I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12818

The Employment Effects of Ethnic Politics

Francesco Amodio
Giorgio Chiovelli
Sebastian Hohmann

DECEMBER 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12818

The Employment Effects of Ethnic Politics

Francesco Amodio
McGill University, CIREQ and IZA

Giorgio Chiovelli
Universidad de Montevideo

Sebastian Hohmann
Stockholm School of Economics SITE

DECEMBER 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Employment Effects of Ethnic Politics*

This paper studies the labor market consequences of ethnic politics in African democracies. We combine geo-referenced data from 15 countries, 32 parliamentary elections, 62 political parties, 243 ethnic groups, 2,200 electoral constituencies, and 400,000 individuals. We implement a regression discontinuity design that compares individuals from ethnicities connected to parties at the margin of electing a local representative in the national parliament. We find that having a local ethnic politician in parliament increases the likelihood of being employed by 2-3 percentage points. We hypothesize that this effect originates from strategic interactions between ethnic politicians and traditional leaders, the latter retaining the power to allocate land and agricultural jobs in exchange for votes. The available evidence supports this hypothesis. First, the employment effect is concentrated in the historical homelands of ethnicities with strong pre-colonial institutions. Second, individuals from connected ethnicities are more likely to be employed in agriculture, and in those countries where customary land tenure is officially recognized by national legislation. Third, they are also more likely to identify traditional leaders as partisan, and as being mainly responsible for the allocation of land. Evidence shows that ethnic politics shapes the distribution of productive resources across sectors and ethnic groups.

JEL Classification: J15, J70, O10, P26, Q15

Keywords: ethnic politics, employment, democracy, traditional leaders, Africa

Corresponding author:
Francesco Amodio
Department of Economics
McGill University
855 Sherbrooke Street West
Montreal, Quebec H3A 2T7
Canada

E-mail: francesco.amodio@mcgill.ca

* We thank Dylan Munson for superb research assistance. We are thankful to the following people for helpful comments and discussion: Francisco Alvarez-Cuadrado, Kate Baldwin, Rui Castro, Matteo Cervellati, Elena Esposito, Paola Giuliano, Franque Grimard, Sonia Laszlo, Benjamin Marx, Stelios Michalopoulos, Hannes Mueller, Elias Papaioannou, Markus Poschke, Paolo Surico, and all seminar participants at Bocconi, Brock, Carlos III, EIEF, IECON, London Business School, McGill, PUC Chile, Stanford SITE, Universidad de Montevideo, University of Ottawa, CSAE 2017, Montreal Applied Micro Day, Montreal Workshop on Political Economy of Development, NEUDC 2016. Errors remain our own. A previous version of this paper was previously circulated as "Ethnic Favoritism in Democracy: The Political Economy of Land and Labor in South Africa." We gratefully acknowledge financial support from McGill University and FRQSC Quebec - Project 2019-NP-253220.

1 Introduction

More than half of Africa's population lives under the influence of traditional ethnic authorities (Michalopoulos and Papaioannou 2015). Recent studies show that contemporary economic differences across ethnic groups are related to the strength of ethnic institutions (Gennaioli and Rainer 2007; Michalopoulos and Papaioannou 2013), and that ethnic inequality is a strong correlate of underdevelopment in Africa (Alesina, Michalopoulos, and Papaioannou 2015). The role of ethnic-specific institutions in the continent is reinforced by the weakness of state institutions. Severe lack of infrastructural power (Mann 1984) undermines the capacity of the state to enforce policy throughout its territory and have control of rural areas (Herbst 2000; Michalopoulos and Papaioannou 2014), where traditional leaders continue to assert their authority over land allocation, tenure and redistribution (Economic Commission Africa 2007). In African democracies, state-level formal institutions and traditional ethnic institutions strategically interact, forming the basis of ethnic politics. What are the economic consequences of these strategic interactions? Does ethnic politics affect the allocation of resources? Providing the answer to these questions is crucial in the study of African development.

This paper studies the labor market effects of ethnic politics and its impact on the allocation of resources, namely labor, between and within ethnic groups and sectors. We build a unique, comprehensive dataset that combines geo-referenced data on parliamentary elections at the constituency level with information on individuals across Sub-Saharan Africa from the Demographic and Health Survey (DHS). Our analysis leverages data across 15 countries, 32 parliamentary elections, 62 political parties, 243 ethnic groups, more than 2,200 electoral constituencies and 400,000 individuals. We identify links between ethnic groups and political parties in each country, exploiting the information on political affiliation and vote intention available in the Afrobarometer.

We use these data to test whether the labor market outcome of individuals belonging to different ethnic groups changes with electoral results. More precisely, we compare individuals from ethnicities linked to political parties which do or do not gain a local representative in the national assembly. We focus on those groups and parties whose vote shares at the constituency level put them at the margin of gaining a local representative in parliament. That is, we adopt a *regression discontinuity design* that compares individuals from ethnicities linked to political parties that gain a local representative or not by a narrow vote margin. We find evidence that electoral results affect labor market outcomes differentially according to ethnicity. Individuals connected to ethnic politicians in parliament are 2 to 3 percentage points more likely to be employed, a 3.5 to 5.4% increase over the mean left of the election-winning threshold. This positive effect is stronger when the party the elected politician belongs to supports the central government.

We then turn to investigating the mechanism that is responsible for these employment effects.

We focus on the interactions between formally appointed political leaders and traditional ethnic power structures. We argue that these two levels of governance interact in a strategic way. Political parties recruit traditional ethnic leaders as political brokers to mobilize the individuals under their authority for political support and voting (Baldwin 2014, 2015; de Kadt and Larreguy 2018). In exchange, traditional leaders from connected ethnicities retain their *de facto* powers, first and foremost the one to allocate agricultural land. It follows that, when gaining a local representative in parliament, co-ethnics enjoy a competitive advantage in the agricultural land and labor market.[1] The positive employment effect of being linked to the winning party originates from these strategic interactions between politicians and traditional leaders resulting in preferential access to land and agricultural jobs for connected voters.

A number of quantitative and qualitative empirical results support this hypothesis. First, we find that the employment effect is concentrated in those constituencies located in the historical homelands of ethnicities with strong, more centralized pre-colonial institutions, which map into higher mobilization capacity on behalf of traditional leaders. Second, we find that the employment effect is concentrated in the agricultural sector while finding no evidence of increased employment in the public sector or in manufacturing. This is consistent, on one hand, with preferential land access being the relevant channel and, on the other hand, with the lack of state capacity and infrastructure needed to sustain more direct and conventional forms of patronage (Colonnelli, Prem, and Teso 2018). Third, the effect is stronger in those countries where customary land tenure systems are officially recognized. By formalizing the chiefs' customary power at constitutional level, chiefs interact with elected politician at the same institutional level, making the recognition of the latter fundamental to retain their role in the allocation of land. Finally, evidence from Afrobarometer survey data shows that electoral outcomes affect individuals' assessment of the role of traditional leaders. Individuals from ethnicities connected to parties that gain a local representative by a narrow margin are systematically more likely to identify the traditional leader as partisan, and mainly responsible for allocating land.

Our paper builds upon and contributes to several strands of the literature. First, it speaks to the literature on ethnic voting (Posner 2005). Eifert, Miguel, and Posner (2010) find evidence that political competition reinforces ethnic identification. Acemoglu, Reed, and Robinson (2014) and Baldwin (2014) show, for Sierra Leone and Zambia respectively, how local chiefs trade their ability to mobilize local communities in exchange for public good provision from the national government. We contribute to this literature by focusing on the allocative consequences of ethnic politics in African democracies and its impact on the labor market. We provide causal evidence on cross-national scale using local level electoral results from 15 different countries. We also explore further the interaction between political parties and ethnic chiefs in African

[1] Evidence from developed and democratic countries shows that firms connected to politicians enjoy a certain degree of competitive advantage (Goldman, Rocholl, and So 2013; Cingano and Pinotti 2013). The same is true for individuals on the labor market (Gagliarducci and Manacorda 2017).

democracies. In doing this, we speak to the literature on electoral clientelism (see Stokes, Dunning, Nazareno, and Brusco 2013; Robinson and Verdier 2013).[2] Clientelism emerges in equilibrium if both politician and voters can commit to their promises.[3] Reciprocity (Finan and Schechter 2012) and social (ethnic) networks (Robinson and Verdier 2013; Young and Turner 1985) facilitate clientelism. Ethnic chiefs can solve the commitment problem by holding both politicians and voters accountable (Mamdani 1996). de Kadt and Larreguy (2018) document the importance and effectiveness of traditional leaders in the former Bantustans of South Africa in increasing the electoral support of the ANC. While our analysis builds on the role of traditional leaders as "vote brokers," we highlight the potential costs linked to the mobilization of ethnicity. In this sense, our findings point to the presence of allocative distortions that come with decentralization in the presence of strong ethnic institutions.

Our results also speak to the literature on ethnic favoritism. This can be referred to as a situation in which a disproportionate amount of benefits accrue to individuals who share the same ethnicity of the individual or groups who hold political power.[4] Several studies have shown evidence of ethnic favoritism in the provision of public goods such as roads, schools, and hospitals (Franck and Rainer 2012; Kramon and Posner 2014). Burgess, Jedwab, Miguel, Morjaria, and Padró i Miquel (2015) show that expenditure on roads and the length of paved roads built are twice and five times higher respectively in those Kenyan districts that share the ethnicity of the president as compared to other districts. In line with the theoretical predictions of Lizzeri and Persico (2004), they do not find evidence of ethnic pork-barrel when democratic transitions occur. Using data from 130 countries and nighttime lights as a proxy for economic wellbeing, Mueller and Tapsoba (2016) confirm this finding while also qualifying the intensive margin of the relationship between access to power and ethnic favoritism. In contrast, De Luca, Hodler, Raschky, and Valsecchi (2018) find that nighttime light becomes differentially more intense in the political leaders' ethnic homelands, and even more so around election time. Finally, Dickens (2018) exploits within-group variation across split ethnic groups in Africa to show that ethnic favoritism occurs more along territorial lines than at the individual level.

Our contribution to this body of research is twofold. First, we show that, in African democracies, favoritism of co-ethnics can still take place through preferential access to land and agricultural jobs. In doing this, we depart from the previous literature which focuses on public good provision. Second, we provide suggestive evidence that these allocative effects originate from

[2] See also Larreguy, Marshall, and Querubin (2016) and Fergusson, Larreguy, and Riaño (2018) studying the emergence of clientelism and its implications in Mexico.

[3] Stokes (2007) eloquently summarizes the two-sided commitment problem leading to the emergence of clientelism: "*A voter who receives a bag of food with the understanding (implicit or explicit) that she will return the favor with a vote can easily renege on the deal on election day, especially when she is protected by the secret ballot.*" At the same time, "*a party that before an election promises patronage in exchange for votes may well forget its promise afterwards.*"

[4] Alesina, Michalopoulos, and Papaioannou (2015) stress that ethnic inequality may lead to political inequality along ethnic lines creating discriminatory policies of one group towards another ethnic enclave.

strategic interactions between politicians and traditional ethnic leaders.

Finally, our paper is linked to the empirical literature on African state capacity. The inability of projecting power from the centre to the periphery and rural areas is key in the challenge to improve state capacity in Africa (Herbst 2000).[5] Michalopoulos and Papaioannou (2014) find evidence that state capacity is negatively correlated with distance from the capital. At the same time, there is a growing appreciation of the role of ethnic chiefs for the functioning of local economies in rural Africa (Michalopoulos and Papaioannou 2015; Acemoglu, Reed, and Robinson 2014). Our contribution to these literatures is to shed light on the relationship between national state institutions and local ethnic chiefs. As democratization takes place in several African countries and national level institutions improve, would one expect the role of local chiefs to weaken or strengthen? The prevailing view is that the strength of these local ethnic-specific actors is a legacy of past institutional weaknesses of African countries. In this respect, improvements in institutions should make local chiefs irrelevant. Yet, our findings show that the role of local chiefs and their authority in resource allocation can be enhanced by multi-party politics (Eggen 2011).

The remainder of the paper is organized as follows. Section 2 provides information on data sources and how we build our final dataset. Section 3 describes our main empirical and estimation strategy while Section 4 presents the first set of results. Section 5 focuses on the strategic interactions between politicians and ethnic leaders as a mechanism, and provides the corresponding evidence. Section 6 concludes.

2 Data

We assembled our final dataset by combining and harmonizing several sources of information. This section describes the main features of these data. The Supplementary Appendix to the paper provides further details on the sample definition and the specific variables we use in the empirical analysis.

Elections The first piece of information pertains to electoral results. We started our data collection by accessing the Constituency-Level Election Archive (CLEA) (Kollman, Allen, Caramani, Backer, and Lublin 2018). This is a repository of detailed election results at the constituency level for lower chamber and upper chamber legislative elections from around the world. We first restricted the sample to all Sub-Saharan African countries. In our analysis, we exploit variation in the political affiliation of local representatives in the national parliament. We thus consider only those electoral systems where electoral results at the constituency level map directly into representation in the national assembly. For instance, we consider single-

[5] A related recent literature investigates the challenges of nation building in ethnically diverse countries (Depetris-Chauvin, Durante, and Campante 2018; Bazzi, Gaduh, Rothenberg, and Wong 2019).

member plurality district voting, where one representative is elected in each constituency, and the candidate achieving the relative majority of votes is elected. We also consider proportional systems that allocate a given number of seats in each constituency based on vote shares within the constituency. By contrast, we exclude pure proportional systems where the nationwide allocation of seats to parties is determined by vote shares at the country level, and constituencies only serve logistical and vote counting purposes. We also further restrict the sample to those countries and years for which, as we explain later, we can retrieve information from other sources on the ethnic connotation of parties and individual employment following the election year.

To increase coverage, we collected, processed, and cross-validated information from additional sources. We used data from Elections in Africa (Nohlen, Bernard, and Krennerich 1999) to complement information on Ghanaian and Liberian parliamentary elections. We relied on Carr (2017) to retrieve information for the following countries: Burkina Faso, Kenya, Mali, and Nigeria. We obtained additional electoral constituency-level data from National Electoral Commissions (Burkina Faso) and the European Commission (Kenya).

We then collected information on each party's support for the central government in the aftermath of the elections. We retrieve this data from the Inter-Parliamentary Union (2018), the world organization of parliaments. Whenever missing, we derived this information from a handful of other sources including the Freedom House and Wikipedia.

A fundamental innovation of our empirical analysis is the spatial mapping of individual-level observations into electoral constituencies, which allows us to investigate the relationship between electoral results and individual outcomes. The paucity of digitized maps of electoral constituencies for the countries in our analysis makes this a daunting task. We pursued two strategies to overcome this data limitation. First, we collected and harmonized shapefiles of sub-national administrative divisions to match them with the electoral boundaries. To check for potential inconsistency we cross-checked those maps with both encyclopaedic sources and archival maps from both the British Library and the US Library of Congress (e.g. Ghana and Kenya). Second, when no shapefile existed, we retrieved and digitized ourselves the relevant electoral maps from several sources including Population Census publications (e.g. Uganda) and Electoral Commission's reports (Liberia, Malawi, and Sierra Leone).

The first column of Table A.1 in Appendix A shows the set of countries and election years for which geo-referenced information on election constituencies is available and are thus included in our final election dataset. The second column shows the total number of constituency-level observations available for each country across the different election years. Our final election dataset contains information on election results for 4,721 constituency-level observations across 32 elections in 15 Sub-Saharan African countries. For each constituency and election year, the data provide information on each political party running in the constituency and the

corresponding number of votes. We can thus calculate the vote share of each party, and, knowing the electoral rule, whether any candidate from that party and constituency was ultimately elected to the national assembly.

Ethnic Parties A key component of our analysis is the identification of links between ethnic groups and political parties in each country. We do so using information on the political affiliation of individuals belonging to different ethnic groups. We retrieve this information from Afrobarometer, a research network that conducts public attitude surveys on democracy, governance, economic conditions, and related issues in Africa (Afrobarometer 2015). In our analysis, we use all waves (1-6) available at the time of writing.

Identifying links between ethnic groups and political parties involves several steps. First, we harmonize ethnic group names in Afrobarometer with those in the Demographic Health Survey, our source of information for individual outcomes. Second, we harmonize party names in Afrobarometer with those in our election dataset. Third, we attach a party "label" to each ethnic group based on a quantitative criterion. Specifically, let N and N_p be respectively the total number of individuals in a given country, and the number of voters for party p in that country. Let instead N_e and N_e^p denote respectively the total number of individuals from ethnic group e and the number of voters for party p among them. Our baseline measure aims at capturing over-representation of particular ethnic groups among the voters of a given party. That is, we compute for each ethnic group and country a measure

$$\hat{p}^e = \arg\max_p \left[\frac{N_e^p}{N_e} - \frac{N_p}{N} \right] \tag{1}$$

and we assign party label "p" to ethnic group "e" if party p is (the most) over-represented, relative to its overall vote share in the country, among voters from ethnic group e. Section B.1 of the Supplementary Appendix provides further details on the labelling algorithm. To further validate the algorithm, we use the qualitative information contained from a variety of external sources from which we derived qualitative information.

Employment and Other Individual Characteristics We gather individual-level information on labor market outcomes and sociodemographic characteristics from the Demographic and Health Survey (DHS) (ICF 2017). These are nationally representative household-level surveys that are carried out in developing countries around the world. We considered those Sub-Saharan African countries and waves for which geocoded data are available, so that they can be matched according to geography and time to electoral constituencies and election results from the most recent elections. The geographic information available in DHS corresponds to the geographical coordinates of each DHS cluster (group of villages or urban neighborhoods) in the sample.

DHS data provide information on whether the surveyed individual within the household is

currently working. It also provides harmonized information on occupation from which we can derive employment across sectors. We map occupations into sectors following McMillan and Harttgen (2014), who use DHS employment data to study structural change in Africa. We also use the most detailed information on occupation to identify public sector workers, the vast majority of them being in the education and health sector.

Importantly, the data also report the ethnic group each individual belongs to. By harmonizing the ethnic classification in DHS with the one in Afrobarometer, we can match individuals in the DHS sample to political parties based on the linking procedure described in the previous section. Throughout the empirical analysis, we also use information on other individual socioeconomic characteristics available in the DHS such as gender and education.

Ethnic Institutions We retrieve information on pre-colonial ethnic institutions using the "Jurisdictional Hierarchy beyond the Local Community" index (Gennaioli and Rainer 2007; Michalopoulos and Papaioannou 2013). This variable, which is available for 534 African ethnic groups from the Ethnographic Atlas (Murdock 1967), identifies the jurisdictional structure above the local community level for each ethnicity. In other words, this index measures the degree of political centralization of pre-colonial societies. The degree of political centralization ranges from zero to four. A zero scores indicates stateless societies lacking any form of political authority beyond the local; one refers to petty chiefdoms; a score of two identifies paramount chiefdoms; and three and four are the scores assigned to ethnic groups characterized by pre-modern states as political organization. We attach this information to each electoral constituency in our dataset by matching geographically the constituency centroid to historical ethnic homelands in the Murdoch map (Michalopoulos and Papaioannou 2013, 2014). We therefore assign to each constituency the level of pre-colonial centralization of the ethnic group that is historically dominant in the area.

The richness of Afrobarometer allows us to also recover information on the role of traditional leader within countries, and establish a relationship with electoral outcomes. In order to do so, we use a confidential geo-referenced version of Afrobarometer. The questionnaire administered in 2008 – the fourth wave of the survey – includes questions on traditional leaders such as whether they should be partisan, whether they are responsible for allocating land, maintaining law and order, or governing the community (Michalopoulos and Papaioannou 2015). We use the answers to these questions in the exploration of the mechanisms behind the baseline results. In doing this, we are forced to restrict our sample to a subset of 8 out of the 15 countries that comprise our main dataset.[6]

[6]We can retrieve information on the role of traditional leaders from the 2008 Afrobarometer for: Burkina Faso, Ghana, Kenya, Liberia, Malawi, Nigeria, Senegal, and Zambia.

Land Tenure Systems We gather information on land tenure systems for all countries in our dataset from Bruce (1996). This report belongs to the Land Tenure Center's "Africa Land Tenure Profiles" project undertaken at the request of USAID. For all countries in the three regions of Greater Horn, Southern Africa, and West Africa, the report provides information on the nature of the land tenure system in place in 1996, focusing on the duality between statutory and customary tenure. In our analysis, we will focus on whether the national legislation in each country explicitly recognizes customary land tenure (Herbst 2000).

Village Characteristics We use the geographical information in the DHS data to derive a number of characteristics at the cluster or village level. We do this by spatially matching DHS locations to other datasets. In particular, we use information on latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. In exploring the relevant mechanism, we also use information from satellite data on the share of land within a 1km radius from the village that is devoted to cropland, or mixed cropland and natural vegetation. We use this information as recorded in the year after elections take place. The Supplementary Appendix B.2 provides a detailed explanation of sources and variable definitions, and summary statistics at this level.

3 Empirical Strategy

Do electoral results shape individual outcomes according to ethnicity? To answer this question, we compare individuals from ethnicities linked to political parties which gain or do not gain a local representative in the national assembly. However, those ethnicities and parties that gain a representative may potentially be very different from those that do not along a number of observable and unobservable dimensions which may themselves be related to individual outcomes. This undermines a causal interpretation of a simple comparison across the two groups. To address this issue, we focus on those ethnic groups and political parties that are at the margin of gaining a local representative in parliament. We implement a regression discontinuity design (Imbens and Lemieux 2008; Lee and Lemieux 2009a) and compare individuals from ethnicities linked to political parties that are within a narrow vote margin from gaining or not a representative. The key identification assumption is that average individual characteristics do not change discontinuously at the vote threshold that is relevant for the linked party to gain a local representative.

Let c_{dnt} be the relevant vote share threshold to gain a local representative in constituency d of country n in the the elections prior and closest to year t. Let instead X_{ednt} be the vote share of the party linked to ethnicity e in the constituency. Let D_{ednt} be the treatment dummy, equal to one if the party gained a local representative in the constituency, i.e. $D_{ednt} = \mathbb{1}(X_{ednt} \geq c_{dnt})$.

Consider the following local linear regression specification

$$y_{iednt} = \alpha + \beta\ D_{ednt} + \gamma\ (1 - D_{ednt}) \times (X_{ednt} - c_{dnt}) + \delta\ D_{ednt} \times (X_{ednt} - c_{dnt}) + u_{iednt} \tag{2}$$

where y_{iednt} is the outcome of individual i from ethnicity e surveyed in constituency d in country n and year t. $X_{ednt} - c_{dnt}$ is the adjusted vote share of the party linked to ethnicity e. The terms $\gamma\ (1 - D_{ednt}) \times (X_{ednt} - c_{dnt})$ and $\delta\ D_{ednt} \times (X_{ednt} - c_{dnt})$ capture any systematic linear change in individual outcome with the party vote share on the left and right side respectively of the relevant threshold. Our coefficient of interest is β, which captures a discontinuous change in the average of individual outcomes associated with gaining a local representative in the national assembly.

A necessary condition for implementing this specification is having a complete mapping from ethnicities to parties. Our estimation sample is therefore restricted to those individuals that belong to ethnic groups for which our linking procedure described in Section 2 identifies a link with a political party. The labelling algorithm links a party to the ethnic group that has the largest gap between the share of voters from that ethnic group for that party and the party's overall vote share in the country. Although highly unlikely, this gap could theoretically be zero for all parties. This would happen if the ethnic group in question voted exactly as the population average. More realistically, the gaps might be very small for many parties. We therefore impose a minimum cutoff on the size of the gap used to attribute links. Increasing the size of the cutoff means fewer false positives – erroneously labelling ethnic groups that happen to be over-represented only by a small margin among the voters of a party. There is a cost, however, in that the number of false negatives – failing to label ethnic groups that should be labelled – will increase. Failing to label an ethnic group also reduces the total number of observations, since we can only compute the running variable (see below) for ethnicities that are linked to a party. Our baseline estimates are obtained using a labeling cutoff of 0.01. Appendix A reports the results obtained when imposing a cutoff of 0.05, which are essentially unchanged.

The final dataset counts 409,604 individuals across 243 ethnic groups linked to 62 different political parties across 15 countries and 32 elections. The third column of Table A.1 in Appendix A shows the number of constituency-year observations in the election dataset that we match with individual-level observations and are therefore relevant for our analysis, for a total number of 3,254.[7]

Definition of the Running Variable The vote share threshold c_{dnt} and the associated adjusted vote share $X_{ednt} - c_{dnt}$ is defined differently across constituencies and electoral systems. The

[7]Panel A of Table 1 shows the summary statistics for this sample of all variables used in the empirical analysis.

vast majority of countries in our sample adopt a single-member plurality rule, where one representative – the one obtaining the relative majority of votes – is elected in each constituency. We can therefore define the adjusted vote share as follows. For the winning party, it is the difference between its vote share and that of the second party. For all other parties, it is defined as the difference between their vote shares and that of the winning party. In proportional systems in our sample, a given number of seats are assigned proportionally to vote shares within constituencies. The adjusted vote share is defined differently depending on whether the system follows the d'Hondt – such as Mozambique – or the Hare method – Benin, Burkina Faso – to allocate seats. The relevant threshold is the one associated with the last relevant quotient (d'Hondt) or the vote share of the last party gaining a representative (Hare).

To better understand the variation that we exploit for identification, the last three columns of Table A.1 in Appendix A show the number of constituency-year observations in the election dataset for which we observe any party with a vote margin within 20%, 10%, and 5% respectively from the relevant vote margin threshold associated with gaining a representative in parliament. When considering a vote margin as narrow as 5%, we can still leverage variation in the variables of interest across 419 constituency-year observations. Yet, all our estimates consider a bandwidth of around 20%, thus keeping a number of constituency-year observations as high as 1,456 in the estimation sample.[8] Figure 1 depicts the spatial distribution of constituency-year observations that are included in our final sample. It does so separately for West Africa (a), East Africa (b), and Southern Africa (c). We highlighted in black those constituency-years where we observe two or more parties in our dataset with a vote margin of less than 20% percent from the relevant threshold for winning a seat in parliament. There is significant spatial variation both across and within countries. The average number of parties in contested constituency-year observations is 1.64 with a standard deviation of 0.7.[9]

Estimation We estimate the coefficient of interest β from equation 2 using different methods. The first one is OLS – $\hat{\beta}_{OLS}$. This requires selecting the estimation bandwidth, which in turns defines the estimation sample. We also rely on the latest advances in regression discontinuity methods and also implement the estimators proposed by Imbens and Kalyanaraman (2012) and Calonico, Cattaneo, and Titiunik (2014) – $\hat{\beta}_{IK}$ and $\hat{\beta}_{CCT}$ respectively. These are both local polynomial nonparametric estimators with data-driven bandwidth selectors and bias-correction techniques. We implement the bandwidth selector proposed by Calonico, Cattaneo, and Titiunik (2014) and implement all estimators including OLS using that bandwidth.

[8] Panels B and C of Table 1 provide the summary statistics for the subsample of individual observations from ethnic groups linked to political parties within 20% of the vote margin threshold that is relevant to elect a local representative to the national assembly, right and left of the threshold respectively.

[9] Notice that, in order to be part of our final sample, a party has to be linked with an ethnicity in DHS. Therefore, the final dataset does not necessarily has information on all parties running in a given constituency.

Balance Tests The validity of regression discontinuity designs rests on the assumption of there being no manipulation of the running variable around the threshold. To support this assumption, we test for the null hypothesis of continuity of the density of the running variable. Figure A.1 in Appendix A shows graphically the distribution of the density function on both sides of the threshold, together with 95% confidence intervals. The p-value from a McCrary (2008) test of equality of the value of the density function on the left and right side of the threshold is equal to 0.4761, and equal to 0.4263 using the more recent test developed by Cattaneo, Janssonz, and Ma (2018). We therefore cannot reject the hypothesis of no discontinuity in the density of the running variable at the relevant vote share threshold, which is suggestive of the absence of manipulation.[10]

We also test for any discontinuous change in the value of covariates at the threshold. We consider a wide range of both village-level and individual characteristics. Tables A.2 and A.3 in Appendix A report the estimated coefficient of interest that we obtain when estimating equation 2 and having as dependent variable these covariates. We report estimates we obtain using the three estimators discussed above. With the exception of a few OLS estimates, none of the estimated coefficients are statistically significant, which is what we expect if local randomization is effectively implemented at the relevant threshold. Figure A.3 in Appendix A provides a graphical representation of the data and the local linear regression fit, showing the absence of any meaningful discontinuities in the value of covariates at the threshold.

4 Results

4.1 Descriptive Evidence

Panel A of Table 1 shows the summary statistics in the full sample of all the variables we use in the empirical analysis. Panels B and C of Table 1 do the same for the subsample of individuals from ethnic groups linked to parties within 20% of the vote margin threshold that is relevant to elect a local representative to the national assembly, on the right and left of the threshold respectively. 58.7% of the individuals in the full sample report to be working at the time of the DHS interview. Panel B and C indicate that the average probability of employment is 1.4 percentage points higher for individuals from ethnicities linked to parties that gain a representative compared to individuals linked to losing parties. The probability of working in the agricultural sector is 2.3 percentage points higher, while that of working in the manufacturing or service sectors is either lower or the same as the one of individuals linked to parties that do not gain a representative. The same is true for the probability of reporting an occupation that is associated with the public sector, such as teachers or health professionals.

[10] Figure A.2 in Appendix A shows graphically the distribution of the density function on both sides of the threshold together with 95% confidence intervals across different subsamples over which we later implement our main regression specification.

Panel D of Table 1 shows the summary statistics in the sample that we derive from the 2008 Afrobarometer, which provides information on individuals' beliefs and attitudes towards traditional leaders. On the one hand, only 1.1% of individuals on average believe that traditional leaders should be partisan in politics. On the other hand, the authority of local chiefs is widely recognized. 38% of respondents recognize the traditional leader as the main authority that is responsible for the allocation of land, and 57% believe the chief is responsible for governing the community.

Next, we provide some insights into the extent of political competition in the electoral constituencies in our sample and its correlates. To this end, we keep from our sample only those individual DHS observations that belong to the year that is closest in time to the one of elections. We derive ethnic diversity measures and summary statistics for the 2,293 unique electoral constituencies in our sample, and correlate them with measures of political competition. The number of parties running in the average electoral district is equal to 4.72. Of these, an average of 1.19 parties achieve a vote share that is within a 20% vote margin of gaining a representative in parliament. We can use vote shares to derive a Herfindahl-type index of political competition. Its average value across constituencies is 0.41.

Table A.4 in Appendix A shows the coefficient estimates we obtain when regressing measures of political competition in the constituency over a number of district-level characteristics. We start by investigating the relationship between number of parties running in the constituency and ethnic diversity as measured by the *fractionalization* and *polarization* indices.[11] When included separately as the only regressor in columns 1 and 2, their coefficient is positive and significant. This suggests that political competition is higher in more ethnically diverse constituencies. Yet, when we include the two of them together in column 3, only the coefficient of the fractionalization index retains its significance. This does not change when conditioning on the full set of country fixed effects in column 4. This indicates that the number of parties running for elections is higher in more ethnically fractionalized constituencies, where the number of ethnic groups is larger. In column 5, we add a set of variables that capture other economic and demographic characteristics of the electoral constituency. The number of parties is higher in more populated, more educated, and less rural constituencies. The share of individuals that are employed is positively correlated with the number of parties, although not significantly so. The share of individuals that report to be working in agriculture and services – including subsistence self-employment – is negatively and significantly correlated with the number of parties running for election. In column 6, we replace as dependent variable the Herfindahl-type index of political competition that we calculate using party vote shares. The index value is higher when vote shares are more concentrated, and political competition is lower. Coefficient signs are largely consistent with the ones obtained when looking at the raw number of

[11] See Montalvo and Reynal-Querol 2005 for an extensive discussion of the two indices and their relationship with the incidence of civil war.

running parties. In particular, ethnic fractionalization is significantly and positively correlated with political competition.

Before implementing the empirical strategy illustrated in the previous section, we present two examples that illustrate the kind of variation that we exploit for identification. Figures 2 shows a cluster of constituencies in Sierra Leone where the Sierra Leone People's Party (SLPP) gained or did not gain a local representative in the national assembly in the 2007 elections. The left map reports the adjusted vote share of the SLPP across these constituencies. In 2007, the SLPP lost by a narrow margin in the two constituencies of Kakua and Panga Kabonde while winning in the other constituencies in the cluster with a vote margin of less than 20%. The right map reports the employment rate in 2008 of individuals belonging to the Mende ethnic group, which the labelling algorithm identifies as linked to the SLPP. The employment rate of Mendes is lowest in the those constituencies where the SLPP lost.

Figure 3 focuses on the case of two Ghanaian constituencies and the electoral performance of the New Patriotic Party (NPP) across the two elections of 2004 and 2012. The left maps report the adjusted vote share of the NPP across these constituencies. The NPP gained a representative from both districts in 2004, but lost the one from Nkoranza in 2012. The right map reports the employment rate in 2008 and 2014 respectively of individuals belonging to the Akan ethnic group, which the labelling algorithm identifies as linked to the NPP. The employment rate of Akans decreases in the district that the NPP lost in 2012 while increasing in the district that the party retained. The analysis that follows aims to investigate these patterns in a systematic way.

4.2 Employment Effects

Table 2 shows the first set of coefficient estimates that we obtain by implementing the specification in equation 2. The dependent variable is a dummy equal to one if the individual is employed. For each specification, we report different estimates of β using the same optimal bandwidth (Calonico, Cattaneo, and Titiunik 2014). The bottom row reports the number of individual-level observations that are effectively used to produce these estimates, meaning those for which the vote margin of the associated party is within the selected bandwidth. Column 1 reports unconditional estimates of β. These range from 2 to 3.8 percentage points. Yet, none of them is statistically significant. Notice that our sample includes individuals from 15 countries over a period of around 20 years of DHS. It therefore comes as no surprise that the variation in employment rates that is not explained by the included regressors is large, and so are the estimated standard errors. This can be clearly seen in the left graph of Figure 4, which groups observations in 30 bins left and right of the vote margin threshold and plots average employment probabilities per bin.

Column 2 of Table 2 reports the estimates that we obtain when exploiting variation within each country and year. Point estimates are remarkably similar to the one in column 1, and

are now statistically significant at the 10% level. This is indeed what we would expect if our identification strategy is valid and local randomization is effectively implemented at the threshold: including the full set of country-year fixed effects does not change meaningfully the estimated coefficients, but decreases the amount of unexplained variation in the dependent variable, thus decreasing standard errors.

In column 3, we net out average differences in employment rate across ethnicities. Once again, point estimates do not change in any meaningful way, but all of them are now significant at the 5% level. Results are unchanged when including the full set of controls at the village level discussed above. Including individual-level controls in column 5 slightly reduces the significance of the conventional and bias-corrected estimates, leaving again their magnitude substantially unaltered. In column 6, we implement our most demanding specification, which exploits variation within constituencies across individuals belonging to different ethnicities and over time. The magnitude of all estimates decreases slightly, but they remain significant at the 10% level.

Overall, results from Table 2 shows that individuals from ethnic groups linked to parties that gain a local representative in parliament are 2 to 3 percentage points more likely to be employed compared to others. This represents a 3.5 to 5.4% increase over the average employment probability left of the threshold.[12] Figure 4 provides a graphical representation of the discontinuity. The centre and right graphs show conditional probabilities obtained after netting out average differences across country-year pairs and ethnicities (centre) and also constituencies (right) respectively. The three graphs altogether show how the inclusion of fixed effects decreases the amount of unexplained variation while leaving the size of the discontinuity at the threshold essentially unchanged.[13]

4.3 Sorting and Placebo Estimates

One possible concern with these estimates pertains to the direction of causality. It may be that those parties that are linked to ethnicities that have higher (lower) employment probabilities are systematically more likely (not) to gain a representative by a narrow margin. Notice that the focus on a sufficiently narrow margin and the higher weights attached to observations closer to the threshold would by itself in principle address this point, unless this sorting pattern still occurs in the vicinity of the threshold.

We take this concern seriously. We restrict the sample to those countries, constituencies, and political parties that run in more than one election in our sample – thus excluding a priori Ivory

[12] Table 1 in Appendix A shows that the average employment probability left of the threshold within the 20% bandwidth is equal to 55.9%.

[13] As discussed earlier, Tables A.11 to A.15 in Appendix A report all the estimates we obtain when using an overrepresentation cutoff of 0.05 in identifying links between ethnic groups and political parties. All results are very similar to the ones we obtained with our baseline cutoff choice of 0.01.

Coast, Mali, Mozambique and Uganda, for which we have data on one election round only. Table A.5 in Appendix A shows that our main findings hold in this restricted sample. We then implement the same baseline specification as before, but, in constructing the main regressors, we consider the vote margin obtained by the party in the *following* election. That is, we test whether the employment probability of individuals changes discontinuously when the party linked to their ethnic group gains a representative by a narrow margin in elections that occur after the individual was surveyed. This amounts to a placebo test: if no sorting occurs around the threshold along dimensions related to pre-existing employment rates, the estimated β from this specification should be close to zero and insignificant.

Table A.6 in Appendix A reports the corresponding placebo estimates. Indeed, none of them is significant, and the majority of them are close to zero in magnitude. We interpret this as further evidence of a lack of manipulation around the threshold, and additional support for the validity of our identification approach.

4.4 Party-level Heterogeneity

Individuals belonging to ethnicities connected to the elected member of parliament are systematically more likely to be employed. The extent to which this is the case likely depends on the amount of political power held by the politician. In other words, the effect may depend on whether the politician is aligned with the government that forms in the aftermath of the election.

For each election in our dataset, we exploit information on whether each party supports the government. We then test whether the employment effects in Table 2 are heterogeneous along this dimension. Table 3 reports the corresponding results. Columns 1 to 3 present the coefficient of interest from equation 2 as estimated from the subsamples of individual observations belonging to ethnicities connected to parties that do not support the government. The estimated coefficient is for the most part statistically insignificant and economically insignificant in the most demanding specifications. A different story emerges for those individuals belonging to ethnicities connected to parties that support the government. The estimates in columns 4 to 6 show that individuals connected to parties that support the government are 5 to 8 percentage points more likely to be employed when gaining a local representative in parliament compared to the others.[14]

5 The Mechanism

Several authors have highlighted the salience of traditional leaders in African politics.[15] In the rural parts of the continent, to gain the support of voters means to recognize and engage with

[14] Figure A.2 in Appendix A shows graphically the distribution of the density function on both sides of the threshold across the two subsamples, together with 95% confidence intervals.

[15] See Michalopoulos and Papaioannou (2014) and Baldwin (2014) for a review.

traditional chiefs. The interaction between politicians and traditional leaders takes different forms. Hagberg (2007) describes how the arrival of democracy in Burkina Faso strengthened the involvement of traditional leaders in party politics. He explains how chiefs act as "big electors" able to mobilize voters in support of one given party. Similarly, de Kadt and Larreguy (2018) discuss the crucial role of Zulu traditional leaders as political brokers for the Inkhata Freedom Party before 2009 in South Africa. In the case of Zambia, Baldwin (2013) reports that voters tend to align their voting decision to that of their chief if they perceive him to be collaborating with the politician running for the parliamentary election. In some cases, the role of chiefs as political brokers is formalized. This is the case for Cameroon in the 2002 elections, when the South West Chiefs Conference stated that the chiefs would act as campaign managers for the Cameroon People's Democratic Movement (Fonchingong 2005).

In order to secure the support of traditional leaders, politicians cede control over resources. Baldwin (2014) explains that "because traditional leaders are generally thought to favor their indigenous subjects in the allocation of land, the [politicians'] promise to devolve power to these leaders effectively signals a commitment to empowering indigenous community members." Boone (2009) explains how such a mobilization strategy has been put in place by political candidates in the Ivory Coast.

Importantly, the relationship between politicians and traditional chiefs is not exclusive to the election period. Given the mismatch between their limited mandate and the lifetime position of chiefs, politicians are aware that they will need to keep on interacting with chiefs to gain their support in future elections. This repeated interaction framework allows the chiefs to hold the elected politicians accountable for delivering their electoral promises. Lindberg (2010) reports that "chiefs exert considerable powers over MPs from rural constituencies" and they "overwhelmingly hold MPs accountable for community development." Baldwin (2013) stresses that, in the period between democratic elections in Zambia, "non-elected traditional leaders have continued to influence governance in rural areas."

This large body of qualitative evidence informs our hypothesis regarding the mechanism behind the employment effects presented in the previous section. We argue that the positive employment effect of being linked to the winning party originates from the interactions between formally appointed political leaders and traditional ethnic power structures. These two levels of governance interact in a strategic way. Political parties demand votes and supply *de facto* empowerment to traditional leaders. The latter mobilize the individuals under their authority for political support and voting. When gaining a local representative, traditional leaders from connected ethnicities retain their *de facto* powers, first and foremost the one to allocate agricultural land. As a result, co-ethnics enjoy a competitive advantage in the agricultural land and labor market, gaining preferential access to land and agricultural jobs. The analysis that follows provides evidence that supports this hypothesis.

5.1 Pre-Colonial Ethnic Institutions

Embedded in our argument is the statement that traditional ethnic leaders are key in mobilizing co-ethnics, and building and sustaining the relationship between a given ethnic base and political parties. If that is the case, we should expect the benefits of having a local ethnic politician in parliament to be higher where pre-colonial ethnic institutions are stronger. To test this hypothesis, we implement the baseline employment regression specification and estimate the coefficient of interest separately in two subsamples as defined by the strength of pre-colonial ethnic institutions.

Table 4 shows the corresponding results. The employment effect of gaining a local ethnic representative in the parliament is concentrated in constituencies that are located in ethnic homelands with traditionally high levels of pre-colonial political centralization. In this case, point estimates range from 4 to 6 percentage points, double the magnitude of the main average effect. The effect is virtually zero in constituencies located in ethnic homelands with low levels of pre-colonial centralization. None of these estimates change meaningfully in magnitude with the inclusion of village-level and individual-level controls.[16]

These results are consistent with those in Table 2 in showing that those estimates are a weighted average between a positive effect in areas with high levels of pre-colonial centralization and a zero or no effect in other areas. We interpret this as evidence that traditional ethnic power structures interact with formal politicians and shape the agency relationship between the latter and their ethnic base, consistent with the hypothesis we put forward concerning the main mechanism that generates our baseline findings.

5.2 Employment by Sector

Using the information on the sector of employment, we can detect whether the employment effect identified above is concentrated in a particular sector. Column 1 of Table 5 report the coefficient estimates that we obtain when replacing as dependent variables a dummy equal to one if the individual is employed in the agricultural sector. Estimates are conditional on country-year, ethnicity, and constituency fixed effects. In column 2, we also include as regressors the full set of village-level and individual-level characteristics. In columns 3 to 6, we replace as dependent variable a dummy equal to one if the individual is employed in the manufacturing or service sector.

Results show that the employment effects of gaining a local ethnic representative are concentrated in the agricultural sector. The effect is virtually indistinguishable from zero for manufacturing and service employment. The point estimates suggest that the effect on agricultural

[16]Figure A.2 in Appendix A shows graphically the distribution of the density function on both sides of the threshold across the two subsamples, together with 95% confidence intervals.

employment is around 2 percentage points, a 7.6% increase over the average agricultural employment rate left of the threshold.[17]

In columns 7 and 8, we replace as dependent variable a dummy equal to one if the individual reports to have an occupation that we categorize as belonging to the public sector, such as teachers and health professionals. Individuals belonging to ethnicities linked to parties that gain a local representative are not systematically more likely to have public sector occupations. The point estimate is not only insignificant at standard levels, but negligible in magnitude.

These results are once again consistent with our hypothesis that identifies preferential access to land and agricultural jobs as the main channel for the main employment effect. It also speaks to the inability for politicians to sustain more direct forms of patronage (Colonnelli, Prem, and Teso 2018) in settings where the public sector is small, formal states are weak and lack infrastructural power (Mann 1984).

5.3 Ethnic Disfavoritism and Land Use

A situation in which existing land is reallocated away from non co-ethnics and assigned to co-ethnics is different from one in which new land is cleared and assigned to co-ethnics. Differentiating between the two scenarios is important in setting the ground for the evaluation of the welfare consequences of this phenomenon.

We gain insights into what the relevant scenario is in two ways. First, we identify those ethnic groups which can potentially be *dis*-favored by political parties. We identify *negative* links between ethnic groups and political parties using the same method described in Section 2, but adopting an *underrepresentation* criterion. That is, we assign party label "p" to ethnic group "e" if party p is (the most) *under*-represented, relative to its overall vote share in the country, among voters from ethnic group e.[18] We can then implement the same specification in equation 2 to test whether the labor market outcomes of individuals belonging to ethnicities that are underrepresented among the voters of a given party change discontinuously when that party gains a representative in parliament. Table 6 presents the coefficient estimates that we obtain when having as dependent variables a dummy equal to one if the individual is employed in the agricultural (columns 1 and 2), manufacturing (columns 3 and 4), or service sector (columns 5 and 6). As in Table 5, all estimates are conditional on country-year, ethnicity, and constituency fixed effects. Point estimates in column 1 and 5 provide evidence that individuals belonging to ethnicities that are underrepresented among the voters of the winning party reallocate away

[17] Table 1 in Appendix A shows that the average agricultural employment probability left of the threshold within the 20% bandwidth is equal to 26.2%.

[18] We impose a labelling cutoff of 0.05. The trade-off involved in the choice of the labelling cutoff is specular to the one we face when attaching party labels based on over-representation, which we discuss in Section 3 above. Tables A.7 and A.8 in Appendix A reports the results obtained when imposing a cutoff of 0.01, which are similar to the ones we report in Table 6 and Table A.14 in Appendix A.

from agriculture towards the service sector – which includes non-farming entrepreneurship. Most of the relevant estimates are not robust to the inclusion of village-level and individual-level characteristics as controls in the second column of each dependent variable. This evidence of reallocation across sectors is consistent with the results from Table A.14 of Appendix A showing no discontinuous change in the overall probability of employment. Taken together with the results from Table 2 and 5, we regard them as supportive of the hypothesis that ethnic favoritism takes place through a reallocation of existing land which favors connected ethnicities rather than through cropland expansion that makes the latter available to such ethnicities.

We further validate this interpretation of results using satellite data on land use. To test for cropland expansion, we test whether the share of land devoted to cropland in a given village changes with the ethnic links of local representatives in the national parliament. We implement the regression specification in equation 2, but having now as unit of observation a DHS cluster. We then assign to each cluster the vote margin of the party that is linked to the ethnicity that has the highest population share in the cluster, and define the regressors accordingly. The dependent variable is the share of land within a 1km buffer area around the cluster that is classified as cropland. In this case, β captures whether the share of cropland changes discontinuously when the majoritarian ethnic group in the village elects a local representative in the national parliament. Table A.9 in Appendix A shows the corresponding OLS coefficient estimates, using all available observations. With the exception of column 1, all estimates of the coefficient of interest are not statistically significant.[19]

5.4 Land Tenure Systems

Our hypothesis is that traditional leaders from ethnicities connected to elected members of parliament are empowered to allocate agricultural land, and this generates the employment effects identified above. The extent to which this happens likely depends on whether the land allocation implemented by traditional leaders is formally recognized by national legislation.

To test for this hypothesis, we use information on whether the national legislation in each country explicitly recognizes customary land tenure. We then implement separately our main regression specification across two subsamples as defined by whether the observations belong to countries whose national legislation officially recognizes customary land tenure or not.[20]

Table 7 shows the corresponding set of coefficient estimates. The employment effect of gaining a local ethnic representative in the parliament is concentrated in those countries whose national legislation officially recognizes customary land tenure. This supports our hypothesis that the competitive advantage on the labor market of individuals from ethnicities connected to the win-

[19] Figure A.5 shows a graphical representation of the (absence of) discontinuity across different specifications.

[20] Figure A.2 in Appendix A shows graphically the distribution of the density function on both sides of the threshold across the two subsamples, together with 95% confidence intervals.

ning party originates from positive discrimination in access to agricultural land as allocated by traditional leaders, as this happens only in those countries where such an allocation has legal recognition.

5.5 Role of Traditional Leaders

We provide additional evidence supporting the mechanism using Afrobarometer survey data. In the 2008 round, surveyed individuals are asked a number of questions about local politics and the authority of traditional leaders in society. We can therefore test whether the answers to such questions change with electoral outcomes in a discontinuous manner. Our final Afrobarometer dataset includes 8 countries and around 5,000 individuals – all of those for whom we identify links between their ethnic group and political parties running in the constituencies where they are surveyed. We lack the density of observations around the threshold that is needed to implement the most advanced regression discontinuity estimators. We therefore report only OLS estimates of β, which can therefore be interpreted only as suggestive evidence.

In the survey, individuals are asked whether traditional leaders should be partisan. We implement the regression specification in equation 2, but replacing as dependent variable a dummy that is equal to one if the individual answers positively this question. Table 8 shows the corresponding coefficient estimates. Gaining a local representative in parliament is associated with a positive significant and discontinuous increase in the likelihood of indicating that traditional leaders should be partisan. Point estimates are invariant to the inclusion of constituency and ethnicity fixed effects, and village-level and individual-level controls, ranging between 1.3 and 2 percentage points. We interpret this as evidence that the agency relationship between traditional ethnic leaders and politicians is strengthened when a local representative is elected.

The Afrobarometer questionnaires also ask about the authority that the individual considers responsible for the allocation of land. We construct a dummy that is equal to one if the individual identifies the traditional leader as responsible, and implement again the same regression specification. Table 9 reports the corresponding estimates of β. Depending on the specification, point estimates range from 5.8 to 9.4 percentage points, a 15 to 24% increase over the sample mean. Having a local ethnic politician in parliament is associated with a discontinuous increase in the probability of identifying the traditional leader as responsible for the allocation of land.[21] We interpret this as consistent with the mechanism we suggest: as local ethnic politicians are

[21] Table A.10 in Appendix A shows additional results that we obtained considering other variables such as the role of the traditional leader in maintaining law and order or governing the community. Interestingly, we find some evidence that having a local ethnic politician in parliament is associated with a decrease in the probability of identifying the traditional leader as responsible for governing the community, suggesting a pattern of division of power between traditional leaders and formally appointed politicians. This finding seems to point out the existence of a "duality" in the governance structure, split between formally appointed political leaders and traditional ethnic chiefs. Figure A.4 in Appendix A provides a graphical representation of these results.

elected to parliament, traditional ethnic leaders are empowered to allocate agricultural land and favor co-ethnics. This differential access to land maps into the higher employment probabilities in the agricultural sector that we previously documented.

6 Conclusions

This paper studies the labor market effects of ethnic politics and its impact on the allocation of resources, namely labor, between and within ethnic groups and sectors. Combining geo-referenced data across 15 countries on elections, ethnicities, and individual outcomes, we show that individuals from ethnicities connected to parties that elect a local representative in the national parliament enjoy a competitive advantage in the local labor market, and in the agricultural sector in particular. Our argument is that strategic interactions between formally appointed politicians and traditional leaders trigger differential access to land and agricultural jobs across individuals belonging to different ethnicities. We present multiple pieces of evidence that are consistent with this hypothesis. First, the employment effects are concentrated in the historical homelands of ethnicities with strong pre-colonial institutions. Second, individuals from connected ethnicities are more likely to be employed in agriculture, and in those countries where customary land tenure is officially recognized by national legislation. Third, they are also more likely to identify traditional leaders as partisan, and as being mainly responsible for the allocation of land.

Our results call for additional work on this topic. In particular, our study is silent on the welfare consequences of the allocative effects of ethnic politics that we document. It is not clear whether the market distortions introduced by the strategic interactions between politicians and ethnic leaders increase or decrease overall efficiency, as this will depend on the distribution of skills across ethnic groups and sectors. The quest for appropriate data to answer this question motivates our future research program.

References

Acemoglu, D., T. Reed, and J. A. Robinson (2014). Chiefs: Economic development and elite control of civil society in sierra leone. *Journal of Political Economy 122*(2), 319–368.

Afrobarometer (2015). *Afrobarometer Data*. Available at http://www.afrobarometer.org.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2015). Ethnic inequality. *Journal of Political Economy 123*(3), 547–724.

Baldwin, K. (2013). Why vote with the chief? political connections and public goods provision in zambia. *American Journal of Political Science 57*(4), 794–809.

Baldwin, K. (2014). When Politicians Cede Control of Resources: Land, Chiefs, and Coalition-Building in Africa. *Comparative Politics 46*, 253–271.

Baldwin, K. (2015). *The Paradox of Traditional Chiefs in Democratic Africa*. Cambridge University Press.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2019). Unity in diversity? how intergroup contact can foster nation building. *American Economic Review 109*(11), 3978–4025.

Boone, C. (2009). Electoral populism where property rights are weak: Land politics in contemporary sub-saharan africa. *Comparative Politics 41*(2), 183–201.

Bruce, J. W. (1996). *Country Profiles of Land Tenure: Africa*. Land Tenure Center.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Padró i Miquel (2015). The value of democracy: Evidence from road building in kenya. *American Economic Review 105*(6), 1817–51.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs. *Econometrica 82*, 2295–2326.

Carr, A. (2017). *Election Archive*. Available at http://psephos.adam-carr.net.

Cattaneo, M. D., M. Janssonz, and X. Ma (2018). Simple Local Polynomial Density Estimators. Mimeo.

Cingano, F. and P. Pinotti (2013). Politicians At Work: The Private Returns And Social Costs Of Political Connections. *Journal of the European Economic Association 11*(2), 433–465.

Colonnelli, E., M. Prem, and E. Teso (2018). Patronage and Selection in Public Sector Organizations. Documentos de Trabajo 016723, Universidad del Rosario.

de Kadt, D. and H. A. Larreguy (2018). Agents of the regime? traditional leaders and electoral politics in south africa. *The Journal of Politics 80*(2), 382–399.

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

Depetris-Chauvin, E., R. Durante, and F. R. Campante (2018). Building Nations Through Shared Experiences: Evidence from African Football. NBER Working Papers 24666, National Bureau of Economic Research, Inc.

Dickens, A. (2018). Ethnolinguistic favoritism in african politics. *American Economic Journal: Applied Economics 10*(3), 370–402.

Economic Commission Africa (2007). Relevance of african traditional institutions of governance. Technical report.

Eggen, O. (2011). Chiefs and everyday governance: Parallel state organisations in malawi. *Journal of Southern African Studies 37*(2), 313–331.

Eifert, B., E. Miguel, and D. N. Posner (2010). Political competition and ethnic identification in africa. *American Journal of Political Science 54*(2), 494–510.

Fergusson, L., H. Larreguy, and J. F. Riaño (2018). Political competition and state capacity: Evidence from a land allocation program in mexico. Technical report, LACEA Working Paper.

Finan, F. and L. Schechter (2012). Vote-buying and reciprocity. *Econometrica 80*(2), 863–881.

Fonchingong, C. C. (2005). Exploring the politics of identity and ethnicity in state reconstruction in cameroon. *Social Identities 11*(4), 363–380.

Franck, R. and I. Rainer (2012). Does the leader's ethnicity matter? ethnic favoritism, education, and health in sub-saharan africa. *American Political Science Review 106*, 294–325.

Friedl, M. A., D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment 114*, 168–182.

Gagliarducci, S. and M. Manacorda (2017). Politics in the family nepotism and the hiring decisions of italian firms. *American Economic Journal: Applied Economics*. Forthcoming.

Gennaioli, N. and I. Rainer (2007). The modern impact of precolonial centralization in africa. *Journal of Economic Growth 12*(3), 185–234.

Goldman, E., J. Rocholl, and J. So (2013). Politically connected boards of directors and the allocation of procurement contracts. *Review of Finance 17*(5), 1617–1648.

Hagberg, S. (2007). *Traditional Chieftaincy, Party Politics, and Political Violence in Burkina Faso*, pp. 131–153. New York: Editors: Buur, Lars and Kyed, Helene Maria. Palgrave Macmillan US.

Herbst, J. (2000). *States and power in Africa: Comparative lessons in authority and control*. Princeton University Press.

ICF (2017). *Demographic and Health Surveys (various)*. Funded by USAID. Rockville, Maryland: ICF [Distributor].

Imbens, G. and K. Kalyanaraman (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *Review of Economic Studies 79*(3), 933–959.

Imbens, G. W. and T. Lemieux (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics 142*(2), 615–635.

Inter-Parliamentary Union (2018). *New Parline: the IPU's Open Data Platform.* https://data.ipu.org/ (accessed August 1, 2018).

Jedwab, R. and A. Storeygard (2018). The Average and Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa. *mimeo*.

Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. Ehrlich, and J. Sachs (2004). A global index representing the stability of malaria transmission. *American Journal of Tropical Medicine and Hygiene 70*(5), 486–98.

Kollman, K., H. Allen, D. Caramani, D. Backer, and D. Lublin (2018). *Constituency-Level Elections Archive, Release 11*. Produced and distributed by Ann Arbor, MI: Center for Political Studies, University of Michigan, http://www.electiondataarchive.org.

Kramon, E. and D. N. Posner (2014). Ethnic Favoritism in Primary Education in Kenya. mimeo.

Larreguy, H., J. Marshall, and P. Querubin (2016). Parties, brokers, and voter mobilization: How turnout buying depends upon the party's capacity to monitor brokers. *American Political Science Review 110*(1), 160–179.

Lee, D. S. and T. Lemieux (2009a). Regression Discontinuity Designs in Economics. Working Papers 1118, Princeton University, Department of Economics, Industrial Relations Section.

Lee, D. S. and T. Lemieux (2009b). Regression Discontinuity Designs in Economics. Working Papers 1118, Princeton University, Department of Economics, Industrial Relations Section.

Lindberg, S. I. (2010). What accountability pressures do mps in africa face and how do they respond? evidence from ghana. *The Journal of Modern African Studies 48*(1), 117–142.

Lizzeri, A. and N. Persico (2004). Why did the elites extend the suffrage? democracy and the scope of government, with an application to britain's "age of reform". *The Quarterly Journal of Economics 119*(2), 707–765.

Mamdani, M. (1996). *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton University Press.

Mann, M. (1984). The autonomous power of the state : its origins, mechanisms and results. *European Journal of Sociology / Archives Européennes de Sociologie / Europäisches Archiv für Soziologie 25*(2), 185–213.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

McMillan, M. and K. Harttgen (2014). What is driving the 'African Growth Miracle'? *NBER Working Paper 20077*. Citation Key: McMillanHarttgen14.

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

Michalopoulos, S. and E. Papaioannou (2014). National institutions and subnational development in africa. *The Quarterly journal of economics 129*(1), 151–213.

Michalopoulos, S. and E. Papaioannou (2015). On the ethnic origins of african development: Chiefs and precolonial political centralization. *The Academy of Management Perspectives 29*(1), 32–71.

Montalvo, J. G. and M. Reynal-Querol (2005). Ethnic polarization, potential conflict, and civil wars. *American Economic Review 95*(3), 796–816.

Mueller, H. and A. Tapsoba (2016). Access to power, political institutions and ethnic favoritism.

Murdock, G. (1967). *Ethnographic Atlas*. University of Pittsburgh Press.

Natural Earth (2018). 1:10m populated places.

Nohlen, D., T. Bernard, and M. Krennerich (1999). *Elections in Africa: A Data Handbook*. Oxford: Oxford University Press.

Nunn, N. and D. Puga (2012). Ruggedness: The Blessing of Bad Geography in Africa. *The Review of Economics and Statistics 94*(1), 20–36.

Posner, D. N. (2005). *Institutions and ethnic politics in Africa*. Cambridge University Press.

Ramankutty, N., J. Foley, J. Norman, and K. McSweeney (2002). The global distribution of cultivable lands:current patterns and sensitivity to possible climate change. *Global Ecology and Biogeography 11*(5), 377–392.

Robinson, J. A. and T. Verdier (2013). The political economy of clientelism. *The Scandinavian Journal of Economics 115*(2), 260–291.

Stokes, S. C. (2007). Political clientelism. In *The Oxford handbook of political science*.

Stokes, S. C., T. Dunning, M. Nazareno, and V. Brusco (2013). *Brokers, voters, and clientelism: The puzzle of distributive politics*. Cambridge University Press.

U.S. Geological Survey (1996). GTOPO30. *Sioux Falls, SD: U.S. Geological Survey Center for Earth Resources Observation and Science*.

Young, C. and T. E. Turner (1985). *The rise and decline of the Zairian state*. University of Wisconsin Pres.

Tables and Figures

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	Observations
Panel A: Overall Sample					
Vote Margin (Adjusted)	0.094	0.387	-0.992	1	406592
Elected Local Ethnic Politician	0.587	0.492	0	1	406592
Employed	0.577	0.494	0	1	409604
Employed in Agriculture	0.287	0.452	0	1	409604
Employed in Manufacturing	0.115	0.319	0	1	409604
Employed in Services	0.199	0.4	0	1	409604
Employed in Public Sector	0.017	0.131	0	1	409604
Age	29.049	10.12	15	64	409604
Rural	0.637	0.481	0	1	409604
Female	0.722	0.448	0	1	409604
Completed Primary School	0.334	0.472	0	1	409585
Secondary Education or Higher	0.351	0.477	0	1	409604
Latitude	1.683	11.13	-26.817	16.656	409604
Longitude	11.349	19.497	-17.498	41.877	409604
Distance from Improved Roads (km)	16.78	30.567	0	577.4	409604
Distance from Cities (km)	34.145	29.497	0.078	578.515	409604
Elevation (m)	545.854	552.033	-4	3224.667	409501
Terrain Ruggedness	53.072	74.376	0	1311.546	409604
Agricultural Suitability	0.389	0.213	0	0.987	393254
Malaria Suitability	14.955	10.045	0	37.609	409604
Panel B: Observations Within 20% Right of the Threshold					
Vote Margin (Adjusted)	0.094	0.056	0	0.2	72784
Elected Local Ethnic Politician	1	0	1	1	72784
Employed	0.573	0.495	0	1	72784
Employed in Agriculture	0.289	0.453	0	1	72784
Employed in Manufacturing	0.108	0.311	0	1	72784
Employed in Services	0.201	0.4	0	1	72784
Employed in Public Sector	0.018	0.133	0	1	72784
Age	29.18	10.149	15	64	72784
Rural	0.657	0.475	0	1	72784
Female	0.722	0.448	0	1	72784
Completed Primary School	0.37	0.483	0	1	72779
Secondary Education or Higher	0.352	0.478	0	1	72784
Latitude	-0.586	11.102	-26.046	16.532	72784
Longitude	15.676	18.542	-16.73	41.83	72784
Distance from Improved Roads (km)	20.248	36.621	0	577.4	72784
Distance from Cities (km)	34.442	32.608	0.078	578.515	72784
Elevation (m)	592.29	544.453	-1.25	3224.667	72740
Terrain Ruggedness	63.453	91.787	0	1311.546	72784
Agricultural Suitability	0.421	0.211	0	0.987	69775
Malaria Suitability	14.249	10.173	0	37.609	72784
Panel C: Observations Within 20% Left of the Threshold					
Vote Margin (Adjusted)	-0.097	0.055	-0.2	0	69979
Elected Local Ethnic Politician	0	0	0	0	69979
Employed	0.559	0.497	0	1	69979
Employed in Agriculture	0.262	0.439	0	1	69979
Employed in Manufacturing	0.118	0.323	0	1	69979
Employed in Services	0.199	0.4	0	1	69979
Employed in Public Sector	0.02	0.139	0	1	69979
Age	28.926	10.033	15	64	69979
Rural	0.631	0.483	0	1	69979
Female	0.724	0.447	0	1	69979
Completed Primary School	0.367	0.482	0	1	69976
Secondary Education or Higher	0.364	0.481	0	1	69979
Latitude	0.315	10.717	-25.977	16.532	69979
Longitude	13.799	19.311	-16.723	41.249	69979
Distance from Improved Roads (km)	18.034	31.315	0	316.133	69979
Distance from Cities (km)	33.117	27.365	0.137	162.633	69979
Elevation (m)	614.114	557.529	0	2867.333	69979
Terrain Ruggedness	57.062	75.945	0.5	925.859	69979
Agricultural Suitability	0.38	0.221	0	0.925	67319
Malaria Suitability	13.524	9.759	0	35.064	69979
Panel D: Afrobarometer					
Vote Margin (Adjusted)	0.05	0.344	-0.99	0.993	4899
Elected Local Ethnic Politician	0.553	0.497	0	1	4899
Traditional Leader:					
Should Be Partisan	0.011	0.102	0	1	4922
Allocates Land	0.384	0.486	0	1	4922
Maintains Law and Order	0.071	0.256	0	1	4922
Governs the Community	0.566	0.496	0	1	4922
Age	35.851	14.288	18	99	4830
Female	0.502	0.5	0	1	4922
Rural	0.61	0.488	0	1	4922
Latitude	2.461	9.805	-17.858	14.957	4922
Longitude	8.334	17.484	-17.452	40.11	4922

Notes. The table reports the summary statistics of all variables used in the empirical analysis, described in Section 2 and with additional details in Supplementary Appendix B.2. Panel A shows the summary statistics for the overall final sample, while Panel B and C provide summary statistics for the subsample of observations within an adjusted vote margin of 20% right and left respectively of the vote margin threshold that is relevant for the party connected to the ethnicity the individual belongs to for gaining a local representative from the constituency in the national assembly.

Table 2: Effect on Employment

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0206	0.0242*	0.0258**	0.0278**	0.0240**	0.0155
	(0.0261)	(0.0147)	(0.0129)	(0.0120)	(0.0118)	(0.0109)
$\hat{\beta}_{IK}$	0.0304	0.0291*	0.0268**	0.0253**	0.0221*	0.0150*
	(0.0288)	(0.0160)	(0.0125)	(0.0118)	(0.0114)	(0.0077)
$\hat{\beta}_{CCT}$	0.0384	0.0340*	0.0290**	0.0291**	0.0245*	0.0169*
	(0.0342)	(0.0185)	(0.0146)	(0.0134)	(0.0134)	(0.0092)
Bandwidth	0.193	0.197	0.213	0.257	0.214	0.193
Country-Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	139183	140920	151738	169114	146051	133802

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table 3: Effect on Employment and Affiliation with Government

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
	Non-Government Party			*Goverment Party*		
$\hat{\beta}_{OLS}$	0.0348*	0.0363**	0.0176	0.0300	0.0578**	0.0536**
	(0.0179)	(0.0172)	(0.0161)	(0.0273)	(0.0235)	(0.0218)
$\hat{\beta}_{IK}$	0.0211	0.0184	0.0091	0.0586*	0.0760**	0.0717**
	(0.0165)	(0.0159)	(0.0146)	(0.0353)	(0.0344)	(0.0318)
$\hat{\beta}_{CCT}$	0.0221	0.0196	0.0100	0.0648	0.0853**	0.0798**
	(0.0191)	(0.0184)	(0.0170)	(0.0409)	(0.0398)	(0.0367)
Bandwidth	0.216	0.239	0.211	0.151	0.134	0.138
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	Yes	No	Yes	Yes
Individual Controls	No	No	Yes	No	No	Yes
Observations	77804	84067	74838	30663	27699	28226

Notes. * p-value$<$ s 0.1; ** p-value$<$0.05; *** p-value$<$0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Columns 1 to 3 are estimated using the subsample of parties that do not support the central government in the aftermath of elections. Columns 4 to 6 report the same estimates using the subsample of parties that do not support the central government in the aftermath of elections. Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table 4: Effect on Employment and Pre-Colonial Ethnic Institutions

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
	Low Level of Centralization			*High Level of Centralization*		
$\hat{\beta}_{OLS}$	0.0140	0.0158	0.0068	0.0304	0.0204	0.0276
	(0.0196)	(0.0184)	(0.0174)	(0.0220)	(0.0231)	(0.0207)
$\hat{\beta}_{IK}$	0.0025	0.0024	-0.0036	0.0401	0.0508**	0.0436*
	(0.0185)	(0.0178)	(0.0168)	(0.0245)	(0.0250)	(0.0229)
$\hat{\beta}_{CCT}$	-0.0022	-0.0020	-0.0084	0.0489*	0.0598**	0.0520**
	(0.0211)	(0.0203)	(0.0192)	(0.0275)	(0.0280)	(0.0258)
Bandwidth	0.160	0.162	0.152	0.193	0.183	0.177
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	Yes	No	Yes	Yes
Individual Controls	No	No	Yes	No	No	Yes
Observations	46267	46269	44079	40090	36455	34970

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Columns 1 to 3 are estimated using the subsample of constituencies located in ethnic group homelands with a degree of political centralization equal to 0 or 1. Columns 4 to 6 report the same estimates using the subsample of constituencies located in ethnic group homelands with a degree of political centralization higher than 1. Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table 5: Effect on Employment by Sector

	Agriculture		Manufacturing		Services		Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	0.0239*	0.0197	0.0046	-0.0015	-0.0056	-0.0036	-0.0002	-0.0009
	(0.0132)	(0.0126)	(0.0074)	(0.0073)	(0.0068)	(0.0065)	(0.0025)	(0.0025)
$\hat{\beta}_{IK}$	0.0208**	0.0167**	0.0034	0.0023	-0.0047	-0.0026	-0.0002	-0.0003
	(0.0088)	(0.0079)	(0.0046)	(0.0045)	(0.0046)	(0.0042)	(0.0016)	(0.0015)
$\hat{\beta}_{CCT}$	0.0231**	0.0184**	0.0047	0.0039	-0.0045	-0.0019	-0.0006	-0.0007
	(0.0103)	(0.0093)	(0.0056)	(0.0054)	(0.0056)	(0.0051)	(0.0019)	(0.0019)
Bandwidth	0.193	0.195	0.205	0.221	0.270	0.268	0.237	0.229
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	139144	134337	145754	150565	181891	174806	166843	156162

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working in agriculture (columns 1 and 2), manufacturing (columns 3 and 4), services (columns 5 and 6), and the public sector (columns 7 and 8) $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table 6: Disfavoritism - Effect on Employment by Sector

	Agriculture		Manufacturing		Services		Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	-0.0202	-0.0037	-0.0008	0.0040	0.0240**	0.0201**	0.0001	0.0033
	(0.0132)	(0.0127)	(0.0076)	(0.0081)	(0.0105)	(0.0096)	(0.0038)	(0.0039)
$\hat{\beta}_{IK}$	-0.0138*	-0.0051	-0.0020	-0.0028	0.0125**	0.0073	0.0022	0.0023
	(0.0077)	(0.0077)	(0.0045)	(0.0049)	(0.0057)	(0.0050)	(0.0018)	(0.0018)
$\hat{\beta}_{CCT}$	-0.0161*	-0.0068	-0.0033	-0.0043	0.0142**	0.0078	0.0027	0.0027
	(0.0086)	(0.0087)	(0.0056)	(0.0059)	(0.0069)	(0.0061)	(0.0021)	(0.0021)
Bandwidth	0.164	0.168	0.245	0.233	0.247	0.276	0.266	0.246
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	77887	74270	107106	97218	108011	111145	114980	100482

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2, but considering the vote share of the party that is (the most) under-represented, relative to its overall vote share in the country, among voters from ethnic group e, and using a cutoff of 0.05. The dependent variable is a dummy equal to one if the individual reports to be working in agriculture (columns 1 and 2), manufacturing (columns 3 and 4), services (columns 5 and 6), and the public sector (columns 7 and 8) $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table 7: Effect on Agricultural Employment and Land Tenure System

	Agricultural Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
	Customary Land Tenure Not Recognized			*Customary Land Tenure Recognized*		
$\hat{\beta}_{OLS}$	-0.0287	-0.0106	0.0159	0.0372**	0.0403**	0.0377**
	(0.0206)	(0.0196)	(0.0164)	(0.0155)	(0.0165)	(0.0156)
$\hat{\beta}_{IK}$	0.0072	0.0072	0.0039	0.0291**	0.0295**	0.0228*
	(0.0078)	(0.0072)	(0.0064)	(0.0142)	(0.0147)	(0.0129)
$\hat{\beta}_{CCT}$	0.0048	0.0054	0.0021	0.0326**	0.0331*	0.0256*
	(0.0093)	(0.0084)	(0.0075)	(0.0165)	(0.0173)	(0.0151)
Bandwidth	0.240	0.272	0.291	0.183	0.189	0.181
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	Yes	No	Yes	Yes
Individual Controls	No	No	Yes	No	No	Yes
Observations	75915	81318	86765	72824	70264	68067

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working in agriculture (columns 1 and 2), manufacturing (columns 3 and 4), services (columns 5 and 6), and the public sector (columns 7 and 8) $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Columns 1 to 3 are estimated using the subsample of countries where customary land tenure is not recognized by the national legislation. Columns 4 to 6 report the same estimates using the subsample of countries where customary land tenure is not recognized by the national legislation. Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table 8: Partisanship of Traditional Leader

	Traditional Leader Should be Partisan				
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}_{OLS}$	0.0136**	0.0159**	0.0178**	0.0176**	0.0196***
	(0.0054)	(0.0078)	(0.0081)	(0.0081)	(0.0075)
Constituency FE	No	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes
Individual Controls	No	No	No	No	Yes
Observations	4899	4880	4869	4869	4778
R^2	0.0017	0.0609	0.0916	0.0929	0.0926

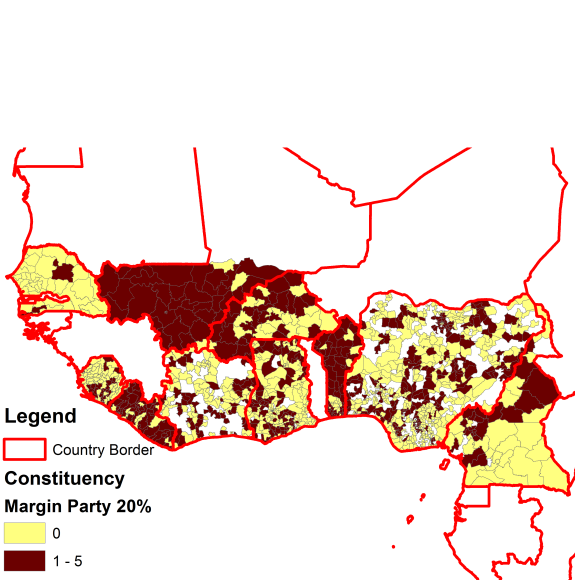
Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the 2008 Afrobarometer. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual affirms that traditional leaders should be partisan. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude and longitude. Individual controls include age, a dummy for respondents in rural areas, and a dummy for female. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table 9: Traditional Leaders and Land Allocation

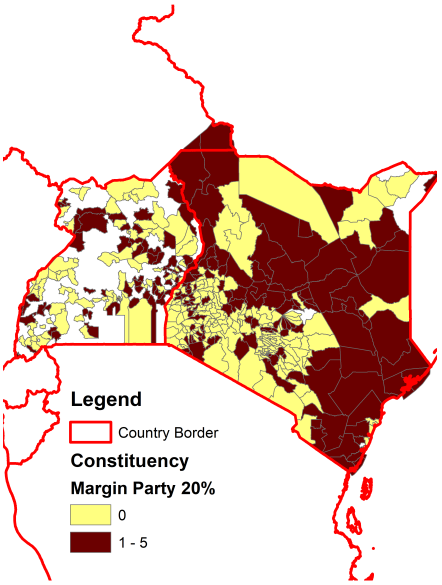
	Primary Responsible for Land Allocation				
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}_{OLS}$	0.0864	0.0936***	0.0652**	0.0653**	0.0584*
	(0.0564)	(0.0247)	(0.0320)	(0.0320)	(0.0313)
Constituency FE	No	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes
Individual Controls	No	No	No	No	Yes
Observations	4899	4880	4869	4869	4778
R^2	0.0082	0.3125	0.3201	0.3201	0.3259

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the 2008 Afrobarometer. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual indicates that the traditional leader is mainly responsible for the allocation of land. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude and longitude. Individual controls include age, a dummy for respondents in rural areas, and a dummy for female. Standard errors are clustered at the level of ethnic group, constituency and election year.

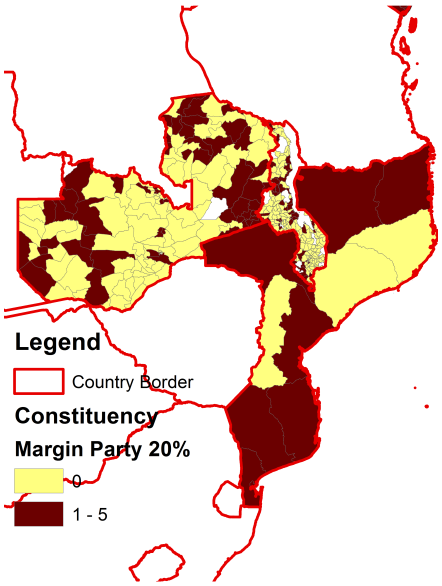
Figure 1: Spatial Distribution of Contested Constituencies



(a) West Africa



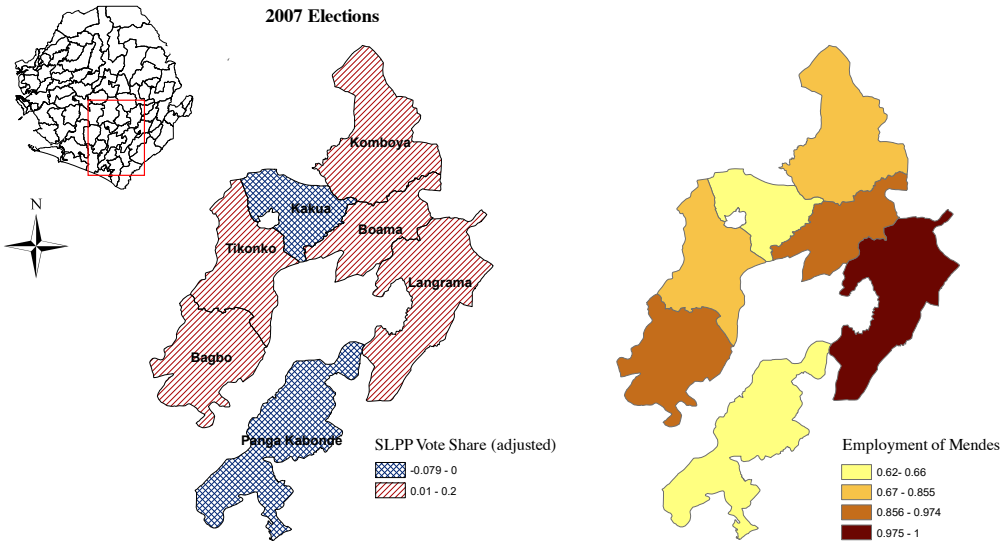
(b) East Africa



(c) Southern Africa

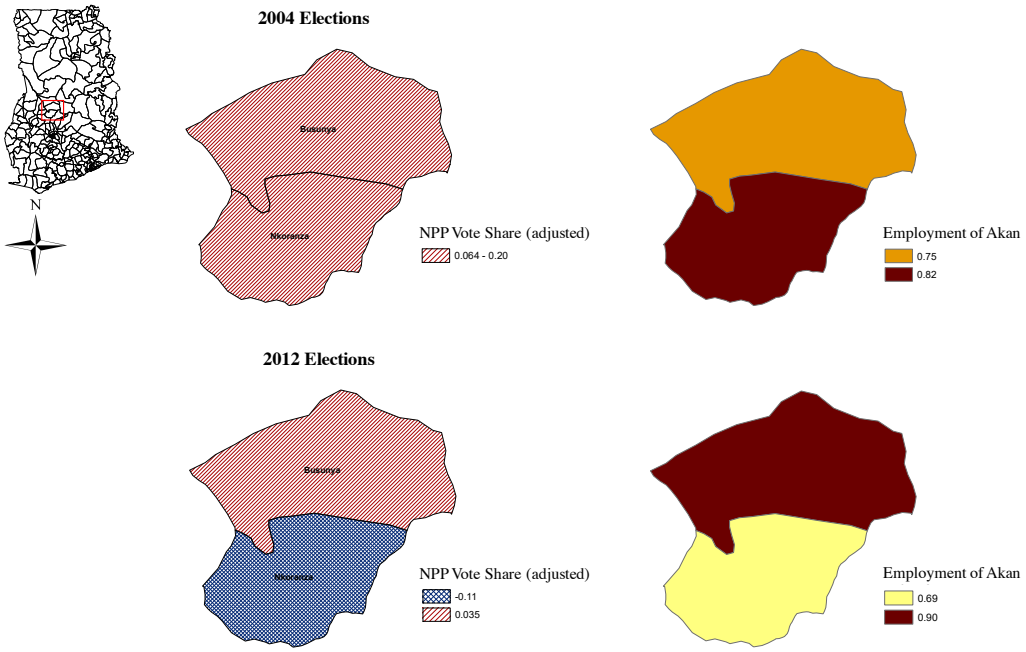
Notes. The figures show the spatial distribution of constituency-year observations for West Africa (a), East Africa (b), and Southern Africa (b). We highlighted in brown those constituency-years where two parties or more score a vote margin of less of 20% percent from the relevant threshold for winning a seat in parliament.

Figure 2: Example in Sierra Leone – 2007 Elections



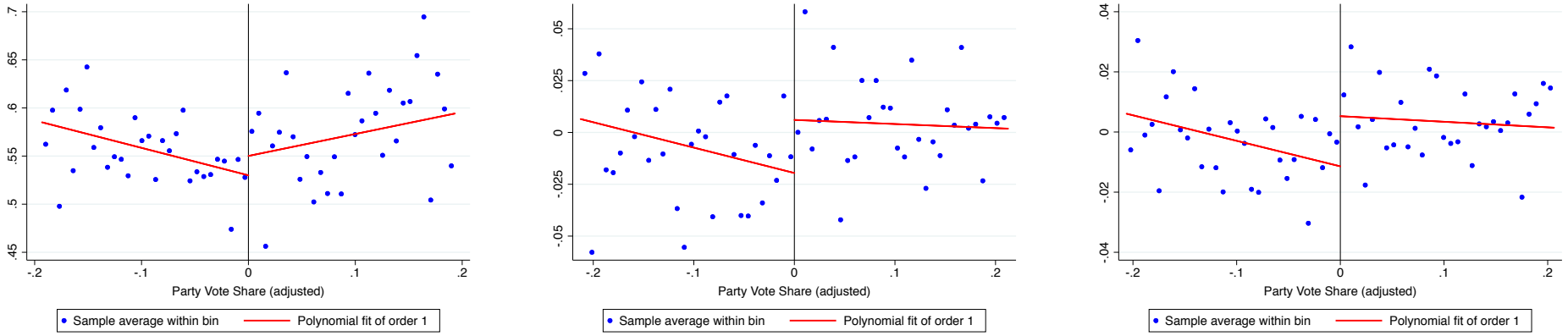
Notes. The figure shows a cluster of constituencies in Sierra Leone where the Sierra Leone People's Party (SLPP) gained or did not gain a local representative in the national assembly in the 2007 Elections. The left map reports the adjusted vote share of the SLPP across these constituencies. The right map reports the employment rate in 2008 of individuals belonging to the Mende ethnic group, which the labelling algorithm identifies as linked to the SLPP. The two districts where in 2007 the SLPP lost by a narrow margin – Kakua and Panga Kabonde – are those where in 2008 the employment rate of Mendes is lowest.

Figure 3: Example in Ghana – 2004 and 2012 Elections



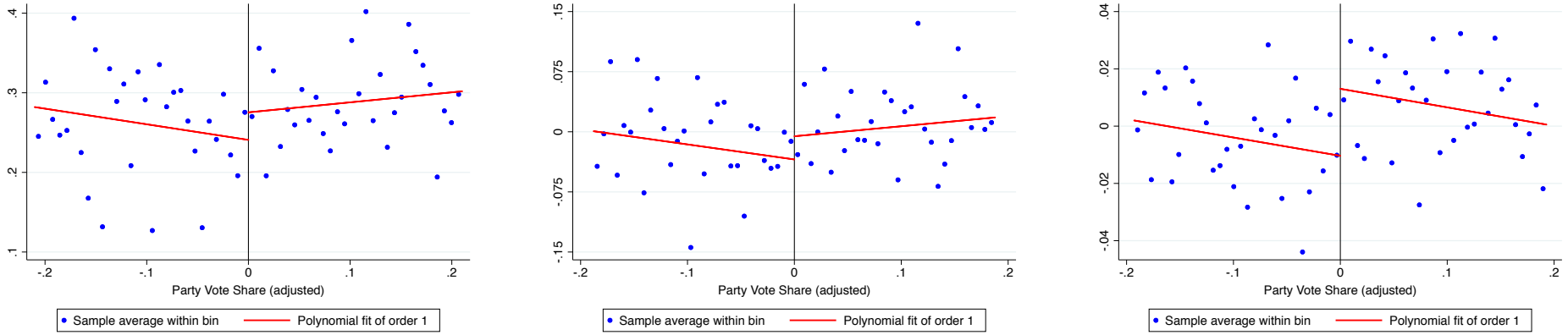
Notes. The figure shows a cluster of constituencies in Ghana where the New Patriotic Party (NPP) gained or did not gain a local representative in the national assembly in the 2004 and 2012 Elections. The left maps report the adjusted vote share of the NPP across these constituencies. The right map reports the employment rate in 2008 and 2014 of individuals belonging to the Akan ethnic group, which the labelling algorithm identifies as linked to the NPP. The employment rate of Akans decreases in the district that the NPP lost in 2012 – Nkoranza – while increasing in the district that the party retained – Busunya.

Figure 4: Effect on Employment



Notes. The figures plot the relationship between individual employment probabilities as derived from the DHS and electoral outcomes. The figures provide a graphical representation of the local linear regression fit on both sides of the threshold that determines whether the ethnicity the individual belongs to is linked to a party that gains a local representative in the national assembly. It also plots a scatterplot showing employment probabilities within 30 bins right and left of the threshold. The first graph shows unconditional probabilities, the second shows residual probabilities net of country-year and ethnicity fixed effects. The third also nets out constituency fixed effects.

Figure 5: Effect on Agricultural Employment



Notes. The figures plot the relationship between individual probabilities of employment in the agricultural sector as derived from the DHS and electoral outcomes. The figures provide a graphical representation of the local linear regression fit on both sides of the threshold that determines whether the ethnicity the individual belongs to is linked to a party that gains a local representative in the national assembly. It also plots a scatterplot showing employment probabilities within 30 bins right and left of the threshold. The first graph shows unconditional probabilities, the second shows residual probabilities net of country-year and ethnicity fixed effects. The third also nets out constituency fixed effects.

A Appendix

A.1 Additional Tables and Figures

Table A.1: Constituencies and Contestation

Countries Election Years	Constituency Observations	Merged with DHS	Any Party Within $\leq 20\%$	Any Party Within $\leq 10\%$	Any Party Within $\leq 5\%$
Benin 1999, 2011	48	44	28 (0.64)	21 (0.48)	9 (0.20)
Burkina Faso 2002, 2007, 2012	104	102	54 (0.53)	35 (0.34)	23 (0.23)
Cameroon 2002, 2007	150	81	32 (0.40)	13 (0.16)	4 (0.05)
Ghana 1996, 2000, 2004, 2012	905	699	352 (0.50)	187 (0.27)	96 (0.14)
Ivory Coast 2011	185	129	46 (0.36)	22 (0.17)	11 (0.09)
Kenya 2002, 2007, 2013	706	414	185 (0.45)	107 (0.26)	62 (0.15)
Liberia 2005, 2011	137	130	101 (0.78)	61 (0.47)	33 (0.25)
Malawi 1999, 2004, 2009, 2014	759	567	256 (0.45)	136 (0.24)	64 (0.11)
Mali 2013	54	42	39 (0.93)	35 (0.83)	23 (0.55)
Mozambique 2009	22	11	8 (0.72)	6 (0.55)	4 (0.36)
Nigeria 2007, 2011	715	359	139 (0.39)	71 (0.20)	33 (0.09)
Senegal 2007, 2012	81	70	11 (0.16)	8 (0.11)	1 (0.01)
Sierra Leone 2007, 2012	222	215	44 (0.20)	21 (0.10)	14 (0.07)
Uganda 2011	335	132	66 (0.50)	39 (0.30)	18 (0.14)
Zambia 2006, 2011	298	259	95 (0.37)	48 (0.19)	24 (0.09)
Total	4721	3254	1456 (0.45)	810 (0.25)	419 (0.13)

Notes. The first column shows the set of countries and election years in our sample, for which geo-referenced information on election constituencies is available. The second column shows the total number of constituency-level observations available for each country in total across the different election years. The third column shows the number of constituency-year observations in the election dataset that we match with individual-level observations and are therefore relevant for our analysis. The last three columns show the number of constituency-year observations in the election dataset that are relevant for identification, meaning for which we observe any party with a vote margin within 20%, 10%, and 5% respectively from the relevant vote margin threshold. The number in parenthesis indicates which fractions of matched constituency-year observations (second column) these represent.

Table A.2: Test of Balance - Village Characteristics

	Latitude		Longitude		Dist. to Roads		Dist. to Cities		Elevation		Ruggedness		Suitability		Malaria	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
$\hat{\beta}_{OLS}$	0.0911	-0.0073	-0.0587	-0.0037	1.8011	0.6998	4.6194*	-0.3256	-61.4392**	-11.3130**	-10.1735**	-3.8886	0.0134	-0.0048	0.3173	0.0369
	(0.1052)	(0.0096)	(0.1076)	(0.0119)	(1.8100)	(0.5164)	(2.5707)	(0.5166)	(24.3887)	(4.7363)	(5.0727)	(2.7642)	(0.0152)	(0.0040)	(0.4030)	(0.1126)
$\hat{\beta}_{IK}$	0.0487	0.0079	-0.0650	0.0162	3.1179	0.2487	2.3289	-0.3478	-44.2127	-3.1174	-2.2318	-0.6687	0.0092	-0.0035	0.1212	-0.0441
	(0.1259)	(0.0113)	(0.1019)	(0.0124)	(2.0972)	(0.3986)	(2.6978)	(0.3773)	(26.0370)	(2.9885)	(4.6617)	(1.6780)	(0.0141)	(0.0023)	(0.4217)	(0.0830)
$\hat{\beta}_{CCT}$	0.0273	0.0086	-0.0816	0.0167	3.6599	0.3101	2.3605	-0.3200	-49.2146	-3.2380	-2.2256	-0.2940	0.0096	-0.0040	0.1715	-0.0565
	(0.1562)	(0.0132)	(0.1181)	(0.0144)	(2.5551)	(0.4836)	(3.2344)	(0.4580)	(31.0622)	(3.3127)	(5.4860)	(1.9139)	(0.0160)	(0.0026)	(0.4711)	(0.0950)
Bandwidth	0.238	0.189	0.221	0.168	0.306	0.171	0.231	0.293	0.219	0.206	0.146	0.156	0.260	0.209	0.218	0.205
Country/Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	167269	136975	156708	123589	201817	125435	163136	194741	155567	147091	110363	116494	169939	143776	154189	145754

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable each village-level covariate as indicated in the column header. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.3: Test of Balance - Individual Characteristics

	Age		Rural		Female		Primary School		Secondary School or Higher	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\hat{\beta}_{OLS}$	0.1488	0.0066	0.0361	-0.0001	0.0042	0.0022	-0.0008	-0.0023	-0.0267	-0.0005
	(0.1355)	(0.1203)	(0.0367)	(0.0141)	(0.0069)	(0.0063)	(0.0103)	(0.0072)	(0.0181)	(0.0101)
$\hat{\beta}_{IK}$	0.2100	0.0436	0.0510	0.0125	0.0119	0.0027	0.0005	0.0002	-0.0245	-0.0078
	(0.1564)	(0.0771)	(0.0409)	(0.0091)	(0.0076)	(0.0042)	(0.0110)	(0.0048)	(0.0196)	(0.0064)
$\hat{\beta}_{CCT}$	0.2030	0.0353	0.0614	0.0153	0.0146	0.0032	-0.0003	0.0006	-0.0282	-0.0085
	(0.1899)	(0.0937)	(0.0469)	(0.0107)	(0.0086)	(0.0051)	(0.0131)	(0.0057)	(0.0231)	(0.0077)
Bandwidth	0.219	0.309	0.183	0.274	0.139	0.235	0.256	0.260	0.188	0.211
Country/Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	155435	203764	134049	184554	106124	165737	175172	176569	136892	151348

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable each village-level covariate as indicated in the column header. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.4: Correlates of Political Competition

	Number of Parties					HHI
	(1)	(2)	(3)	(4)	(5)	(6)
Fractionalization	0.8131***		1.3793***	1.2005***	0.0932	-0.0947***
	(0.2149)		(0.4265)	(0.2867)	(0.3144)	(0.0314)
Polarization		0.4724**	-0.5774	-0.4015	0.0925	0.0243
		(0.1896)	(0.3757)	(0.2519)	(0.2540)	(0.0254)
Share of Employed					0.5562	-0.0938**
					(0.3830)	(0.0383)
Share of Agric.					-1.0275***	0.0536
					(0.3454)	(0.0345)
Share of Services					-2.3737***	0.0372
					(0.4159)	(0.0415)
Population					0.0012***	-0.0000
					(0.0004)	(0.0000)
Share of Rural Pop.					-0.3881***	-0.0217
					(0.1447)	(0.0145)
Primary School					1.0297***	-0.0120
					(0.2784)	(0.0278)
Secondary School					1.8123***	-0.0288
					(0.2509)	(0.0251)
Country FE	No	No	No	Yes	Yes	Yes
Observations	2293	2293	2293	2293	2293	2293
R^2	0.0062	0.0027	0.0072	0.5908	0.6171	0.2341

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is the electoral constituency in an election year. In columns 1 to 5, the dependent variable is the number of parties running in the constituency. In column 6, the dependent variable is a Herfindahl-type index of political competition calculated using vote shares by party. The independent variables included are derived from the individual-level DHS observations in the full sample, keeping only those belonging to the year that is after and closest to the election year.

Table A.5: Effect on Employment - Restricted Placebo Sample

	Employment Dummy				
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}_{OLS}$	0.0612	0.0641**	0.0446**	0.0406*	0.0275
	(0.0410)	(0.0275)	(0.0224)	(0.0244)	(0.0237)
$\hat{\beta}_{IK}$	0.0729	0.0696**	0.0465**	0.0518**	0.0426*
	(0.0465)	(0.0331)	(0.0226)	(0.0251)	(0.0250)
$\hat{\beta}_{CCT}$	0.0759	0.0729*	0.0457*	0.0509*	0.0429
	(0.0546)	(0.0400)	(0.0262)	(0.0290)	(0.0291)
Bandwidth	0.206	0.239	0.254	0.202	0.186
Country-Year FE	No	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes
Individual Controls	No	No	No	No	Yes
Observations	36359	41830	43187	34138	32038

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. The unit of observation is an individual as surveyed in the DHS. Sample is restricted to those observations used in the placebo subsample, i.e. for whom it is possible to locate the adjusted vote share of the same party in the same constituency in the next election in our dataset. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.6: Effect on Employment - Placebo Estimates

	Employment Dummy				
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}_{OLS}$	0.0037	0.0118	0.0155	-0.0048	-0.0091
	(0.0458)	(0.0229)	(0.0223)	(0.0224)	(0.0203)
$\hat{\beta}_{IK}$	0.0385	-0.0058	0.0230	0.0126	0.0020
	(0.0419)	(0.0232)	(0.0227)	(0.0242)	(0.0223)
$\hat{\beta}_{CCT}$	0.0485	-0.0057	0.0260	0.0160	0.0032
	(0.0488)	(0.0259)	(0.0260)	(0.0278)	(0.0256)
Bandwidth	0.152	0.268	0.173	0.198	0.209
Country-Year FE	No	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes
Individual Controls	No	No	No	No	Yes
Observations	26342	43395	29797	32566	34127

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. The placebo adjusted vote share is equal to the one obtained by the same party in the same constituency in the next election in our dataset. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.7: Disfavoritism - Effect on Employment

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0400	0.0366	0.0078	0.0047	0.0106	-0.0049
	(0.0301)	(0.0233)	(0.0176)	(0.0176)	(0.0162)	(0.0105)
$\hat{\beta}_{IK}$	0.0334	0.0305	0.0075	0.0093	0.0101	0.0009
	(0.0344)	(0.0283)	(0.0195)	(0.0194)	(0.0187)	(0.0058)
$\hat{\beta}_{CCT}$	0.0307	0.0337	0.0029	0.0060	0.0056	-0.0010
	(0.0410)	(0.0332)	(0.0218)	(0.0217)	(0.0210)	(0.0065)
Bandwidth	0.268	0.216	0.236	0.233	0.245	0.214
Country-Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	137297	116614	124966	116590	120503	109090

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2, but considering the vote share of the party that is (the most) under-represented, relative to its overall vote share in the country, among voters from ethnic group e. The dependent variable is a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.8: Disfavoritism - Effect on Employment by Sector

	Agriculture		Manufacturing		Services		Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	-0.0131	-0.0078	0.0037	0.0098	0.0145	0.0066	0.0008	0.0025
	(0.0124)	(0.0129)	(0.0074)	(0.0076)	(0.0096)	(0.0088)	(0.0028)	(0.0028)
$\hat{\beta}_{IK}$	-0.0060	-0.0007	-0.0033	-0.0035	0.0102**	0.0053	0.0015	0.0013
	(0.0063)	(0.0060)	(0.0043)	(0.0046)	(0.0050)	(0.0044)	(0.0014)	(0.0013)
$\hat{\beta}_{CCT}$	-0.0064	-0.0008	-0.0050	-0.0054	0.0117*	0.0061	0.0019	0.0016
	(0.0073)	(0.0070)	(0.0050)	(0.0054)	(0.0060)	(0.0053)	(0.0016)	(0.0015)
Bandwidth	0.244	0.223	0.201	0.191	0.222	0.233	0.276	0.283
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	127366	111802	109817	98481	118449	116564	140933	136129

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2, but considering the vote share of the party that is (the most) under-represented, relative to its overall vote share in the country, among voters from ethnic group e. The dependent variable is a dummy equal to one if the individual reports to be working in agriculture (columns 1 and 2), manufacturing (columns 3 and 4), services (columns 5 and 6), and the public sector (columns 7 and 8) $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.9: Cropland

	Share of Cropland				
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}_{OLS}$	0.0970***	0.0304	0.0185	0.0244	-0.0195
	(0.0262)	(0.0208)	(0.0203)	(0.0213)	(0.0158)
Country-Year FE	No	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	Yes
Observations	10629	10629	10134	10116	9533
R^2	0.0056	0.2308	0.3079	0.3649	0.6297

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is a DHS cluster. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable the share of land allocated to cropland within a 1km buffer area around the DHS cluster. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). Ethnicity and thus links with parties are identified based on the highest ethnic group population share in each village. Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.10: Role of Traditional Leader

	Should Be Partisan		Allocates Land		Maintains Law and Order		Governs the Community	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	0.0178**	0.0196***	0.0652**	0.0584*	0.0101	0.0146	-0.0607	-0.0718*
	(0.0081)	(0.0075)	(0.0320)	(0.0313)	(0.0198)	(0.0202)	(0.0432)	(0.0428)
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Ethnicity FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4869	4778	4869	4778	4869	4778	4869	4778
R^2	0.0916	0.0926	0.3201	0.3259	0.1396	0.1452	0.1970	0.2022

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the 2008 Afrobarometer. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual indicates that the traditional leader should be partisan (columns 1 and 2), is the main responsible for the allocation of land (columns 3 and 4), maintains law and order (columns 5 and 6), and governs the community (columns 7 and 8). $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude and longitude. Individual controls include age, a dummy for respondents in rural areas, and a dummy for female. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.11: Effect on Employment - Overrepresentation cutoff at 0.05

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{IK}$	0.0220	0.0330**	0.0261**	0.0262 **	0.0223**	0.0133*
	(0.0254)	(0.0145)	(0.0122)	(0.0117)	(0.0112)	(0.0077)
$\hat{\beta}_{CCT}$	0.0244	0.0360**	0.0280*	0.0301**	0.0245*	0.0148
	(0.0308)	(0.0173)	(0.0144)	(0.0136)	(0.0133)	(0.0093)
Bandwidth	0.231	0.229	0.236	0.274	0.240	0.227
Country-Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	145553	145227	147947	156746	143112	136874

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the overrepresentation method discussed in Section 2 and using a cutoff of 0.05. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.12: Effect on Employment and Pre-Colonial Institutions - Overrepresentation Cutoff at 0.05

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
	Low Level of Centralization			*High Level of Centralization*		
$\hat{\beta}_{OLS}$	0.0151	0.0129	0.0107	0.0131	0.0022	0.0060
	(0.0217)	(0.0204)	(0.0189)	(0.0223)	(0.0233)	(0.0222)
$\hat{\beta}_{IK}$	-0.0018	0.0023	-0.0010	0.0168	0.0252	0.0174
	(0.0203)	(0.0192)	(0.0176)	(0.0245)	(0.0245)	(0.0227)
$\hat{\beta}_{CCT}$	-0.0090	-0.0044	-0.0067	0.0225	0.0309	0.0232
	(0.0229)	(0.0217)	(0.0201)	(0.0284)	(0.0280)	(0.0266)
Bandwidth	0.142	0.145	0.147	0.188	0.182	0.185
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	Yes	No	Yes	Yes
Individual Controls	No	No	Yes	No	No	Yes
Observations	38721	38942	38949	32226	29232	30053

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the overrepresentation method discussed in Section 2 and using a cutoff of 0.05. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Columns 1 to 3 are estimated using the subsample of constituencies located in ethnic group homelands with a degree of political centralization equal to 0 or 1. Columns 4 to 6 report the same estimates using the subsample of constituencies located in ethnic group homelands with a degree of political centralization higher than 1. Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.13: Effect on Employment by Sector - Overrepresentation Cutoff at 0.05

	Agriculture		Manufacturing		Services		Public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}_{OLS}$	0.0274**	0.0215	0.0069	0.0027	-0.0075	-0.0093	0.0005	-0.0016
	(0.0134)	(0.0132)	(0.0078)	(0.0079)	(0.0076)	(0.0070)	(0.0029)	(0.0028)
$\hat{\beta}_{IK}$	0.0164*	0.0132	0.0043	0.0043	-0.0038	-0.0025	-0.0003	-0.0006
	(0.0093)	(0.0085)	(0.0050)	(0.0049)	(0.0049)	(0.0044)	(0.0018)	(0.0017)
$\hat{\beta}_{CCT}$	0.0186*	0.0152	0.0052	0.0053	-0.0041	-0.0021	-0.0008	-0.0011
	(0.0110)	(0.0101)	(0.0061)	(0.0061)	(0.0060)	(0.0054)	(0.0022)	(0.0021)
Bandwidth	0.194	0.195	0.210	0.221	0.269	0.280	0.234	0.228
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	No	Yes	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	123361	118308	133403	133342	161071	160772	147336	138094

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the DHS. Links between ethnic groups and political parties are derived by implementing the overrepresentation method discussed in Section 2 and using a cutoff of 0.05. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual reports to be working in agriculture (columns 1 and 2), manufacturing (columns 3 and 4), services (columns 5 and 6), and the public sector (columns 7 and 8) $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.14: Disfavoritism - Effect on Employment - Underrepresentation Cutoff at 0.05

	Employment Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}_{OLS}$	0.0402	0.0107	0.0128	0.0098	0.0144	-0.0014
	(0.0325)	(0.0192)	(0.0210)	(0.0211)	(0.0200)	(0.0123)
$\hat{\beta}_{IK}$	0.0430	0.0135	0.0070	0.0078	0.0095	0.0020
	(0.0321)	(0.0222)	(0.0231)	(0.0235)	(0.0233)	(0.0066)
$\hat{\beta}_{CCT}$	0.0416	0.0134	0.0048	0.0065	0.0078	0.0006
	(0.0360)	(0.0251)	(0.0259)	(0.0262)	(0.0261)	(0.0076)
Bandwidth	0.189	0.239	0.214	0.209	0.212	0.239
Country-Year FE	No	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	No	Yes
Observations	86897	105893	96972	89295	90041	99249

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the DHS. The table reports the estimate of β that we obtain when estimating equation 2, but considering the vote share of the party that is (the most) under-represented, relative to its overall vote share in the country, among voters from ethnic group e, and using a cutoff of 0.05. The dependent variable is a dummy equal to one if the individual reports to be working. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). The estimation bandwidth is the one obtained using the selector proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Individual controls include age, a dummy for respondents in rural areas, a dummy for female, a dummy for whether the individual completed primary school, a dummy for secondary school or higher. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.15: Cropland - Overrepresentation cutoff at 0.05

	Share of Cropland				
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}_{OLS}$	0.1234***	0.0483**	0.0269	0.0186	-0.0198
	(0.0250)	(0.0204)	(0.0190)	(0.0183)	(0.0160)
Country-Year FE	No	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes
Constituency FE	No	No	No	No	Yes
Observations	9638	9638	9151	9135	8567
R^2	0.0105	0.2367	0.3215	0.3741	0.6430

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is a DHS cluster. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable the share of land allocated to cropland within a 1km buffer area around the DHS cluster. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). Ethnicity and thus links with parties are identified based on the highest ethnic group population share in each village. Links between ethnic groups and political parties are derived by implementing the overrepresentation method discussed in Section 2 and using a cutoff of 0.05. Village controls include latitude, longitude, distance from improved roads, distance from the closest urban centre with population of 50,000 or more, elevation, terrain ruggedness, agricultural suitability, and malaria suitability. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.16: Partisanship of Traditional Leader - Overrepresentation Cutoff at 0.05

	Traditional Leader Should Be Partisan				
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}_{OLS}$	0.0142**	0.0199**	0.0254***	0.0260***	0.0273***
	(0.0058)	(0.0084)	(0.0074)	(0.0074)	(0.0069)
Constituency FE	No	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes
Individual Controls	No	No	No	No	Yes
Observations	4268	4247	4234	4234	4173
R^2	0.0018	0.0693	0.0992	0.1001	0.1002

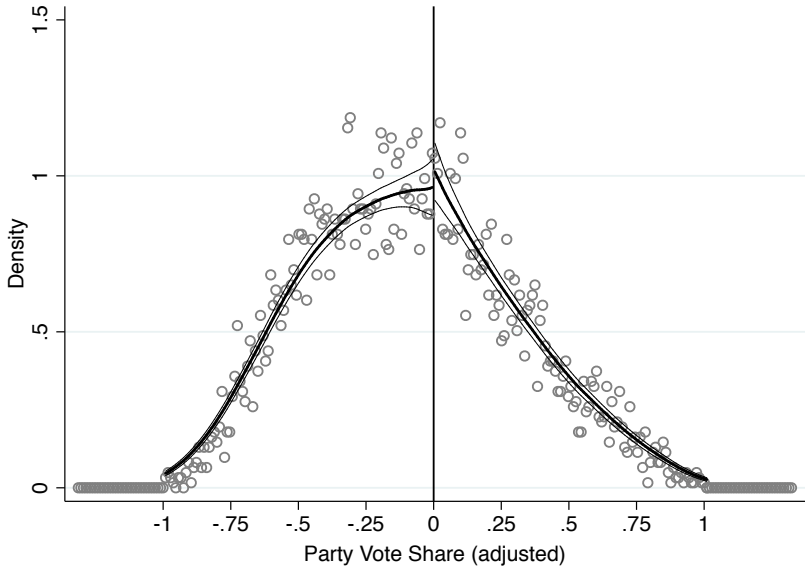
Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the 2008 Afrobarometer. Links between ethnic groups and political parties are derived by implementing the overrepresentation method discussed in Section 2 and using a cutoff of 0.05. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual affirms that traditional leaders should be partisan. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude and longitude. Individual controls include age, a dummy for respondents in rural areas, and a dummy for female. Standard errors are clustered at the level of ethnic group, constituency and election year.

Table A.17: Traditional Leaders and Land Allocation - Overrepresentation Cutoff at 0.05

	Primary Responsible for Land Allocation				
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}_{OLS}$	0.0736	0.0850***	0.0538*	0.0524	0.0497
	(0.0591)	(0.0279)	(0.0323)	(0.0323)	(0.0318)
Constituency FE	No	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	Yes
Village Controls	No	No	No	Yes	Yes
Individual Controls	No	No	No	No	Yes
Observations	4268	4247	4234	4234	4173
R^2	0.0227	0.3334	0.3412	0.3413	0.3473

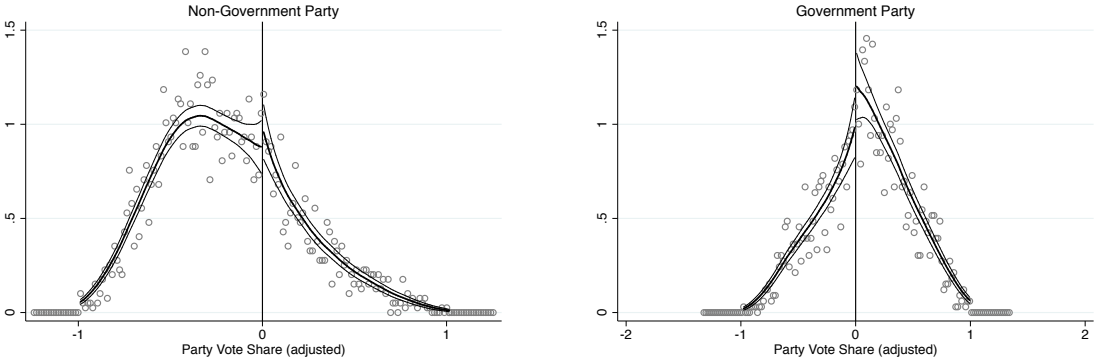
Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is an individual as surveyed in the 2008 Afrobarometer. Links between ethnic groups and political parties are derived by implementing the overrepresentation method discussed in Section 2 and using a cutoff of 0.05. The table reports the estimate of β that we obtain when estimating equation 2 and having as dependent variable a dummy equal to one if the individual indicates that the traditional leader is mainly responsible for the allocation of land. $\hat{\beta}_{OLS}$ is obtained using OLS, $\hat{\beta}_{IK}$ is obtained using the estimator proposed by Imbens and Kalyanaraman (2012), and $\hat{\beta}_{CCT}$ is obtained using the estimator proposed by Calonico, Cattaneo, and Titiunik (2014). Village controls include latitude and longitude. Individual controls include age, a dummy for respondents in rural areas, and a dummy for female. Standard errors are clustered at the level of ethnic group, constituency and election year.

Figure A.1: Density Plot - McCrary Test

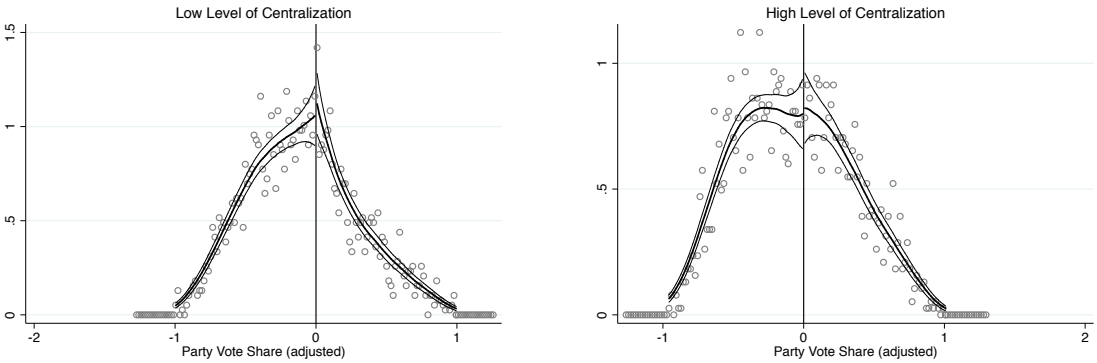


Notes. The figure shows the distribution of the density function of the running variable as estimated separately on both sides of the threshold, together with 95% confidence intervals. The p-value from a test of equality of the value of the density function on the left and right side of the threshold is equal to 0.4761 using the test by McCrary (2008), and equal to 0.4263 using the test by Cattaneo, Janssonz, and Ma (2018). We therefore cannot reject the hypothesis of no discontinuity in the density of the running variable at the threshold.

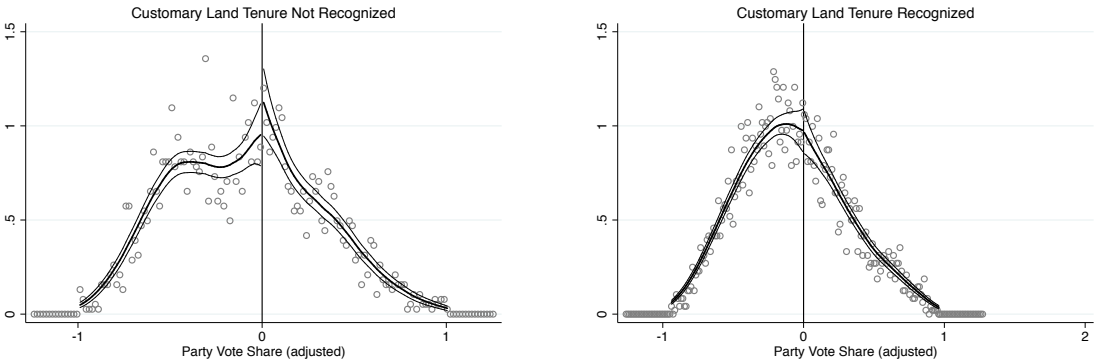
Figure A.2: Density Plot - McCrary Test by Subsample



Notes. The figures show the distribution of the density function of the running variable as estimated separately on both sides of the threshold, together with 95% confidence intervals. The two are estimated separately in the subsample of parties that do not and do support the central government in the aftermath of elections respectively. For the non-government party subsample, the p-value from a test of equality of the value of the density function on the left and right side of the threshold is equal to 0.5281 using the test by McCrary (2008), and equal to 0.0725 using the test by Cattaneo, Janssonz, and Ma (2018). For the government party subsample, the p-value is equal to 0.0180 and 0.4833 respectively.

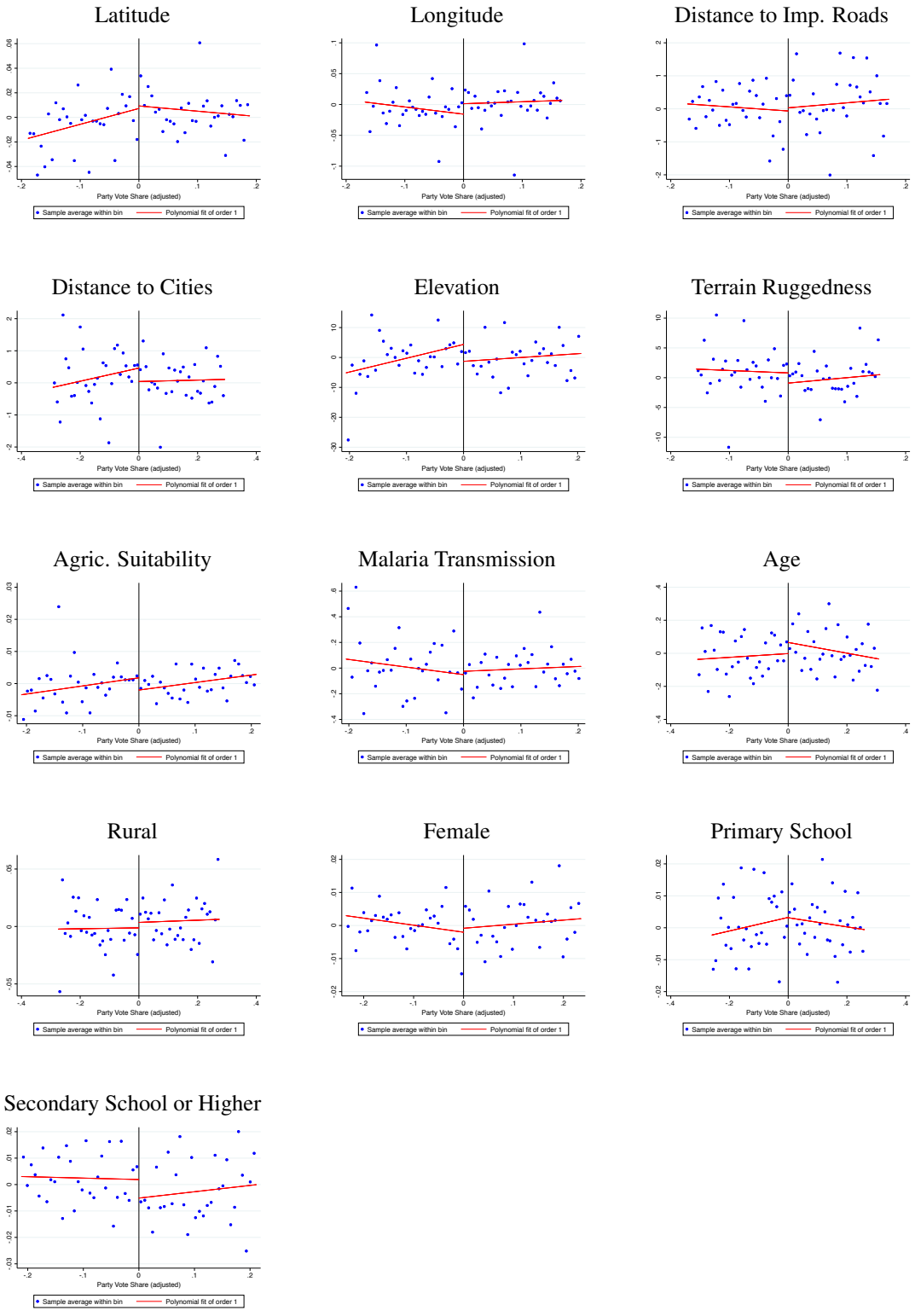


Notes. The figures show the distribution of the density function of the running variable as estimated separately on both sides of the threshold, together with 95% confidence intervals. The two are estimated separately in the subsample of constituencies that belong to ethnic homelands with low and high level of centralization respectively. For the low centralization subsample, the p-value from a test of equality of the value of the density function on the left and right side of the threshold is equal to 0.7448 using the test by McCrary (2008), and equal to 0.7699 using the test by Cattaneo, Janssonz, and Ma (2018). For the high centralization subsample, the p-value is equal to 0.7154 and 0.9088 respectively.



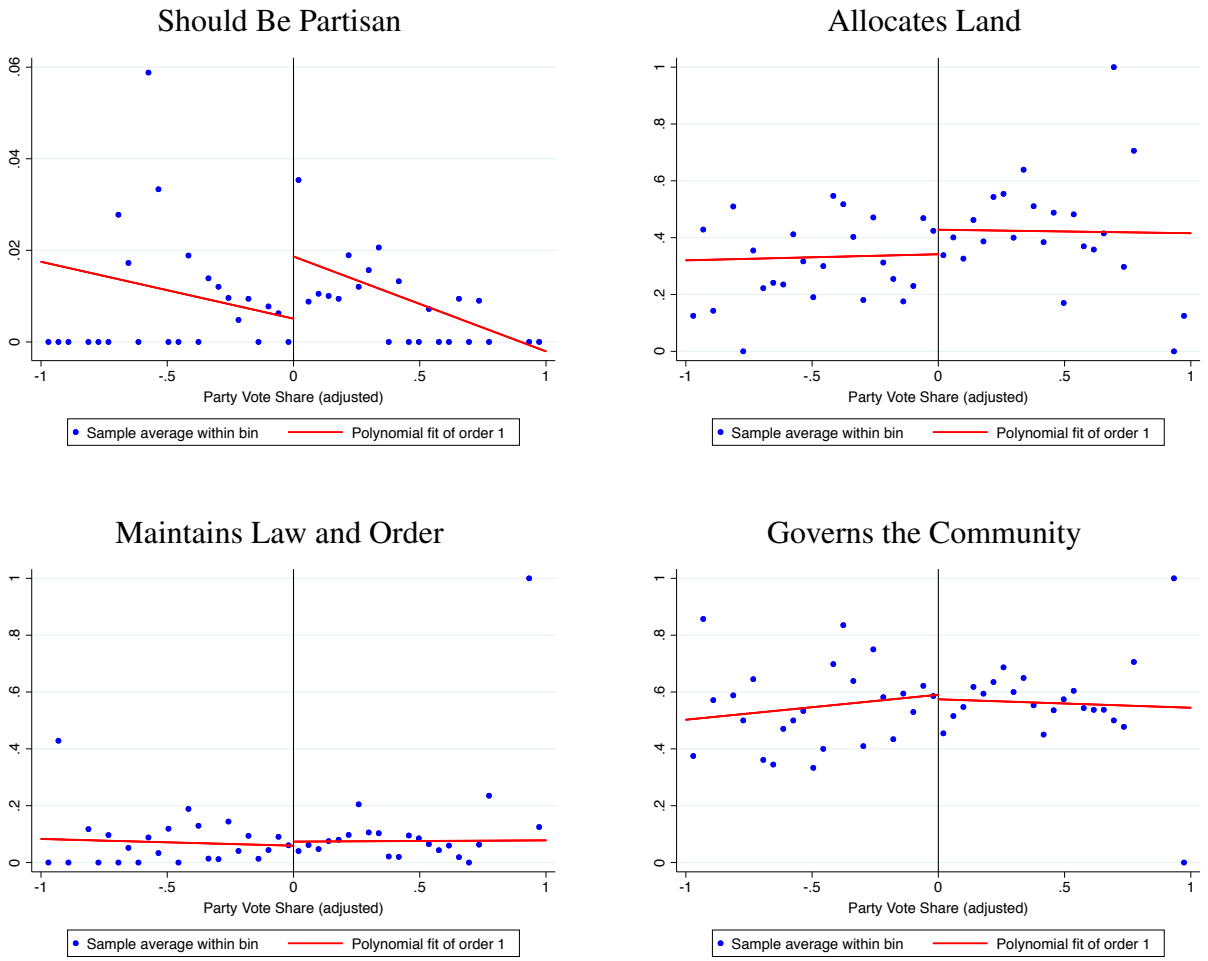
Notes. The figures show the distribution of the density function of the running variable as estimated separately on both sides of the threshold, together with 95% confidence intervals. The two are estimated separately in the subsample of countries where customary land tenure is not and is officially recognized by the national legislation respectively. For the first subsample, the p-value from a test of equality of the value of the density function on the left and right side of the threshold is equal to 0.4460 using the test by McCrary (2008), and equal to 0.6779 using the test by Cattaneo, Janssonz, and Ma (2018). For the second subsample, the p-value is equal to 0.7892 and 0.4972 respectively.

Figure A.3: Balancedness at the Threshold



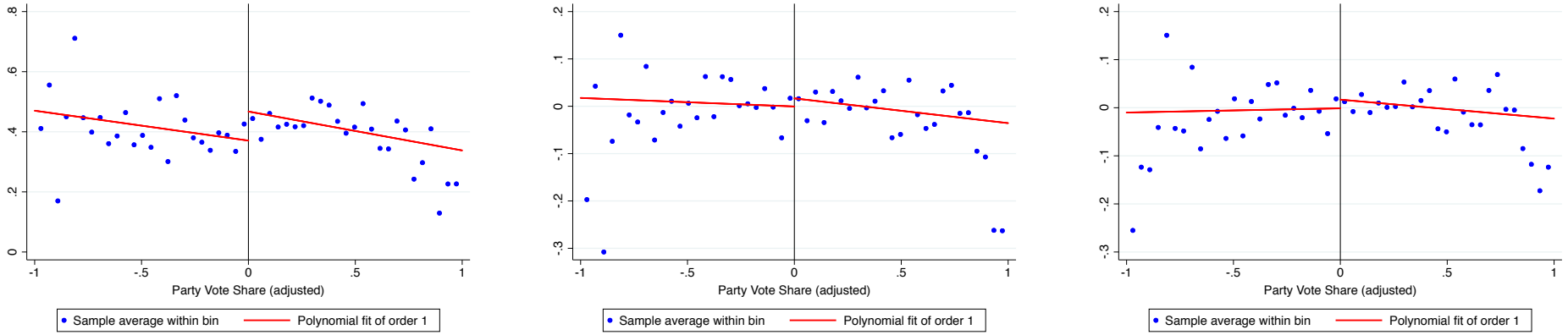
Notes. The figures provide a graphical representation of the local linear regression fit on both sides of the threshold. It also plots a scatterplot showing the average value of each residual covariate net of country-year, ethnicity, and constituency fixed effects in 30 bins right and left of the threshold. The figures show the absence of any meaningful discontinuities in the value of covariates.

Figure A.4: Role of Traditional Leader



Notes. The figures plot the relationship between a set of individual reported beliefs about the role of traditional leaders and electoral outcomes. These variables are obtained from the 2008 Afrobarometer survey, and are dummies equal to one if the individual indicates that the traditional leader should be partisan, is mainly responsible for the allocation of land, maintains law and order, and governs the community. The figures provide a graphical representation of the local linear regression fit on both sides of the threshold that determines whether the ethnicity the individual belongs to is linked to a party that gains a local representative in the national assembly. It also plots a scatterplot showing the unconditional average value of each dummy variable in 25 bins right and left of the threshold.

Figure A.5: Share of Cropland



Notes. The figures plot the relationship between the share of cropland within a 1km buffer area around the DHS cluster and electoral outcomes. The figures provide a graphical representation of the local linear regression fit on both sides of the threshold that determines whether the ethnicity the majority of individuals in the cluster belong to is linked to a party that gains a local representative in the national assembly. It also plots a scatterplot showing the average value of each dummy variable in 30 bins right and left of the threshold. The first graph shows unconditional averages, the second shows average values of residuals net of country-year fixed effects. The third also nets out ethnicity fixed effects.

B Supplementary Appendix

B.1 Details on the labelling algorithm

Political parties may function as the representatives of special interests, such as as ethnic groups. Which interest groups are represented by which party is typically well understood by participants of any given political system, but assigning a "group label" to a party still involves a certain amount of subjectivity. To overcome this challenge and tie our hands, we use survey-data that has information on individuals' ethnic identification as well as the parties individuals voted for (both self-reported) to "learn" the ethnic affiliation of parties from data. The data we use come from Afrobarometer, waves 1-6. Table B.1 gives an example.

Table B.1: Example of ethnicity-voting data from Afrobarometer

country	Afrobarometer round	ethnic group	party vote
Ghana	5	Akan	NPP
Ghana	5	Akan	NDC
Ghana	5	Ewe	NDC

These data can be used in two ways.

(a) We can try to predict which ethnic groups are disproportionately likely to be among the voters for some party to assign ethnicity-labels to parties

(b) We can try to predict which party is disproportionately likely to be the one chosen by the voters of a particular ethnic group to assign party-labels to ethnicities.

This appendix describes different methods for both of these tasks and compares them in practise using our data.

Harmonizing ethnicity names Reported ethnicities in Afrobarometer sometimes fall into many (50+ per country) individual groups. Many of these groups are tiny (3-5 individuals only in the data) and are subgroups of other, larger groups. If groups are very small, predicted labels may be very sensitive to outliers, e.g. if all individuals sampled from that group happen to vote for one party, which, given the clustered design of Afrobarometer, is more likely to happen. Moreover, the DHS ethnicity data is less granular than the Afrobarometer data. Since we observe our outcome of interest (employment) for the DHS data, we can only use the level of detail on ethnic groups available in the DHS data. As a first step, we therefore harmonize groups appearing in Afrobarometer to their DHS equivalent and assign smaller subgroups to their larger ethnic kin. We create ethnicity link files for each country to automate the matching. We do not change the level of ethnic detail on the DHS side but we do harmonize spellings of groups where they differ between DHS survey rounds. For Afrobarometer, we use an individual's language (in the variables "What is your home language" (round 1) and "Language of respondent" (round 2), where explicit ethnicity information is not available. For rounds 3-6, we have the variable "Tribe or ethnic group", and only use the language variable where this variable is missing.

Harmonizing party names Afrobarometer records the affiliation to political parties in great detail. In rounds 1-2, respondents are asked "Do you feel close to any particular party?" "If yes, which party?" In rounds 3-6, they are additionally asked "If the election were held tomorrow,

which party would you vote for?" We use the direct voting question where available and use the closeness to party only where the answer to the direct voting question is not available. We then match all answers given in Afrobarometer to the names of parties in our voting data. There are several parties appearing in Afrobarometer that do not appear in our voting data. In running the algorithms, we therefore constrain the set of parties to be that in our election data.

Missing data We exclude from the dataset used for the labelling individuals from whom either ethnicity or party information is missing.

B.1.1 The algorithms

(a) Assigning ethnicity labels to parties

Vector-distance over-representation The first algorithm assigns ethnicity labels to parties by checking which ethnicities are over-represented relative to their national population share among the voters of each party. First, define the following notation:

- N number of individuals in the country
- N^e number of individuals of ethnic group e
- N_p number of votes cast for party p
- N_p^e number of votes cast by individuals of ethnic group e for party p

It carries out the following steps

1) Compute ethnicity-party vote share $s_p^e = \frac{N_p^e}{N_p}$, party vote share $s_p = \frac{N_p}{N}$, and ethnicity population shares $s^e = \frac{N^e}{N}$
2) Compute ethnicity-party-over-representation measure $\beta_p^e = s_p^e - s^e$
3) Keep only party×ethnicity observations with $\beta_p^e > \overline{\beta}$ (this cutoff can be varied).
4) Sort the labels by β_p^e such that the ethnicity with the highest β_p^e is the first label of each party.

The end result is a set of labels for each party. As an alternative to simple over-representation, we also compute a measure of "percentage" over-representation, replacing β_p^e with $\gamma_p^e = \frac{s_p^e - s^e}{s^e}$. Intuitively, when an ethnicity is small, the simple difference $s_p^e - s^e$ can never be very large. Normalizing by group size puts all ethnicities on an equal footing, but makes the measure assign large "meaning" to very small groups that are over-represented among voters of certain parties.

Dummy regressions This algorithm predicts the ethnic affiliation of parties by

1) running a set of $k = 1, \ldots, K$ regressions (one set of K regressions for every party p, where each individual regression includes only an ethnicity dummy for ethnic group k):

$$\mathbb{I}(\text{vote for party } p)_{is} = F(\alpha_s, E^k; \boldsymbol{\theta}), \tag{1}$$

where E^k is a dummy equal to one if individual i observed in survey(-year) s belongs to ethnicity k and zero otherwise, α_s is a survey fixed effect, $\boldsymbol{\theta}$ is a vector of coefficients,

and $F()$ is a function. The code implements this regression as either LPM or Probit.

2) Obtain the estimated coefficient $\hat{\theta}_p^k$ on E^k as well as its standard error, and compute the absolute value of the t-statistic as $t_p^k \equiv \frac{\hat{\theta}_p^k}{\widehat{se}\left(\hat{\theta}_p^k\right)}$.

3) Discard party×ethnicities with $|\hat{\theta}_p^k| < t_{0.05,N-n_s-1}$, where $t_{0.05,N-n_s-1}$ is the critical value for a two-sided t-test from a distribution with $N - n_s - 1$ d.f. (number of individuals minus number of survey years minus 1; corresponds to significance at the 5% level).

4) Among the significant party×ethnicities, retain only those with $\hat{\theta}_p^k > \overline{\theta}$ (the cutoff can be varied).

5) Sort the labels by $\hat{\theta}_p^k$ such that the ethnicity with the highest $\hat{\theta}_p^k$ is the first label of each party.

The end result is a set of labels for each party.

(b) Assigning party labels to ethnicities

Vector-distance over-representation This algorithm works similarly to the vector-distance over- representation algorithm assigning ethnicity labels to parties. It asks: "Are parties over-represented relative to their overall vote share among the voters from certain ethnic groups?" and carries out the following steps:

1) Compute party-ethnicity vote share $s_e^p = \frac{N_p^e}{N_e}$, and party vote share $s^p = \frac{N_p}{N}$

2) Compute party-ethnicity over-representation measure $\beta_e^p = s_e^p - s^p$

3) Keep only ethnicity×party observations with $\beta_e^p > \overline{\beta}$ (this cutoff can be varied).

4) Sort the labels by β_e^p such that the party with the highest β_e^p is the first label of each ethnicity.

The end result is a set of party-labels for each ethnic group. As for assigning ethnicity-labels to parties, we also compute a version of the labels based on normalized over-representation measure $\gamma_e^p = \frac{s_e^p - s^p}{s^p}$

Dummy regressions This algorithm predicts the party affiliation of ethnic groups by

1) running a set of regressions (one for every ethnicity e):

$$\mathbb{I}(\text{individual belongs to ethnicity } e)_i = F(\alpha_s, P^k; \boldsymbol{\theta}), \tag{2}$$

where P^k is a dummy equal to one if individual i observed in survey(-year) s voted for party k and zero otherwise, α_s is survey fixed effect, $\boldsymbol{\theta}$ is a vector of coefficients, and $F()$ is a function. The code implements this regression as either LPM or Probit.

2) Obtain the estimated $\hat{\theta}_e^k$ on P^k as well as its standard error, and compute the absolute value of the t-statistic as $t_e^k \equiv \frac{\hat{\theta}_e^k}{\widehat{se}\left(\hat{\theta}_e^k\right)}$.

3) Discard ethnicity×parties with $|\hat{\theta}_e^k| < t_{0.05,N-n_s-1}$, where $t_{0.05,N-n_s-1}$ is the critical value for a two-sided t-test from a distribution with $N - n_s - 1$ d.f. (number of individuals

minus number of survey years minus 1; corresponds to significance at the 5% level).

4) Among the significant ethnicity×parties, retain only those with $\hat{\theta}_e^k > \overline{\theta}$ (the cutoff can be varied).

5) Sort the labels by $\hat{\theta}_e^k$ such that the party with the highest $\hat{\theta}_e^k$ is the first label of each ethnicity.

The end result is a set of party-labels for each ethnic group.

B.1.2 Comparing the algorithms

Our baseline results are based on the simple over-representation measure, direction (b). There are two reasons for choosing (b). The first is practical. Ideally, we would like to obtain the party affiliation of each individual in the DHS data based on their ethnicity. In going in direction (a), one ethnic group may be "the most" over-represented group for several parties, while other groups are the most over-represented group for no party. For those latter groups, we would then not have a label. The second reason is more conceptual. For direction (a), an ethnic group with 5 percent population share may represent 7 percent of the voters for a certain party. At the same time, a party may capture a greater share of the voters from that ethnic group.

B.2 Village Characteristics

B.2.1 Variable definitions

distance to cities is computed using GIS software as the geodesic distance from the DHS cluster coordinates to the closest city. Cities are from Natural Earth (2018). For how "cities" are defined, see
https://www.naturalearthdata.com/downloads/10m-cultural-vectors/10m-populated-places/.

distance to road is the geodesic distance from the DHS cluster coordinates to the closest point on a colonial road (in existence in 1960). Computed using GIS software. Data on colonial roads come from Jedwab and Storeygard (2018).

elevation is computed using GIS software from cell-level data on elevation at a 30 arc-second resolution from U.S. Geological Survey (1996) as the elevation value at the DHS cluster coordinates.

terrain ruggedness is computed using GIS software from cell-level data on elevation at a 30 arc-second resolution from U.S. Geological Survey (1996). Given the grid cell data, picture a 3×3 block of 9 cells and let $e_{r,c}$ be the elevation of the cell in row r, column c of the grid. Following Nunn and Puga (2012), we compute ruggedness as $\sqrt{\sum_{i=r-1}^{r+1}\sum_{i=c-1}^{c+1}(e_{i,j}-e_{r,c})^2}$, that is, the square root of the sum of all the squared differences in elevation between the middle cell and the surrounding 8 cells.

malaria stability is mean stability of malaria transmission in the constituency. The latter variable is computed, using GIS software, as the within-constituency zonal statistic of a raster provided by Kiszewski et al. (2004), which we resample to a resolution of 30 arc-seconds prior to computing the statistic.

agricultural suitability is mean agricultural suitability in the constituency. The latter variable is computed, using GIS software, as the within-constituency zonal statistic of a raster provided by Ramankutty et al. (2002), which we resample to a resolution of 30 arc-seconds prior to computing the statistic.

cropland share are computed using data from MODIS (Friedl et al. (2010)). Data are available for all years from 2001 to 2012. We consider only those years and countries for which our election datasets records a parliamentary election in the previous year. This restricts our dataset to 14 countries – all but Mali – and 28 elections out of 32. MODIS cell-level data classify land cover according to five schemes. We use the first of these, the "IGBP global vegetation classification scheme". Starting from the raw data, we use GIS software to reclassify a cell as "1" if the original classification scheme classified it as "croplands" or "cropland/natural vegetation mosaic". Otherwise we reclassify the cell as "0". As a second step, we compute the within polygon cropland share as the mean zonal statistic of all cells within a polygon. We compute these shares for two sets of polygons. The DHS 1-km buffer cropland shares are mean zonal statistics inside circular polygons of 1 km radius centered on the DHS village coordinates. For further detail on the MODIS data, see https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mcd12q1.

Table B.2: Summary Statistics of DHS Cluster/Village Characteristics

Variable	Mean	Std. Dev.	Min	Max	Observations
Latitude	0.651	10.971	-26.817	16.656	13457
Longitude	13.984	19.456	-17.498	41.877	13457
Distance from Improved Roads (km)	18.242	33.915	0	577.4	13457
Distance from Cities (km)	33.28	29.771	0.078	578.515	13457
Elevation (m)	588.217	573.283	-4	3224.667	13453
Terrain Ruggedness	56.652	76.701	0	1311.546	13457
Agricultural Suitability	0.406	0.215	0	0.987	12888
Malaria Suitability	14.487	10.434	0	37.609	13457
Share of Cropland	0.413	0.385	0	1	10650

Notes. The table reports the summary statistics of all village-level variables used in the empirical analysis. Unit of observation is a DHS cluster/village.