Transborder Ethnic Kin and Regional Prosperity : Evidence from Night-time Light Intensity in Africa*

Christophe Muller[†], Pierre Pecher[†]

March 6, 2017

Abstract

This study investigates the consequences of cross-border ethnic linkages affecting local income in Africa. We show that kinship connections to transborder ethnic clans affect the allocation of public good spending, measured by luminosity in an ethnic group's homeland, controlling for many other factors. We estimate spatial panel models that relate politico-ethnic variables to luminosity measured by satellite imaging from 1992 to 2012. After correcting for unobserved factors, we find that having more politically dominant transborder ethnic kin groups significantly increases economic activity measured by luminosity in the corresponding ethnic homeland. Finally, in contrast to prior evidence, we show that ethnic groups sharing power are not necessarily richer when accounting for spatial effects.

JEL Classification: D72, R11, O43.
Keywords: Local Development, Ethnic Networks, Institutions.

*This research was accomplished thanks to the financial support of the A*MIDEX project 'Social, Demographic and Democratic Transitions in MENA'. We are grateful to participants at the 3rd TMENA Meeting and at seminar presentations at UCLouvain and Aix-Marseille University for interesting and useful comments.

†Aix-Marseille Univ, CNRS, EHESS, Ecole centrale de Marseille, GREQAM, France

1 Introduction

Do cross-border relations linking ethnic groups between neighbouring countries generate externalities from central government participation with measurable consequences in terms of local prosperity? This is the question studied in this paper. One can think of several non-exclusive channels for these externalities, like e.g. technology diffusion, human capital accumulation, trade, violence or institutions. Concerning these linkages, a small literature studies the consequences in terms of risk of civil conflict resulting from the presence of transborder ethnic kin bonds and argues that it facilitates the collective action to take up arms.[1] Yet, as Gleditsch (2007) and Gurses (2015) sustain, these ethnic bonds should not necessarily increase the probability of conflict, but might instead give rise to preferable outcomes for the allied faction, by shifting the balance of power in their favour.

In this paper, we investigate this conjecture by looking at the African continent as a whole. In particular, we assess whether connections to ethnic groups included in central governments in neighbouring countries give a comparative edge in the distribution of public goods. The empirical relevance of these ethnic ties is bolstered by papers exploring the issue of partitioned ethnicities on the African continent, like Michalopoulos and Papaioannou (2013a) on long-run comparative development and Michalopoulos and Papaioannou (2016) on social unrest.[2] Compared to this, we study short-term fluctuations in prosperity consequent to changes in the political configuration occuring in neighbouring countries, in a spatial panel setting. Our work builds on Hodler and Raschky (2014) who are the firsts to assemble a panel data set of night-time light intensity for subnational units to study the effect of the political leader's birthplace, although we broaden the scope of the effect by considering political participation by ethnic constituencies instead of the sole position of president and by moving away from the apportionment in administrative regions by considering ethnic homelands.

The economic consequences of transborder ethnic kin relations are our main focus but there are other political economy effects involving ethnicity and it is crucial to control for them when studying ethnic interactions.[3] For example, political leaders tend to favour their

[1] This may happen thanks to financial and military support and by providing a safe haven for the retreat of the rebels. See for instance Haynes (2016), Aidt and Leon (2015) and Cederman et al. (2009).

[2] Michalopoulos and Papaioannou (2016) argue that the drawing of boundaries by colonial powers way back in the nineteenth century was unrelated to the location of historic homelands of African ethnicities and that the resulting separations can thus be regarded as random. They use this accident as a quasi-natural experimental setting and show that the prevalence of conflicts is higher over the long period in the regions inhabited by split groups. In Michalopoulos and Papaioannou (2013a), they show that institutional differences between countries in the homeland of split ethnicities do not have a significant effect on comparative development measured by luminosity.

[3] Throughout the paper, we use the convention of naming the related group abroad the kin group, or we use the acronym TEK for Transborder Ethnic Kin. We prefer to reserve the term 'group' to describe the internal situation. A note of clarification is perhaps necessary at this stage to fix ideas. In this interpretation where we

ethnic group in the distribution of public goods. Burgess et al. (2015) for instance find that the districts sharing the leader's ethnicity receive up to twice as much road construction expenditures in Kenya.[4] Our framework incorporates this by controlling for the internal political status of the ethnic groups.

The mechanism that we test here rests on the presence of a network of ethnic ties connecting populations between modern nations on the African continent.[5] This situation was due historically to the division of land by colonial borders that were subsequently upheld at independence in spite of their incompatibility with traditional kingdoms and their the artificiality amplified by their permeable and non-signalled nature. The reciprocity linking the ethnicities of split homelands is still strong nowadays because of the shared language and culture and give rise to many kinds of interactions and loyalties, helped by the fact that borders are mainly porous. 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 monopoly over power to the detriment of the northern Tuareg and Moors groups. What were the consequences of this evolution in neighbouring Guinea for the powerless Malinke and Peul groups, in a country where the Susu group was dominant at that time. We mitigate endogeneity concerns thanks to the fact that this type of variation allows us to control for many observed and unobserved factors by including spatial unit fixed-effects and country-year fixed effects in our specifications.[6]

In the context of weakly institutionalized states, the theoretical models of ethnic bargaining predict that the stake allotted to a faction is proportional to its expected utility in case of strife.[7] By reducing the cost of conflict of a side compared to the others and thus improving

look at recent outcomes variations due to recent changes, populations separated by a modern state border thus constitute distinct ethnic entities, even though historically they descend from the same lineage. For example, the Shona-Ndau group in Mozambique is not from the same group as the Shona group in Zimbabwe. We would rather say that the Shona from Zimbabwe is the Transborder Ethnic Kin of the Shona-Ndau in Mozambique. In some sense, this nomenclature better reflects the political situation on the ground than that based on historical kingdoms, that exist nowadays more in the form of solidarities and traditional bonds.

[4]This finding is confirmed by other studies like Franck and Rainer (2012) on health and education in a sample of African countries and Kramon and Posner (2013) on various outcomes simultaneously. Dreher et al. (2015) show that Chinese official development aid is disproportionately affected towards regions of the leader's birthplace.

[5]This network is documented in the Ethnic Power Relations Database, which is our main source of information on the location and political status of ethnic groups. A code is associated to each ethnic group and the links are identified by the list of codes corresponding to the ethnic friends, for each group.

[6]The spatial unit fixed effects control for all elements that are constant over time in a particular region like the geographic situation (distance to the coast or capital), the type of settlements, the natural environment or the presence of subsoil resources. Country-year fixed effects control for all elements that apply to a particular country in a particular year, like occurence of civil conflict, coups, foreign military intervention and flows of official development aid.

[7]Francois et al. (2014, 2015); Bidner et al. (2014); Driscoll (2008, 2012). The ruler in Francois et al. (2015) distributes cabinet minister positions to elite members of rival ethnicities in order to secure his leadership and maximizes the expected gains tied to his office. See also the literature on neopatrimonialism (Kelsall, 2011;

outside options, the presence of a dominant transborder ethnic kin might tilt redistributive politics favourably for them. A contribution of this paper is to construct a novel measure of political dominance by ethnic friends abroad to quantify this effect.

Beyond this particular effect, societies located next by next may affect each other through a myriad of other channels. For instance, a group in power could decide the construction of a road improving the connection with their kin who would benefit economically thanks to trade or other newly available opportunities. Jedwab and Moradi (2016) and Storeygard (2016) estimate the consequences on regional development in Africa of railroad and road connections and show that these infrastructures have lasting effects, by allowing new profitable activities thanks to lower transport costs to international markets. Likewise, a dominant and richer group could offer greater latitude for lucrative smuggling activities or could give back a greater amount of remittances, which constitute a pervasive and direct way to transmit economic wealth between co-ethnics (Yang, 2011; Giuliano and Ruiz-Arranz, 2009). We acknowledge these other kinds of interactions and our transborder index typically measures the total of all these possible ethnic impacts.

Research Question Our goal is to answer the following questions. Do cross-border ethnic linkages generate externalities affecting the equilibrium distribution of income between territories in African countries? In particular, how does the control of the state apparatus by transborder ethnic kin groups in a neighbouring country affect the allocation of public spending and regional economic development (proxied by night-time light intensity) in an ethnic group's base region? What were, for instance the consequences for the Shona-Ndau group in Mozambique of the fact that their related kin, the Shona group became politically dominant in Zimbabwe when Mugabe ceased to share power with the Whites and Ndebele-Kalanga-Tongo in 2000?[8] We estimate this effect, while controlling for numerous other factors, including political participation, meteorological conditions and civil conflicts.

Results To investigate these issues, we estimate standard panel models and spatial panel models that relate the politico-ethnic variables to 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 2929 spatial units and 21 consecutive years from 1992 to 2012, we find that having more politically dominant transborder ethnic kin groups significantly increases economic activity measured by luminosity in the corresponding ethnic base regions. Our Spatial Autoregressive (SAR) estimates of the coefficient of the variable Ethnic Kin Dominant (EKD) imply that the average neighbouring groups gains around 2% in light intensity, which corresponds to an approximate 0.6% rise in income in the long run, whenever a transition to dominance occurs. This effect remains statistically significant over

Cammack, 2007; Clapham, 1985) to understand the internal logic of the functioning of these states in the grey zone, between fully-fledged democracies and autocracies.

[8]In 2000, the government conducted its Fast Track Land Reform programme, a policy of forced redistribution of land from the minority white population to the majority black population.

a broad range of alternative specifications.

Our SAR models with spatial unit fixed-effects and country-year fixed effects integrate the consequences of these ethnic externalities on the one hand, and of spatially correlated unobservables on the other, by including a spatial lag of the dependent variable. In this setting, any remaining variability captured by our transborder kin index would then have to be attributed specifically to ethnically related factors. We find that even with a substantial fraction of the variability captured by the spatial lag of the dependent variable, the transborder impact (EKD) remains positive and significant. Thus, ethnic groups with connections to politically dominant tribes are not merely richer because their neighbours are richer, but also because other causes operate, possibly through a change in the internal balance of power, or other ethnic channels.

We break down the SAR estimates of our variable of interest in direct and indirect components and find that they are in line with our previous findings. In addition to that, we perform a series of robustness checks to assess the solidity of our findings. Beyond using many controls and robust standard errors, we (i) vary the type of spatial models used, (ii) change the definitions of geographic locations and (iii) the assumptions made in the construction of our spatial weights matrix and (iv) transborder power indices.

Related Literature Our work is related to the literature using night-time light as a measure of development. In a celebrated seminal paper, Henderson et al. (2012) introduce this data in the economics field as a measure of economic growth less prone to errors in countries that notoriously report highly contaminated GDP figures. The global coverage and fine-grain of this data are manifest advantages. Magee and Doces (2015) for instance, estimate how much autocracies generally exaggerate their income. Alesina et al. (2016) construct measures of ethnic inequality and show that they provide better predictors of slow development than traditional diversity measures in a cross-country analysis. Michalopoulos and Papaioannou (2013a) link comparative development proxied by night light with the ethnicities of the historic Murdock map that were partitioned by contemporary borders. They demonstrate that differences in institutions across modern states have no effect in terms of local development in these split homelands. Michalopoulos and Papaioannou (2013b) show that pre-colonial centralization positively affects modern wealth. The closest papers to ours are Hodler and Raschky (2014) and Hodler et al. (2015). The first of these papers assesses the effect of the leader's birthplace on regional favouritism measured by changes in night-time light intensity in administrative regions, while Hodler et al. (2015) evaluate the effect of the leader's ethnicity on ethnic favouritism in ethnic homelands. Both these papers find positive correlations.

We also relate to the huge theoretical and empirical literatures on economic growth on the one hand e.g. Barro and Sala-i Martin (2004) and to the institutional economics literature that studies the root causes of development, on the other hand (Acemoglu et al., 2001; Glaeser et al., 2004). A number of authors concentrate on the consequences of autocracy, democracy or other types of power sharing arrangements and transitions between these political states.[9] However, no evidence of kin solidarity related to power participation is present in the literature.

We focus our attention on Africa, 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 politics there (Posner, 2004). Finally, 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).

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

2 Empirical Model

2.1 Panel Data Model

Our first empirical specification is based on a standard panel model with the measure of development as dependent variable, as in Hodler et al. (2015). We express it observation by observation in a first step :

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

$\text{Log(Lights)}_{i,t}$ is our dependent variable, the logarithm of total light intensity emitted by spatial unit i in year t. $\text{EKD}_{i,t-1}$ 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 $t-1$. We choose to use all explanatory variables lagged by one year instead of the contemporary values to allow for delays and mitigate the occurrence of simultaneitiy issues in the estimations. EKD is constructed using the information contained in the Transborder Ethnic Kin supplementary material of the Ethnic Power Relations database (Vogt et al., 2015). For each ethnic group, we can identify their transborder ethnic connections that result generally from ethnic partitioning. EPR-TEK contains the political status of these groups at each point in

[9] Meyersson (2015) studies the economic consequences of coups and Papaioannou and Siourounis (2008) challenges the view that democratizations do not have a favourable impact on growth considering consolidated transitions only. Mitton (2016) confronts the role played by geography and institutions in a panel of sub-nations in 101 countries.

time in their respective countries. Consequently, we are able to compute this time-varying index reflecting political transitions in the sample.

The vector $X_{i,t-1}$ contains $k-1$ other control variables, encompassing a count variable of transborder ethnic friends with the political status 'in power' instead of dominant and indicator variables of political status in domestic politics for dominance and inclusion in the government. We control also for population, meteorological factors (precipitation and temperatures) and for the occurrence of civil conflicts.

The spatial unit fixed-effects are represented by α_i and control for all elements that do not vary over time. We can think for instance about many geographical factors like the presence of mountains or rivers and the distance to the coastline or capital city. $\delta_{c(i),t}$ stands for the country-year fixed-effects i.e. a set of parameters associated with binary variables for each combination of country and year taking the value one in observations corresponding to that particular combination and zero elsewhere.[10] These variables control for a wide range of observed and unobserved factors like major institutional changes in a country, or shifts in its international relations with aid donors or commercial partners. These situations would thus not create bias in our estimates. Finally, $\epsilon_{i,t}$ is an error term such that

$$E[\epsilon_{i,t}|EKD_{i,t-1}, X_{i,t-1}, \alpha_i, \gamma_{c(i),t}] = 0 \qquad \text{for } i = 1, \ldots, N \text{ and } t = 1, \ldots, T \tag{2}$$

which is a standard strict exogeneity assumption. By piling all observation numbered by $i = 1, \ldots, N$ of each cross-section in a certain year t in vectors and matrices, we obtain the matrix expression of the model :

$$Y_t \;=\; EKD_{t-1} \cdot \gamma + X_{t-1} \cdot \beta + \alpha_0 + \delta_{c,t} + \epsilon_t \qquad \text{for } t = 1, \ldots, T \tag{3}$$

where we define Y_t to be equal to $\text{Log(Lights)}_t = (y_{1,t}, y_{2,t}, \ldots, y_{N,t})'$, a $N \times 1$ vector describing the logarithm of luminosity for all spatial units in year t. Our main independent variable of interest stays identical : EKD_{t-1}, the number of neighbouring kin groups that are politically dominant at $t-1$, of dimension $N \times 1$. Similarly, X_{t-1} is a $N \times (k-1)$ matrix of time-varying controls, again lagged one year. α_0 is a $N \times 1$ vector of spatial unit fixed-effects. $\delta_{c,t}$ stands for the country-year fixed effects i.e. it is a $N \times 1$ vector that contains constants common to groups of spatial units by country.[11] (γ, β) is a $k \times 1$ parameter vector to be estimated.

[10] The notation $c(i)$ is a function that applies the index of a spatial unit to the index the corresponding country that contains it.

[11] In some specifications, we will replace the country-year fixed-effects by year fixed-effects, δ_t, that is a constant for each year in the sample.

2.2 Spatial Panel Data Model

One lingering concern with our model so far is the possible occurrence of spatially correlated omitted variables on the one hand, and spatial heterogeneity and spillover effects on the other. To deal with these issues, specifications including linear combinations of the dependent variable, the independent variables and error terms of the neighbouring units are described in Anselin (1988). Methods for maximum likelihood estimation, testing and inference of the panel versions of these models have recently been developed.[12] To take spatial interactions into account, we extend equation (3) by adding a spatially lagged dependent variable. As in the the previous subsection, we proceeded stepwise by presenting first the model in observation by observation form which is:

$$\begin{aligned} Y_{i,t} &= \rho \sum_{j=1}^{N} w_{i,j} Y_{j,t} + EKD_{i,t-1} \cdot \gamma + X_{i,t-1} \cdot \beta + \alpha_i + \delta_{c(i),t} + \epsilon_{i,t} \\ &\quad \text{for } i = 1, \ldots, N \text{ and } t = 1, \ldots, T \end{aligned} \tag{4}$$

The new term in equation (5) compared to equation (2) involves a single regressor coupled with the parameter ρ, denominated the spatial autoregressive parameter (Kelejian and Prucha, 2010). The associated variable is a linear combination of the dependent variable, with weights described by the $w_{i,j}$'s. We will discuss in greater detail how these are established below. For now, let's just recognize that a positive value indicates that units i and j are neighbours in some sense and that the weight is equal to zero in the contrary.[13]. These weights are constant in time and non-stochastic. The $\epsilon_{i,t}$'s are the error terms and we assume that their conditional distribution

$$\epsilon_{i,t} \Big| \sum_{j=1}^{N} w_{i,j} Y_{j,t}, EKD_{i,t-1}, X_{i,t-1}, \alpha_i, \gamma_{c(i),t} \text{ is NID}(0, \sigma_i^2).$$

By piling up equation (5), using the $N \times N$ spatial weights matrix W, the general form of our spatial model reads:

$$Y_t = \rho W \cdot Y_t + EKD_{t-1} \cdot \gamma + X_{t-1} \cdot \beta + \alpha_0 + \delta_{c(i),t} + \epsilon_t \qquad \text{for } t = 1, \ldots, T \tag{5}$$

(γ, β) is a $k \times 1$ parameter vector to be estimated, again. We refer to $W \cdot Y_t$ as the spatial lag of Y_t. In our baseline estimations, W is a $N \times N$ spectral-normalized matrix of spatial

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

[13] By the way, a unit is never neighbour to itself that is $w_{i,i} = 0 \forall i$

weights, but our findings are robust to alternatives. This specification allows for spatial dependence between geographically close units through the explained variable.

A concurrent explanation to our finding about the effect of dominant transborder kin groups could be that it is in fact a consequence of spatial spillovers, since kin ethnicities are often located in neighbouring areas. In particular, since the related groups are often located on both sides of the partitioning country border, a positive effect could result from interactions distinct from a shift of power in internal politics. There are various reasons why economic activity tend to be clustered in space.

Trade models (Tinbergen, 1962) show that bilateral flows of goods and services are the largest between nearby places. Likewise, opportunities in terms of job creation and incentives to educate may arise from positive externalities between close units. For example, a worker might choose to migrate across the border to access a better job and remit cash to his family back at home. Similarly, a student could choose to pursue his education longer in view of possibilities offered by the firms located across the border. Conflicts also tend to diffuse across space. Therefore, investment in a region could be thwarted by concerns that situations of insurgency or unrest could spread. This is why security somewhere might imply that surrounding places perform better economically.

However, all these mechanisms are not specifically related to ethnic ties linking populations on the two sides of the border and it should be possible to distinguish ethnic connections from spatial ones by incorporating both kinds of effects in the regression. As a consequence our estimation strategy explicitly models those interactions by adding a spatially lagged dependent variable to our equation, meaning that a linear combination of the dependent variable in the nearby geographic units is added as a control. We expect this variable to capture a large fraction of the variability in the data.

In this setting, the variable counting the number of dominant friendly groups captures the effect when these ethnic ties are present above what would normally prevail. Our main preferred explanation of the consequence of this change in the political power of the friendly groups is in terms of threat towards the central government.

Michalopoulos and Papaioannou (2016) show that this explanation is empirically relevant, as conflicts are disproportionately present in the homelands of split groups. However, we cannot entirely dismiss other explanations that are the consequence of the ethnic ties and the fact that the other group becomes dominant. The normal externalities mentioned above could be stronger with ethnic friends even if this does not undermine our message, that buttresses the importance of transborder ethnic relations.

Because of the presence of the spatially lagged dependent variable, the coefficients γ and β cannot be interpreted straightforwardly. A change in the r^{th} explanatory variable associated with $beta_r$, generates not only a direct effect on the dependent variable in unit i, and but also an indirect effect through its impact on the others units $j \neq i$, and the feedback effect of these units on unit i. In our understanding shared and that of Elhorst (2013), the spatial direct effect subsumes the effects which corresponds to what would happen in the absence of neighbours and the feedback loops from unit i to itself through all possible paths across the network.

An expression of all marginal effects taken together, can be obtained by pre-multiplying equation (5) by the matrix $(I_n - \rho W)^{-1}$. The matrix that summarizes these effects is[14]

$$S_r(W) \;=\; (I_n - \rho W)^{-1} I_n \gamma_r. \tag{6}$$

The complex structure of the network reflected in W implies that the direct and indirect effects may be heterogeneous across units. However, to simplify reporting, we collapse the effect into a summary statistic over all units. The formula for these statistics are given in the Appendix A. Essentially, for the r^{th} explanatory variable, the estimate of the direct effect is the average of the diagonal elements of $S_r(W)$, while the indirect effect is the average over all off-diagonal elements of this matrix. It is important to keep in mind that the indirect effects are stated as the effect on the whole set of other units. This number must be divided by the average number of contiguous locations to obtain the average indirect effect for a typical unit different from the original one. The estimators of these indicators as well as their standard errors have been derived in Yu et al. (2008) and Lee and Yu (2010). Estimates of the spatial coefficients ρ as well as their estimated standard errors are also available.[15]

3 Econometric Issues

3.1 Panel Data Model

We specify a yearly panel, first to exploit all the information contained in the 22 years of light data, and second, to match the timing of changes in political status of the ethnic groups.[16]

Our two political variables, $DOM_{n,t-1}$ and $EKD_{n,t-1}$, are suspect of endogeneity. For instance, some unobserved characteristics at the ethnic group level, like social norms and local institutions, could be correlated to both the participation of the ethnic group in the national government and the group's economic activity. This situation may cause our estimates to be inconsistent as they would also reflect the impact of both the political inclusion and the unobserved element. Reverse causality is a concern here as well, because richer populations could be able to marshal the necessary means for their political action and to provide support to their neighbouring related kin. For example, a richer region might be more likely to become included in the government. Likewise, a poorer ethnic group may generate larger

[14] The details of the computations are in Appendix A.

[15] It is important to check that these autoregressive coefficients are below one for the stability of the model. However, this is not a problem in any of the estimations reported in this paper.

[16] Random-Effects are rejected by the results of Hausman tests in favour of Fixed-Effects specification.

external effects on neighbouring kin groups inducing them to enter the political arena. Even though, the treatment of endogeneity is not very advanced in this literature mainly because of the difficulty to find good instruments. The difference in scale between lights measured in small geographic areas and ethnic power participation that happens and the national level mitigates the issue of reverse causality. About omitted variable bias, endogeneity must come from factors that change over time.[17]

Using a fixed-effects framework is likely to reduce much the bias due to omitted variables. Indeed, all variables constant over time are perfectly captured by the fixed effects. Only the remaining co-movements between our explanatory variables and unobservables must be dealt with.

3.2 Spatial Model

Our grid is similar to that of Alesina et al. (2016), i.e. a mesh fitted together with the meridians and parallels of the globe, albeit six times wider. This choice is guided by limitations in the number of observations of our spatial estimator. We use a grid because some ethnic regions span over extended territories and we want our fixed-effects to have a clear interpretation, distinguishing for instance coastal areas from the hinterland.

The estimation of the model in equations (5) requires us to specify a spatial weights matrix. The estimation of our spatial model involves the maximization of likelihoods, which are cumbersome and voracious in machine-time. We use an spectral-normalized contiguity matrix. This way, we restrict the autoregressive parameter to be within to the interval $[-1, 1]$.[18]

An alternative would have been to insert a time-lagged dependent variable in the model, like in Yu et al. (2008), but this is not feasible in practice. This issue is related to the presence of many fixed effects in our equation.[19] Finally, the reason why we may not use the transformation proposed by Lee and Yu (2010) is because it is only applicable in models without year-FE, which are important here due to the nature of the data. This is because an internal setting of the measurement device changes how lights are amplified over time, moving the reference point while not the amplitude. This reference point varies from year to year, but this information is not available.

[17] To alleviate the endogeneity concern, Hodler et al. (2015) use a time trend and an indicator specific to the ethnic region that will later have one of their member become the leader and find that these variables are not significant. The instruments must be correlated with the endogenous regressors and uncorrelated with the regressand.

[18] Row normalization is defined above. Each row is divided by the row total. In spectral normalization, each element is divided by the modulus of the largest eigenvalue of the matrix. Both ways of doing result in a matrix that has eigenvalues not above 1 and estimates of the spatial coefficients necessarily between -1 and +1.

[19] Still, we report one spatial estimation with a time-lagged dependent variable and a spatially dependent variable for illustrative purpose. We do not choose this estimation as our benchmark because it is the only one that successfully converged among a variety of attempts.

4 Data

4.1 Sources and Units of Observation

Picture of overlapping ethnicities.

Our dataset is constructed by combining geo-referenced information using a Geographic Information System, commonly referred to as a GIS, which is a computer program specialized in the treatment of geographic information. This data comes in the form of shapefiles for areas, lines and points or in the form of rasters, which are chequered images of pixels that each have a numerical value. The GIS allows us to perform operations of combination and computation on these files and to extract numerical values. For example, the light data comes in the form of a raster and it is possible to calculate the sum of the value associated with the pixels falling in a certain region defined by a shapefile, be it a country, or the homeland of an ethnic group.

The construction of our unit of observation is illustrated in Figure II for the Democratic Republic of Congo with a simplified example, neglecting all ethnic groups except the Mongo and the Luba Kasai. Panel (a) of the figure displays the homeland of the Mongo group.[20] Panel (b) adds in the homeland of the Luba Kasai group. Because these areas intersect, our procedure creates a different unit there. Also, because our objective is to estimate useful spatial models, we cannot afford to have regions that span over too large areas. Thence, we add a grid in a manner similar to Alesina et al. (2016) in the picture. The grid that we use is represented in panel (c) for DRC. Our grid is six times wider than in Alesina et al. (2016). This choice is guided by sample size considerations, as the spatial models are difficult to estimate when the sample size is too large.

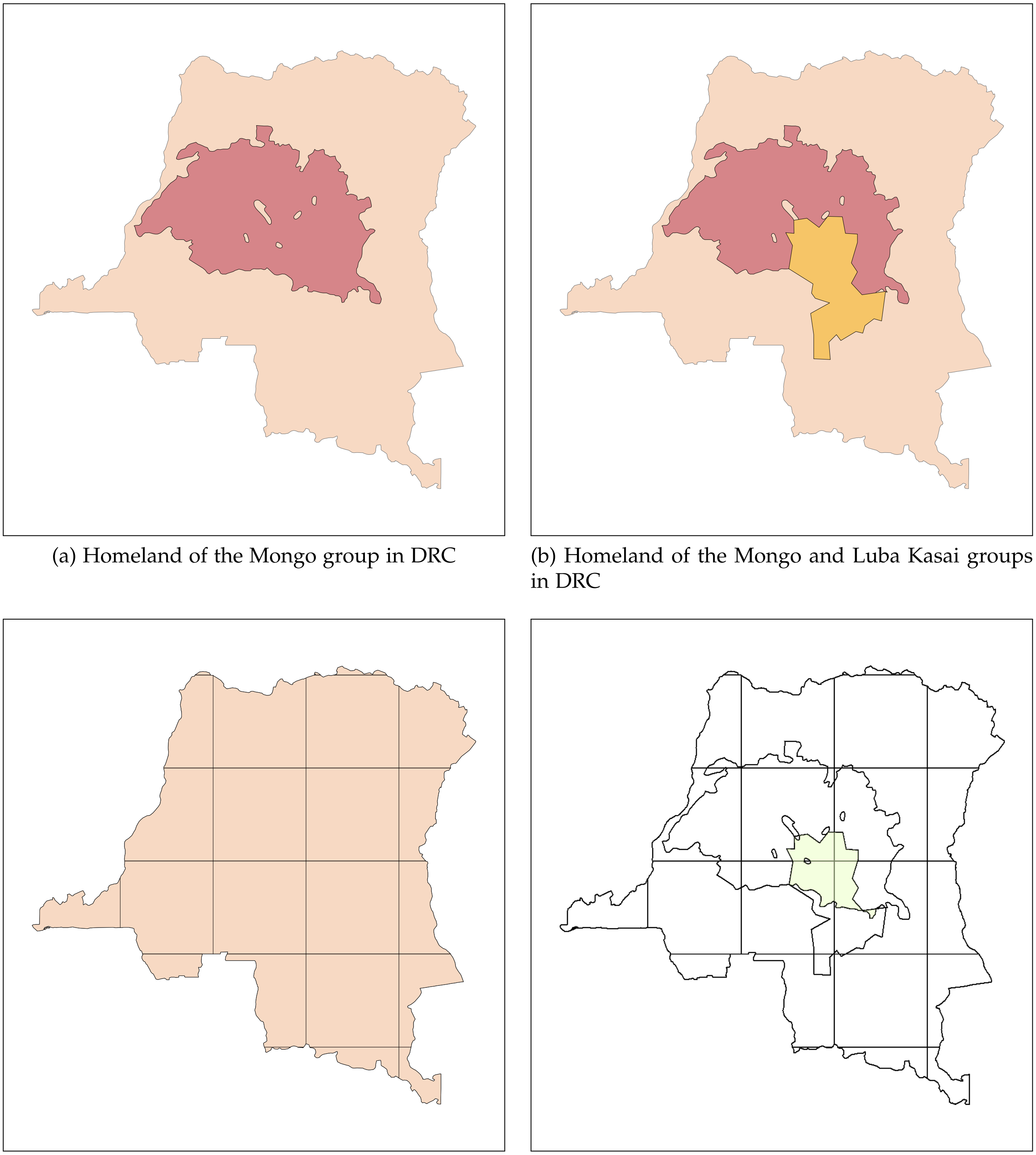
At the opposite, making sure that the zones are not too small reduces measurement errors, when enough light pixels fall into each area. This procedure also guarantees that the cells are time invariant and allows the construction of a perfectly balanced panel, a condition necessary for the estimation of the spatial models.[21] Table II (d) shows the resulting units in this simplified case.

To construct our subdivision, we thus use a 2x2 or a 4x4 degrees grid covering the whole African continent that we intersect with the contemporary country borders and the 238

[20]The EPR list of the 14 ethnic groups of DRC is : Mongo, Tetela-Kusu, Lunda-Yeke, Luba Kasai, Azande-Mangbetu cluster, Bakongo, Luba Shaba, Lulua, Tutsi-Banyamulenge, Ngbandi, Mbandja, Ngbaka, Luba Kasai and Other Kivu groups.

[21]The ethnic regions are valid for certain time periods. In constructing our grid, we use the regions of all years at the same time. Also, to avoid very small areas to be generated at the borders of countries and ethnic regions due to an imperfect superposition of the geographic data from different sources, we force the vertices of our polygons to be at least 20 kilometres apart. This choice is guided by sample size considerations but also reduces measurement errors in the dependent variable, that could occur if too few light pixels fall into the geographic unit of observation.

Figure II: ***The Construction of the Geographic Units of Observation.***



(a) Homeland of the Mongo group in DRC

(b) Homeland of the Mongo and Luba Kasai groups in DRC

(c) Grid of Parralels and Meridians

(d) Resulting Grid of Cells

relevant ethnic regions from the GeoEPR dataset in the whole period under consideration, 1992-2012. Our units of analysis are the ensuing 2929 disjoint areas forming a partition of the territory, in the baseline case.

4.2 Dependent Variable

Log Light The dependent variable of our equations (**??**) and (**??**), 'Log Light', is calculated as the deviation from the country-year mean of 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 light intensity at night, available from the National Oceanic and Atmospheric Administration. The initial sources are the satellite measures of 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 due for instance to the bright half of the lunar cycle, daytime lights in summer months, forest fires and northern and southern lights. An average over all valid days, with measures taken from 8:30 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 nowadays available for the years up to 2013.

Henderson et al. (2012) discuss the interest of these data for economic analyses. First, it is well-known that GDP figures are poorly reported in developing countries. Measures of light intensity for their part cannot be manipulated and thus escape the measurement bias occurring in national accounts. Second, the large fraction of informal economic activity in those countries is not reflected in typical growth statistics. Henderson et al. (2012) demonstrate that night light data are a useful proxy for economic growth and track its short-term fluctuations. Their significantly estimated elasticity of GDP with respect to light intensity is 0.3. Finally, 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 Christopher D. Elvidge and Ghosh (2014) to mitigate measurement errors in light intensities. We apply a mask to suppress gas flares.

4.3 Explanatory Variables of Interest

Ethnic Kin Dominant (EKD) This variable is a count taking integer values between 0 and 3 in our sample. It describes the number of kin groups that are politically active in neighbouring countries. Countries are neighbours here in the sense that their borders touch each other, in the baseline case, or are within a 100km or 1000km radius of each other in alternative specifications. Specifically, we use the definitions of political statuses proposed by Vogt et al. (2015).

The last version of the EPR data describes 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 that are spread over several countries. To construct our variable 'Dominant Transborder Ethnic Kin', we first count, for each ethnic group, the number of their transborder ethnic kin that are coded politically dominant, while restricting the scope to neighbouring countries to exclude the link between Arabs in Morocco and Arabs in Egypt, for instance. Then, for each cell, we sum these numbers of TEK's to obtain our variable.

Dominant Ethnic Group (DOM) This is a binary variable taking the value 1 at t if the ethnic group of the cell is the only one represented in the central government of the country at time t, and 0 otherwise. We adopt the definition of political representation in the sense of the EPR data source (Vogt et al., 2015) i.e. whenever the group has monopoly over central power, or is sole in charge with the other groups enjoying only token representation. Our geographical grid is constructed using the International Conflict Research's GeoEPR dataset (Wucherpfennig et al., 2011), which is a geocoded version of the EPR dataset. It contains the regional bases of the politically relevant ethnic groups across space and time.

This methodology has the advantage of allowing us to identify the ethnicity of the population living in each cell and each year. GeoEPR is build on the Geographic Representation of Ethnic Groups (GREG) dataset (Weidmann et al., 2010) that digitalizes 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 nineteen-sixties, 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 historic 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 like in the Atlas and GREG. This is of particular importance for our analysis, where the hypothesized transmission mechanisms are political.[22]

These data enable us to identify the access of the groups to central state power delineated by executive positions.[23] An ethnic group is coded included in the central state power 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 may be important for distributive politics.[24]

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

[23]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 details.

[24]For constructing of our political variables, all the groups of a cell are pooled together as in Alesina et al.

Figure III represents the ethnic regions for the whole continent. Unshaded regions are where no politically relevant ethnic groups are based. Where two or three groups overlap, we code the variable 'Dominant Ethnic Group' in cell i and year t as one if at least one of the corresponding groups is coded included in central government in year t, and zero otherwise.

Our sample spans over 48 countries. Some countries of the sample are not included in the EPR data: Tunisia, Somalia, Burkina Faso, Equatorial Guinea, Lesotho and Swaziland. However, we still include them in the analysis because a compact region is necessary for the spatial estimation. We checked that the results of the panel estimation stay valid when we exclude these countries. However, we still include these points because we need them for the spatial analysis and because they do not prevent the estimation thanks to presence of fixed-effects.

(2016). This does not matter much since most cell contain no more than one group. In our fixed-effects framework, this approach allows us to identify the impact analogously to a difference-in-difference approach between cells with a unique ethnic group and shared cells.

4.4 Controls

Here is the list of other controls that we use in our estimations:

Population In order construct this variable, we use the Gridded Population of the World that provides the geo-referenced population estimates between 1990 and 2012 by the Socioeconomic Data and Applications Center.[25] The variable used is the logarithm of the total number of people in the spatial unit in that year.

Meteorological Conditions We obtain monthly rainfall and temperature data from the National Oceanic and Atmospheric Administration.[26] 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 mean value over the pixels of the temperature and rainfall rasters for each year to construct our 'Mean Temperature' and 'Total Precipitations' variables. Aidt and Leon (2015) stresses the importance of these meteorological factors in agriculture-based African economies.

Violence For our measures of violence and instability, we use the Georeferenced Event Dataset from the Uppsala Conflict Data Program.[27] This academic project collects and distributes the most accurate information on violent conflicts since 1975. The variable 'Incident Count' is the number of civil violence incidents that have occurred within each area and each year. Our procedure of variable construction circumvents the problem of missing data in macro times series (Elhorst, 2003).

5 Results

5.1 Baseline Estimations

In Table I, we report the estimation results of equation (3), our panel Fixed Effects model linking log light intensity to the dominance by ethnic groups and to the number of politically dominant neighbouring kin, and of equation (5), our Spatial Panel Fixed Effects model linking the same variables. All estimations reported in Table I have robust estimators of standard errors clustered at the unit level. In columns (1) to (4), we use the dataset constructed with a four by four grid and in columns (5) to (8), a two by two grid instead. All spatial models are estimated using a spectrally normalized queen-neighbours weights matrix. Columns 1 and 5 contain Fixed effects estimates. Columns 2 to 4 and 5 to 8 contain SAR estimates, where the transborder ethnic kin relationships are limited to direct neighbour countries (in 2 and 6), countries within a radius of 100 km (in 3 and 7) and countries within a radius of 1000

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

[26]UDel-AirT-Precip data provided by the NOAA/OAR/ESRLPSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/.

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

km (in 4 and 8). All estimations control for spatial unit fixed-effects and country-year fixed effects using deviations from the country-year mean (see equations (3), (**??**) and (5) above). All columns in this table use the whole sample of African countries.

As mentioned before, our main independent variable of interest in this investigation is 'Ethnic Kin Dominant$_{t-1}$' (EKD). This variable receives significantly positive estimated coefficients at 1% with the 2x2 fishnet and at 5% with the 4x4 fishnet, except in column 4 where the 'Ethnic Kin Dominant' loses significance and 'Ethnic Kin in Power' becomes significant.

In our prefered estimation in column (6), the coefficient estimate for our variable of interest is equal to 0.024, slightly lower than the FE estimate of 0.029. 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 given by the factor of 0.3 computed by Henderson et al. (2012). In the next table, we explicitely compute 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.

5.2 Direct and Indirect Effects

In Table II, we report the estimates of the direct and indirect effects of the spatial model expressed in equations (5).[28] In columns 1 to 3, we use the dataset constructed with a four by four grid and in columns 4 to 6, a two by two grid instead. All spatial models are estimated using a spectrally normalized queen-neighbours weights matrix. The total effects are broken down in direct and indirect components (see equations (13) and (14) in Appendix A). All columns contain SAR estimates, where the transborder ethnic kin relationships are limited to direct neighbour countries (in 1 and 4), countries within a radius of 100 km (in 2 and 5) and countries within a radius of 1000 km (in 3 and 6). All estimations control for spatial unit fixed-effects and country-year fixed effects using deviations from the country-year mean (see equations (3), (**??**) and (5) above). All columns use the whole sample of African countries.

The results presented in this table perfectly validate our previous conclusions. The direct component are in the line with the estimated coefficient of the previous table and the significance levels are alike.

5.3 Robustness

In what follows, we discuss various robustness checks. Those estimations are reported in Tables III and IV and we comment mainly on our variable of interest, Ethnic Kin Dominant.

[28] Robust standard errors clustered at the spatial unit level in parenthesis.

In Table III, columns 1 to 4, we use the dataset constructed with a four by four grid and in columns 5 to 8, a two by two grid instead. The equation of the Spatial Error Model is

$$\begin{aligned} Y_t &= EKD_{t-1} \cdot \gamma + X_{t-1} \cdot \beta + \alpha_0 + u_t \\ u_t &= \epsilon_t + \lambda W \cdot u_t \qquad \text{for } t = 1, \ldots, T \end{aligned} \tag{7}$$

The equation of the Spatial Durbin Model is

$$Y_t = \rho W \cdot Y_t + EKD_{t-1} \cdot \gamma + X_{t-1} \cdot \beta + \nu W \cdot Z_{t-1} \beta_2 + \epsilon_t \qquad \text{for } t = 1, \ldots, T \tag{8}$$

The term $\nu W \cdot \tilde{Z}_{t-1} \beta_2$ in the Spatial Durbin Model represents spatially lagged explanatory variables i.e. $\tilde{Z}_{t-1}$ is a subset of $\tilde{X}_{t-1}$ The SEM and SDM (columns 1,2,5 and 6) models are estimated using a spectrally normalized queen-neighbours weights matrix. In columns 3 and 6, we estimate a SAR with a rook-contiguity weights matrix instead. In columns 4 and 8 we estimate a SAR with a row-normalized weights matrix instead. The transborder ethnic kin relationships are limited to direct neighbour countries. All estimations control for spatial unit fixed-effects and country-year fixed effects using deviations from the country-year mean (see equations (3), (**??**) and (5) in the text). All columns use the whole sample of African countries. The results presented in this Table III corroborate our previous results.

In Table IV, we use the restricted sample of Sub-Saharan African countries in all columns. Columns 1 and 2, we use the dataset constructed with a four by four grid and in columns 3 and 4, a two by two grid instead. All spatial models are estimated using a spectrally normalized queen-neighbours weights matrix. Columns 1 and 3 contain Fixed effects estimates. Columns 2 and 4 contain SAR estimates, where the transborder ethnic kin relationships are limited to countries within a radius of 1000 km. All estimations control for spatial unit fixed-effects and country-year fixed effects using deviations from the country-year mean (see equations (3), (**??**) and (5) above).

In Table IV, where we restrict the sample to Sub-Saharan Africa, the transborder political variable that becomes significant is 'Ethnic Kin in Power'. This variables receives a positive and significant coefficient in specification (3) and (4).

6 Conclusion

In the context of the African continent, where ethnicity is politically salient and where historic ethnic groups have been divided across borders designed by colonial powers, we study comparative development proxied by luminosity measured from space. In particular, we examine whether the political dominance by its transborder related kin positively affects the redistribution of resources towards an ethnic group's base region. This study follows the literature on ethnic favouritism (Hodler et al., 2015; Hodler and Raschky, 2014) and explores the question of its transnational effect, which is a contribution of the paper as the question had not been considered before in this context.

Using various panel estimation strategies, with spatial externalities, we find some evidence that this variable has positive effect, statistically significant and possibly substantial, with gains in income around 0.6% depending on the specification. To investigate whether the found interactions between ethnicities are not confounded with other possible spatial externalities, like trade or the diffusion of technologies and ideas, we estimate Spatial Autoregressive (SAR) models with fixed-effects.

Our new dataset constructed using the geo-localized version of the Ethnic Power Relations database allows us to estimate spatial panel data models. It contains information for 2929 geographical cells for the period 1992 to 2012. The most recent version of the EPR data describes the network of international ethnic bonds. We use the total number of politically dominant ethnic kin groups in neighbouring countries as an explanatory variable in addition to the power status.

We interpret this effect to be the consequence of changes in balance of power between ethnicities of the situation abroad with the importance of transnational threats documented in Michalopoulos and Papaioannou (2016) in mind, along other possible ethnic channels.

We investigate the robustness of this result to a series of controls and specification checks, including meteorological factors, restrictions to sub-samples and the inclusion of country-times-year fixed effects.

In a possible extension of this work we could consider the question whether this phenomenon is an African particularity or if the data suggests that it is also active in other parts of the world. Some weaknesses of this work stem from the quality of the data used to identify the location and political participation of ethnic groups. Even though we use the best data available, this is still prone to measurement errors and debate.

Eventually, the policy implications for international organizations and foreign intervention are straightforward and only underline an already accepted concept. It is important to have a global point of view and to anticipate the repercussions in neighbouring countries before implementing foreign intervention or aid policies.

7 Appendix

A : Direct and Indirect Effect in SAR and SAC Models.

For further reference, we make the structure of the matrix of regressors explicit in (9). In addition, in this appendix, we change the notation and we collapse the variables of interest in this matrix.

$$X_{n,t-1} = \begin{pmatrix} x_{1\ t-1,1} & \dots & x_{1\ t-1,k} \\ \vdots & \ddots & \\ x_{N\ t-1,1} & & x_{N\ t-1,k} \end{pmatrix} \tag{9}$$

The r^{th} column contains the values of the r^{th} independent variable for all N units at time $t-1$. The total effect of each covariate can be decomposed into a direct and an indirect component. The direct effect is the partial effect of a change in $x_{it-1,r}$ on y_{it}, in general including feedback loops.

The indirect effect is the partial effect of a change in $x_{it-1,r}$ on y_{jt} with $j \neq i$, also comprising interactive channels.

We adapt the calculus in LeSage and Pace (2009) with our panel specification with unit and year fixed-effects. This is straightforward because the fixed-effects do not play a role in the partial derivatives. The fact that our explanatory variable is lagged leaves the interpretation unchanged because the spatial effects are conducted through Y_{nt} and the disturbance term U_{nt} that are both present here as well.

Solving (5) for Y_{nt} yields equation (10).

$$Y_{nt} = (I_n - \rho W_n)^{-1}(X_{nt}\gamma + \mathbf{c}_{n0} + \alpha_{t0}\iota_n + U_{nt}) \tag{10}$$

Using formula (4), equation (10) can be rewritten as

$$\begin{aligned} Y_{nt} &= \sum_{r=1}^{k} \begin{pmatrix} S_r W_{11} & S_r W_{12} & \dots & S_r W_{1n} \\ S_r W_{21} & S_r W_{22} & & \\ \vdots & \vdots & \ddots & \\ S_r W_{n1} & S_r W_{n2} & & S_r W_{nn} \end{pmatrix} X_{.r,nt-1} \\ &\quad + (I_n - \rho W_n)^{-1}(\mathbf{c}_{n0} + \alpha_{t0}\iota_n + U_{nt}), \end{aligned} \tag{11}$$

where $X_{.r,nt-1}$ is the r^{th} column of $X_{n,t-1}$. We can see in this expression that the derivative of y_{it} with respect to $x_{jt-1,r}$ with $j \neq i$ may differ from zero. Moreover, the derivative of y_{it} with respect to $x_{it-1,r}$ is in general distinct from γ_r because of the feedback loops to the neighbours and back to region i. The following equations (13) to (14) elicit the explicit expressions of the decomposition of the total effect into direct and indirect effects, as in LeSage and Pace (2009). These formulae are valid for both the SAR and SAC models because the disturbances do not affect the partial derivatives.

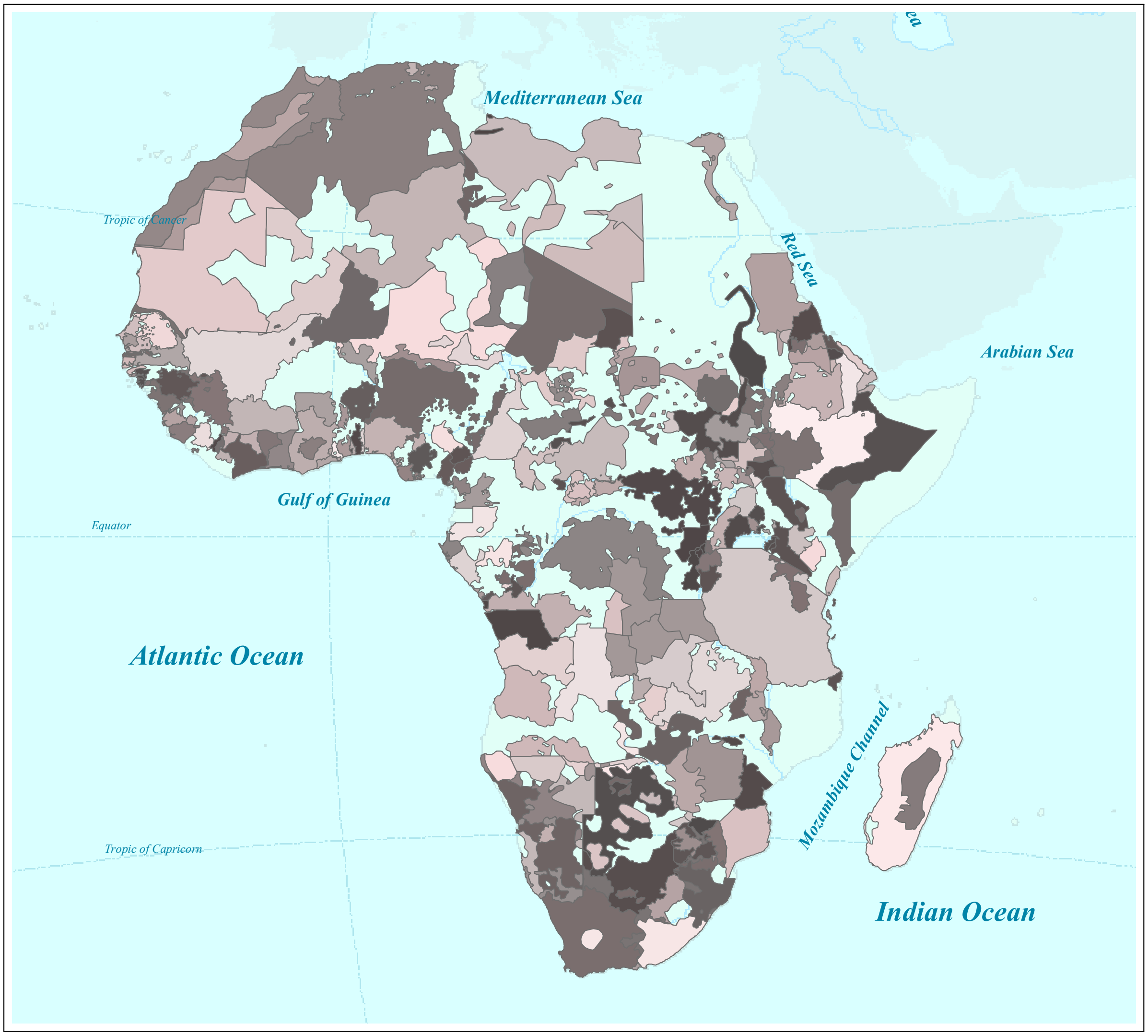
$$\overline{M}_{(r)total} = \frac{\iota_n' S_r(W_n)\iota_n}{n} \tag{12}$$

$$\overline{M}_{(r)direct} = \frac{tr(S_r(W_n))}{n} \tag{13}$$

$$\overline{M}_{(r)indirect} = \overline{M}_{(r)total} - \overline{M}_{(r)direct}, \tag{14}$$

where ι_n denotes a vector of ones of dimension N. The direct effect is equal to the mean over all spatial units of the diagonal elements. The indirect effect is the mean over all spatial units of the off-diagonal elements. The total impact here is $\frac{\gamma_r}{1-\rho}$.

Figure III: Map of Ethnic Groups Homelands in Africa



note: Source is Vogt et al. (2015). In total, there are 238 politically represented ethnic groups. The EPR information exist for all African countries except Tunisia, Somalia, Burkina Faso, Equatorial Guinea, Lesotho and Swaziland.

Table I Baseline Results

	Dependent variable is Log $Light_{i,t}$							
	2x2 Fishnet						4x4 Fishnet	
	(1) FE	(2) SAR	(3) SAR	(4) FE	(5) SAR	(6) SAR	(7) FE	(8) SAR
Ethnic Kin $Dominant_{t-1}$	0.028	0.026	0.021	0.029	0.027	0.027	0.025	0.023
	(0.009)***	(0.009)***	(0.008)***	(0.009)***	(0.008)***	(0.007)***	(0.011)**	(0.011)**
Ethnic Kin in $Power_{t-1}$	0.009	0.006	0.004	0.000	-0.001	-0.003	0.006	0.004
	(0.002)***	(0.002)***	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
$Dominant_{t-1}$	0.011	0.006	0.004	-0.006	-0.007	-0.008	-0.033	-0.034
	(0.009)	(0.009)	(0.010)	(0.013)	(0.013)	(0.013)	(0.020)*	(0.019)*
Ethnic Group in $Power_{t-1}$	-0.006	-0.006	-0.006	-0.009	-0.008	-0.009	-0.007	-0.006
	(0.005)	(0.005)	(0.005)	(0.005)*	(0.005)*	(0.005)*	(0.007)	(0.007)
$Log(Pop)_{t-1}$	0.108	0.094	0.096	-0.009	-0.010	-0.010	-0.017	-0.017
	(0.016)***	(0.015)***	(0.015)***	(0.018)	(0.017)	(0.018)	(0.027)	(0.027)
$Violence_{t-1}$	0.004	0.002	0.002	0.006	0.005	0.004	0.009	0.008
	(0.003)	(0.003)	(0.003)	(0.003)**	(0.002)*	(0.002)*	(0.003)***	(0.003)***
$Rainfall_{t-1}$	0.374	0.245	0.245	0.085	0.024	0.024	0.007	-0.025
	(0.054)***	(0.054)***	(0.054)***	(0.046)*	(0.049)	(0.049)	(0.026)	(0.030)
$Temperature_{t-1}$	-0.004	-0.002	-0.003	-0.003	-0.002	-0.002	-0.003	-0.002
	(0.001)***	(0.001)**	(0.001)**	(0.001)***	(0.001)**	(0.001)**	(0.002)*	(0.002)
Spatial ρ		0.15	0.15		0.09	0.09		0.06
		(0.000)***	(0.000)***		(0.000)***	(0.000)***		(0.000)***
Year FE	Y	Y	Y	N	N	N	N	N
Country*Year FE	N	N	N	Y	Y	Y	Y	Y
Number of Units	2929	2929	2929	2929	2929	2929	1762	1762
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 in parenthesis. In columns 1 to 6, we use the dataset constructed with a two by two grid and in columns 7 and 8, a four by four grid instead. All spatial models are estimated using a spectrally normalized queen-neighbours weights matrix. Columns 1, 4 and 7 contain Fixed effects estimates. Columns 2, 3, 5, 6, 8 contain SAR estimates, where the transborder ethnic kin relationships are limited to direct neighbour countries (in 2, 5 and 8) or countries within a radius of 1000 km (in 3 and 6). All estimations control for spatial unit fixed-effects. We add year fixed effects in columns 1 to 3 and country-year fixed effects in columns 4 to 8. All columns use the whole sample of African countries.

TABLE II DIRECT AND INDIRECT EFFECTS

		Dependent variable is Log $\text{Light}_{i,t}$		
		2x2 Fishnet		
		(1)	(2)	(3)
		SAR	SAR	SAR
Direct	Ethnic Kin Dominant_{t-1}	0.026	0.026	0.021
		(0.008)***	(0.008)***	(0.007)***
	Ethnic Kin in Power_{t-1}	0.006	0.006	0.003
		(0.002)***	(0.002)***	(0.003)
Indirect	Ethnic Kin Dominant_{t-1}	0.003	0.003	0.003
		(0.001)***	(0.001)***	(0.001)***
	Ethnic Kin in Power_{t-1}	0.001	0.001	0.000
		(0.000)**	(0.000)***	(0.000)
Total	Ethnic Kin Dominant_{t-1}	0.029	0.029	0.024
		(0.010)***	(0.009)***	(0.008)***
	Ethnic Kin in Power_{t-1}	0.007	0.007	0.004
		(0.003)***	(0.002)***	(0.003)
Main	Ethnic Kin Dominant_{t-1}	0.026	0.026	0.021
		(0.009)***	(0.009)***	(0.008)***
	Ethnic Kin in Power_{t-1}	0.006	0.006	0.004
		(0.002)***	(0.002)***	(0.003)
	Spatial ρ	0.15	0.15	0.15
		(0.000)***	(0.000)***	(0.000)***
	Number of Units	2929	2929	2929
	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 in parenthesis. In all columns, we use the dataset constructed with a two by two grid. All spatial models are estimated using a spectrally normalized queen-neighbours weights matrix. The total effects are broken down in direct and indirect components (see equation (13) and (14) in Appendix A). All columns contain SAR estimates, where the transborder ethnic kin relationships are limited to direct neighbour countries (in 1), countries within a radius of 100 km (in 2) and countries within a radius of 1000 km (in 3). All estimations control for spatial unit fixed-effects and year fixed effects. All columns use the whole sample of African countries.

Table III Robustness Checks

	Dependent variable is Log Light$_{i,t}$ 4x4 Fishnet			
	(1) SEM	(2) SDM	(3) SAR	(4) SAR
Ethnic Kin Dominant$_{t-1}$	0.026	0.027	0.028	0.018
	(0.010)***	(0.008)***	(0.008)***	(0.007)**
Ethnic Kin in Power$_{t-1}$	0.007	-0.000	-0.001	0.001
	(0.003)**	(0.003)	(0.003)	(0.003)
Dominant$_{t-1}$	0.003	-0.012	-0.007	0.007
	(0.013)	(0.014)	(0.013)	(0.011)
Ethnic Group in Power$_{t-1}$	-0.003	0.003	-0.009	-0.005
	(0.005)	(0.006)	(0.005)*	(0.004)
Log(Pop)$_{t-1}$	0.008	-0.009	-0.010	-0.003
	(0.019)	(0.017)	(0.018)	(0.015)
Violence$_{t-1}$	0.002	0.004	0.005	0.003
	(0.002)	(0.002)*	(0.002)**	(0.002)*
Rainfall$_{t-1}$	0.001	0.028	0.035	0.074
	(0.045)	(0.048)	(0.049)	(0.036)**
Temperature$_{t-1}$	-0.003	-0.002	-0.003	-0.001
	(0.002)*	(0.001)*	(0.001)***	(0.001)
Spatial ρ	0.70	0.09	0.09	0.60
	(0.003)***	(0.000)***	(0.000)***	(0.002)***
Weights Matrix	W	W	W_{row}	W_{row}
Year FE	N	N	N	
Country*Year FE	Y	Y	Y	
Number of Units	2929	2929	2929	2929
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 in parenthesis. In all columns, we use the dataset constructed with a two by two grid . The SEM and SDM (columns 1,2) models are estimated using a spectrally normalized queen-neighbours weights matrix. In column 3 we estimate a SAR with a rook-contiguity weights matrix instead. In columns 4 we estimate a SAR with a row-normalized weights matrix instead. The transborder ethnic kin relationships are limited to direct neighbour countries. All estimations control for spatial unit fixed-effects and country-year fixed effects. All columns use the whole sample of African countries.

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.

Aidt, T. S. and Leon, G. (2015). The Democratic Window of Opportunity: Evidence from Riots in Sub-Saharan Africa. *Journal of Conflict Resolution*.

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

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. NATO Asi Series. Series E, Applied Sciences. Springer.

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.

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.

Cammack, D. (2007). The Logic of African Neopatrimonialism: What Role for Donors? *Development Policy Review*, 25(5):599–614.

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.

Christopher D. Elvidge, Feng-Chi Hsu, K. E. B. and Ghosh, T. (2014). *National Trends in Satellite-Observed Lighting: 1992–2012*, pages 97 –120. Global Urban Monitoring and Assessment through Earth Observation.

Clapham, C. (1985). *Third World Politics: An Introduction*. Routledge.

Dreher, A., Fuchs, A., Hodler, R., Parks, B. C., Raschky, P. A., and Tierney, M. J. (2015). Aid on Demand: African Leaders and the Geography of China's Foreign Assistance. CESifo Working Paper Series 5439, CESifo Group Munich.

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.

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. (2003). Specification and Estimation of Spatial Panel Data Models. *International Regional Science Review*, 26(3):244–268.

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

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

Franck, R. and Rainer, I. (2012). Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education, and Health in Sub-Saharan Africa. *American Political Science Review*, 106:294–325.

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.

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.

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*.

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.

Hodler, R., Raschky, P. A., Valsecchi, M., and De Luca, G. (2015). Ethnic Favoritism: An Axiom of Politics? CESifo Working Paper Series 5209, CESifo Group Munich.

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.

Kelsall, T. (2011). Rethinking the Relationship between Neopatrimonialism and Economic Development in Africa. *IDS Bulletin*, 42(2):76–87.

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.

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

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

Magee, C. S. P. and Doces, J. A. (2015). Reconsidering Regime Type and Growth: Lies, Dictatorships, and Statistics. *International Studies Quarterly*, 59(2):223–237.

Meyersson, E. (2015). 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.

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

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

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*.

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.