(De facto) Historical Ethnic Borders and Land Tenure in Sub-Saharan Africa*

Emilio Depetris-Chauvin†ⓡ and Ömer Özak‡

Abstract

We study the role of proximity to historical ethnic borders in determining individual land ownership in Sub-Saharan Africa. Following an instrumental variable strategy, we document that individuals have a lower likelihood of owning land near historical ethnic borders. In particular, the likelihood of owning land decreases by 15 percentage points, i.e., about 1/3 of the mean rate of landownership, for rural migrants who move from 57km (90^{th} percentile) to 2 km (10^{th} percentile) from the border. This result aligns with the view that competition for land is stronger and property rights are weaker close to historical ethnic borders in Sub-Saharan Africa.

Keywords: Land ownership, Borders, Property Rights, Historical Homelands, Development, Africa, Voronoi Tessellation, Thiessen Tessellation

JEL Classification: D74, N57, O13, O17, O43, P48, Q15, Q34

*We wish to thank Giorgio Chiovelli, James Fenske, Raphael Franck, Pauline Grosjean, Moshe Hazan, Federico Masera, Moti Michaeli, Nathan Nunn, Elias Papaioannou, Dominic Rohner, Assaf Sarid, Yannay Spitzer, Mathias Thoenig, Leonard Wantchekon, and David Weil, as well as participants in multiple seminars and conferences. Emilio Depetris-Chauvin acknowledges financial support from CONICYT, FONDECYT Regular 1201036.

†Instituto de Economía, Pontificia Universidad Católica de Chile, Santiago de Chile. E-mail: edepetris@uc.cl. ORCID 0000-0002-4236-7869.

‡Department of Economics, Southern Methodist University, IZA and GLO. PO Box 0496, Dallas, TX 75275-0496. E-mail: ozak@smu.edu. Tel: (214) 768-2755. Fax: (214) 768-1821. ORCID 0000-0001-6421-2801.

"Landownership inequalities and landlessness are growing, and in Côte d'Ivoire, Kenya, Liberia, and Southern Africa, they are high enough to undermine shared growth and social cohesion, as has happened in Zimbabwe."

Securing Africa's Land for Shared Prosperity
The World Bank

1 Introduction

The study of land tenure in Sub-Saharan Africa is critical for comprehending the region's economic development. First, Sub-Saharan Africa possesses about half of the world's total usable uncultivated land, which holds enormous potential for agricultural growth (Byamugisha, 2013). Second, despite facing significant productivity challenges, agriculture remains a vital source of income for many households in the region. Third, economic theory suggests that land ownership is strongly linked to enhanced investment and higher agricultural productivity (yet, the empirical evidence for this effect is mixed, see, e.g., Goldstein and Udry, 2008; Fenske, 2011; Atwood, 1990; Jayne et al., 2010). Fourth, an unequal distribution of land ownership poses a threat to both political and economic stability in the region (Byamugisha, 2013). Therefore, understanding the main drivers of land tenure is crucial for developing effective policies and strategies that can unlock the region's agricultural potential and promote inclusive and sustainable economic growth. Paramount among the determinants of land tenure in Sub-Saharan Africa are ethnicities and their cultural norms, rules, and institutions. Indeed, land allocation and property rights are typically influenced in varying degrees by ethnic customary laws (Boone, 2014).[1]

In this paper, we explore a particular dimension of ethnicity for the determination of individuals' land ownership: proximity to historical ethnic borders, i.e., the borders of ancestral ethnic homelands in precolonial times. In particular, we hypothesize that individuals living closer to these historical ethnic borders are less likely to own land. Our analysis builds upon Depetris-Chauvin and Özak (2023), where we show that the vague nature of these historical ethnic borders is an underlying cause of contemporary land disputes. Our central premise is that these ethnic homelands are vital to group identities today, as they highlight ancestral land ownership (Horowitz et al., 1985; Fearon and Laitin, 2011). However, the demarcation and enforcement of these borders in precolonial Africa were not strictly implemented due to the abundance of land and the scarcity of population (Fanso, 1984; Herbst, 2000). Nonetheless, the post-colonial era saw a dramatic change as the population grew exponentially, leading to a scarcity of land, increased competition for resources, and the rising value of marginal lands (Boone, 2017; Herbst, 2000). Thus, we argue that the fuzzy and porous nature of these historical borders acts as a catalyst for the emergence of land disputes. Moreover, soft historical borders impede strong ethnic and personal property rights, resulting in overlapping claims on land that can muddle individuals' land

[1]Note that the term "customary" pertains to the land ownership privileges that are enshrined and transmitted in these property systems based on ethnicity, as well as to the role they assign to customary leaders in the resolution and implementation of land rights (Boone and Nyeme, 2015). As Berry (1993) claims, "people's ability to exercise claims to land remains closely linked to membership in social networks and participation in both formal and informal political processes" (p. 104).

ownership.[2] Our framework suggests that individuals who live close to historical ethnic borders may face larger competition for land, and may also have less secure ownership, which should result in lower land tenure rates. In this paper, we empirically support this hypothesis by showing that individuals' land ownership decreases in locations close to historical ethnic borders.

To test our hypothesis we employ data from several waves of the Demographic and Health Surveys (DHS, from now on) and information on the spatial distribution of historical ethnic homelands from Murdock (1959). To determine the causal effect of proximity to historical ethnic borders on land ownership, we exploit two primary sources of variation. First, we compare the land ownership status of individuals at varying distances to borders within the same ethnic homeland and within the same year to account for differences in local ethnic-level institutions and time-varying shocks. Second, we incorporate an instrumental variable strategy to exploit exogenous variation and address potential measurement errors in Murdock's map, as well as potential bias from omitted variables in an simple OLS estimate.

Our instrument exploits variation from the location of potential ethnic borders generated by a plausibly exogenous ethno-spatial partition of Africa that we formalized in Depetris-Chauvin and Özak (2023). Specifically, our theoretical model of ethnic border formation predicts that the location of ethnic borders in a homogeneous world, in which ethnicities do not differ in their geographical, institutional, cultural, linguistic, historical, and ethnic characteristics, generates a Voronoi partition of the world. I.e., ethnic borders partition the world in such a way that an ethnicity's homeland contains all locations closest to its center of gravity compared to that of any other ethnicity. In our analysis below we show that conditional on ethnic homeland $\times$ year fixed effects, our instrument does not predict a wide range of potential drivers of land ownership, including individual and location-specific characteristics. We then present strong evidence to support our main hypothesis: land ownership systematically decreases in the proximity of historical ethnic borders. Importantly, we find that this effect is substantially larger for migrants who live in rural areas.

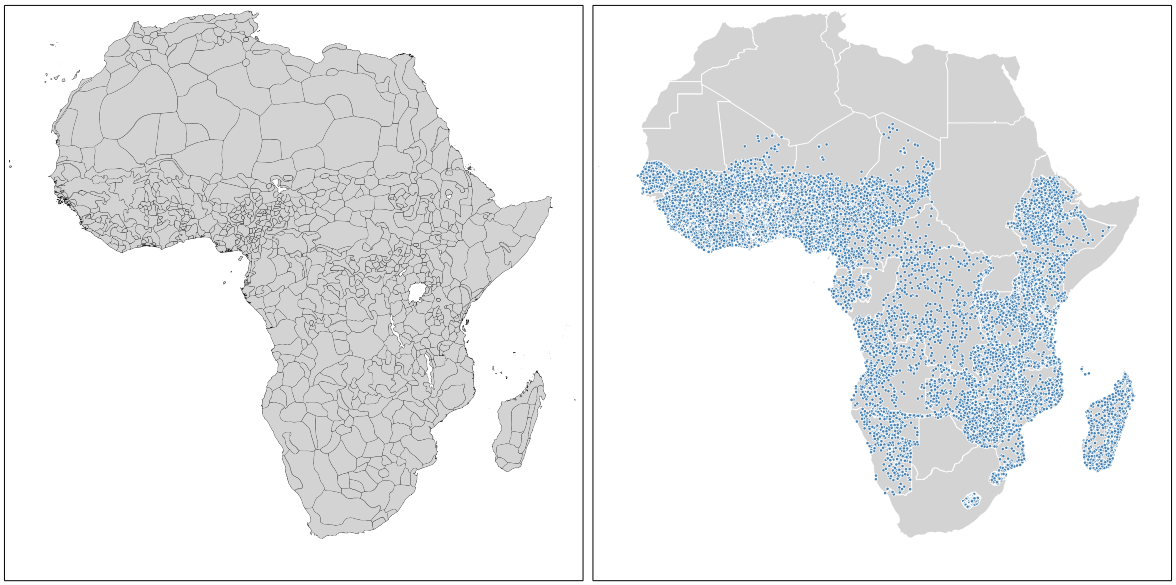
This research advances our understanding of land tenure in Sub-Saharan Africa by identifying the role of historical ethnic borders. It provides a novel mechanism linking ethnic level institutions to land tenure, conflict, and economic development. In particular, it contributes to the literature on the interaction of ethnicity and landownership in Africa (Bubb, 2013; Boone and Nyeme, 2015). It additionally contributes to the literature on the persistent effects of precolonial ethnic characteristics on African conflict and development (Moscona et al., 2020; McGuirk and Nunn, 2020; Fenske, 2013).

2 Data and Empirical Strategy

Our aim is to analyze how land ownership systematically changes in the proximity of historical ethnic borders. To do so, we use georeferenced data from 22 waves of the Demographic Health Surveys (DHS) in 30 Subsaharan African countries (ICF, 2004-2017, 2019) and information on the spatial distribution of historical ethnic homelands (Murdock, 1959). DHS provide georeferenced location data

[2]Berry (2001) argues that land "is subject to multiple, overlapping claims and ongoing debate over these claims' legitimacy"

for respondents which allows us to compute the distance of each respondent to the closest historical ethnic border, as well as the geographical characteristics of the location, and assign each respondent to an historical ethnic homeland. Our sample consists of individuals situated across 602 distinct ethnic homelands. Among the data collected in these surveys, the respondents were asked whether they owned land.[3] The DHS Program distributes surveys separately for women and men; being the dataset larger and more comprehensive for the former. We then focus our analysis on about 570,000 women (albeit our results hold regardless of the gender of the individual interviewed as shown in the appendix). It is worth noting that, when in cohabitation, women are also asked separately about ownership of the land for their partners or other family members. Therefore, our measure of land ownership reflects ownership of the land of any household member.[4]



(a) Historical Ethnic Borders in Africa (Murdock Map)

(b) DHS Locations

Figure 1: Historical Ethnic Borders and Location of Individuals

To identify the location of the historical ethnic borders and compute our measure of proximity to them, we use the geocoded version of the so-called Murdock map introduced in Nunn (2008). This map presents the location of ethnic homelands in Africa at the eve of colonization (Murdock, 1959) and has been widely and effectively used in economics, history, anthropology, and political science. Although potentially mismeasured, it has been shown that it captures deeply rooted ethnic-level relevant information (Michalopoulos and Papaioannou, 2013; Moscona et al., 2020).[5] Moreover,

[3]We use two questions from the DHS to compute a dummy variable indicating individuals' land ownership status. This dummy variable takes a value of 1 if the respondent reports owning the land either alone or jointly (variable v745b in DHS), or working their own land (variable v740), and 0 otherwise.

[4]The intersection of gender, state and customary laws, and land ownership is a vibrant topic of research in the development literature. In particular, it has been suggested that some of these laws are biased against female land ownership (Goldstein and Udry, 2008; Doss, 2006). Our analysis accounts for potential biases due to these differences by controlling for ethnic homeland (and also country) fixed effects.

[5]The map is available at https://worldmap.harvard.edu/data/geonode:Murdock_EA_2011_vkZ

in Depetris-Chauvin and Özak (2023), we show that these historical borders have persisted and remain relevant to ethnic groups. Figure 1 depicts the distribution of ethnic homelands in Murdock's map and locations of surveyed individual in DHS.

In order to analyze how proximity to historical ethnic borders impacts land ownership we estimate different versions of the following equation:

$$Ownership_{i,l(h),t} = \alpha + \beta Distance_l + \gamma' X_i + \delta' G_l + \Theta_{h,t} + \epsilon_h, \tag{1}$$

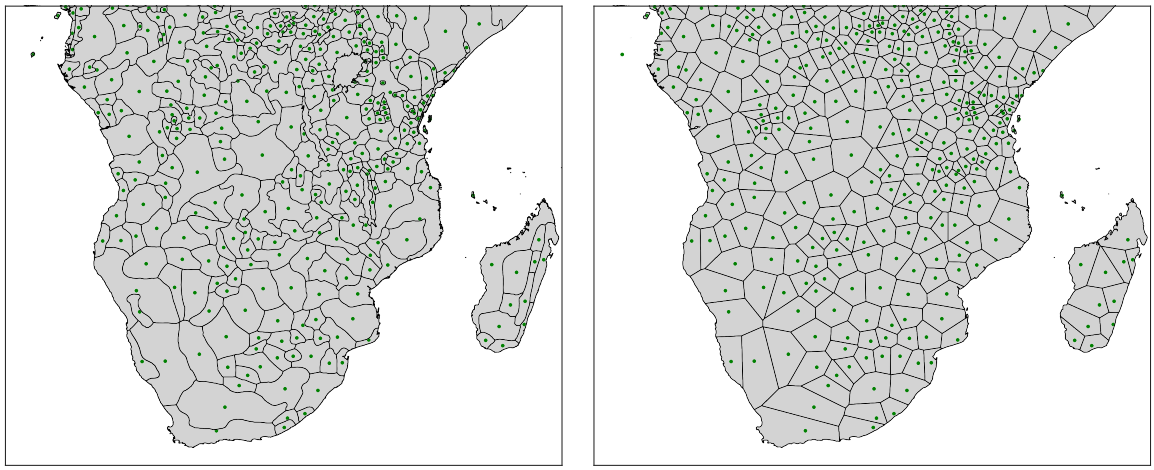
where $Ownership_{i,l(h),t}$ is our measure of land ownership as reported by individual i living in location l situated in historical ethnic homeland h and interviewed in the year t. $Distance_l$ is the logged distance from the location of the respondent to the closest historical ethnic border. X_i is the vector of basic individual controls (i.e., education, age, and squared age). The vector G_l is a set of geographic and climatic variables including absolute latitude, longitude, elevation above sea level, mean temperature, mean precipitation, and the caloric agricultural suitability of the land. We computed these measures for a 25km buffer around the location l. Further, $\Theta_{h,t}$ refer to a full set of ethnic homeland $\times$ year fixed effects. The presence of ethnic homeland $\times$ year fixed effects means that, when estimating equation (1), we identify our main effect of interest by comparing respondents interviewed close to the borders with all other respondents interviewed far from the border, but within the same ethnic homeland and during the same round of the DHS. Therefore, we are accounting for all observable and unobservable factors specific to an ethnic homeland (and even to deeply-rooted fixed characteristics of the main ethnic group), in a short period of time (say a year), that may affect land ownership. Finally, ϵ_h is an error term, which we allow to be heteroskedastic and correlated at the ethnic homeland level.[6]

We are interested in the coefficient β, describing the impact of proximity to historical ethnic borders on land ownership. Simply estimating (1) via OLS is not enough, however, as the coefficient of interest may still be biased for multiple reasons, even after conditioning on our set of control variables. Indeed, there are several potential concerns when giving a causal interpretation to our estimated β. First, given the historical nature of the measure of ethnic borders the main independent variable in our analysis may be mismeasured thus introducing a downward bias in our estimate. Second, a more plausible concern is that historical drivers of inter-ethnic interaction (e.g., conflict and trade) may codetermine the location of historical ethnic borders and land ownership potentially generating biases in any direction in our OLS estimation. Indeed, the observed association estimated in (1) may be governed by omitted geographical, institutional, cultural, linguistic, historical, and ethnic factors. To mitigate these concerns we exploit an instrumental variable approach.

Our instrument accounts for the distance to the potential location of historical ethnic borders

[6]We performed various robustness analyses: (i) additionally accounting for country $\times$ year fixed effects, (ii) using the full set of individuals, which included gender as an additional individual control, (iii) varying clustering methods to account for spatial autocorrelation. The results were qualitatively unchanged. Given the large samples in columns (1)-(3), we required more than 1.5TB of memory to perform the spatial autocorrelation robustness checks. For this reason we were only able to perform corrections for spatial autocorrelation for columns (4)-(5), which were the largest manageable samples in SMU's high-performance computer cluster. The results in those two columns suggest that clustering at the ethnic homeland level is more conservative and provides an upper bound to spatial clustering at these distance cutoffs. We complemented these analyses by also clustering at the DHS cluster, and at the region within country level. These two approaches deliver smaller standard errors suggesting that our main approach is the most conservative one.

and is constructed based on the basic prediction of our ethnic border formation model in Depetris-Chauvin and Özak (2023). Our theoretical model predicts that the location of ethnic borders in a homogeneous world, in which ethnicities do not differ in their geographical, institutional, cultural, linguistic, historical, and ethnic characteristics, generates a Voronoi partition of the world (i.e., the homeland of a particular ethnicity is composed of regions that are closer to its center than to the center of any other ethnicity). This predicted partition is based solely on the location of the centers and some notion of distance, and is independent of any local characteristics of the region where the border is located. It is worth noting that there are various potential locations that could be taken as the center of an ethnic group, including its most important city, its most densely populated location, or its earliest populated location. However, due to the lack of data or endogeneity concerns associated with using these characteristics, we use locations that are plausibly exogenous and have high predictive power. Specifically, for our main analyses, we use the geographical centroid of each historical ethnic homeland, which is identified by the average latitude and longitude of all points in the homeland, as the centers for the construction of the Voronoi partition.[7] Once we predict the potential location of the ethnic borders (i.e, Voronoi borders), we compute our instrument as the minimum distance from a given location to these borders. We refer to our instrument as the logged distance to a Voronoi border.



(a) Historical Borders & Centroids of Ethnic Homelands

(b) Potential Borders & Centroids of Ethnic Homelands

Figure 2: Historical Ethnic Borders, Centroids and Potential (Voronoi) Borders in Africa

To illustrate a Voronoi partition, Figure 2(a) depicts for each ethnic group in the southern part of Africa its historical ethnic border and centroid. Additionally, Figure 2(b) depicts the centroids and the unique potential (Voronoi) ethnic borders associated with them. Visual inspection suggests a positive correlation between the locations of historical and potential ethnic borders, suggesting that the proximity to potential borders predicts the proximity to historical ethnic borders. We explore this association more formally below.

As with any valid instrument, conditional on our full set of controls, our instrumental variable must

[7]See Depetris-Chauvin and Özak (2023) for further discussion of the properties of these Voronoi borders, as well as robustness to choice of central locations and methods of construction.

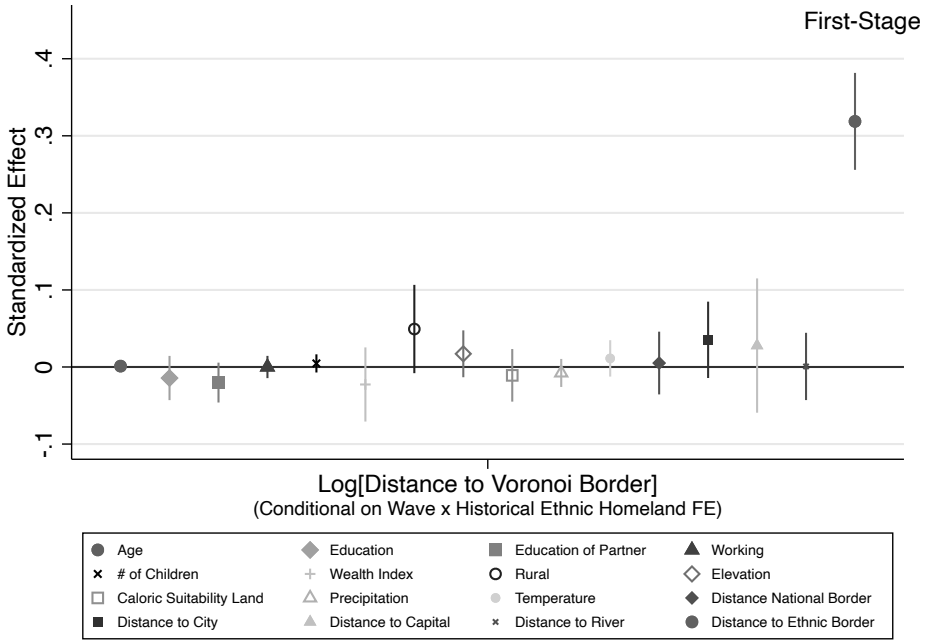


Figure 3: Plausible Exogeneity and Relevance of Instrument

be correlated with distance to historical ethnic borders (relevance condition), and uncorrelated with any unobserved characteristic of a locality (or of individuals) that may affect land ownership in a systematic way (exogeneity condition). In Figure 3, we show that, conditional on ethnic homeland $\times$ year fixed effects, our instrument does not predict a wide range of potential drivers of land ownership. This includes individual characteristics such as age, education, partner's education, work status, number of children in household, wealth, and rural status, as well as location measures such as distances to national borders, cities, rivers, and the country's capital (to account for the potential remoteness of these locations), and factors important for agricultural productivity, such as land quality, mean precipitation, mean temperature, and elevation.[8] The small standardized associations between our instrument and these covariates provide strong evidence of the plausible exogeneity of our instrument. Additionally, the same figure depicts the first-stage (standardized) coefficient showing a very strong and significant association between our instrument and distance to the closest historical ethnic border. In particular, increasing the distance to the closest Voronoi border by one standard deviation is associated with a 0.33 standard deviation increase in the distance to the nearest historical ethnic border. As we discuss below, this empirical relationship changes virtually little as we introduce an expanded set of controls.

3 Results

In Table 1, we present our main analysis in three panels. Panel A shows the first-stage relationship whereas Panels B and C show estimations of equation (1) by OLS and Instrumental Variables, re-

[8]Moreover, when using both males and females, our instrument does not predict gender of the respondent.

spectively. We focus on four different samples: all women for whom landownership data is available (columns 1 and 2), the subset living in rural areas (column 3), the subset who are migrants (column 4), and the subset of migrants living in rural areas (column 5).

Two main patterns emerge from Panel A (First-Stage). First, regardless of the sample we focus on, we find a very strong and statistically significant relationship between the logged distance from the location of the respondent to the closest historical ethnic border and our instrument based on a Voronoi partition. In particular, the elasticity of distance to border with respect to our instrument is 0.2-0.3, while the Kleibergen-Paap rk Wald F-statistic ranges between 43 and 103 for all the specifications we estimate, suggesting a strong instrument. Second, upon comparing columns 1 and 2, it becomes evident that the inclusion of individual level and location specific controls has a negligible effect on the first-stage.

OLS results in Panel B of Table 1 show a positive conditional association (albeit somewhat statistically weak under the standard levels of confidence) between distance to historical ethnic borders and land ownership. Nonetheless, this suggests that, when comparing individuals living in the same homeland and interviewed in the same year, those living further from the historical ethnic border have a greater likelihood of land ownership. As discussed above, there are several potential concerns when giving a causal interpretation to our OLS estimates. Therefore we follow our instrumental variable approach and present our results in Panel C of Table 1.

Our IV estimates for the whole sample of women (columns 1 and 2) suggest that proximity to the historical ethnic border strongly reduces the likelihood of land ownership. Interestingly, accounting for the set of individual and location specific controls alters little our main estimate. The estimated effect is economically important and implies that moving an individual from a location in the lowest decile of distance (i.e., 2km) from the border to one in the highest decile of distance (i.e., 57km) from the border, will increase their probability of land ownership by 7 percentage points (which is equivalent to 1/5 of the mean land ownership in our sample). When compared with its OLS counterpart, this IV point estimate is roughly three times larger. This inflation in the IV coefficient is consistent with our presumption that attenuation bias due to measurement error in our historical ethnic borders from Murdock's map was likely to be sizable. Moreover, omitted historical factors that could have influenced the location of ethnic boundaries and the underlying factors that facilitated land ownership (including land abundance or proximity to a central political authority) may further introduce a bias towards zero in our OLS estimates.

Our estimated effect increases by roughly 50% when we focus on the sample of females living in rural areas (column 3). Moreover, we find qualitatively similar results for the sample of female migrants (column 4), and female migrants living in rural areas (column 5). While the results are quantitatively larger for the subgroup of migrants, they are also somewhat less precise due to significantly smaller sample sizes which result in less variation within fixed effects (since we have fewer observations in each fixed effect, and also a smaller number of fixed effects). In particular, the estimated effect doubles for the subgroup of migrant women living in rural areas: moving an individual from the lowest decile of distance (i.e., 1km) from the border to the highest decile of distance (i.e., 55km) from the border will increase their probability of land ownership by 15 percentage points (which represents almost 1/3 of the

Table 1: Historical Ethnic Borders and Land Ownership

	Panel A: First-Stage				
	Log[Distance Historical Ethnic Border]				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Voronoi Border]	0.33***	0.31***	0.24***	0.22***	0.20***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Adjusted-R^2	0.41	0.43	0.36	0.37	0.34
	Panel B: OLS - Land Ownership				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Historical Ethnic Border]	0.007*	0.006*	0.004*	0.009	0.007
	(0.004)	(0.003)	(0.002)	(0.007)	(0.005)
Adjusted-R^2	0.17	0.25	0.26	0.23	0.22
	Panel C: IV - Land Ownership				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Historical Ethnic Border]	0.023***	0.021***	0.035***	0.031	0.051*
	(0.008)	(0.008)	(0.009)	(0.021)	(0.026)
First-stage F-statistic	99.04	102.57	86.69	43.71	43.07
Wave × Historical Ethnic Homeland FE	Yes	Yes	Yes	Yes	Yes
Geographical Controls	No	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	Yes	Yes
Observations	569971	569971	397469	109115	81807
Mean Land Ownership Rate	0.36	0.36	0.42	0.43	0.50
Δ Land Ownership (90-10 percentile change distance)	0.07	0.07	0.10	0.09	0.15

Notes: Heteroskedasticity robust standard error estimates clustered at the ethnic homeland-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests. Geographical controls include absolute latitude, longitude, elevation, precipitation, temperature, and caloric suitability for agriculture (CSI). Individual controls include age and its squared, and education.

mean land ownership in this subgroup). The results are quantitatively similar but statistically stronger when including males in the regression (which increases the sample size by about 1/3 as shown in the appendix). This result is consistent with the findings presented in Fenske (2010), who demonstrated that migrants in Côte d'Ivoire have weaker land property rights compared to local residents.

4 Concluding Remarks

There is an extensive literature on land ownership in Sub-Saharan Africa. While most of it focuses on the effects of land tenure on development, little attention has been paid to studying its determinants. Our paper provides causal evidence for a significant driver of land tenure in the region. In particular, we explore the role of proximity to historical ethnic borders in determining individual land ownership.

Employing an instrumental variable strategy, we document that individuals are less likely to own land near historical ethnic borders.

Consistent with a large and vibrant literature, our results support the view that historical ethnic borders have persisted and remained relevant in contemporary Sub-Saharan Africa. Yet, unlike the previous literature, our analysis highlights the direct impact of these precolonial boundaries on socio-economic outcomes. We argued that the effect of these de facto borders could originate in their fuzzy and porous nature. This begs the question whether the interaction of these borders with geographical, institutional, and legal features may change their effect. In particular, it may be essential to understand how their effect might change as they become better defined (e.g, by becoming de jure borders).

References

Atwood, D. A. (1990). Land registration in africa: The impact on agricultural production, *World development* **18**(5): 659–671.

Berry, S. (1993). *No Condition Is Permanent: The Social Dynamics of Agrarian Change in Sub-Saharan Africa*, Univ. Wisconsin Press, Madison.

Berry, S. (2001). *Chiefs Know Their Boundaries: Essays on Property, Power and the Past in Asante, 1896–1996*, Heinemann, Portsmouth, NH.

Boone, C. (2014). *Property and political order in Africa: Land rights and the structure of politics*, Cambridge University Press.

Boone, C. (2017). Sons of the soil conflict in africa: Institutional determinants of ethnic conflict over land, *World Development* **96**: 276 – 293.

Boone, C. and Nyeme, L. (2015). Land institutions and political ethnicity in africa: Evidence from tanzania, *Comparative Politics* **48**(1): 67–86.

Bubb, R. (2013). The evolution of property rights: state law or informal norms?, *The Journal of Law and Economics* **56**(3): 555–594.

Byamugisha, F. F. K. (2013). *Securing Africa's Land for Shared Prosperity*, The World Bank.
URL: *https://elibrary.worldbank.org/doi/abs/10.1596/978-0-8213-9810-4*

Depetris-Chauvin, E. and Özak, Ö. (2023). (De facto) Historical Ethnic Borders and Contemporary Conflict in Africa, *Technical report*, Working Paper.

Doss, C. (2006). The effects of intrahousehold property ownership on expenditure patterns in ghana, *Journal of African economies* **15**(1): 149–180.

Fanso, V. G. (1984). African traditional and european colonial boundaries: Concepts and functions in inter-group relations with special reference to south western cameroon, *Journal of the Historical Society of Nigeria* **12**(3/4): 23–43.

Fearon, J. D. and Laitin, D. D. (2011). Sons of the soil, migrants, and civil war, *World Development* **39**(2): 199 – 211. Ethnicity and Ethnic Strife.

Fenske, J. (2010). L'étranger: Status, property rights, and investment incentives in côte d'ivoire, *Land Economics* **86**(4): 621–644.
URL: *https://le.uwpress.org/content/86/4/621*

Fenske, J. (2011). Land tenure and investment incentives: Evidence from west africa, *Journal of Development Economics* **95**(2): 137–156.
URL: *https://www.sciencedirect.com/science/article/pii/S030438781000043X*

Fenske, J. (2013). Does land abundance explain african institutions?, *The Economic Journal* **123**(573): 1363–1390.

Galor, O. and Özak, Ö. (2015). Land productivity and economic development: Caloric suitability vs. agricultural suitability, *SMU Working Paper Series* .

Galor, O. and Özak, Ö. (2016). The agricultural origins of time preference, *American Economic Review* **106**(10).

GLOBE Task Team and others (ed.) (1999). *The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version 1.0.*, National Oceanic and Atmospheric Administration, National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80303, U.S.A.

Goldstein, M. and Udry, C. (2008). The profits of power: Land rights and agricultural investment in ghana, *Journal of Political Economy* **116**(6): 981–1022.
URL: *https://doi.org/10.1086/595561*

Herbst, J. (2000). *States and Power in Africa: Comparative Lessons in Authority and Control*, stu - student edition edn, Princeton University Press.

Horowitz, D., L, H. and Horowitz, P. (1985). *Ethnic Groups in Conflict*, Political science. Sociology: Area Studies, University of California Press.

ICF (2004-2017). Demographic and Health Surveys (various) [Datasets]. Funded by USAID. Rockville, Maryland: ICF [Distributor].

ICF (2019). The DHS Program Spatial Data Repository. Funded by USAID. spatialdata.dhsprogram.com. [Accessed May, 19, 2019].

Jayne, T., Mather, D. and Mghenyi, E. (2010). Principal challenges confronting smallholder agriculture in sub-saharan africa, *World Development* **38**(10): 1384 – 1398. The Future of Small Farms Including Special Section: Impact Assessment of Policy-Oriented International Agricultural Research (pp. 1453–1526).

McGuirk, E. F. and Nunn, N. (2020). Transhumant pastoralism, climate change, and conflict in africa, *Technical report*, National Bureau of Economic Research.

Michalopoulos, S. and Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary african development, *Econometrica* **81**(1): 113–152.

Moscona, J., Nunn, N. and Robinson, J. A. (2020). Segmentary lineage organization and conflict in sub-saharan africa, *Econometrica* **88**(5): 1999–2036.

Murdock, G. P. (1959). *Africa: its peoples and their culture history*, McGraw-Hill, New York.

Nunn, N. (2008). The long-term effects of africa's slave trades, *The Quarterly Journal of Economics* **123**(1): 139–176.

Appendix (For Online Publication)

A Variable Definition and Sources

- **Land Ownership**: Dummy variable indicating individuals' land ownership status. This variable takes a value of 1 if the respondent reports owning the land either alone or jointly (variable v745b in DHS), or working their own land (variable v740 in DHS), and 0 otherwise.
- **Female**: This variable takes a value of 1 if respondent in DHS is a woman, 0 otherwise.
- **Age**: Age, in years, reported by respondent in DHS (variable v012 in DHS).
- **Education**: Education reported by respondent in DHS (variable v149 in DHS).
- **Education of Partner**: Ordinal variable indicating husband/partner's maximum education level (variable v701 in DHS). This variable takes value of 0 if partner has no education, 1 if primary education, 2 if secondary education and 3 if higher education.
- **Working**: This variable takes a value of 1 if respondent is currently working, 0 otherwise (variable v714 in DHS).
- **Rural**: This variable takes a value of 1 if respondent lives in a rural area, 0 otherwise (variable v102 in DHS).
- **Wealth Index**: A 5-point scale ordinal variable (variable v190 in DHS). The variables takes value of 1 for poorest, 2 for poorer, 3 for middle, 4 for richer, and 5 for richest.
- **Number of Children**: Number of total children ever born (variable v201 in DHS).
- **Absolute latitude**: The absolute value of the latitude of the DHS cluster.
- **Longitude**: The longitude of the DHS cluster.
- **Log[Distance to Historical Ethnic Border]**: logged distance, in kilometers, from respondent's location to nearest Historical Ethnic Border. Author's computations.
- **Log[Distance to Voronoi Border]**: logged distance, in kilometers, from respondent's location to nearest border of Voronoi partition. Author's computations.
- **Other Distances Variables**: logged distances, in kilometers, from respondent's location to large cities, country's capital, national borders, and major rivers. Author's computations.
- **Mean Elevation**: The mean elevation of a homeland in km above sea level, calculated using geospatial elevation data taken from GLOBE Task Team and others (1999). Computed for a 25km buffer around the location of the DHS cluster. Author's computations.
- **Caloric Suitability**: Pre-1500CE Caloric suitability is the potential caloric output in a region as reported in Galor and Özak (2015) and Galor and Özak (2016). Computed for a 25km buffer around the location of the DHS cluster. Author's computations.
- **Climate variables (temperature and precipitation)**: Mean of climatic characteristics (e.g., temperature and precipitation) constructed using v3.2 of the Climatic Research Unit (CRU) database. Computed for a 25km buffer around the location of the DHS cluster. Author's computations.

B Robustness

B.1 Extended Set of Fixed Effects

Table B.1: Historical Ethnic Borders and Land Ownership
Robustness - Extended Set of Fixed Effects

	Panel A: First-Stage				
	Log[Distance Historical Ethnic Border]				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Voronoi Border]	0.33***	0.31***	0.24***	0.22***	0.20***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Adjusted-R^2	0.41	0.43	0.36	0.37	0.34
	Panel B: OLS - Land Ownership				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Ethnic Border]	0.007*	0.006*	0.004*	0.009	0.007
	(0.004)	(0.003)	(0.002)	(0.007)	(0.005)
Adjusted-R^2	0.17	0.25	0.26	0.23	0.22
	Panel C: IV - Land Ownership				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Ethnic Border]	0.023***	0.022***	0.035***	0.030	0.051*
	(0.008)	(0.008)	(0.009)	(0.021)	(0.026)
First-stage F-statistic	99.94	101.71	88.12	43.86	43.43
Wave × Ethnic Homeland FE	Yes	Yes	Yes	Yes	Yes
Wave × Country FE	Yes	Yes	Yes	Yes	Yes
Geographical Controls	No	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	Yes	Yes
Observations	569971	569971	397469	109115	81807
Mean Land Ownership Rate	0.36	0.36	0.42	0.43	0.50
Δ Land Ownership 90-10 percentile change distance	0.07	0.07	0.10	0.09	0.15

Notes: Heteroskedasticity robust standard error estimates clustered at the ethnic homeland-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests. Geographical controls include absolute latitude, longitude, elevation, precipitation, temperature, and caloric suitability for agriculture (CSI). Individual controls include age and its squared, and education.

B.2 All Males and Females

Table B.2: Historical Ethnic Borders and Land Ownership
Robustness - Sample of Females and Males

	Panel A: First-Stage				
	Log[Distance Historical Ethnic Border]				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Voronoi Border]	0.33***	0.31***	0.24***	0.22***	0.20***
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Adjusted-R^2	0.41	0.43	0.36	0.38	0.34
	Panel B: OLS - Land Ownership				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Ethnic Border]	0.008**	0.007**	0.005***	0.010	0.008*
	(0.004)	(0.003)	(0.002)	(0.006)	(0.005)
Adjusted-R^2	0.13	0.25	0.26	0.21	0.21
	Panel C: IV - Land Ownership				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Ethnic Border]	0.023***	0.021***	0.031***	0.020	0.046**
	(0.007)	(0.007)	(0.008)	(0.017)	(0.023)
First-stage F-statistic	112.79	115.18	95.27	44.46	41.73
Wave × Ethnic Homeland FE	Yes	Yes	Yes	Yes	Yes
Geographical Controls	No	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	Yes	Yes
Observations	848720	848720	589694	162567	115950
Mean Land Ownership Rate	0.38	0.38	0.45	0.42	0.49
Δ Land Ownership 90-10 percentile change distance	0.07	0.07	0.09	0.06	0.14

Notes: Heteroskedasticity robust standard error estimates clustered at the ethnic homeland-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests. Geographical controls include absolute latitude, longitude, elevation, precipitation, temperature, and caloric suitability for agriculture (CSI). Individual controls include age and its squared, and education.

B.3 Alternative Clustering - Spatial Autocorrelation

Table B.3: Historical Ethnic Borders and Land Ownership
Robustness - Alternative Clustering using DHS Clusters/Regions

	Panel A: First-Stage				
	Log[Distance Historical Ethnic Border]				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Voronoi Border]	0.33***	0.31***	0.24***	0.22***	0.20***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
	[0.03]	[0.03]	[0.03]	[0.03]	[0.03]
Adjusted-R^2	0.41	0.43	0.36	0.38	0.34
	Panel B: OLS - Land Ownership				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Ethnic Border]	0.007***	0.006***	0.004***	0.009***	0.007***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
	[0.003]	[0.002]	[0.002]	[0.005]	[0.004]
Adjusted-R^2	0.17	0.25	0.26	0.23	0.22
	Panel C: IV - Land Ownership				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Ethnic Border]	0.023***	0.021***	0.035***	0.031***	0.051***
	(0.003)	(0.004)	(0.006)	(0.010)	(0.012)
	[0.007]	[0.007]	[0.009]	[0.017]	[0.021]
First-stage F-statistic	1243.05	1163.49	560.43	311.01	207.33
Wave × Ethnic Homeland FE	Yes	Yes	Yes	Yes	Yes
Geographical Controls	No	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	Yes	Yes
Observations	569971	569971	397469	109115	81807
Mean Land Ownership Rate	0.36	0.36	0.42	0.43	0.50
Δ Land Ownership 90-10 percentile change distance	0.07	0.07	0.10	0.09	0.15

Notes: Heteroskedasticity robust standard error estimates clustered at the DHS cluster level are reported in parentheses, and clustered at the region within countries level in brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests. Geographical controls include absolute latitude, longitude, elevation, precipitation, temperature, and caloric suitability for agriculture (CSI). Individual controls include age and its squared, and education.

Table B.4: Historical Ethnic Borders and Land Ownership
Robustness - Spatial Autocorrelation

	Panel A: First-Stage Log[Distance Historical Ethnic Border]				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Voronoi Border]	0.33***	0.31***	0.24***	0.22***	0.20***
	(0.033)	(0.031)	(0.026)	(0.033)	(0.031)
				[0.029]	[0.027]
				((0.027))	((0.027))
				[[0.032]	[[0.029]
				([0.031])	([0.027])
Adjusted-R^2	0.41	0.43	0.36	0.38	0.34
	Panel B: OLS - Land Ownership				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Ethnic Border]	0.007*	0.006*	0.004*	0.009	0.007
	(0.004)	(0.003)	(0.002)	(0.007)	(0.005)
				[0.005]	[0.004]
				((0.006))	((0.004))
				[[0.007]	[[0.005]
				([0.008])	([0.005])
Adjusted-R^2	0.17	0.25	0.26	0.23	0.22
	Panel C: IV - Land Ownership				
	Full Sample		Rural	Migrants	Rural Migrants
	(1)	(2)	(3)	(4)	(5)
Log[Distance to Ethnic Border]	0.023***	0.021***	0.035***	0.031	0.051*
	(0.008)	(0.008)	(0.009)	(0.021)	(0.026)
				[0.017]	[0.021]
				((0.019))	((0.023))
				[[0.019]	[[0.023]
				([0.020])	([0.024])
First-stage F-statistic	1243.05	1163.49	560.43	311.01	207.33
Wave × Ethnic Homeland FE	Yes	Yes	Yes	Yes	Yes
Geographical Controls	No	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	Yes	Yes
Observations	569971	569971	397469	109115	81807
Mean Land Ownership Rate	0.36	0.36	0.42	0.43	0.50
Δ Land Ownership 90-10 percentile change distance	0.07	0.07	0.10	0.09	0.15

Heteroskedasticity robust standard error estimates clustered at the ethnic homeland-level are reported in parentheses, spatial auto-correlation corrected standard errors with distance cutoffs at 50 (brackets), 100 (double parenthesis), 200 (double brackets), and 500km (parenthesis and brackets) are shown below; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests. Given the large samples in columns (1)-(3), we would require more than 1.5TB of memory to perform similar robustness checks. For this reason we were only able to perform corrections for spatial autocorrelation for columns (4)-(5), which were the largest manageable samples in SMU's high-performance computer cluster. The results in those two columns suggest that clustering at the ethnic homeland level is more conservative and provides an upper bound to spatial clustering at these distance cutoffs. Geographical controls include absolute latitude, longitude, elevation, precipitation, temperature, and caloric suitability for agriculture (CSI). Individual controls include age and its squared, and education.