pre-treatement period. Even though the confidence intervals overlap just after the treament, they are disjoint in the final sub-period. Although this figure does not provide any statistical evidence of the causal effect of the treatment, it indicates that the phenomenon is present in the data, at least to some extent. Moreover, it offers a diagnosis check that the conditions of parallel trends for applying the method are not rejected in the pre-treatment period.

Here, we investigate whether the post-treatment gap is statistically significant using a difference-in-differences approach. Namely, we compute the statistic:

$$\mathrm{DiD} = \sum_{i \in \mathcal{T}} \left(\sum_{t \in \mathcal{P}_i(1)} \frac{Y_{i,t}}{n_{i,1}} - \sum_{t \in \mathcal{P}_i(0)} \frac{Y_{i,t}}{n_{i,0}} \right) - \sum_{i \in \mathcal{C}} \left(\sum_{t \in \mathcal{P}_i(1)} \frac{Y_{i,t}}{n_{i,1}} - \sum_{t \in \mathcal{P}_i(0)} \frac{Y_{i,t}}{n_{i,0}} \right),$$

i.e., the average difference in the outcome variable after and before the transition in the treated group minus that for the control group, where $\mathcal{T}$ is the set of spatial units belonging to the treatment group, $\mathcal{C}$ the set of spatial units of the control group, $\mathcal{P}_i(1)$ the set of time indices of the post-treatment period in spatial unit i, $\mathcal{P}_i(0)$ the set of time indices of the pre-treatment period in spatial unit i, $Y_{i,t}$ is 'Log Light' in spatial units i and year t. Eventually, $n_{i,p}$ is the number of observations corresponding to spatial unit i and period p in the original panel dataset, where the index p denotes the pre-treatment ($p = 0$) or post-treatment period ($p = 1$).

Alternatively, we can estimate this statistic of the causal effect of the treatment by running the following OLS regression:

$$\overline{\text{Log Light}}_{i,p} = \beta_1 \text{Treatment}_{i,p} + \beta_2 \text{Period}_{i,p} + \beta_3 \text{Treatment}_{i,p} * \text{Period}_{i,p} + \beta_4 X_{i,p} + \epsilon_{i,p}, \quad (12)$$

for $i = 1, \ldots, N$ and $p = 0, 1$ where

$$\overline{\text{Log Light}}_{i,p} = \sum_{t \in \mathcal{P}_i(p)} \frac{Y_{i,t}}{n_{i,p}},$$

is the mean of the variable 'Log Light' in spatial unit i and all the years of period p. The explanatory variable 'Treatment' equals one for the spatial units belonging to the treated group (irrespective of the period) and zero otherwise, 'Period' equals zero for the pre-treatment period and one for the post-treatment period. $\epsilon_{i,p}$ is a centred error term. $X_{i,p}$ stands for additional controls at the spatial unit-period level in some specifications (log-means of population, temperatures, and precipitations per square kilometre).

The ordinary least squares estimates of the coefficients of (12) without additional controls are presented in Table XIII, Column 1, based on the data of the new two-period panel. The estimated coefficient of the interacted term, which corresponds to the average treatment effect on the treated, is -0.0261. Remarkably, it is of opposite sign and approximately equal magnitude as the coefficient estimated for the variable 'Ethnic Kin Dominant' in Tables V, IX and XII, which seems to somewhat back up these results.

Robust asymptotic standard error estimators clustered at the spatial unit level, which are displayed in parentheses, indicate a treatment significant at the five per cent level. To better account for small sample biases, and possible miss-specifications, 1,000 bootstrap iterations are performed, each time reconstructing the two-period panel from the new random sample of spatial units, and re-estimating equation (12).[38] The estimated parameter values are recorded and the standard errors of the distribution of these parameters are estimated using the standard deviation of the 1,000 bootstrap estimates. The bootstrap standard errors are displayed in square brackets in Table XIII, Column 1, and confirm the significance levels of the asymptotic clustered standard errors.

Column 2 contains estimates with additional controls for Log mean population, Log of mean precipitations per square kilometre, and Log of mean temperature, and the asymptotic standard errors in parentheses are clustered at the spatial unit level, and those in square brackets are bootstrap, as in Column 1. The coefficient of the interactive term becomes larger in absolute value than previously, -0.0479 and the R-square statistic increases to 0.615. Adding these controls does not undermine the significance of the ATT estimate, and confirm our previous findings. Irrespective of the method used to compute the standard error, the coefficient of the interacted term remains negative and significant

[38] As previously, this sample is clustered by spatial unit, i.e., when some spatial units are selected at random, all years of those spatial units are also selected, and stratified at the country level, i.e., the bootstrap sample contains precisely the same number of spatial units per country as the original sample.

at the five per cent level. This implies that, on average, the ethnic homeland of the group that loses a Dominant Ethnic Kin group emits substantially less luminosity in the following years.

6 Conclusion

In Africa, where ethnicity is politically salient and where borders that were designed by colonial powers have divided historical ethnic groups, we study comparative development as approximated by luminosity measured from space. We build on the literature on ethnic favouritism (Hodler and Raschky, 2014; De Luca et al., 2018) and we extend it by investigating transnational effects. We provide robust empirical evidence, for the first time, that the political predominance of transborder related kin in its national government strengthens the economic prosperity of a given ethnic group's homeland in another country.

We constructed a new dataset by matching the geolocalised version of the Ethnic Power Relations database with geolocalised luminosity data, and several other geolocalised databases. This dataset contains panel information for 2,929 geographical spatial units over the period 1992 to 2012.

Using various panel estimation strategies, which incorporate spatial externalities, we find robust evidence that the total number of politically dominant ethnic kin groups in neighbouring countries has a positive statistically significant effect. To control for other competing spatial externalities, such as trade or the diffusion of technologies and ideas, we confirm the findings with diverse dynamic spatial autoregressive models with fixed effects for spatial units and country$\times$year pairs. Finally, the causal nature of the effect is supported by adjusted difference-in-differences results.

In a possible extension of this work, one could test whether the elicited role of the political status of transborder ethnic kin for local development extends to other parts of the world, beyond Africa. However, a better capture of the socio-eco-political mechanisms that are involved in the estimated effect is what is really needed to a full understanding of the phenomenon. To do this, collecting detailed and geolocalised data on various economic flows, migration and information exchanges seems to be a prerequisite. Only in these conditions would it be thinkable to precisely pin down what is behind the interaction of faraway ethnic solidarities with local variations in emissions of light.

Eventually, only when a clear understanding of the mechanism at play is achieved will it become possible to derive serious policy implications. However, the highlighted interactions suggest that a global point of view on development processes cannot be avoided. In particular, it seems useful to attempt to anticipate the economic repercussions in neighbouring countries before implementing foreign interventions or aid policies that have political consequences.

Table I: **Descriptive Statistics (Pooled Data)**

Variable	Mean	Std. Dev.	Min.	Max.	N
Log(Light)_t	8.776	1.833	0	13.019	61,509
Light_t	22,009	33,303	1	451,100	61,509
Log(Pop)_{t-1}	10.166	2.524	0	17.451	61,509
Pop_{t-1}	274,626	988,301	1	3.79e+07	61,509
Dominant_{t-1}	0.091	0.287	0	1	61,509
$\text{Ethnic Kin Dominant}_{t-1}$	0.194	0.58	0	3	61,509
Rainfall_{t-1}	0.023	0.04	0	0.37	61,509
Temperature_{t-1}	24.395	3.802	6.846	31.781	61,509
Violence_{t-1}	0.065	0.246	0	1	61,509

Table II: **Descriptive Statistics (Between- and Within-Variations)**

Variable	Between			Within		
	Std. Dev.	Min	Max	Std. Dev.	Min	Max
Log(Light)_t	1.797	3.97	12.698	0.366	0.740	10.150
Log(Pop)_{t-1}	2.519	0	17.298	0.168	8.872	11.047
Dominant_{t-1}	0.267	0	1.000	0.107	-0.814	1.043
$\text{Ethnic Kin Dominant}_{t-1}$	0.565	0	2.714	0.117	-0.713	1.096
Rainfall_{t-1}	0.040	0	0.313	0.006	-0.076	0.164
Temperature_{t-1}	3.757	0	30.776	0.588	-3.696	27.081
Violence_{t-1}	0.149	0	1.000	0.196	-0.887	1.017

Table III: **Transitions of Ethnic Kin Dominant$_t$**

		Ethnic Kin Dominant$_t$				
		0	1	2	3	Total
Ethnic Kin Dominant$_{t-1}$	0	51,809	104	0	0	51,913
		99.80	0.20	0.00	0.00	100.00
	1	174	2,493	0	0	2,667
		6.52	93.48	0.00	0.00	100.00
	2	0	36	3,319	0	3,355
		0.00	1.07	98.93	0.00	100.00
	3	0	0	43	602	645
		0.00	0.00	6.67	93.33	100.00
	Total	51,983	2,633	3,362	602	58,580
		88.74	4.49	5.74	1.03	100.00

Note: This table presents the transition proportions and counts between years $t-1$ and t for the variable 'Ethnic Kin Dominant'. The number of transitions across distinct states occurring in the sample is 357.

Table IV: **Transitions of Dominant$_t$**

		Dominant$_t$		
		0	1	Total
Dominant$_{t-1}$	0	53,108	68	53,176
		99.87	0.13	100
	1	242	5,162	5,404
		4.48	95.52	100
	Total	53,350	5,230	58,580
		91.07	8.93	100

Note: This table presents the transition proportions and counts between years $t-1$ and t for the variable 'Dominant'. The number of transitions across distinct states occurring in the sample is 310.

Table V: **Baseline Results**

	Dependent variable is Log $Light_{i,t}$ 2x2 Fishnet			
	(1) FE	(2) SAR	(3) FE	(4) SAR
Ethnic Kin $Dominant_{t-1}$	0.0311	0.0287	0.0340	0.0295
	(0.0101)***	(0.0095)***	(0.0095)***	(0.0087)***
$Dominant_{t-1}$	0.0088	0.0045	-0.0104	-0.0108
	(0.0091)	(0.0095)	(0.0132)	(0.0130)
$Log(Pop)_{t-1}$	0.1098	0.0954	-0.0094	-0.0102
	(0.0164)***	(0.0154)***	(0.0180)	(0.0175)
$Violence_{t-1}$	0.0050	0.0026	0.0062	0.0047
	(0.0031)	(0.0029)	(0.0025)**	(0.0024)*
$Rainfall_{t-1}$	0.3641	0.2372	0.0822	0.0220
	(0.0546)***	(0.0546)***	(0.0464)*	(0.0488)
$Temperature_{t-1}$	-0.0036	-0.0024	-0.0029	-0.0020
	(0.0010)***	(0.0010)**	(0.0010)***	(0.0010)**
Spatial ρ		0.1475		0.0910
		(0.0001)***		(0.0001)***
Year FE	Y	Y	N	N
Country*Year FE	N	N	Y	Y
Number of Units	2,929	2,929	2,929	2,929
Number of Years	21	21	21	21

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust standard errors clustered at the spatial unit level are in parentheses. We use the dataset constructed with a two-by-two degrees grid. All spatial models are estimated using a spectrally normalised queen-neighbours weights matrix. Columns 1 and 3 contain the Fixed-Effects estimates. Columns 2 and 4 contain SAR estimates, where the transborder ethnic kin relationships are limited to direct neighbour countries. All estimations control for spatial unit fixed-effects. We add year fixed-effects in Columns 1 and 2 and country×year fixed-effects in Columns 3 and 4. All columns are based on the whole sample of African countries.

Table VI: **Correlates of the Spatial Unit Fixed-Effects**

	Estimated spatial unit fixed-effects	
	$\hat{\alpha}_{i,FE}$ (1)	$\hat{\alpha}_{i,SAR}$ (2)
Mean Dominant	0.060	0.078
	(0.029)**	(0.034)**
	[0.033]*	
Mean Ethnic Kin Dominant	0.048	0.038
	(0.011)***	(0.015)***
	[0.017]***	
Mean Log Population	0.029	0.041
	(0.005)***	(0.005)***
	[0.010]***	
Log(Surface)	0.957	0.843
	(0.005)***	(0.006)***
	[0.005]***	
Mean Temperature	-0.027	-0.029
	(0.002)***	(0.002)***
	[0.157]***	
Mean Rainfall	-1.673	-3.083
	(0.156)***	(0.213)***
	[0.002]***	
R^2	0.9687	0.9453
Number of Observations	2,929	2,929

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. OLS estimations in both columns. The dependent variable is the spatial unit fixed-effect extracted from the baseline panel estimation, equation (2), in Column 1, and from the baseline spatial panel (SAR) estimation, equation (4) in Column 2. The independent variables are means by spatial unit over the whole period. Robust standard errors are in parentheses and bootstrap standard errors are in square brackets.

Table VII: **Simulations**

	All Observations	Ethnic Network		Ethnic Outsiders	
		Favoured	Non-Favoured	Favoured	Non-Favoured
	(1)	(2)	(3)	(4)	(5)
	Experiment 1 : Largest Ethnic Group In Power				
Log Light	8.7762	8.4008	8.8962	8.8368	8.7604
	(1.8315)	(1.7697)	(1.9415)	(1.8478)	(1.8153)
Log Light+Δ Log Light	8.7934	8.5408	8.9740	8.8988	8.7602
	(1.8329)	(1.7715)	(1.9439)	(1.8509)	(1.8156)
Δ Log Light	0.0172	0.1400	0.0779	0.0620	-0.0002
	(0.0352)	(0.0029)	(0.0214)	(0.0226)	(0.0079)
Number of Observations	61,509	1,071	5,499	7,938	47,001
	Experiment 2 : Second Largest Ethnic Group In Power				
Log Light	8.7762	8.7022	8.8787	8.5790	8.7949
	(1.8315)	(1.8511)	(1.8031)	(1.8530)	(1.8288)
Log Light+Δ Log Light	8.7840	8.8579	8.9073	8.6271	8.7850
	(1.8296)	(1.8835)	(1.8091)	(1.8557)	(1.8234)
Δ Log Light	0.0079	0.1557	0.0286	0.0482	-0.0100
	(0.0694)	(0.0897)	(0.1270)	(0.0389)	(0.0401)
Number of Observations	61,509	3,210	5,486	6,093	46,720

Note: This table shows the summary statistics of the simulation experiments, by categories of observations. 'Favoured' refers to the spatial units of the groups that directly benefit from the counterfactual change, as opposed to the 'Non-Favoured'. 'Ethnic Network' refers to the spatial units of the groups that are ethnically related to the groups that directly benefit from the counterfactual change, as opposed to the 'Ethnic Outsiders'. The table displays the mean and standard deviation of the predicted value of the dependent variable before the experiment, after the experiment and the change between the two situations. The top panel displays the values for Experiment 1, in which the largest ethnic group of the country becomes dominant. The bottom panel displays the values for Experiment 2, in which the second largest ethnic group of the country becomes dominant.

Table VIII: **Simulation & Channels**

Δ Log Light	Ethnic Network		Ethnic Outsiders	
	Favoured	Non-Favoured	Favoured	Non-Favoured
	Experiment 1 : Largest Ethnic Group In Power			
(1) $\rho = 0$	0.1345	0.0741	0.0594	-0.0009
	(0.0000)	(0.0201)	(0.0216)	(-0.0009)
(2) $\gamma = 0$	0.0695	0.0004	0.0617	-0.0005
	(0.0015)	(0.0009)	(0.0225)	(-0.0005)
(3) $\beta_1 = 0$	0.0705	0.0775	0.0003	0.0003
	(0.0015)	(0.0214)	(0.0010)	(0.0003)
(4) $\gamma = 0, \rho = 0$	0.0670	0.0000	0.0594	-0.0009
	(0.0000)	(0.0000)	(0.0216)	(-0.0009)
(5) $\beta_1 = 0, \rho = 0$	0.0675	0.0741	0.0000	0.0000
	(0.0000)	(0.0201)	(0.0000)	(0.0000)
	Experiment 2 : Second Largest Ethnic Group In Power			
(1) $\rho = 0$	0.1496	0.0267	0.0463	-0.0098
	(0.0856)	(0.1213)	(0.0375)	(-0.0098)
(2) $\gamma = 0$	0.0630	-0.0121	0.0561	-0.0045
	(0.0199)	(0.0273)	(0.0271)	(-0.0045)
(3) $\beta_1 = 0$	0.0927	0.0407	-0.0079	-0.0054
	(0.0812)	(0.1089)	(0.0313)	(-0.0054)
(4) $\gamma = 0, \rho = 0$	0.0611	-0.0122	0.0544	-0.0046
	(0.0190)	(0.0259)	(0.0262)	(-0.0046)
(5) $\beta_1 = 0, \rho = 0$	0.0885	0.0389	-0.0082	-0.0053
	(0.0778)	(0.1043)	(0.0304)	(-0.0053)

Note: This table displays summary statistics of the simulation experiments with parameter restrictions by categories of observations. 'Favoured' refers to the spatial units of the groups that directly benefit from the counterfactual change, as opposed to the 'Non-Favoured'. 'Ethnic Network' refers to the spatial units of the groups that are ethnically related to the groups that directly benefit from the counterfactual change, as opposed to the 'Ethnic Outsiders'. The table displays the mean and standard deviation of the change in the predicted value of Log Light before/after the experiment, in five cases of restrictions on the parameters. The top panel displays the values for Experiment 1, in which the largest ethnic group of the country becomes dominant. The bottom panel displays the values for Experiment 2, in which the second largest ethnic group of the country becomes dominant.

Table IX: **Robustness Checks**

	Dependent variable is Log $\text{Light}_{i,t}$							
	2x2 Fishnet						4x4 Fishnet	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	SEM	SDM	SAR	SAR	SAR	SAR	FE	SAR
Ethnic Kin Dominant_{t-1}	0.0257	0.0309	0.0284	0.0179	0.0207	0.0274	0.0242	0.0227
	(0.0097)***	(0.0092)***	(0.0083)***	(0.0070)**	(0.0077)***	(0.0075)***	(0.0113)**	(0.0109)**
Dominant_{t-1}	0.0006	-0.0148	-0.0106	0.0047	0.0033	-0.0123	-0.0364	-0.0369
	(0.0127)	(0.0141)	(0.0130)	(0.0105)	(0.0095)	(0.0130)	(0.0197)*	(0.0193)*
Log(Pop)_{t-1}	0.0078	-0.0103	-0.0102	-0.0030	0.0962	-0.0107	-0.0178	-0.0172
	(0.0194)	(0.0175)	(0.0176)	(0.0148)	(0.0154)***	(0.0175)	(0.0273)	(0.0267)
Violence_{t-1}	0.0025	0.0047	0.0051	0.0034	0.0023	0.0046	0.0089	0.0079
	(0.0021)	(0.0024)**	(0.0024)**	(0.0020)*	(0.0029)	(0.0024)*	(0.0029)***	(0.0028)***
Rainfall_{t-1}	0.0019	0.0208	0.0327	0.0726	0.2408	0.0225	0.0063	-0.0257
	(0.0449)	(0.0481)	(0.0490)	(0.0358)**	(0.0546)***	(0.0489)	(0.0261)	(0.0299)
Temperature_{t-1}	-0.0032	-0.0021	-0.0028	-0.0013	-0.0026	-0.0022	-0.0028	-0.0022
	(0.0017)*	(0.0010)**	(0.0011)***	(0.0008)	(0.0010)**	(0.0011)**	(0.0016)*	(0.0016)
Spatial ρ	0.7037	0.0917	0.0876	0.5961	0.1486	0.0917		0.0643
	(0.0027)***	(0.0088)***	(0.0001)***	(0.0019)***	(0.0001)***	(0.0001)***		(0.0001)***
Weights Matrix	W	W	W_{rook}	W_{row}	W	W		W
Year FE	N	N	N	N	N	N	N	N
Country*Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Number of Units	2,929	2,929	2,929	2,929	2,929	2,929	1,762	1,762
Number of Years	21	21	21	21	21	21	21	21

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust standard errors clustered at the spatial unit level are in parentheses. All estimations control for spatial unit fixed-effects and country×year fixed-effects are included. In Columns 1 to 6, we use a two-by-two degrees grid. The SEM and SDM (Columns 1 and 2) models are estimated using a spectrally normalised queen-neighbours weights matrix. The Durbin variable Z_{t-1} in Column 2 is 'Dominant'. In Column 3, we estimate a SAR with a rook-contiguity weights matrix instead. In Columns 4, we estimate a SAR with a row-normalised weights matrix instead. The transborder ethnic kin relationships are limited to direct neighbour countries in all columns, except in Columns 5 and 6 for which we extend to countries within a radius of 100 km and 1,000 km instead, respectively. In Columns 7 and 8, we report the estimates of the FE and SAR models with the sample constructed with a four-by-four degrees grid instead. All Columns are based on the whole sample of African countries.

Table X: **Heterogeneous Effects: Interactions with Institutions and Legal Origin**

	Dependent variable is Log Light$_{i,t}$ 2x2 Fishnet			
	(1) FE	(2) SAR	(3) FE	(4) SAR
	Autocracy		French LO	
Ethnic Kin Dominant$_{t-1}$	0.0470	0.0417	0.0325	0.0310
	(0.0099)***	(0.0095)***	(0.0094)***	(0.0090)***
Dominant$_{t-1}$	-0.0280	-0.0231	-0.0052	-0.0057
	(0.0198)	(0.0195)	(0.0056)	(0.0063)
	Democracy		U.K. LO	
Ethnic Kin Dominant$_{t-1}$	0.0187	0.0186	0.0065	0.0017
	(0.0069)***	(0.0067)***	(0.0172)	(0.0174)
Dominant$_{t-1}$	-0.0092	-0.0106	-0.0138	-0.0141
	(0.0121)	(0.0120)	(0.0215)	(0.0211)
Log(Pop)$_{t-1}$	-0.0096	-0.0104	-0.0094	-0.0102
	(0.0180)	(0.0175)	(0.0180)	(0.0175)
Violence$_{t-1}$	0.0064	0.0051	0.0060	0.0047
	(0.0025)**	(0.0024)**	(0.0025)**	(0.0024)*
Rainfall$_{t-1}$	0.0854	0.0256	0.0823	0.0216
	(0.0461)*	(0.0486)	(0.0464)*	(0.0488)
Temperature$_{t-1}$	-0.0030	-0.0021	-0.0029	-0.0020
	(0.0010)***	(0.0010)**	(0.0010)***	(0.0010)*
Spatial ρ		0.0895		0.0911
		(0.0001)***		(0.0001)***
Year FE	N	N	N	N
Country*Year FE	Y	Y	Y	Y
Number of Units	2,929	2,929	2,929	2,929
Number of Years	21	21	21	21

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust standard errors clustered at the spatial unit level are in parentheses. All estimations are controlled for spatial unit fixed-effects and country×year fixed-effects. We use a two-by-two degrees grid in all columns and a spectrally normalised queen-neighbours weights matrix in Columns 2 and 4. In Columns 1 and 2, we interact the Dominant and Ethnic Kin Dominant variables with an autocracy and a democracy indicators. In Columns 3 and 4, we interact them with legal origin indicators.

Table XI: **Extended Dynamics:**
Lagged Dependent Variable and Spatially and Temporally Lagged Dependent Variable

	Dependent variable is Log $Light_{i,t}$ 2x2 Fishnet							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FE	SAR	SAR	SAR	FE	SAR	SAR	SAR
Ethnic Kin $Dominant_{t-1}$	0.0169	0.0158	0.0310	0.0155	0.0185	0.0165	0.0347	0.0166
	(0.0054)***	(0.0055)***	(0.0098)***	(0.0054)***	(0.0056)***	(0.0059)***	(0.0093)***	(0.0059)***
$Dominant_{t-1}$	0.0020	-0.0032	-0.0051	-0.0028	-0.0111	-0.0107	-0.0240	-0.0108
	(0.0037)	(0.0042)	(0.0091)	(0.0040)	(0.0063)*	(0.0067)	(0.0127)*	(0.0066)
$Log(Pop)_{t-1}$	0.0725	0.0638	0.1397	0.0606	0.0022	0.0048	-0.0000	0.0034
	(0.0139)***	(0.0157)***	(0.0189)***	(0.0161)***	(0.0099)	(0.0101)	(0.0203)	(0.0100)
$Violence_{t-1}$	0.0038	0.0010	0.0023	0.0008	0.0023	-0.0000	0.0032	-0.0000
	(0.0016)**	(0.0015)	(0.0026)	(0.0015)	(0.0016)	(0.0015)	(0.0021)	(0.0015)
$Rainfall_{t-1}$	0.2469	0.1707	0.2597	0.0942	0.1032	0.0737	0.0702	0.0268
	(0.0361)***	(0.0337)***	(0.0562)***	(0.0421)**	(0.0276)***	(0.0293)**	(0.0498)	(0.0327)
$Temperature_{t-1}$	-0.0022	-0.0021	-0.0019	-0.0020	-0.0012	-0.0007	-0.0019	-0.0008
	(0.0006)***	(0.0005)***	(0.0010)**	(0.0005)***	(0.0006)*	(0.0006)	(0.0010)*	(0.0006)
$Log(Light)_{t-1}$	0.6185	0.6742		0.7009	0.5856	0.6527		0.6626
	(0.1051)***	(0.1109)***		(0.1192)***	(0.1229)***	(0.1295)***		(0.1337)***
Spatial ω			0.0459	0.0459			0.0492	0.0492
			(0.0000)***	(0.0000)***			(0.0000)***	(0.0000)***
Spatial ρ		0.0831	0.1174	0.1889		0.0514	0.0558	0.1119
		(0.0001)***	(0.0001)***	(0.0003)***		(0.0001)***	(0.0001)***	(0.0002)***
Year FE	Y	Y	Y	Y	N	N	N	N
Country*Year FE	N	N	N	N	Y	Y	Y	Y
Number of Units	2,929	2,929	2,929	2,929	2,929	2,929	2,929	2,929
Number of Years	20	20	20	20	20	20	20	20

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust standard errors clustered at the spatial unit level are in parentheses. Columns 1 and 5 contain the Fixed-Effects estimates with a lagged dependent variable. Columns 2 to 4 and 5 to 8 contain the SAR estimates (with spatially-lagged dependent variable), with an additional lagged dependent variable (in 2 and 5), a temporally and spatially-lagged dependent variable (in 3 and 6) or both (in 4 and 8). The transborder ethnic kin relationships are limited to direct neighbour countries. All estimations are controlled for spatial unit fixed-effects. We add year fixed-effects in Columns 1 to 4 and country×year fixed-effects in Columns 5 to 8. A two-by-two degrees grid is used. All spatial models are estimated using a spectrally normalised queen-neighbours weights matrix.

Table XII: **Ethnic Network Transmission**

	Dependent variable is Log $\text{Light}_{i,t}$ 2x2 Fishnet		
	(1) SAR	(2) SDM	(3) SDM
Ethnic Kin Dominant_{t-1}	0.0406	0.0361	0.0295
	(0.0094)***	(0.0090)***	(0.0095)***
Dominant_{t-1}	-0.0045	-0.0068	-0.0376
	(0.0130)	(0.0130)	(0.0286)
Log(Pop)_{t-1}	-0.0112	-0.0114	-0.0103
	(0.0177)	(0.0174)	(0.0175)
Violence_{t-1}	0.0064	0.0051	0.0049
	(0.0025)***	(0.0024)**	(0.0024)**
Rainfall_{t-1}	0.0661	0.0141	0.0182
	(0.0465)	(0.0491)	(0.0486)
Temperature_{t-1}	-0.0027	-0.0019	-0.0021
	(0.0010)**	(0.0010)*	(0.0010)**
Spatial ρ		0.0859	0.0913
		(0.0087)***	(0.0089)***
Spatio-ethnic δ	0.0357	0.0242	0.0037
	(0.0059)***	(0.0058)***	(0.0030)
Country*Year FE	Y	Y	Y
Number of Units	2,929	2,929	2,929
Number of Years	21	21	21

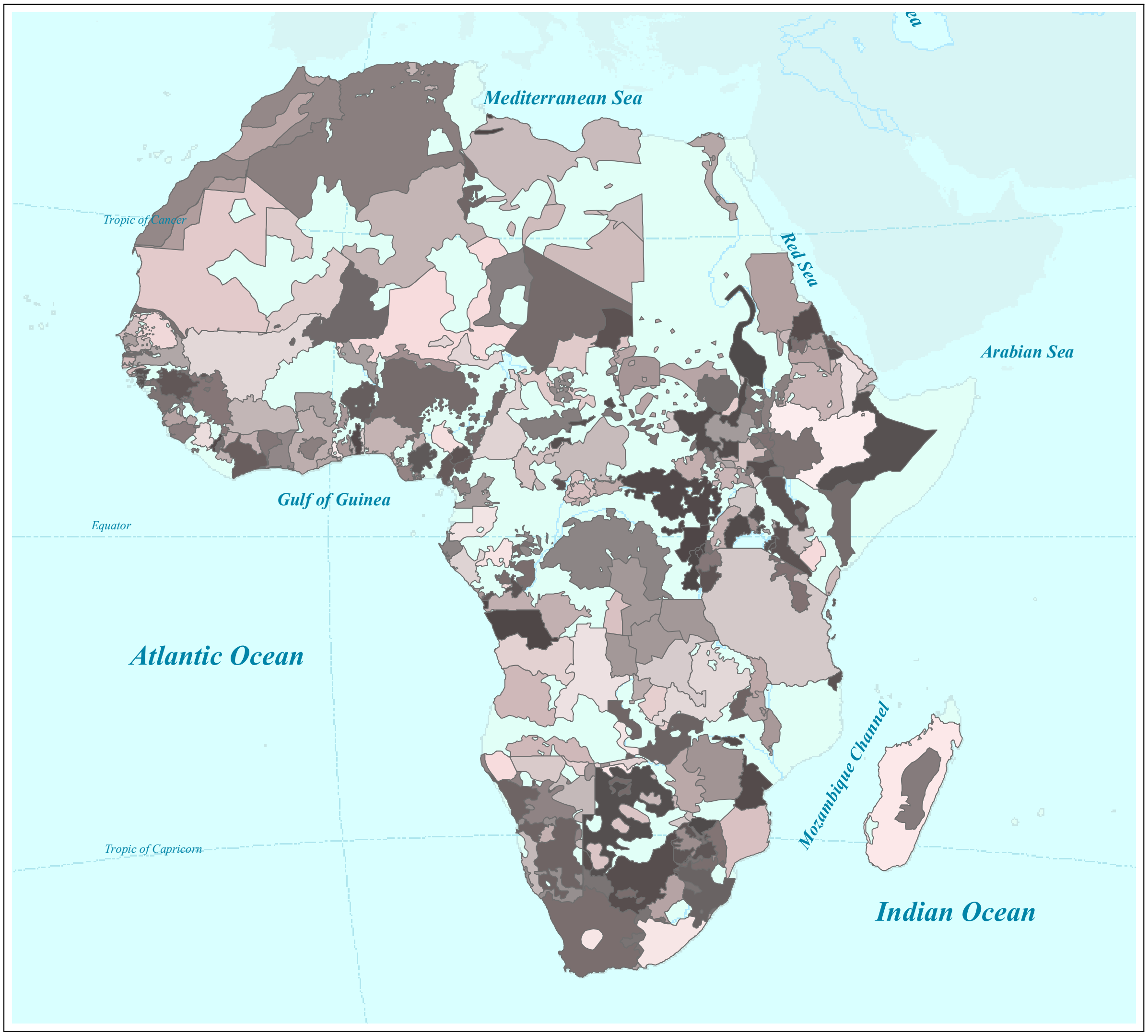
Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust standard errors clustered at the spatial unit level are in parentheses. Column 1 contains SAR estimates and Columns 2 and 3, SDM estimates. The new spatio-ethnic weights matrix is based on contiguity between spatial units inhabited by the same ethnic group or a transborder ethnic kin and is row-normalised because of the sparsity of the matrix. The variable in interaction with W_{Ethnic} is log-luminosity (in Columns 1 and 2), or the product of log-luminosity and lagged local political dominance (in Column 3). All estimations are controlled for spatial unit fixed effects and country$\times$year fixed effects. A two-by-two degrees grid is used.

Table XIII: **Difference-In-Differences Estimations**

	Dependent variable is the Mean of Log $\text{Light}_{i,p}$ (1)	(2)
Interaction	-0.0261	-0.0479
	(0.0108)**	(0.0195)**
	[0.0108]**	[0.0190]**
Treatment	-0.0938	-0.5168
	(0.1257)	(0.0797)***
	[0.1197]	[0.0740]***
Period	0.5545	0.3675
	(0.0040)***	(0.0087)***
	[0.0030]***	[0.0088]***
R^2	0.023	0.615
Additional Controls	No	Yes
Number of Units	1,346	1,346
Number of Periods	2	2

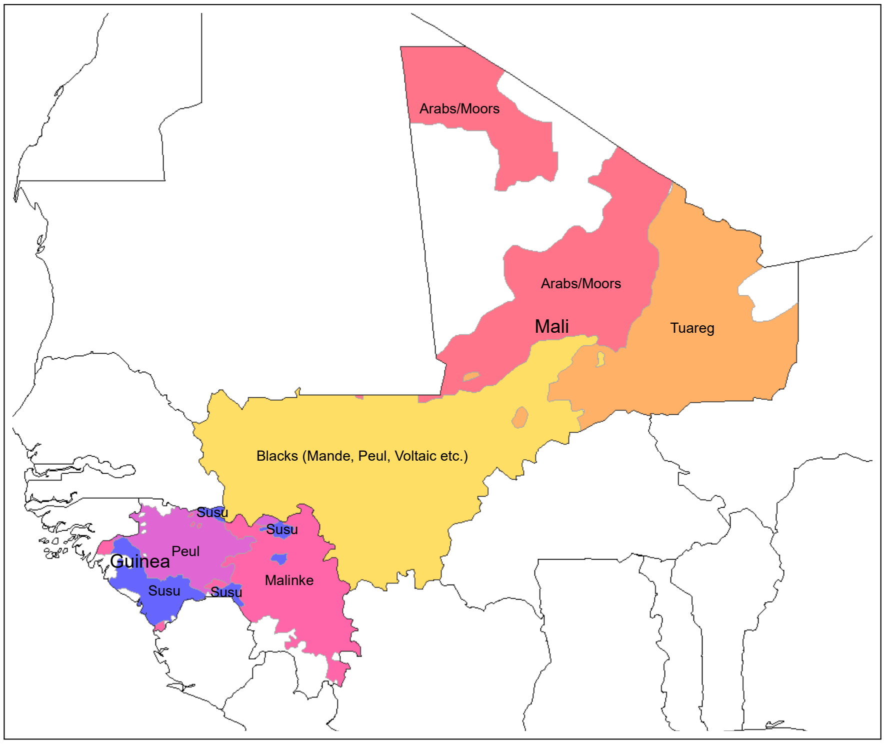
Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Ordinary Least Squares estimates of equation (12) with a synthetic two-period panel constructed from the original dataset. The index p stands for the period. $p = 0$ in the pre-treatment period and $p = 1$ in the post-treatment period. Standard errors clustered at the spatial unit level are in parentheses and bootstrap standard errors are in square brackets. The treatment corresponds to a decrease in the number of dominant transborder ethnic kin groups. Additional controls for Log mean population, Log of mean temperature, and Log of mean precipitations per square kilometre are used in Column 2.

Figure I: **Map of The Ethnic Groups' Homelands in Africa**



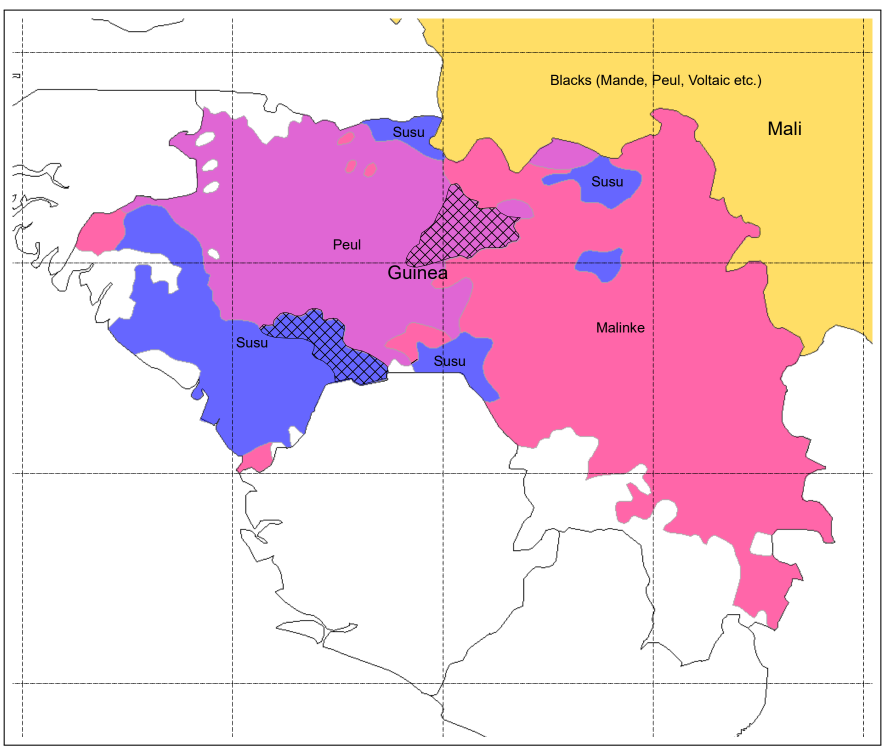
Note: Source is Vogt et al. (2015). In total, there are 194 politically represented ethnic groups. The EPR information exists for all African countries, except for Tunisia, Somalia, Burkina Faso, Equatorial Guinea, Lesotho, and Swaziland.

Figure II: **Map of Ethnic Groups' Homelands in Guinea and Mali**



Note: Guinea is populated by the Susu, the Peul, and the Malinke. The Arabs and the Moors, the Tuareg, and the Mande, Peul and Voltaic coalition live in Mali.

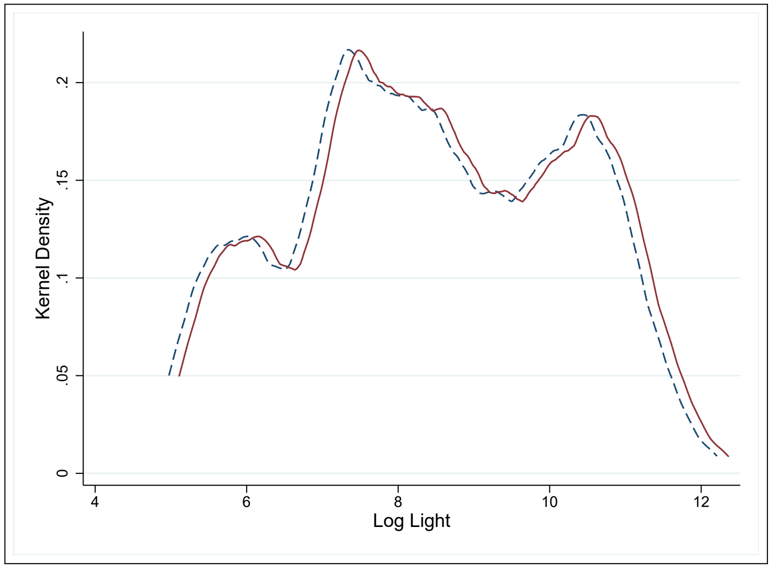
Figure III: **Subdivision in Spatial Units in Guinea**



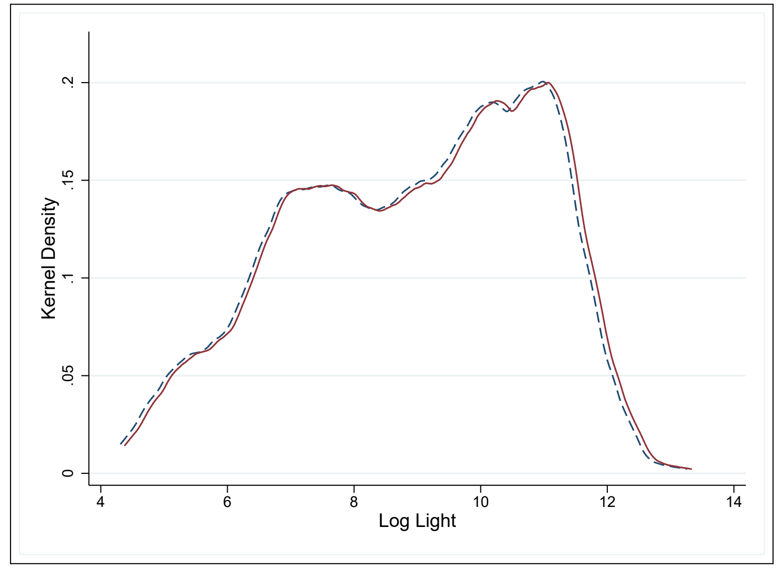
Note: Politically Relevant Ethnic Groups in Guinea. The crossed regions represent overlaps between the Susu and Peul homelands in the west, and between the Peul and Malinke homelands in the east.

Figure IV: **Kernel Density of the Simulated Predicted Values : Experiment 1**

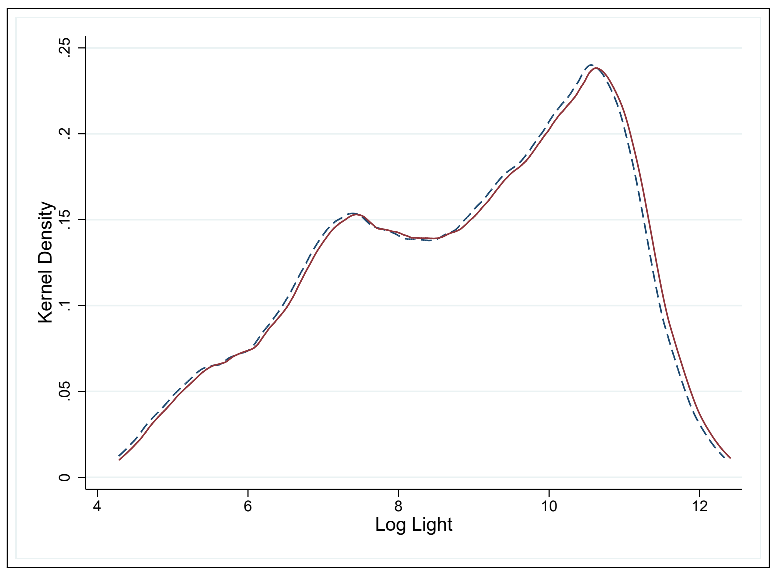
(1) **Ethnic Network & Favoured Groups**



(2) **Ethnic Network & Non-Favoured Groups**



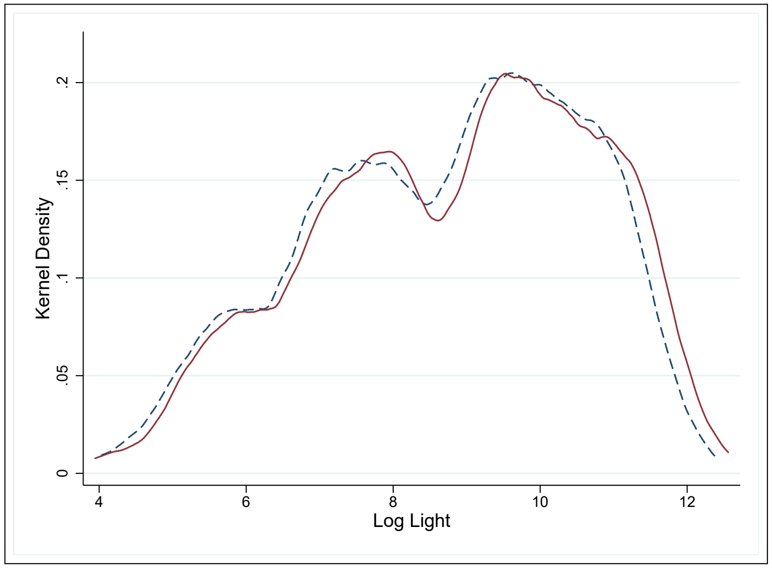
(3) **Ethnic Outsiders & Favoured Groups**



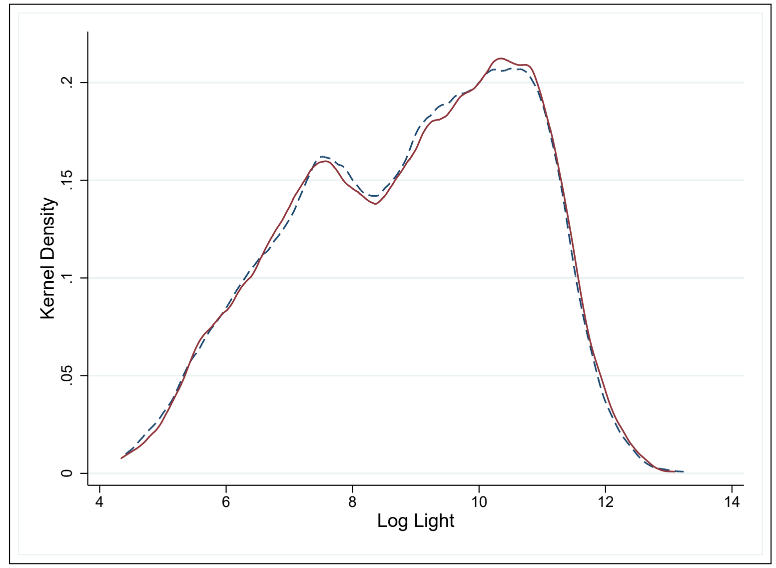
Note: The solid line represents the kernel density estimates of the simulated predicted values in Experiment 1. The dashed line represents the kernel density estimates of the predicted values in the baseline case by categories. Sub-Figure (1) contains the observations of the Ethnic Network & Favoured Groups, Sub-Figure (2) the Ethnic Network & Non-Favoured Groups, and Sub-Figure (3) the Ethnic Outsiders & Favoured Groups. An Epanechnikov kernel with 0.25 bandwidth is used.

Figure V: **Kernel Density of the Simulated Predicted Values : Experiment 2**

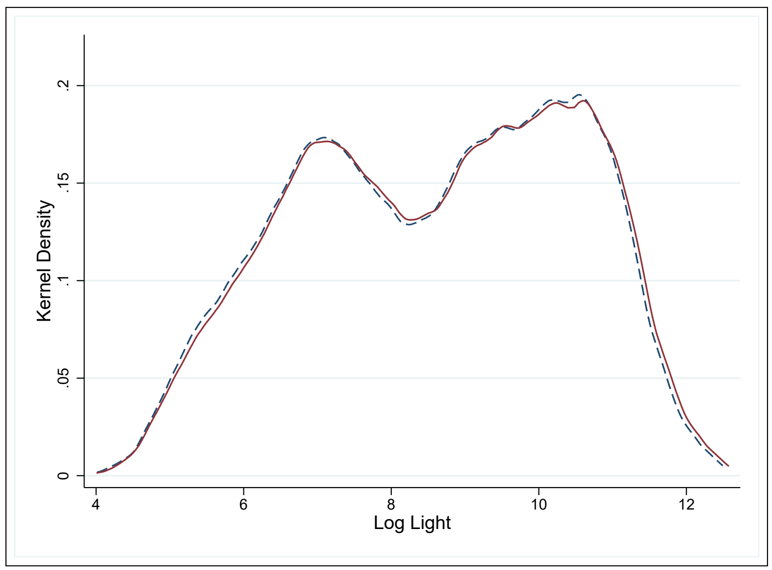
(1) **Ethnic Network & Favoured Groups**



(2) **Ethnic Network & Non-Favoured Groups**

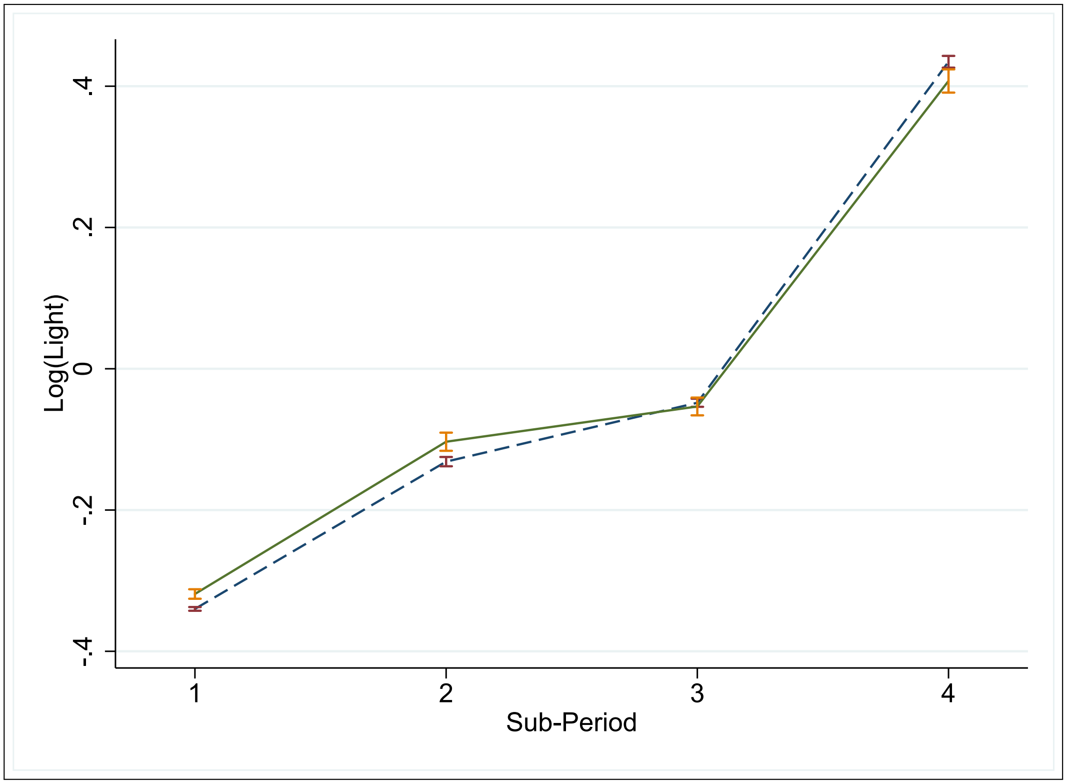


(3) **Ethnic Outsiders & Favoured Groups**



Note: The solid line represents the kernel density estimates of the simulated predicted values in Experiment 2. The dashed line represents the kernel density estimates of the predicted values in the baseline case by categories. Sub-Figure (1) contains the observations of the Ethnic Network & Favoured Groups, Sub-Figure (2) the Ethnic Network & Non-Favoured Groups, and Sub-Figure (3) the Ethnic Outsiders & Favoured Groups. An Epanechnikov kernel with 0.25 bandwidth is used.

Figure VI: **Difference-In-Differences : Diagnosis Analysis**



Note: The figure displays of averages of log-luminosity subdivided into Treatment Group/Control Group and four sub-periods, with five per cent confidence intervals. The dashed line is for the Control Group and the solid line is for the Treatment Group. Since the year of treatment is country specific, the periods ex-ante and ex-post the treatment are split into two periods of equal length. The values are in deviation from the time mean at the spatial unit level.

References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 91(5):1369–1401.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2005). Institutions as a Fundamental Cause of Long-Run Growth. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1, Part A, chapter 6, pages 386–472. Elsevier S.B.V.

Aidt, T. S. and Leon, G. (2016). The Democratic Window of Opportunity: Evidence from Riots in Sub-Saharan Africa. *Journal of Conflict Resolution*, 60(4):694–717.

Alesina, A. and La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3):762–800.

Alesina, A., Michalopoulos, S., and Papaioannou, E. (2016). Ethnic inequality. *Journal of Political Economy*, 124(2):428–488.

Ambrus, A., Mobius, M., and Szeidl, A. (2014). Consumption Risk-Sharing in Social Networks. *American Economic Review*, 104(1):149–82.

Baltagi, B. H., Song, S. H., and Koh, W. (2003). Testing panel data regression models with spatial error correlation. *Journal of Econometrics*, 117(1):123 – 150.

Barro, R. and Sala-i-Martin, X. (2004). *Economic Growth*. McGraw-Hill Advanced Series in Economics. McGraw-Hill.

Bennett, M. M. and Smith, L. C. (2017). Advances in using multitemporal night-time lights satellite imagery to detect, estimate, and monitor socioeconomic dynamics. *Remote Sensing of Environment*, 192:176–197.

Berman, N., Couttenier, M., Rohner, D., and Thoenig, M. (2017). This Mine is Mine! How Minerals Fuel Conflicts in Africa. *American Economic Review*, 107(6):1564–1610.

Bidner, C., Francois, P., and Trebbi, F. (2014). A Theory of Minimalist Democracy. Working Paper 20552, National Bureau of Economic Research.

Burgess, R., Jedwab, R., Miguel, E., Morjaria, A., and Padró i Miquel, G. (2015). The Value of Democracy: Evidence from Road Building in Kenya. *American Economic Review*, 105(6):1817–51.

Card, D. and Krueger, A. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4):772–93.

Cederman, L.-E., Girardin, L., and Gleditsch, K. S. (2009). Ethnonationalist Triads: Assessing the Influence of Kin Groups on Civil Wars. *World Politics*, 61(03):403–437.

Cederman, L.-E., Wimmer, A., and Min, B. (2010). Why do ethnic groups rebel? New data and analysis. *World Politics*, 62:87–119.

Cogneau, D. and Dupraz, Y. (2014). Questionable Inference on the Power of Pre-Colonial Institutions in Africa. PSE Working Papers n2014-25.

Collier, P. and Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, 56(4):563–595.

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

De Weerdt, J. and Dercon, S. (2006). Risk-sharing networks and insurance against illness. *Journal of Development Economics*, 81(2):337 – 356.

Di Tella, R. and Schargrodsky, E. (2004). Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack. *American Economic Review*, 94(1):115–133.

Driscoll, J. (2008). Inside The Leviathan : Coup-Proofing After State Failure. Working paper, Stanford University.

Driscoll, J. (2012). Commitment Problems or Bidding Wars? Rebel Fragmentation as Peace Building. *Journal of Conflict Resolution*, 56(1):118–149.

Dubois, P., Jullien, B., and Magnac, T. (2008). Formal and Informal Risk Sharing in LDCs: Theory and Empirical Evidence. *Econometrica*, 76(4):679–725.

Durlauf, S. N. and Johnson, P. A. (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics*, 10(4):365–384.

Elhorst, J. (2013). *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*. SpringerBriefs in Regional Science. Springer Berlin Heidelberg.

Elvidge, C., Hsu, F., Baugh, K., and Ghosh, T. (2014). *National Trends in Satellite-Observed Lighting: 1992–2012*, pages 97 –120. Global Urban Monitoring and Assessment through Earth Observation.

Englebert, P., Tarango, S., and Carter, M. (2002). Dismemberment and suffocation. *Comparative Political Studies*, 35(10):1093–1118.

Ertur, C. and Koch, W. (2007). Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics*, 22(6):1033–1062.

Esteban, J. and Ray, D. (1994). On the measurement of polarization. *Econometrica*, 62(4):819–51.

Fafchamps, M. and Gubert, F. (2007). The formation of risk sharing networks. *Journal of Development Economics*, 83(2):326 – 350.

Francois, P., Rainer, I., and Trebbi, F. (2014). The Dictator's Inner Circle. Working Paper 20216, National Bureau of Economic Research.

Francois, P., Rainer, I., and Trebbi, F. (2015). How Is Power Shared in Africa? *Econometrica*, 83:465–503.

Fujita, M. and Thisse, J.-F. (2013). *Economics of Agglomeration*. Cambridge University Press.

Giovanni, J., Levchenko, A., and Ortega, F. (2015). A global view of cross-border migration. *Journal of the European Economic Association*, 13(1):168–202.

Giuliano, P. and Ruiz-Arranz, M. (2009). Remittances, financial development, and growth. *Journal of Development Economics*, 90(1):144–152.

Glaeser, E. L., Porta, R. L., de Silanes, F. L., and Shleifer, A. (2004). Do Institutions Cause Growth? *Journal of Economic Growth*, 9(3):271–303.

Gleditsch, K. S. (2007). Transnational dimensions of civil war. *Journal of Peace Research*, 44(3):293–309.

Grimard, F. (1997). Household consumption smoothing through ethnic ties: evidence from Cote d'Ivoire. *Journal of Development Economics*, 53(2):391 – 422.

Guidolin, M. and La Ferrara, E. (2007). Diamonds are forever, wars are not: Is conflict bad for private firms? *American Economic Review*, 97(5):1978–1993.

Gurses, M. (2015). Transnational Ethnic Kin and Civil War Outcomes. *Political Research Quarterly*, 68(1):142–153.

Haynes, K. (2016). Diversity and Diversion : How Ethnic Composition Affects Diversionary Conflict. *International Studies Quarterly*, 60(2):258–271.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028.

Hodler, R. and Raschky, P. A. (2014). Regional Favoritism. *The Quarterly Journal of Economics*, 129(2):995–1033.

Ioannides, Y. M. and Datcher Loury, L. (2004). Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, 42(4):1056–1093.

Jack, W. and Suri, T. (2014). Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution. *American Economic Review*, 104(1):183–223.

Jedwab, R. and Moradi, A. (2016). The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa. *The Review of Economics and Statistics*, 98(2):268–284.

Kapoor, M., Kelejian, H. H., and Prucha, I. R. (2007). Panel data models with spatially correlated error components. *Journal of Econometrics*, 140(1):97–130.

Kelejian, H. H. and Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157(1):53–67.

Klenow, P. and Rodriguez-Clare, A. (2005). Externalities and Growth. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1, Part A, chapter 11, pages 817–861. Elsevier S.B.V.

Kramon, E. and Posner, D. N. (2013). Who Benefits from Distributive Politics? How the Outcome one Studies Affects the Answer one Gets. *Perspectives on Politics*, 11(02):461–474.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3):483–99.

La Ferrara, E. (2003). Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana. *American Economic Review*, 93(5):1730–1751.

La Porta, R., Lopez-De-Silanes, F., Shleifer, A., and Vishny, R. W. (1997). Legal determinants of external finance. *The Journal of Finance*, 52(3):1131–1150.

La Porta, R., Lopez-De-Silanes, F., Shleifer, A., and Vishny, R. W. (1999). The quality of government. *Journal of Law, Economics and Organization*, 15(1):222–279.

Lee, L.-f. and Yu, J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154(2):165–185.

Leider, S., Möbius, M. M., Rosenblat, T., and Do, Q.-A. (2009). Directed Altruism and Enforced Reciprocity in Social Networks. *The Quarterly Journal of Economics*, 124(4):1815.

LeSage, J. and Pace, R. (2009). *Introduction to Spatial Econometrics*. Statistics: A Series of Textbooks and Monographs. CRC Press.

Lucas, R. (1990). Why doesn't capital flow from rich to poor countries? *American Economic Review*, 80(2):92–96.

Marshall, M. and Cole, B. (2011). *Conflict, Governance, and State Fragility. Polity IV Global Report 2011*. Center for Global Policy / Center for Systemic Peace.

Meyersson, E. (2016). Political Man on Horseback, Military Coups and Development. Working Paper, Stockholm Institute for Transition Economics.

Michalopoulos, S. and Papaioannou, E. (2013a). National Institutions and Subnational Development in Africa. *The Quarterly Journal of Economics*, 129(1):151–213.

Michalopoulos, S. and Papaioannou, E. (2013b). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, 81(1):113–152.

Michalopoulos, S. and Papaioannou, E. (2016). The Long-Run Effects of the Scramble for Africa. *American Economic Review*, 106(7):1802–48.

Michalopoulos, S. and Papaioannou, E. (2017). Spatial patterns of development: A meso approach. Working Paper 24088, National Bureau of Economic Research.

Mitton, T. (2016). The wealth of subnations: Geography, institutions, and within-country development. *Journal of Development Economics*, 118(C):88–111.

Muller, C. (2018). Difference-in-Differences Over Time. Mimeo, Aix-Marseille School of Economics, Aix-Marseille University.

Munshi, K. and Rosenzweig, M. (2016). Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap. *American Economic Review*, 106(1):46–98.

Papaioannou, E. and Siourounis, G. (2008). Democratisation and Growth. *The Economic Journal*, 118(532):1520–1551.

Pinkovskiy, M. L. (2017). Growth discontinuities at borders. *Journal of Economic Growth*, 22(2):145–192.

Posner, D. N. (2004). The Political Salience of Cultural Difference: Why Chewas and Tumbukas Are Allies in Zambia and Adversaries in Malawi. *American Political Science Review*, 98(4):529–545.

Storeygard, A. (2016). Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa. *Review of Economic Studies*, 83(3):1263–1295.

Tinbergen, J. (1962). An Analysis of World Trade Flows. *Shaping the World Economy*.

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L.-E., Hunziker, P., and Girardin, L. (2015). Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family. *Journal of Conflict Resolution*, 59(7):1327–1342.

Weidmann, N. B., Rød, J. K., and Cederman, L.-E. (2010). Representing ethnic groups in space: A new dataset. *Journal of Peace Research*, 47(4):491–499.

Wucherpfennig, J., Weidmann, N. B., Girardin, L., Cederman, L.-E., and Wimmer, A. (2011). Politically Relevant Ethnic Groups across Space and Time: Introducing the GeoEPR Dataset. *Conflict Management and Peace Science*, 28(5):423–437.

Yang, D. (2011). Migrant remittances. *Journal of Economic Perspectives*, 25(3):129–52.

Yu, J., de Jong, R., and Lee, L.-F. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics*, 146(1):118–134.