

Cash Crops, Print Technologies and the Politicization of Ethnicity in Africa*

Yannick I. Pengl[†] Philip Roessler[‡] Valeria Rueda[§]

June 25, 2021

Abstract

What are the origins of the ethnic landscapes in contemporary states? Drawing on a pre-registered research design, we test the impact of dual socioeconomic revolutions that spread across Africa in the 19th and 20th centuries—export agriculture and print technologies. We argue these changes transformed ethnicity via their effects on politicization and boundary-making. Print technologies strengthened imagined communities, leading to more salient—yet porous—ethnic identities. Cash crop endowments increased groups' mobilizational potential but with more exclusionary boundaries to control agricultural rents. Using historical data on cash crops and African language publications, we find that groups exposed to these historical forces are more likely to be politically relevant in the post-independence period, and their members report more salient ethnic identities. We observe heterogeneous effects on boundary-making as measured by inter-ethnic marriage; relative to cash crops, printing fostered greater openness to assimilate linguistically-related outsiders. Our findings not only illuminate the historical sources of ethnic politicization, but mechanisms shaping boundary formation.

*Acknowledgments: This research is supported by the US National Science Foundation (award #1628498) and the Swiss National Science Foundation (grant #P0EZP1_159076). We thank Paola Galano Toro, Vanessa Kellerhals, Benjamin Füglistler, Lukas Dick, Carlos Mairoce, and Julian Seitlinger (ETH Zurich), Sidhart Bhushan and Hedda Roberts (Oxford), as well as Layla Abi-Falah, Aaron Spitler, and Henry Young (W&M) for excellent research assistance. Earlier versions of our research design were presented at WGAPE, LSE, March 2018; Annual Meeting of the American Political Science Association, Boston, September 2018; Princeton University Comparative Politics Colloquium, October 2018; Annual Meeting of the AEHN Bologna, October 2018. First results and early paper versions presented at the Annual Meeting of the American Political Science Association, Washington DC, September 2019; Political Economy Workshop, Zurich, October 2019; Annual Meeting of the Swiss Political Science Association, February 2020; virtual APSA 2020; Zurich Workshop in Empirical Political Economy, September 2020; ASREC annual conference, November 2020; Leiden Workshop in Political Science, October 2020; Workshop on the Politics of Favoritism, ZEW Mannheim, February 2021. We are grateful to participants for their suggestions and feedback. Special thanks to three anonymous reviewers, Matthew Gichohi, Corinne Bara, Joan Ricart-Huguet, Leila Demarest, Dan Posner, Tim Phillips, Carl Müller-Crepon, and Lars-Erik Cederman for comments and discussions.

[†]International Conflict Research, ETH Zürich. ORCID: 0000-0002-0373-0952

[‡]Department of Government, William & Mary. ORCID: 0000-0002-4935-6468

[§]School of Economics, University of Nottingham, and CEPR. ORCID: 0000-0002-4967-3172

Introduction

What are the origins of the ethno-political landscapes that shape contemporary states? A voluminous literature points to the impact ethnicity—social identity based on shared descent and culture—has on politics and the allocation of state resources. From the provision of public services to civil war, ethnicity is found to structure a wide-range of political and economic processes (Horowitz, 1985; Chandra, 2004; Habyarimana et al., 2009; Roessler, 2016). In this paper we address the question of what drives ethnic politicization in the first place—that is, why politics revolves around some cultural groups and not others. Despite a rich qualitative and historical literature on the topic (Vail, 1989a; Bates, 1983; Posner, 2005), quantitative studies typically do not engage the endogenous sources of ethno-political divisions that shape policy outcomes. This represents an important limitation, as inferences on the consequences of ethnic politics may be vulnerable to selection problems (Birnie et al., 2015, 2018).¹

We aim to advance knowledge on this question, reporting the results of a pre-registered research design.² We distinguish between two inter-related processes that shape ethnic politics: boundary-making and politicization. The former—the sine qua non of ethnicity (Barth, 1969)—encapsulates the social boundaries that regulate group membership and shape inter-ethnic ties. Following from Weber (1978) and others (Parkin, 1974; Fearon, 1999; Wimmer, 2013; Caselli and Coleman, 2013), we conceive of the porosity of group boundaries as being especially consequential for ethnic politics. Politicization, on the other hand, occurs when members of a cultural group coordinate on their shared identity to compete for state power (Bates, 1983; Fearon, 1999). In accounting for variation in boundary-making and politicization, our framework focuses on periods of significant material and cultural change that potentially strengthened groups' mobilizational capabilities and redefined the markers of group membership.

We study these phenomena across countries in Africa, a region in which ethnicity has structured political competition but only among a subset of ethno-linguistic groups.³ Much existing scholarship on the politicization of ethnicity in Africa points either to the lasting impact of colonialism—via the arbitrary territorial partition of the continent to the imposition of indirect rule (Asiwaju, 1984; Ekeh, 1990; Engleburt, Tarango and Carter, 2002; Mamdani, 1996)—or to the role of contemporary political competition (Posner, 2005). These factors are no doubt important but arguably too widespread to explain significant within-country variation in ethnic identity salience (Vail, 1989a). In addition, colonialism was embedded in larger socio-economic changes. Two of particular importance were the cash crop revolution and the spread of Christianity by missionaries that preceded the “Scramble for Africa” and may have affected

¹Across the ethnic politics literature, many studies model competition for power and resources among a given subset of politically-relevant ethnic groups.

²We pre-registered our research design with Evidence in Governance and Politics (EGAP) on April 30, 2019 after some promising preliminary analyses but *before* merging our publications and cash crop data with Ethnologue language categories and group polygons and, via Ethnologue, to EPR, PREG, Afrobarometer and DHS. We had already seen geographic correlations between cash crop locations and Afrobarometer/DHS outcomes as well as between proximity to missionary printing presses and Afrobarometer identity salience. However, we were in no position to analyze group-level outcomes, the actual publications treatment, or the ethnic specifications described below as all of these require ethnic matches. An anonymized version of our pre-analysis plan can be found here: <http://bit.ly/3qYufI8>

³In our sample of 35 Sub-Saharan African countries, there exist 2303 Ethnologue languages, whereas the Ethnic Power Relations dataset counts 140 groups relevant in the first year countries in the region enter the dataset and another 158 groups relevant through its last year (Vogt et al., 2015).

ethnic identities independently of or in interaction with colonial policy-making. We argue that these fundamental changes have path-dependent effects on contemporary ethnic mobilization and coalition formation despite significant institutional change over the last 150 years.

First, we posit that the spread of cash crops and Christian missions both contributed to the politicization of ethnicity. We hypothesized the transition to commercial export agriculture increased the ethnic politicization of groups endowed with cash crops through a resource channel that bolstered these groups' mobilizational capabilities, but also via competition for land and the enforcement of descent-based property rights regimes. While missionaries also brought about important material changes through investments in new infrastructure and provision of education, perhaps even more important was the communication revolution they unleashed. Intent on spreading the Gospel, missionaries invested heavily in standardizing, writing and printing what were primarily oral languages. This improved treated groups' communication capabilities, while increasing ethnic salience through the strengthening of "imagined communities" (Anderson, 1983)—as the adoption of a standardized language and the consumption of a uniform set of cultural characteristics, texts, and histories enhanced group solidarity.

Even as these dual socio-economic forces increased ethnic politicization, we hypothesized they differently reshaped ethnic boundaries. The "imagined communities" reconstructed through language standardization created an opportunity for the assimilation of outsiders through language and cultural immersion—leaving a legacy of more inclusionary ethnic boundaries. Cash crop agriculture had a very different effect as it was tied to control of the land. In the face of growing demand for access to their agricultural-rich homeland, including from migrant farmers and laborers, ethnicity was employed as a means of "social closure" (Weber, 1978; Parkin, 1974) to regulate land ownership and control agricultural rents—leaving a legacy of more exclusionary ethnic boundaries.⁴

To test these hypotheses, we combine detailed historical data on cash crop production and the diffusion of print and writing technology (as measured by publications in African languages) with contemporary ethnicity data. Our cash crop data is based on a comprehensive historical map on the source locations of exports in late colonial Africa created by Hance, Kotschar and Peterec (1961) and digitized by Roessler et al. (2020). To measure language standardization and its dissemination through printing, we compile a novel dataset of historical African language publications from Rowling and Wilson (1923) and Mann and Sanders (1994). Together, these two bibliographic sources cover approximately 10,000 titles in 370 distinct African languages.

We employ group-level and individual-level indicators to measure ethnic politicization. At the group-level, we use the Ethnic Power Relations (EPR) (Vogt et al., 2015) and the Politically Relevant Ethnic Groups (PREG) datasets (Posner, 2004a) to measure which ethnic groups or coalitions have been active in competition for state power in the post-independence period. At the individual level, we use Afrobarometer Rounds 3–6 that include a question on whether respondents self-identify more in ethnic or national terms. To analyze the hypothesized heterogeneous legacies of cash crops and print technologies on boundary-making and social closure, we employ a behavioral measure of ethnic assimilation: inter- and intra-ethnic marriages from a large sample of couples surveyed by USAID's Demographic and Health Surveys.

⁴See Caselli and Coleman (2013) for a formalization of the link between social closure and ethnicity.

We use linguistic groups identified in the Ethnologue database as our primary unit of analysis to minimize concerns about endogenous sample selection (Laitin, 2000a, p. 142). This enables us to merge our cash crop, publishing, and outcome data, along with a host of control variables, to the Ethnologue groups through spatial overlays or ethnic name matching.⁵ In the survey-based analyses, we use two types of specifications. The first—*geographic* models—are based on the location of individuals and the Ethnologue polygons in which they reside. These models compare people located in different places with and without historical cash crop production and/or missionary publishing. The second—*ethnic* models—are based on survey respondents' affiliation to a given ethnic group rather than place of residence. Thus, they compare individuals residing in the same location but from ethnic groups with differential exposure to historical cash crop production and missionary publishing. This enables separating culturally transmitted attitudes and behaviors from locational effects.

We employ three main methods to mitigate endogeneity concerns. First, we employ location fixed effects in our ethnic-level specifications to address mission selection into areas with favorable locational fundamentals or those populated by already large and more powerful groups. Second, we use our African-language publishing data to analyze the effects of print technologies at the intensive margin (i.e. estimating the effects of the magnitude of publication records among groups with at least one publication). Third, we instrument actual crop production with agro-climatic suitability to address the potentially endogenous uptake of commercial agriculture. We also conduct additional robustness checks to rule out alternative explanations, such as the effects of group size, pre-colonial centralization, indirect rule, ethnic diversity and conversion to Christianity.

We find that groups historically exposed to cash crops or print technologies are significantly more likely to be politically relevant after independence. Groups that cultivated one of five major cash crops through the end of colonialism or with a historical publication in their language are, respectively, 129 (54) and 88 (45) percent more likely than the average group to be politically relevant according to PREG (EPR). These results are robust to instrumenting crops with suitability and when focusing only on the subsample of groups exposed to Christian missions.

At the individual level, we find citizens residing in areas of historical cash crop production or living in Ethnologue polygons with a history of publishing are significantly more likely to self-identify with their ethnic group rather than nationality. Moreover, ethnic salience follows our expectation of cash crops producing location-specific effects among “stayers” and publishing producing broader cultural effects, including among “movers” (i.e. respondents living outside their ancestral ethnic homeland). We do not find evidence, however, that groups treated with cash crops or print technologies have more homogeneous political preferences today.

We find strikingly different effects of cash crops and publishing on the porosity of ethnic boundaries, as measured by observed inter-ethnic marriage rates. Consistent with our expectation that cash crops engendered social closure and less openness to ethnic outsiders, we find inter-ethnic marriage to be significantly lower even with linguistically closely related groups. In contrast, and consistent with the hypothesis that print technologies led to salient but more porous ethnic boundaries, we find null and sometimes positive effects on inter-ethnic marriage with linguistically close ethnic out-

⁵For combining Ethnologue groups with information from EPR, PREG, DHS, and Afrobarometer, we use the publicly available ethnic links coded by Müller-Crepon, Pengl and Bormann (2020).

siders but negative effects on marriages across large linguistic distances. However, in contrast to our expectations, both exposure to cash crops and print technologies are positively associated with contemporary ethnic-based conflict—suggesting that, even as print technologies opened the door to assimilation of culturally proximate outsiders, its politicizing effects ensured these groups have not escaped cycles of ethnic conflict.

Our findings speak to different research streams in the social sciences. Despite a strong consensus on the constructivist nature of ethnicity (Laitin and Posner, 2001; Chandra, 2012), the endogenous sources of ethnogenesis remain understudied. Our paper illuminates the historical role of export agriculture and publishing in Africa. Moreover, our analysis sheds light on the relationship between ethnic politicization and boundary-making (Wimmer, 2013). It is generally assumed that these two processes are reinforcing, leading perhaps to convergence in the types of social boundaries regulating politically-relevant ethnic groups. This may or may not be the case; as we illustrate, even across politicized groups, boundary policing can vary based on path dependent effects of material and cultural changes on assimilationist practices and norms of openness.

In advancing this line of inquiry, we draw on classic theories of group formation—Weber’s (1978) notion of social closure, Anderson’s (1983, pp. 46-47, 7) framework on the ethno-national impact of print technologies, and prominent but conflicting accounts of how economic change transforms ethnic identities (Gellner, 1983; Bates, 1974; Robinson, 2014). To date, there have been few systematic tests of Anderson’s “imagined communities” hypothesis.⁶ We find strong support for a link between print technologies, language standardization and ethno-nationalism in Africa. However, as we explain below, the mechanisms through which these processes reconstructed ethnic identity differed from 19th century Europe where “print capitalism”, bureaucratic “languages of power,” and state-sponsored nation-building fostered national identities—rather than the subnational identities that arose across Africa.

As far as “modernization” is concerned, our results are broadly in line with Bates’ (1974) intuition that competition for economic benefits may deepen ethnic divisions. At the same time, our focus on cash crops produced by African smallholder farmers suggests that rural economic change was just as important as the urban dynamics prominently highlighted in the existing literature (Epstein, 1958; Cohen, 1969).

Finally, our paper employs a pre-registered design to address growing concerns about publication bias and data mining for significant results in historical persistence studies. Beyond guarding against cherry-picking positive findings, pre-registration encourages careful ex ante theorizing and hypotheses development. Pre-registration does not preclude ex-post modifications of the pre-specified analyses, but it does necessitate transparency about any changes made. In this vein, we describe all pre-specified hypotheses and analyses in Supplementary Information IV.

The Determinants of Africa’s Ethnic Landscape

In this section we more fully advance our theoretical argument on the impact of the cash crop and print revolutions on shaping Africa’s modern ethnic landscape. Before addressing each in turn, we first situate our argument within the broader ethnicity

⁶However, see Sasaki (2017), who focuses on the impact of the printing press in Europe on language standardization.

scholarship.

Ethnic Boundary-Making and Politicization

We conceive of a country's ethnic landscape as shaped by two key processes: boundary-making and politicization. The former encompasses the construction and maintenance of social differences (Barth, 1969), in which individuals employ "points of social reference," such as ascriptive, cultural, or other markers, to place themselves and others into groups to "order" the world (Hale, 2004). Boundary-making helps to solidify social groups through the adoption of criteria for membership and their enforcement by in-group members (Wimmer, 2013). Following from Weber (1978), we consider a group's closure or accessibility as one of the most important dimensions of boundary-making (Wimmer, 2013). Politicization, on the other hand, entails members of a given group consciously or subconsciously leveraging their shared identity to coordinate their behavior to access political and economic benefits (Bates, 1983; Fearon, 2006).

Generally, boundary-making and ethnic politicization are theorized to be reinforcing. This is perhaps most starkly illuminated in the civil war literature in which conflict along ethnic lines contributes to the hardening of social boundaries (Fearon and Laitin, 2000; Kalyvas, 2008). Other forms of political competition, such as elections, are also found to increase ethnic salience (i.e., the intensity of one's ethnic identity vis-à-vis others) (Ouchou, 2002; Eifert, Miguel and Posner, 2010)—although this does not necessarily translate into higher degrees of closure.⁷ Likewise, the reverse—that boundary-making facilitates ethnic politicization—is an important assumption in rationalist accounts of ethnic coalition formation that stress the need to exclude outsiders from the returns to collective action (Fearon, 1999).⁸ The reinforcing effects of boundary-making and politicization may suggest some degree of convergence in the structure of social boundaries across politicized groups but as far as we know this has not been empirically assessed.⁹

Existing Literature

What then explains boundary-making and politicization? Following from our conceptual framework, we expect factors shaping boundary-making to drive the construction and enforcement of socially-differentiated groups, whereas factors activating politicization likely work through their effects on group coordination and mobilization. Here we briefly synthesize existing research with a focus on sub-Saharan Africa.

Evolutionary and geographic approaches, respectively, attribute Africa's comparatively high ethnic diversity to the loss of genetic variation as human species migrated from the cradle of humankind (van den Berghe, 1981; Ahlerup and Olsson, 2012)

⁷Salience and closure capture different but potentially reinforcing identity dimensions. The former reflects the importance of an identity to oneself or others—i.e. the likelihood that a given identity and not others will be invoked across different situations (Stryker, 1980). In contrast, closure reflects the degree to which a group is accessible to outside members (Wimmer, 2013). Following from Stryker (1980), we might expect closed groups, in which entry and exit pose higher costs, to correlate with more salient identities.

⁸See also Bates (1983); Posner (2017); Chandra (2006)

⁹We analyze this in Supplementary Information I. We find that politically-relevant groups do tend to have less porous boundaries as measured by inter-ethnic marriage, though these correlations are not particularly strong.

and ecological variation leading to economic and cultural differentiation (Nettle, 1998; Michalopoulos, 2012). What form these groups take and the degree of their politicization then depends on a host of historical, material and institutional factors.

One factor regularly advanced as contributing to political relevance is group size, following the logic that a minimum support base is necessary to sustain viable political coalitions (Bates, 1983; Posner, 2004b, 2005). Beyond size, others point to the importance of groups' socio-political structures, in particular legacies of centralized and hierarchical institutions (Michalopoulos and Papaioannou, 2013).¹⁰ In the context of Africa's multi-ethnic states, historical statehood may have deepened ethno-political cleavages (Paine, 2019).

Other research focuses on economic change and its differential effects on groups across the continent. Ekeh (1990), Nunn (2008), and Nunn and Wantchekon (2011) highlight how the slave trades contributed to ethnic fractionalization and strengthened norms of mistrust. The decline of the slave trades corresponded with the spread of export agriculture (Hopkins, 1973) and Christian missionaries across the continent (Cagé and Rueda, 2016). Bates (1974) mentions both of these factors as examples for spatially concentrated modernization benefits that spurred inter-group inequality, competition, and politicization.¹¹ Other relevant economic changes include mining, railway construction, and perhaps most prominently urbanization. (See also Horowitz, 1985; Vail, 1989a; Cohen, 1969; Nnoli, 1978).

Beyond its material effects, missionaries, export agriculture, and the colonial state had profound cultural impacts. Through historical process-tracing, Ranger (1989) shows how missionary investments in translation and printing of Bibles in vernacular languages "created rather than merely reflected" extant ethno-linguistic divisions. (See also Chimhundu (1992) and Posner (2003)). Berry (1993) and Lentz (2013) point to the effects of the commercialization of agriculture on the reconstruction of social identities, especially the distinction between "natives"—or "sons of the soil"—and "strangers" (Lentz, 2013). Mamdani (1996) argues that the colonial project had much broader cultural effects through social engineering around the "customary." Reinforced through indirect rule and other colonial policies of social control (Eyoh, 1999; Posner, 2005), colonialism sharpened communal identities¹² through ideologies of "tribalism" (Ekeh, 1975) and "autochthony" (Lentz, 2013).¹³

The anti-colonial liberation struggle held the promise to re-imagine social relations and national communities (Fanon, 1963; Ekeh, 1990; Ake, 1993)—and in some cases, such as Nyerere's Tanzania, this was achieved (Miguel, 2004). But, largely, post-colonial competition for state power revolved around ethno-political networks, further deepening ethnic politicization (Horowitz, 1985; Rothchild, 1997; Nnoli, 1998; Roessler, 2016).

¹⁰Koter (2016), in contrast, argues that hierarchical institutions enabled post-independence rulers to target groups with patronage-based policies rather than ethnic appeals, potentially dampening ethnic salience. Also, Dunning and Harrison (2010) find that the historical legacy of cousinage from the Mali Empire has helped to weaken the political effects of ethnicity.

¹¹For an illuminating ethnography on the interactive effects of Christian missionaries and cash crops on ethnic association formation, anti-colonial resistance and political mobilization, see Spear (1997). In the case of the Meru, ethnic mobilization contributed to the development of a broader nationalist movement (Okoth, 2006).

¹²Colonial partition itself, however, may have contributed to stronger national identities among groups divided between two sovereign states (Miles and Rochefort, 1991; Robinson, 2014).

¹³On the genealogy of autochthony and its roots in colonialism, see Ceuppens and Geschiere (2005); Marshall-Fratani (2006).

The advent of multi-party elections with the end of the Cold War, in some cases, transformed ethno-political configurations (Posner, 2005), but often intensified rather than dampened ethnic salience (Oucho, 2002; Eifert, Miguel and Posner, 2010), as well as autochthonous mobilization (Ceuppens and Geschiere, 2005; Marshall-Fratani, 2006). However, there is some evidence that urbanization and demographic change (leading to greater levels of ethnic diversity) as well as democratic institutions are reducing ethnic favoritism (Ichino and Nathan, 2013; Burgess et al., 2015; Kramon et al., 2021).

We build on and extend this literature by developing and systematically testing new hypotheses on how the cash crop and print revolutions shaped processes of ethnic boundary-making and politicization from the 19th century onward.

The Cash Crop Revolution

In the 19th and 20th centuries, African economies underwent an important structural transformation away from the slave trades that dominated exchange for the previous four hundred years to commercial export agriculture (Hogendorn, 1969; Hopkins, 1973; Frankema, Williamson and Woltjer, 2018).¹⁴ The cash crop revolution led to an important spatial shift in economic production to areas suitable for oil palm, groundnuts, cocoa, coffee, and cotton, and enabled millions of African smallholders and traders to benefit from global exchange (Hopkins, 1973). Fuelled by European-financed transportation infrastructure before and during colonialism, these cash crop zones were vertically integrated with export markets but with weak horizontal linkages with the rest of the colony (Rodney, 1972; Hirschman, 1977; Roessler et al., 2020).

Consistent with Bates (1974), we posit that the spatial disparities arising from the cash crop revolution had important path dependent effects on ethnic politicization. The takeoff of export agriculture endowed some groups—those who would be the primary producers of cash crops or the owners of the land on which they were produced—with a common economic niche, much greater wealth potential than others, and clear incentives to defend these advantages in competition with other groups.

A second and closely related channel of ethnic politicization was via the effects of the commercialization of agriculture on land tenure regimes.¹⁵ Many of the most suitable areas experienced an increase in demand for land as waves of farmers, including enterprising migrant farmers (Hill, 1963), adopted cash crops. Labor migration to cash crop areas further increased local diversity, land pressures, and inter-group competition.

The commercialization of agriculture combined with migration-led population growth induced important changes in the social bases of land tenure regimes. In pre-colonial Africa, land rights were contingent on group membership or allegiance to traditional authorities (Berry, 1993). These practices did not change per se with the advent of cash crop agriculture and colonialism. What did, however, were outsiders' eligibility for group membership as ethnic boundaries became more tightly regulated (Boni, 2006;

¹⁴Cash crops would prove a much more important source of colonial exports than minerals. By 1957, across the 35 countries in our dataset, cash crops accounted for 59.4% of total exports (by value) compared to only 22% for minerals (Hance, Kotschar and Peterec, 1961).

¹⁵For important previous work on the socio-political implications of the transition to commercial agriculture, see Colson (1971); Berry (1993); Boone (2014, 2017).

Lentz, 2013). Thus, following from Weber's (1978) idea of social closure (Parkin, 1974), in which social identity is employed as a means to restrict access to economic rents, in the face of rising land values and an in-flux of migrants, ethnic boundaries were more firmly policed to exclude outsiders from land ownership.¹⁶ In line with the idea that ethnic differences are constructed, at least partially, as "a boundary-enforcement device" (Caselli and Coleman, 2013, 162) to control private goods, contestation over land not only made ethnicity more salient, it likely led sons of the soil to emphasize less accessible criteria of group membership such as ascriptive characteristics and ancestral-ties to the land.¹⁷ In a fascinating ethnography of the impact of the spread of cocoa and migrant farmers to the Sefwi homeland (located in present-day western Ghana) from the early 20th century onward, Boni (2006) documents this precise dynamic unfolding—resulting in the "ancestralization of land rights" and more stringent enforcement to prevent migrants from permanently owning land.

We expect these mechanisms to only apply to regions of African smallholder production. Where European companies, settlers, or the colonial state dominated production, land alienation and labor coercion likely undercut local control of agricultural rents, weakened ethnic institutions, and reduced opportunities for ethnic boundary-making.

Christian Missions, Print Technologies, and African Language Publications

As the abolition of the slave trade ushered in cash crop agriculture in Africa, it also gave momentum to the spread of Christian missions across the continent. In their endeavor to spread the Gospel, missionaries spearheaded a communication revolution.

Missionaries translated the Bible and education materials into vernacular languages as a vehicle for conversion (Ranger, 1989; Laitin, 2007; Woodberry, 2012). As most African languages were oral languages, missionaries first invested in language standardization and developing Latin-script writing systems (Ranger, 1989; Posner, 2003). To propagate language knowledge and consumption of the written texts, printing presses were imported to publish Bibles, hymnals and grammar books that were then used in churches and schools (Posner, 2003; Cagé and Rueda, 2016). This communication revolution was most intense in British colonies given the preponderance of Protestant missionaries and the promotion of local languages and culture as part of indirect rule arrangements (Albaugh, 2014).¹⁸

Anderson's (1983) argument on the impact of print capitalism on European nation-building is a valuable reference when considering the effects of Africa's print revolution on ethno-national communities. However, while language standardization and printing underpinned significant social changes in both regions, some mechanisms differed (Ranger, 1989). First, given lower literacy and less integrated markets, in Africa the

¹⁶This process of ethnic boundary hardening was driven from below—as chiefs found themselves under growing pressure from their constituents not to give away too much land to outsiders (Boni, 2006)—but also supported from above—as colonial government's promoted neocustomary land tenure regimes (Boone, 2017; Mamdani, 1996, 104-105).

¹⁷See also Bates (1974, pp. 465-67) on how local elites in cash crop areas used ethnic criteria to restrict access to modernization benefits.

¹⁸In French colonies educational instruction was mandated to be in French. Albaugh (2014) estimates that by 1950 only around 58% of the population in French colonies had their languages transcribed compared to 76-81% in British, Belgian and Portuguese colonies.

consolidation of ethno-linguistic consciousness and politicization did not result from the simultaneous *mass* consumption of newspapers and novels followed by state adoption and enforcement of national languages. Instead, missionary investments in language and printing in Africa instigated much more localized “imagined communities,” which were constructed and sustained by new cultural entrepreneurs (initially indigenous missionaries undertaking the language standardization) and by community members’ exposure to the translated Bibles, conversionary material and other printed texts in vernacular languages. These activities spurred ethnonational “awakenings” similar to what [Anderson \(1983, 73\)](#) describes in Europe—where the “energetic activities of...professional intellectuals were central to the shaping of nineteenth-century European nationalisms.” In Africa, the intelligentsia, of mostly mission-educated linguists, writers and teachers, transmitted ideas of groupness through the churches and schools and, in turn, created new ethnic elites who further promoted the group’s values and solidarity through literature, newspapers, and the formation of cultural associations ([Vail, 1989a](#), 11-12).

Another important difference with Europe was the role of the state. According to [Anderson \(1983, 76\)](#), the expansion of European states increased the importance of official languages and fostered the development of a bureaucratic middle class. At the same time, state-sponsored nationalisms promoted linguistic assimilation and national identities ([Weber, 1976](#)). In contrast, in colonial Africa, the state was run by Europeans with little interest in fostering an African class of bureaucrats, and, instead, focused on thwarting rather than promoting any kind of national identity, fearing the rise of revolutionary movements ([Vail, 1989a](#)).

The Yoruba represent a paradigmatic case of the impact of missionary language investments and publishing on the reconstruction of ethnic identity.¹⁹ With the collapse of the Oyo empire at the end of the 18th century, civil wars and slave raiding divided the Yoruba into rivalrous subgroups ([Adediran, 1984](#)). From the 1840s onward, however, missionaries from the Church Missionary Society (CMS), including freed slaves, such as Samuel Crowther, contributed to the rebuilding of the Yoruba ethnic nation. Intent on spreading Christianity, the CMS missionaries worked on Yoruba orthography, translation, and publishing, even starting a Yoruba newspaper in as early as 1859 ([Falola, 1999](#)). In propagating a standardized language and embracing and promoting the ethnonym ‘Yoruba’, the Christian missionaries boosted Yoruba ethnic *consciousness* ([Peel, 2003](#)). Moreover, as missionaries interpreted Yoruba history and tradition through a Christian lens (most famously Samuel Johnson in *The History of the Yorubas*), ethnogenesis and religious change reinforced each other. Consistent with [Vail \(1989b\)](#), missionary schools contributed to the propagation of standardized Yoruba through instruction in the language, which then produced new elites who served as champions of Yoruba solidarity and nationalism ([Usman and Falola, 2019](#)). This is personified in the life of Obafemi Awolowo, one of Nigeria’s founding fathers. Awolowo, born into one of the first Christian families in Ikenne, was educated in missions before leading a pan-Yoruba cultural association (Egbé Ọmọ Odùduwà) dedicated to “reinventing a common Yorùbá identity” ([Adebanwi, 2014](#)).

The standardization and printing of African languages is therefore expected to have strengthened groups’ ethno-nationalism and their mobilizational capabilities—with the rise of new ethnic elite and the writing and printing technologies they could wield as they competed in the political arena. In addition to strengthening groups’ polit-

¹⁹For other case studies, see [Ranger \(1989\)](#); [Chimhundu \(1992\)](#); [Strommer \(2015\)](#).

ical capacity, the print revolution likely contributed to more expansionary identities than cash crop agriculture as missionaries encouraged language uptake and provided opportunities for outsiders to learn the language via dissemination of language materials, church-related activities, and schooling. The upshot was the construction of more porous ethnic boundaries and assimilationist cultural practices—at least among those who adopted the group’s language.

Hypotheses

Following from our theoretical framework, we pre-registered the hypothesis that groups exposed to cash crops or print technologies are more likely to be politically relevant in the post-independence period. We also expected this to lead to more salient ethnic identities among individual group members. Despite these similar effects on ethnic politicization, we expected differential effects on boundary-making. We hypothesized that the commercialization of agriculture led to the construction of less porous ethnic boundaries than vernacular publishing and predicted lower rates of inter-ethnic marriage for the cash crop than for the publication treatment.²⁰

Data

In this section we describe the various sets of data we assemble to test our hypotheses. We explain the use of Ethnologue to derive units of analysis, describe our historical data on cash crops and African language publishing, and discuss our proxies for ethnic politicization, salience, and boundary-making.²¹

Historical and Geographic Data

Identifying Potentially Relevant Groups For a candidate list of nominal ethnic categories, we use Ethnologue, a reference source on living languages. Ethnologue aims to capture the *complete universe* of languages regardless of their social or political relevance, or demographic size (Simons and Fennig, 2017). Having been compiled from the 1950s onward, Ethnologue may nevertheless miss a few pre-colonial small or extinct ethno-language groups. However, selection issues seem minimal in comparison to datasets like AMAR, EPR, or Murdock (1959, 1967).²² Identifying potentially salient ethnic categories from Ethnologue restricts our focus to ethno-linguistic rather than racial, religious, or regional markers. The analytical consequences of this restriction are minimal since in our Sub-Saharan African sample practically all ethnic categories in EPR, PREG, Afrobarometer, and DHS are equivalent to, or combinations of, language families, languages, or dialects. Another advantage of Ethnologue is that its

²⁰We also pre-registered a set of ancillary hypotheses and analysis on homogeneous political preferences, inter-ethnic trust and ethnic conflict that we report in Supplementary Information IV.

²¹Data and replication scripts for all analyses in this article and the Online Appendix are openly available in the APSR Dataverse (Pengl, Roessler and Rueda, 2021). The replication folder also contains extended Supplementary Information with additional data descriptions and results.

²²AMAR and EPR rely on some indication of social or political relevance as a basis for inclusion. Murdock (1959, 1967) has a much smaller number of groups than Ethnologue. See Laitin (2000b, p. 142) on the advantages of using “language as a proxy for ethnicity.”

companion dataset, the World Language Mapping System (WLMS) provides maps demarcating linguistic homelands, which we leverage to spatially aggregate our cash crop data, survey-based outcome measures, and geographic control variables as described in detail below.

Cash Crops To measure cash crop production, we use a geospatial dataset on the primary commodity revolution in Africa from [Roessler et al. \(2020\)](#), drawing on a historical map produced by [Hance, Kotschar and Peterec \(1961\)](#). The map depicts the source locations of more than 95 percent of exports in 1957 across 38 states in sub-Saharan Africa.²³ Each primary commodity production point represents a value of \$289,270 in 1957 USD. The dataset covers 9 groups of cash crops,²⁴ 20 minerals and metals, and forest, animal and manufactured products.

Our main analysis focuses on the five main cash crops: cocoa, coffee, cotton, palm, and groundnut representing 80% of total cash crop production and no less than half of all exports in 1957 across the countries in our sample. In addition, these five crops were predominantly produced by African smallholders rather than European settlers or on plantations, which makes them more relevant for our stipulated causal mechanism than other resources. Our Supplementary Information (section III.5) presents additional analyses also including other crops and minerals and more precisely coding the mode of production for all country-crop combinations in the Hance data. Figure 1 maps the 4'651 locations that produced one of the five most important export crops.

Print Technologies and Publishing Data To capture exposure to print technologies, we draw on two library databases to construct a record of historical publishing at the language-level.²⁵ In combination with Ethnologue and WLMS, this represents the first ethnically linked and geocoded database of publishing in African languages throughout the colonial period and after independence.

The first source is a 1923 compilation of 2480 publications across 168 languages ([Rowling and Wilson, 1923](#)). It was intended to serve as a reference book for publications by Christian missionaries in Africa, including not just religious texts, but also dictionaries, grammar books, educational materials, and newspapers. It also provides contemporaneous estimates of the number of speakers per included language which we use to normalize the number of publications.

Our second source ([Mann and Sanders, 1994](#)) catalogues “collections of African language texts at SOAS, (...) the African Department of SOAS, the International Institute for African Languages and Cultures (...) and the International Committee on Christian Literature for Africa.” This source complements [Rowling and Wilson \(1923\)](#) especially given its greater temporal coverage. However, [Mann and Sanders \(1994\)](#) exclude grammars and dictionaries, which may have been particularly important for constructing salient ethno-linguistic communities. It is much less comprehensive on early printed materials, as it counts 50% less pre-1925 titles than [Rowling and Wilson \(1923\)](#). We thus use [Rowling and Wilson \(1923\)](#) as the main source in our analysis and present results using [Mann and Sanders \(1994\)](#) in the Online Appendix.

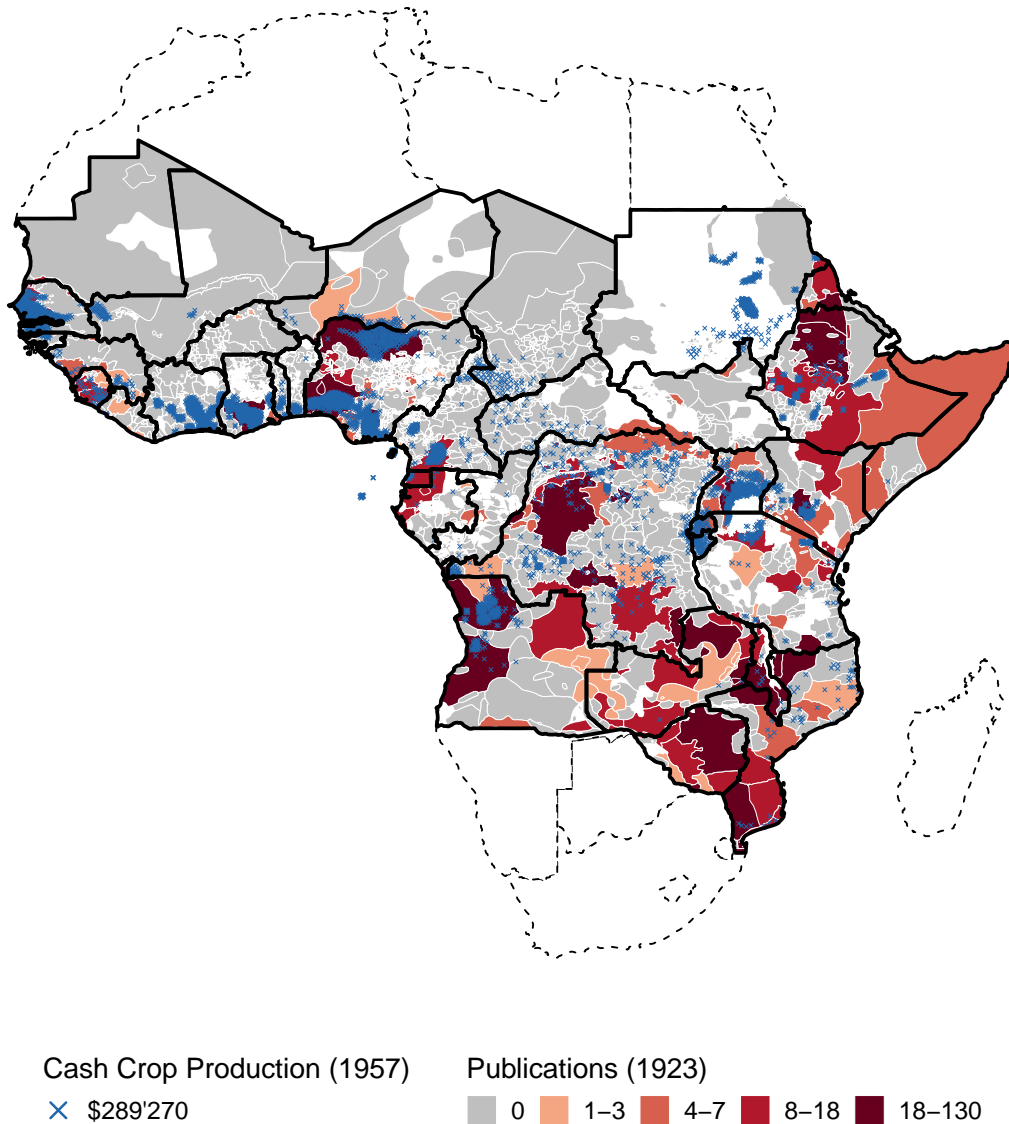
²³It excludes data on the Union of South Africa (including present-day Namibia), Madagascar and other island colonies.

²⁴Cocoa, coffee, cotton, groundnuts, oil palm, stimulants, other food crops, other industrial crops, other oils

²⁵This approach was inspired by [Chaney's \(2016\)](#) work on the Middle East.

The map in Figure 1 shows the total number of publications per ethno-linguistic polygon as listed in [Rowling and Wilson \(1923\)](#).

Figure 1: Publications and Cash Crop locations



Notes: Language homelands are mapped according to Ethnologue. Greyed regions are Ethnologue polygons for which there is no record of publications. Colors indicate the number of publications listed in [Rowling and Wilson \(1923\)](#). Each green circle locates 289,270 USD (1957) of cash crop export value. Solid black country borders describe our sample.

Contemporary Data on Ethnic Identities and Political Relevance

We use several data sources to measure the main outcomes of our study: ethnic politization and boundary-making at the group and individual level.

Group-level Politicization Measures To measure which Ethnologue groups serve as bases for contemporary political mobilization, we match Ethnologue to two expert-coded sources on ethnic groups’ relevance in national-level political competition post-independence: the Politically Relevant Ethnic Groups (PREG) (Posner, 2004a) and the Ethnic Power Relations (EPR) (Vogt et al., 2015) datasets. For each, we code a binary outcome indicating whether the Ethnologue group has a one-to-one match in PREG/EPR (e.g., Yoruba and Yoruba) *or* is a clearly identifiable part of a broader ethnic coalition coded as relevant by the respective dataset (e.g., the Gikuyu language as part of the Kikuyu-Meru-Embu coalition in EPR). All Ethnologue groups without any plausible exclusive or coalition match to the respective dataset are coded zero on the respective PREG or EPR outcome.²⁶

Individual-level Politicization Measures The salience of individual members’ ethnicity vis-à-vis other identities likely varies between and within ethnic groups. To analyze this, we use survey data from rounds 3–6 of Afrobarometer which ask respondents whether they identify more in ethnic or in national terms (Robinson, 2014; Ali et al., 2019). We use a dummy variable of whether a respondent identifies more strongly or even only in ethnic rather than national terms as outcome in our Afrobarometer specifications.

Boundary-Making A key dimension of boundary-making is a group’s accessibility to outsiders. Given the importance of marriage in social relations and group maintenance, many scholars view “endogamy [as] the ultimate measure of the salience of boundaries for intergroup relations” (Hechter, 1978, 304). The underlying assumption is that groups with more exclusionary boundaries are less likely to marry outside their group—and to develop norms against such practices. To calculate ethnic exogamy, we use USAID’s Demographic and Health Surveys (DHS) that includes data on the ethnicity of individuals and their spouses. These measures are described in more detail below.

Analysis I: Ethnic Politicization and Salience

We first report our specifications and results for the effect of cash crops and publishing on ethnic politicization at the group and individual levels.

Group-level Specification and Results

To test group-level effects, we estimate regression equation 1 using OLS.

$$\text{Pol}_{ec} = \beta_0 + \beta_1 \text{Cash Crops}_{ec} + \beta_2 \text{Publications}_e + \mathbf{X}'_{ec} \boldsymbol{\gamma} + \lambda_c + \varepsilon_e \quad (1)$$

²⁶In robustness checks, we also use more restrictive versions and only code Ethnologue groups with exclusive one-to-one matches as 1 and all other groups as 0. Supplementary Information I.3 provides an intuitive example of this distinction and Appendix Figure A1 shows results. We also use AMAR (All Minorities At Risk) to measure groups’ social relevance capturing group consciousness and shared norms and cultural features short of national-level political mobilization (Birnie et al., 2015, 112). See Appendix Figure A2 for results.

Pol_e measures the political relevance of Ethnologue group e in country c , using PREG or EPR. $Cash\ Crops_{ec}$ is a binary measure of historical cash crop cultivation in the Ethnologue polygon. $Publications_{ec}$ indicates whether [Rowling and Wilson \(1923\)](#) lists at least one publication in Ethnologue language e . λ_c represents country fixed-effects. X_{ec} is a set of standard geographic and historical controls including agricultural suitability, tsetse fly and malaria ecology, elevation, ruggedness, average yearly precipitation, average yearly temperature, distances (in logs) to the coast, to navigable rivers, to cities in 1900, to the country capital, historical missions, and to missionary printing presses, as well as absolute longitude and latitude.

Figure 2 reports the estimates of regression 1 when the outcome is a binary variable equal to one if the ethnic group is matched to a politically relevant group or coalition in PREG or EPR. Our baseline results indicate that, conditional on controls, a group with historical cash crop production is roughly 16-17 percentage points more likely to be listed as politically relevant in PREG and EPR (a 129% and 54% increase from the sample mean of the dependent variable, respectively). Similarly, languages with historical publishing are 11-13 percentage points more likely to be listed as politically relevant in PREG and EPR (a 88% and 45% increase from the respective outcome mean).

Potential endogeneity necessitates caution in causally interpreting the correlations reported in Figure 2. One important concern is that our results are driven by geographic or historical determinants of ethnic groups' take-up of cash crops and print publishing.²⁷ We employ several strategies to address this issue.

First, we instrument $Cash\ Crops_{ec}$ with indicators of suitability for cash crop agriculture and estimate the effects using a spatial-2SLS (S2SLS) strategy, following [Betz, Cook and Hollenbach \(2019\)](#). The instrument is the average agro-climatic suitability from the FAO GAEZ database across the five most important African cash crops (cocoa, coffee, cotton, groundnuts, and oil palm) in the homeland of ethnic group e . These suitability scores combine soil and climatic characteristics to predict the ecological potential to grow specific crops in rainfed agricultural systems. To serve as a valid instrument, suitability may only affect outcome variables through its impact on actual cash crop production. We argue that this exclusion restriction likely holds conditional on the rich set of geographic and historical controls in our models, especially general agricultural suitability, temperature, and precipitation which are included to isolate cash crop-specific effects from overall agricultural productivity and its social and political consequences.

The suitability instrument strongly predicts colonial cash crop production in first-stage regressions. The first-stage F-statistic is 13.5 in the EPR and 13.3 in the PREG models. To account for potentially similar spatial patterns in the instrument and outcomes that may threaten the exclusion restriction, the IV models further include a spatial lag of the respective political relevance outcomes instrumented with first and second-order spatial lags of the baseline controls ([Betz, Cook and Hollenbach, 2019](#)). All spatial lags are based on a binary contiguity matrix which defines ethnic group e 's neighbors as all other ethnic polygons within a 100 km centroid distance.²⁸ Line 3 in Figure 2 shows that S2SLS results remain similar to baseline OLS although confidence intervals naturally

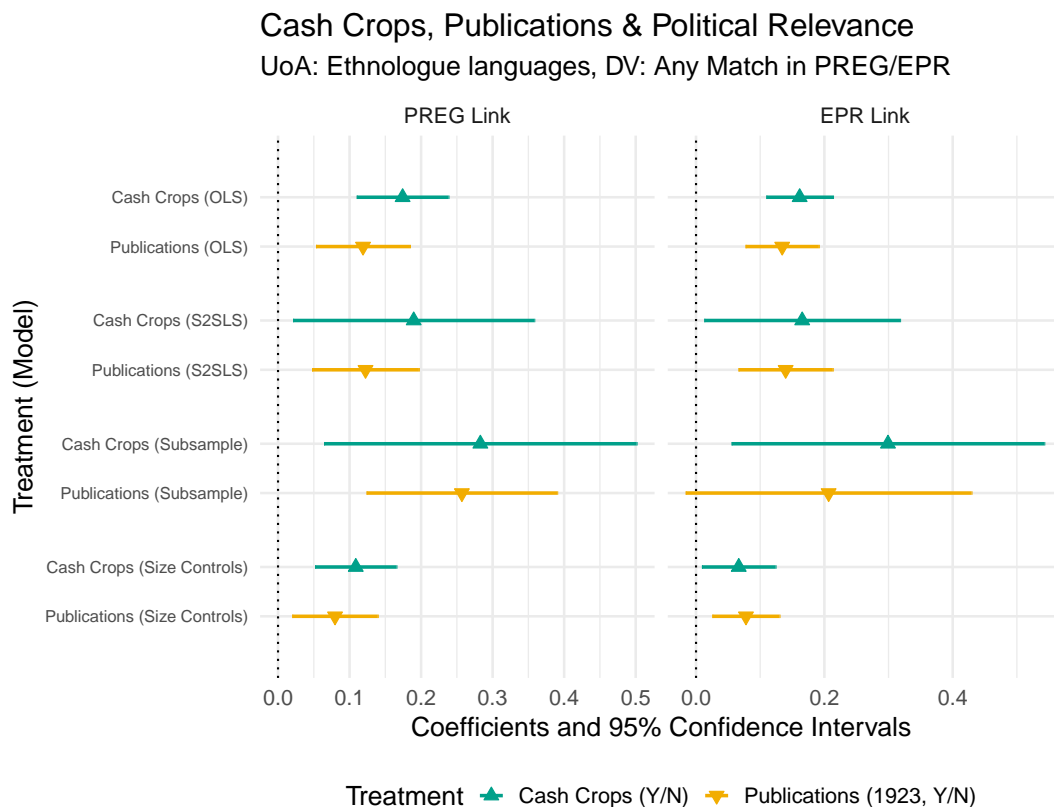
²⁷Figure I.9 in our Supplementary Information shows that groups with cash crops or publications systematically differ from those without on a number of baseline covariates.

²⁸The joint significance of spatially lagged baseline controls in the second first stage (predicting the spatially lagged dependent variable) is high and the respective F statistics remain well above conventional thresholds.

become wider.

A second endogeneity concern is that European missions tended to establish outposts in geographically-favorable areas or those with already more intensive colonial presence (Jedwab, zu Selhausen and Moradi, 2018). Subsetting the analysis to groups exposed to missions makes the analysis sample more comparable in terms of geographic fundamentals and other potential determinants of missionaries' targeting of specific groups and areas. The results, reported in line 6 of Figure 2 remain robust, despite the large reduction in observations and correspondingly large standard errors.

Figure 2: Cash Crops, Print Technologies, and Political Relevance



Notes: These figures summarise the results of eight regression models. The two binary outcomes indicate whether an Ethnologue group is matched to a group or coalition listed as politically relevant in PREG (left-hand panel) or EPR (right-hand panel). Lines 1-2 report effects using binary treatments indicating whether Ethnologue groups were exposed to cash crop production and/or print technologies. In lines 3-4, cash crops are instrumented with the mean agro-climatic suitability for the five most important export crops using the spatial 2SLS approach described in the text. In lines 5-6, the sample is restricted to Ethnologue polygons that experienced missionary activity. Lines 7-8 control for logged historical population per Ethnologue polygon based on HYDE raster data.

Individual-level Specification and Results

To test individual-level effects, we use survey data of expressed ethnic salience and estimate the following equation in a geographic and an ethnic variant:

$$\text{Sal}_{ielcs} = \mu_0 + \mu_1 \text{Cash Crops}_{kcs} + \mu_2 \text{Publications}_{kcs} + \mathbf{W}'_{ilcs} \boldsymbol{\gamma} + \eta_{k'} + \varepsilon_{ielcs} \quad (2)$$

$$k \in \{e, \ell\}, \quad k' \in \{\ell, cs\} \quad (3)$$

Sal_{iel} is a binary Afrobarometer-based survey measure of greater ethnic than national identification. The unit of analysis now is respondent i , who identifies with ethnic group e , residing in survey location ℓ in country c , and is interviewed in Afrobarometer survey round s . We assign Afrobarometer respondent i to ethnic group e based on the language they report speaking at home and use geographic information on Afrobarometer's survey locations ℓ to assign individual respondents to Ethnologue polygons.

In our geographic specifications ($k = \ell$), we use the cash crop production value within a 15 km radius of a survey location as treatment variable $\text{Cash Crops}_{\ell cs}$. The variable $\text{Publications}_{\ell cs}$ is the number of publications in the language of the local Ethnologue polygon normalized by historical group population as provided in [Rowling and Wilson \(1923\)](#). The geographic specifications thus assign treatment variables entirely based on respondents' place of residence and irrespective of their self-reported ethnic identity. The ethnic specifications ($k = e$), on the other hand, only use self-reported ethnic affiliation to assign treatments, regardless of individual locations. More specifically, for all members of group e in country c for survey s , Cash Crops_{ecs} is the value of historical cash crop production per sqm in the ethnic polygon of e , and $\text{Publications}_{ecs}$ is the number of publications in the language of e , again normalized by population.²⁹

Fixed effects η are either at the country-round level (cs) for geographic specifications or at the survey location-level (ℓ) for ethnic specifications. The main motivation for these two specifications is to separate location-specific from culturally-transmitted group-wide effects. Thus, the geographic specifications investigate whether respondents living in areas historically exposed to cash crop production and/or missionary publishing report more salient ethnic identities. The ethnic specifications examine whether members of historically exposed ethnic groups report salient identities, even when compared to respondents from other groups in the same location. Where not absorbed by location fixed effects, geographic and historical controls are the same as those stated in the previous section and always include an estimate of logged historical population from HYDE ([Klein Goldewijk, Beusen and Janssen, 2010](#)). In all Afrobarometer analyses, we also control for individual-level controls include gender, age, education levels, and indicators of standards of living.

Table 1 reports the results of our geographic specifications. A one standard deviation increase in the value of cash crop production around location ℓ increases respondents' ethnic identification by around 1.1% of a standard deviation (approx. 0.4 percentage points or 3% of the outcome mean). Similarly, a one-standard deviation increase in publications per capita increases ethnic identification by around 3.7% of a standard deviation (approx. 1.3 percentage points or 10% of the outcome mean). This effect is robust to intensive-margin comparisons (column 6). The effects are driven by ethnic stayers—individuals who reside in one of the Ethnologue polygons matched to their self-reported ethnic group e . Column 5 shows that restricting the analysis to ethnic leavers (those who reside outside of their ethnic group's homeland) results in a Null effect of print technologies and a significant negative effect of cash crops.

²⁹See Supplementary Information (Figure I.8) for a concrete example.

Table 1: Geographical Persistence in Ethnic Identity

	Geographic-level - Ethnic vs National Id					
	(1)	(2)	(3)	(4)	(5)	(6)
Cash crops USD pkm2	0.012*** (0.004)		0.011** (0.004)		-0.015** (0.007)	0.012** (0.005)
Cash Crops (S2SLS)				-0.0002 (0.007)		
Pubs pth pop (1923)		0.038*** (0.007)	0.038*** (0.007)	0.033*** (0.006)	-0.0005 (0.007)	0.050*** (0.007)
Individual controls	yes	yes	yes	yes	yes	yes
Historical and Geo controls	yes	yes	yes	yes	yes	yes
FE	Country-Round	Country-Round	Country-Round	Country-Round	Country-Round	Country-Round
Ethnic Stayer/Leaver	Both	Both	Both	Both	Leaver	Both
Mean dep. var.	0.1314	0.1314	0.1314	0.1314	0.1226	0.13
Observations	123,883	123,883	123,883	123,883	44,049	88,154
R ²	0.044	0.045	0.045	0.046	0.045	0.050

Notes: $p < 0.1$:*, $p < 0.05$:**, $p < 0.01$:***. The table reports standardized OLS estimates (“beta” coefficients). Standard errors are reported in parentheses and clustered at the location level. The dependent variable is a binary variable flagging whether respondents declare stronger ethnic than national identities. In Column 4, we instrument cash crop production with agricultural suitability to cash crop production using the spatial 2SLS approach described in the text. Column 5 restricts the sample to ethnic leavers. Column 6 restricts the sample to locations with at least one historical publication.

The results of the ethnic specifications are reported in Table 2. Among individuals within the same survey location, ethnic salience is significantly higher among those ethnic groups with a history of publishing. A one standard deviation increase in publications per (est.) thousand people increases respondents’ ethnic identification by around 1.0% of a standard deviation (approximately 0.3 percentage points, or 2.4% of the mean outcome, see columns 2 and 3). In contrast, historical cash crop production now has no significant effect.³⁰

Whereas cash crops increased ethnic salience only among stayers, publishing significantly elevates ethnic identities among movers (Column 4). This cultural mover effect is robust to intensive margin-only comparisons (Column 5). This suggests a culturally transmitted effect of print technologies—the formation of an “imagined community”—which persists even as individuals (or their ancestors) migrate.

Overall, we find that ethnic groups with higher levels of historical cash crop production and publishing are more likely to be politically relevant in the post-independence period and that individuals from these groups report more politically salient ethnic identities. The individual-level ethnic salience results suggest we are capturing two different channels of politicization—one tied to place and the other stemming from cultural transmission. That these correlate, respectively, with localized cash crop production and vernacular publishing increase our confidence that these historical processes were at least part of the causal chain shaping ethnic politicization in Africa.

³⁰These standardized effects are comparable or larger than other controls. For instance, $\hat{\beta}_2$ in Column 3 of Table 2 is six times larger than the absolute effect of a 10% increase in pre-colonial ethnic population, and roughly 20% smaller than the effect of formal primary schooling. Appendix Table A1 compares our coefficients with individual-level proxies used by Robinson (2014). The effect of a one-standard deviation change in our treatments of interest represent, across specifications, 20-56% of the effect of contemporary individual characteristics such as gender or formal employment. Also note that town fixed effects in Table 2 can aggravate attenuation bias (Aydemir and Borjas, 2011).

Table 2: Cultural Persistence in Ethnic Identity

	Ethnic-level - Ethnic vs National Id				
	(1)	(2)	All (3)	(4)	In Biblio (5)
Cash crops USD pkm2	-0.009 (0.008)		-0.010 (0.008)	-0.007 (0.010)	0.019 (0.017)
Pubs pth pop (1923)		0.010* (0.005)	0.010** (0.005)	0.016*** (0.007)	0.015* (0.009)
Individual controls	Yes	Yes	Yes	Yes	Yes
Historical and Geo controls	No	No	No	No	No
Fixed Effect	Town	Town	Town	Town	Town
Ethnic Stayer/Leaver	Both	Both	Both	Leaver	Leaver
Mean dep. var.	0.1314	0.1314	0.1314	0.12	0.12
Observations	120,341	120,854	120,300	44,040	31,241
R ²	0.203	0.202	0.203	0.263	0.268

Notes: $p < 0.1$:*, $p < 0.05$:**, $p < 0.01$:***. The table reports standardized OLS estimates (“beta” coefficients). Standard errors are reported in parentheses and clustered at the location level. The dependent variable is a binary variable equal to one if respondents declare stronger ethnic than national identities. Column 4 restricts the sample to ethnic leavers. Column 5 restricts the sample to ethnic leavers from groups with at least one historical publication.

Analysis II: Ethnic Boundary-Making

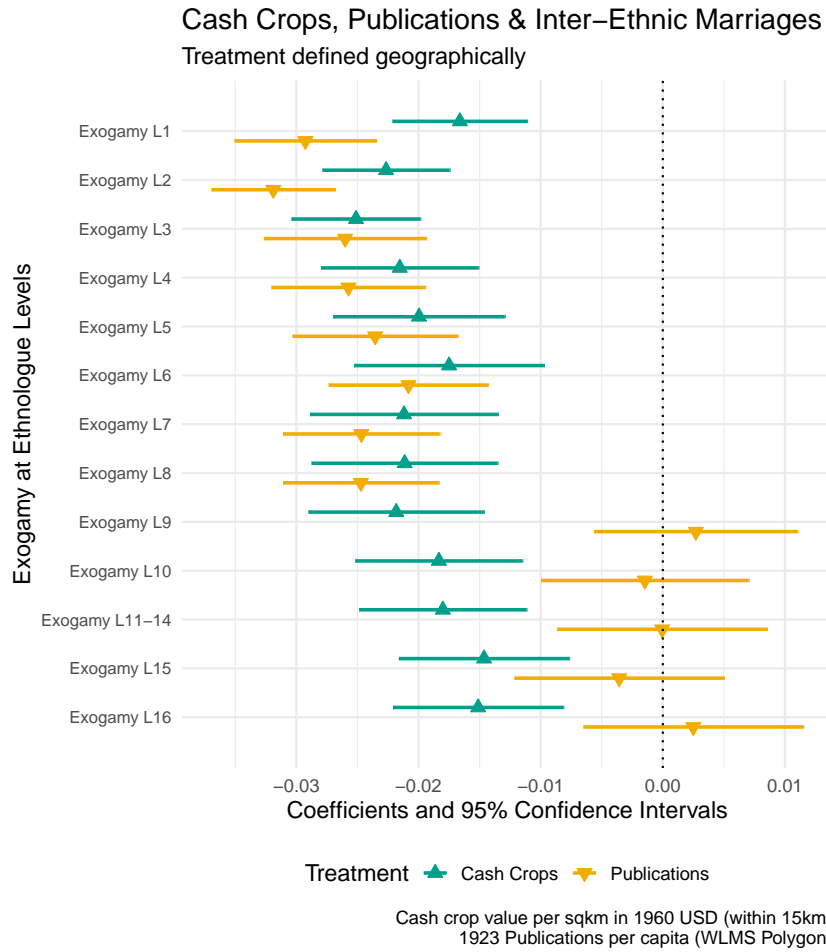
We now turn to analyzing ethnic boundary-making operationalized through inter-ethnic marriage. To measure ethnic exogamy, we take advantage of the couple recodes of the DHS household surveys which capture self-reported ethnic identities of married couples. The empirical specifications are equivalent to the Afrobarometer-based geographic and cultural persistence models above, but now the unit of analysis is interviewed couple i residing in location ℓ in country c with spouses identifying with ethnic group(s) e_f and e_m .

Knowing the appropriate match of practically all raw ethnic categories in DHS on the Ethnologue language tree allows us to analyze inter-ethnic marriages at different levels of ethno-linguistic differentiation.³¹ Ethnologue has 13 levels of language differentiation d in our Sub-Saharan African sample. Differentiation $d = 1$ distinguishes broad language families and as d increases, more closely related ethno-linguistic categories are separated. We therefore define 13 binary outcome variables Sal_{iel}^d indicating if the two spouses in respondent couple i self-report belonging to different ethnic groups at level of differentiation d .

Two examples from Nigeria illustrate the operationalization of our inter-ethnic marriage outcomes. A marriage between a female respondent identifying as Yoruba and a male Hausa respondent is coded as exogamous on all levels of the language tree. The Yoruba language belongs to the Niger-Congo language family, whereas Hausa is an Afro-Asiatic language. These language families are already separate on the first level and therefore Yoruba and Hausa do not share any nodes on the language tree. In contrast, a Yoruba-Igala couple is coded as endogamous on levels 1-6 and as exogamous thereafter. The Yoruba and Igala languages share the first six nodes of the language

³¹See [Cervellati, Chiovelli and Esposito \(2018\)](#) for a similar approach

Figure 3: Geographic Persistence - Cash Crops, Publications & Ethnic Marriages



Notes: The figure reports standardized OLS estimates from 13 regressions with country-round fixed effects. Standard errors are clustered at the survey location level. Each triangle represents the coefficient of geographically assigned cash crops and publications treatments, as described in the text. Bars represent 95% confidence intervals.

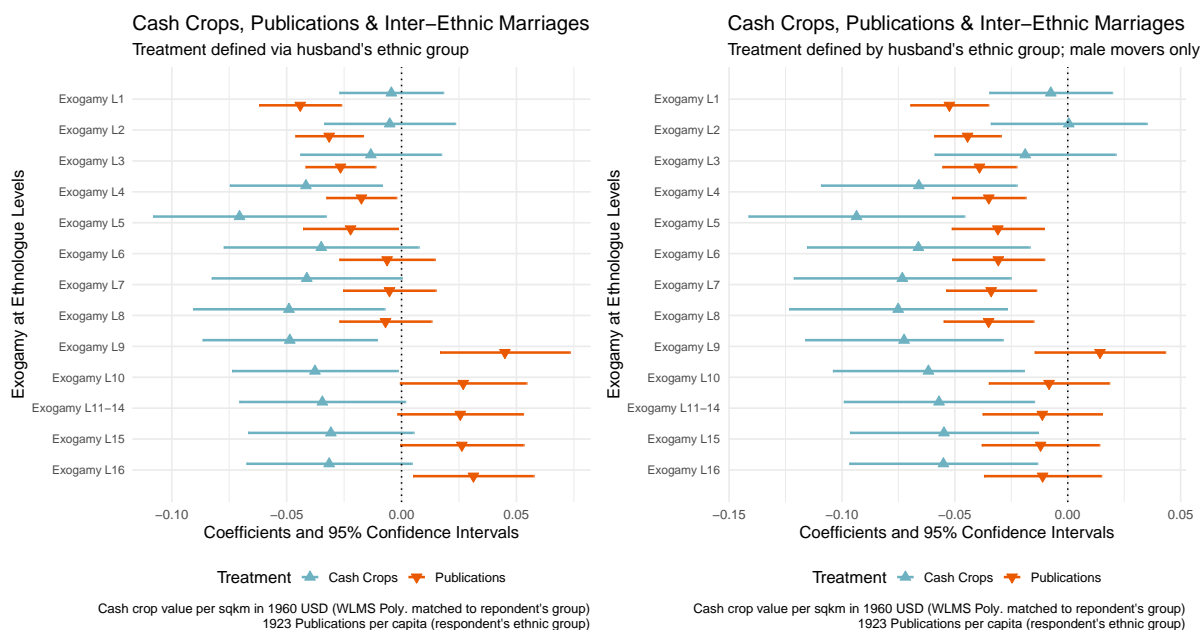
tree but then branch out in different directions.³²

If cash crop agriculture sparked a process of more exclusionary identities, we would expect lower inter-ethnic marriage rates at even the furthest branches of the language tree. A Yoruba respondent from a cash crop region would be similarly less likely to be married to a Hausa as to an Igala speaker. If print technologies led to salient but porous ethnic boundaries, we would expect members of these groups (e.g., Yoruba) to be less likely to choose a spouse from a linguistically distant group (e.g. Hausa), but still open to inter-marry with linguistically related ethnic others (e.g. Igala). We test these hypotheses for both the geographic and ethnic definitions of our treatment, as defined above.

Geographic Persistence. Figure 3 presents coefficient estimates from 13 models based on geographically assigned treatment variables. All 13 exogamy outcomes and both treatment variables are standardized to mean 0 and sd 1 to facilitate comparing coeffi-

³²Figures I.6 and I.7 in the SI schematically illustrate these examples.

Figure 4: Cultural Persistence - Cash Crops, Publications & Ethnic Marriages



Notes: Each triangle represents the standardized OLS estimates (“beta coefficient”) of ethnic-level cash crop and print technology treatments, as described in the text. The left panel is based on analyses of the whole sample while the right panel reports results from models run on the subsample of ethnic movers only. Bars represent 95% confidence intervals.

cient sizes across Ethnologue levels and treatments. The cash crop coefficients in Figure 3 are consistently negative and significant across all linguistic levels of differentiation. Inter-ethnic marriages are between 0.015 and 0.025 standard deviations less likely in locations with one standard deviation higher levels of late colonial cash crop production. While these effect sizes may appear small in standard deviation terms, their coefficients are, again, of similar magnitude as contemporary modernization proxies such as education and formal employment.³³ The coefficients on the publication variable are negative, significant, and somewhat larger in absolute size on levels 1–8 of the Ethnologue language tree. From level 9 onward, publication coefficients drop substantially and become statistically indistinguishable from zero. This pattern supports our theoretical conjecture that African language printing heightened the salience of ethnic identities but, compared to cash crop agriculture, led to more porous boundaries and more assimilation among linguistically close ethnic categories. We show in the Appendix (Figure A6) that, similar to the Afrobarometer analysis above, these geographic effects are driven by ethnic stayers.

Cultural Persistence. Figure 4 summarizes results from models that assign treatment variables by husbands’ ethnic identities and include location fixed effects.³⁴ The left-hand panel reports findings from analyses of the entire sample of couples for which both spouses’ ethnic identity was successfully matched to the Ethnologue language tree, whereas the right-hand panel restricts the sample to ethnic movers only and thus

³³See Appendix Tables A2-A5.

³⁴See Appendix Figure A7 for results when assigning treatments based on wives’ ethnicities.

compares marital choices by husbands outside of their ancestral homeland. These within-location models yield substantively similar results as the geographic persistence analysis above. Effect sizes and the level difference between historical cash crop production and African language publishing appear, if anything, more pronounced.

Robustness & Mechanisms

The empirical results in sections 3-5 suggest that: (i) historical cash crop production and the uptake of print technologies increased groups' mobilizational capabilities and political relevance in the post-independence period; (ii.) these historical forces also have had persistent effects on individual ethnic salience but through different channels—cash crop effects appear tied to land and sites of historical cultivation, and publishing effects stem from cultural transmission among members of the ethno-linguistic group; and (iii.) we observe differential effects on inter-ethnic marriage with linguistically proximate out-groups. Note that in contrast to the Afrobarometer models, we find cultural persistence (ethnic mover) effects of cash crops on ethnic marriages suggesting perhaps that political ethnicity is easier to change than deep-rooted cultural norms about appropriate marital choices.³⁵

In the remainder of this section, we summarize findings from our pre-specified analyses to account for potential endogeneity before presenting additional specifications that address a series of potential alternative explanations that might account for the observed empirical patterns.

Addressing Endogeneity

Across most analyses, we address threats that the effects of historical cash crop production and vernacular language publishing are endogenous to underlying geographic factors or ethnic groups' pre-colonial characteristics. The effects of cash crops on group-level politicization (Figure 2) and inter-ethnic marriages (Appendix Figure A4) are robust to instrumenting cash crop production with indicators of suitability in a spatial-2SLS setup.³⁶ To account for potential selection of missionary and publishing activities into certain areas or groups, we show the results are robust to restricting the analysis to Ethnologue groups with a Christian mission (Figure 2) and publishing at the intensive margin (Column 6 in Table 1; column 5 in Table 2; Figures A4 and A5). To address potential geographic confounders of publishing, the results presented in Table 2 and Figure 4 include location fixed effects. This increases our confidence that geographic confounders do not explain away exogamy patterns nor cultural persistence in ethnic identity.

Alternative Explanations

Group Size If larger groups were more likely to cultivate cash crops or have vernacular publications, our results may pick up their size-based advantages in coalition

³⁵This seems consistent with recent findings that local ethnic minorities face incentives to vote for the local majority candidate rather than one of their own (Ichino and Nathan, 2013).

³⁶Afrobarometer results disappear when using this approach. One explanation is the lower spatial coverage of Afrobarometer which has less than half the number of unique survey locations than DHS. In addition, Afrobarometer was geocoded ex post and location coordinates are probably less accurate.

formation (Bates, 1983; Posner, 2005). We account for this issue in several ways. First, the publications treatment in the survey analyses is normalized by the number of language speakers as estimated by Rowling and Wilson (1923). Second, we use the HYDE population rasters (Klein Goldewijk et al., 2017) to control for pre-colonial population per ethnic polygon across all three analysis sections (see above). As HYDE only imperfectly captures group-level population, Appendix Table B7 and Figure B8 add pre-colonial political centralization as a proxy for pre-colonial group size and political cohesion (Murdock, 1967). Results remain generally robust to accounting for group size, although coefficients get significantly smaller in the group-level political relevance models with the HYDE control.³⁷

Colonizer Effects We also show that the effects of cash crop agriculture and publishing on ethnic politicization and marriage patterns are not mere artifacts of British indirect rule (Ali et al., 2019). The results are reported in section III.4 of our Supplementary Information. We do observe that former French colonies have either zero or dampened publication effects, perhaps a consequence of France’s more hegemonic cultural and linguistic policies in its colonies (Albaugh, 2014; Cogneau and Moradi, 2014). These heterogeneous effects offer additional suggestive evidence of the importance of vernacular language standardization and its propagation through schools and churches as a key mechanism driving ethnic politicization.

Mechanisms We run causal mediation models (Acharya, Blackwell and Sen, 2016) to better gauge the mechanisms through which our historical treatments affect contemporary ethnic salience and exogamy. First, we observe that accounting for modernization proxies such as urbanization, education, and wealth does not explain our findings and, if anything, makes them stronger (Figures B12(a), B12(d), and B13). Second and in line with with Cagé and Rueda (2016), political engagement and public sphere variables from Afrobarometer explain up to 17% of the publications effect. Finally, historical group-level advantages in secondary and higher education account for relatively large shares of the publishing effect on inter-ethnic marriages (15-26% in geographic models, 16-43% in ethnic specifications, see Figure B14). These results, while only suggestive, point to the roles of an early intelligentsia in constructing ethnic identities and of continued political engagement in maintaining them.

Resource Types We expected cash crop agriculture to matter due to local ethnic competition for economic benefits and ethnic elites’ and communities’ strategic boundary-making. This mechanism is unlikely to play out under European-owned plantation or settler agriculture nor in mining regions where there were limited benefits for indigenous farmers or where the colonial state or concession companies regulated access. Consistent with this, we show in Supplementary Information III.5 that our results are mainly driven by smallholder crops predominantly cultivated by African farmers. The effects of historical plantation agriculture and mining are weaker or even point in the opposite direction.

³⁷Figures B10 and B11 further control for ethnic polygon area. Supplementary Information III.1.2 more closely investigates the relationship between group size and publications.

Diversity & Religion One concern about the inter-ethnic marriage results is whether they merely reflect differences in local ethnic diversity. In Supplementary Information III.6, we account for or interact our treatments with local-level ethnic fractionalization scores. The cash crop effects are larger in ethnically diverse locations strengthening our confidence that ethnic competition rather than local-level ethnic homogeneity explains lower exogamy levels.

Another possibility is that the publishing measure is merely picking up the spread of Christianity, which may explain politicization or marital choices. To rule this out, we control for Christian population share in the group-level models, re-run all exogamy models with directed religious couple fixed effects, and use religious denomination dummies in mediation models. Results are nearly identical to our baseline analyses (Supplementary Information III.7).

Conclusion

Our analysis across 35 countries shows that Africa's contemporary ethnic landscape was at least partially shaped by the persistent effects of the cash crop and printing revolutions that spread from the 19th century onward. In line with our hypotheses, geographic variation in cash crop agriculture and the uneven diffusion of print technologies differentially increased groups' mobilizational potential and their capabilities to compete for state power after independence. Our analysis of individual-level identity salience suggests that these two forces affected ethnicity through different channels—with cash crop effects on individual identity salience tied to historic agricultural zones and publishing effects transmitted culturally among language speakers even beyond their ethnic homeland. Beyond self-reported identity salience, we find that these socioeconomic transformations resulted in different patterns of inter-ethnic marriage. Publishing contributed to the construction of more porous boundaries than cash crop agriculture, leading to comparatively higher rates of inter-marriage with linguistically related out-group members. This points to important differences in boundary-policing among politicized groups based on their historical exposure to commercial agriculture and print technologies.

In shedding light on these endogenous processes, we highlight key underlying factors that may confound analyses of contemporary ethnic politics—such as contestation over land and cross-cutting languages.³⁸ These dynamics require greater attention among scholars of ethnic politics and conflict, especially in light of more recent waves of internal migration, climate change, and rising land pressures.³⁹ How these changes affect ethnic boundaries, not least between pastoral and agricultural groups, are important questions for future research.

Our findings also have important implications for understanding the impact of colonialism on ethnicity. Much existing scholarship emphasizes the top-down effects of colonial social engineering and indirect rule on ethnic politicization.⁴⁰ In contrast, our analysis demonstrates the importance of broader social and economic forces, which preceded colonialism and were key drivers of it. Further, our findings suggest that colo-

³⁸On these points, see respectively, [Boone \(2014\)](#) and [Laitin \(2000a\)](#).

³⁹See [Klaus \(2020\)](#) for a recent such example.

⁴⁰See for example [Mamdani \(1996\)](#) and [Posner \(2005\)](#) and more recently [Ali et al. \(2019\)](#), [McNamee \(2019\)](#), and [Müller-Crepon \(2020\)](#).

nialism did not uniformly mold or “fix” ethnic boundaries. Instead, identity (re)construction arose as much from the strategic actions of African farmers, landowners and elites as well as missionaries, culture brokers and ordinary people responding to opportunities and constraints brought about by economic and technological change in the 19th and 20th centuries.

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse:

<https://doi.org/10.7910/DVN/WQEQPN>

FUNDING STATEMENT

This research was generously supported by the Swiss National Science Foundation (grant #P0EZP1_159076, Pengl) and the US National Science Foundation (award #1628498, Roessler).

CONFLICT OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research.

ON HUMAN SUBJECTS

The authors affirm that this research did not involve human subjects.

References

- Acharya, Avidit, Matthew Blackwell and Maya Sen. 2016. "Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects." *American Political Science Review* 110(3):512–529.
- Adebanwi, Wale. 2014. *Yorùbá Elites and Ethnic Politics in Nigeria: Obáfemi Awólowo and Corporate Agency*. Cambridge University Press.
- Adediran, Biodum. 1984. "Yoruba Ethnic groups or a Yoruba Ethnic Group? A Review of the Problem of Ethnic Identification." *África (São Paulo)* 7(7):57–70.
- Ahlerup, Pelle and Ola Olsson. 2012. "The Roots of Ethnic Diversity." *Journal of Economic Growth* 17(2):71–102.
- Ake, Claude. 1993. "What is the Problem of Ethnicity in Africa?" *Transformation* (22).
- Albaugh, Ericka A. 2014. *State-Building and Multilingual Education in Africa*. Cambridge University Press.
- Ali, Merima, Odd-Helge Fjeldstad, Boqian Jiang and Abdulaziz B Shifa. 2019. "Colonial Legacy, State-Building and the Salience of Ethnicity in Sub-Saharan Africa." *The Economic Journal* 129(619):1048–1081.
- Anderson, Benedict O G. 1983. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Asiwaju, Anthony Ijaola. 1984. *Partitioned Africans: Ethnic Relations across Africa's International Boundaries, 1884-1984*. London: C. Hurst.
- Aydemir, Abdurrahman and George J. Borjas. 2011. "Attenuation Bias in Measuring the Wage Impact of Immigration." *Journal of Labor Economics* 29(1):69–112.
- Barth, Fredrik. 1969. *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*. London: Allen and Unwin.
- Bates, Robert H. 1974. "Ethnic Competition and Modernization in Contemporary Africa." *Comparative Political Studies* 6(4):457–484.
- Bates, Robert H. 1983. Modernization, Ethnic Competition, and the Rationality of Politics in Contemporary Africa. In *State Versus Ethnic Claims: African Policy Dilemmas*, ed. Donald Rothchild and Victor A Olorunsola. Boulder, CO: Westview Press pp. 152–171.
- Berry, Sara. 1993. *No Condition is Permanent: The Social Dynamics of Agrarian Change in Sub-Saharan Africa*. Maddison, WI: University of Wisconsin Press.
- Betz, Timm, Scott J. Cook and Florian M. Hollenbach. 2019. "Spatial Interdependence and Instrumental Variable Models." *Political Science Research and Methods* pp. 1–16.
- Birnir, Jóhanna K, David D Laitin, Jonathan Wilkenfeld, David M Waguespack, Agatha S Hultquist and Ted R Gurr. 2018. "Introducing the AMAR (All Minorities at Risk) Data." *Journal of Conflict Resolution* 62(1):203–226.

- Birnie, Jóhanna K, Jonathan Wilkenfeld, James D Fearon, David D Laitin, Ted Robert Gurr, Dawn Brancati, Stephen M Saideman, Amy Pate and Agatha S Hultquist. 2015. "Socially Relevant Ethnic Groups, Ethnic Structure, and AMAR." *Journal of Peace Research* 52(1):110–115.
- Boni, Stefano. 2006. Indigenous Blood and Foreign Labor: The 'Ancestralization' of Land Rights in Sefwi (Ghana). In *Land and the Politics of Belonging in West Africa.*, ed. Richard Kuba and Carola Lentz. Leiden: Brill.
- Boone, Catherine. 2014. *Property and Political Order in Africa: Land Rights and the Structure of Politics*. New York: Cambridge University Press.
- Boone, Catherine. 2017. "Sons of the Soil Conflict in Africa: Institutional Determinants of Ethnic Conflict over Land." *World Development* 96:276–293.
- Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria and Gerard i Miquel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya." *American Economic Review* 105(6):1817–1851.
- Cagé, Julia and Valeria Rueda. 2016. "The Long-Term Effects of the Printing Press in Sub-Saharan Africa." *American Economic Journal: Applied Economics* 8(3):1–31.
- Caselli, Francesco and Wilbur John Coleman. 2013. "On the Theory of Ethnic Conflict." *Journal of the European Economic Association* 11(1):161–192.
- Cervellati, Matteo, Giorgio Chiovelli and Elena Esposito. 2018. "Bite and Divide: Ancestral Exposure to Malaria and the Emergence and Persistence of Ethnic Diversity in Africa." CEPR Working Paper 13437.
- Ceuppens, Bambi and Peter Geschiere. 2005. "Autochthony: Local or Global? New Modes in the Struggle over Citizenship and Belonging in Africa and Europe." *Annual Review of Anthropology* 34:385–407.
- Chandra, Kanchan. 2004. *Why Ethnic Parties Succeed: Patronage and Head Counts in India*. Cambridge: Cambridge University Press.
- Chandra, Kanchan. 2006. "What is Ethnic Identity and Does It Matter?" *Annual Review of Political Science* 9:397–424.
- Chandra, Kanchan. 2012. *Constructivist Theories of Ethnic Politics*. New York: Oxford University Press.
- Chaney, Eric. 2016. "Religion and the Rise and Fall of Islamic Science." Working Paper.
- Chimhundu, Herbert. 1992. "Early Missionaries and the Ethnolinguistic Factor during the 'Invention of Tribalism' in Zimbabwe." *The Journal of African History* 33(1):87–109.
- Cogneau, Denis and Alexander Moradi. 2014. "Borders that Divide: Education and Religion in Ghana and Togo Since Colonial Times." *The Journal of Economic History* 74(03):694–729.
- Cohen, Abner. 1969. *Custom and Politics in Urban Africa: A Study of Hausa Migrants in Yoruba Towns*. Berkeley and Los Angeles: University of California Press.

- Colson, Elizabeth. 1971. The Impact of the Colonial Period on the Definition of Land Rights. In *Colonialism in Africa, 1870-1960: Profiles of Change*, ed. Victor Turner. Vol. III. London: Cambridge University Press pp. 193–215.
- Dunning, Thad and Lauren Harrison. 2010. "Cross-cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali." *American Political Science Review* pp. 21–39.
- Eifert, Benn, Edward Miguel and Daniel N Posner. 2010. "Political Competition and Ethnic Identification in Africa." *American Journal of Political Science* 54(2):494–510.
- Ekeh, Peter P. 1975. "Colonialism and the Two Publics in Africa: A Theoretical Statement." *Comparative Studies in Society and History* 17(1):91–112.
- Ekeh, Peter P. 1990. "Social anthropology and two contrasting uses of tribalism in Africa." *Comparative Studies in Society and History* 32(4):660–700.
- Englebert, Pierre, Stacy Tarango and Matthew Carter. 2002. "Dismemberment and Suffocation: A Contribution to the Debate on African Boundaries." *Comparative Political Studies* 35(10):1093–1118.
- Epstein, Arnold Leonard. 1958. *Politics in an Urban African Community*. Manchester: Manchester University Press.
- Eyoh, Dickson. 1999. Community, Citizenship, and the Politics of Ethnicity in Post-Colonial Africa. In *Sacred Spaces and Public Quarrels: African Cultural and Economic Landscapes*, ed. Paul Tiyambe Zeleza and Ezekiel Kalipeni. Trenton: Africa World Press pp. 271–300.
- Falola, Toyin. 1999. *Yoruba Gurus: Indigenous Production of Knowledge in Africa*. Trenton, NJ: Africa World Press.
- Fanon, Frantz. 1963. *The Wretched of the Earth*. London: Grove Press.
- Fearon, James D. 1999. "Why Ethnic Politics and 'Pork' Tend to Go Together." Working Paper.
- Fearon, James D. 2006. Ethnic Mobilization and Ethnic Violence. In *The Oxford Handbook of Political Economy*, ed. Donald A Wittman and Barry R Weingast. Oxford: Oxford University Press pp. 852–868.
- Fearon, James D and David D Laitin. 2000. "Violence and the Social Construction of Ethnic Identity." *International Organization* 54(4):845–877.
- Frankema, Ewout, Jeffrey Williamson and Pieter Woltjer. 2018. "An Economic Rationale for the West African Scramble? The Commercial Transition and the Commodity Price Boom of 1835–1885." *The Journal of Economic History* 78(1):231–267.
- Gellner, Ernest. 1983. *Nations and Nationalism*. Ithaca, NY: Cornell University Press.
- Habyarimana, James, Macartan Humphreys, Daniel N Posner and Jeremy M Weinstein. 2009. *Coethnicity: Diversity and the Dilemmas of Collective Action*. New York: Russell Sage Foundation.

- Hale, Henry E. 2004. "Explaining Ethnicity." *Comparative Political Studies* 37(4):458–485.
- Hance, William A, Vincent Kotschar and Richard J Peterec. 1961. "Source Areas of Export Production in Tropical Africa." *Geographical Review* 51(4):487–499.
- Hechter, Michael. 1978. "Group Formation and the Cultural Division of Labor." *American Journal of Sociology* 84(2):293–318.
- Hill, Polly. 1963. *The Migrant Cocoa-Farmers of Southern Ghana*. Cambridge, UK: Cambridge University Press.
- Hirschman, Albert O. 1977. "A Generalized Linkage Approach to Development, with Special Reference to Staples." *Economic Development and Cultural Change* 25:67.
- Hogendorn, Jan S. 1969. Colonialism in Africa 1870-1960. In *Colonialism in Africa, 1870-1960: The Economics of Colonialism*, ed. Peter Duignan and Lewis H. Gann. Vol. IV. London: Cambridge University Press.
- Hopkins, Anthony G. 1973. *An Economic History of West Africa*. New York: Columbia University Press.
- Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- Ichino, Nahomi and Noah L Nathan. 2013. "Crossing the Line: Local Ethnic Geography and Voting in Ghana." *American Political Science Review* 107(2):344–361.
- Jedwab, Remi, Felix zu Selhausen and Alexander Moradi. 2018. "The Economics of Missionary Expansion: Evidence from Africa and Implications for Development." Working Paper.
- Kalyvas, Stathis N. 2008. "Ethnic Defection in Civil War." *Comparative Political Studies* 41(8):1043–1068.
- Klaus, Kathleen. 2020. *Political Violence in Kenya: Land, Elections, and Claim-Making*. New York: Cambridge University Press.
- Klein Goldewijk, Kees, Arthur Beusen, Jonathan Doelman and Elke Stehfest. 2017. "Anthropogenic Land Use Estimates for the Holocene – HYDE 3.2." *Earth System Science Data* 9(2):927–953.
- Klein Goldewijk, Kees, Arthur Beusen and Peter Janssen. 2010. "Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1." *The Holocene* 20(4):565–573.
- Koter, Dominika. 2016. *Beyond Ethnic Politics in Africa*. New York: Cambridge University Press.
- Kramon, Eric, Joan Hamory Hicks, Sarah Baird and Edward Miguel. 2021. "Deepening or Diminishing Ethnic Divides? The Impact of Urban Migration in Kenya." *American Journal of Political Science* First View.

- Laitin, David D. 2000a. "Language Conflict and Violence: The Straw that Strengthens the Camel's Back." *European Journal of Sociology/Archives Européennes de Sociologie* 41(1):97–137.
- Laitin, David D. 2000b. "What is a Language Community?" *American Journal of Political Science* 44(1):142–155.
- Laitin, David D. 2007. *Language Repertoires and State Construction in Africa*. New York: Cambridge University Press.
- Laitin, David D and Daniel Posner. 2001. "The Implications of Constructivism for Constructing Ethnic Fractionalization Indices." *APSA-CP: Newsletter of the Organized Section in Comparative Politics of American Political Science Association* 12(1):13–17.
- Lentz, Carola. 2013. *Land, Mobility, and Belonging in West Africa: Natives and Strangers*. Bloomington, IN: Indiana University Press.
- Mamdani, Mahmood. 1996. *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton: Princeton University Press.
- Mann, Michael and Valerie Sanders. 1994. *A Bibliography of African Language Texts*. London: Hans Zell Publishers.
- Marshall-Fratani, Ruth. 2006. "The War of "Who is Who": Autochthony, Nationalism, and Citizenship in the Ivoirian Crisis." *African studies review* pp. 9–43.
- McNamee, Lachlan. 2019. "Indirect Colonial Rule and the Political Salience of Ethnicity." *World Development* 122:142–156.
- Michalopoulos, Stelios. 2012. "The Origins of Ethnolinguistic Diversity." *American Economic Review* 102(4):1508.
- Michalopoulos, Stelios and Elias Papaioannou. 2013. "Pre-colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81(1):113–152.
- Miguel, Edward. 2004. "Tribe or Nation? Nation Building and Public Goods in Kenya Versus Tanzania." *World Politics* 56(3):327–362.
- Miles, William FS and David A Rochefort. 1991. "Nationalism Versus Ethnic Identity in Sub-Saharan Africa." *American Political Science Review* pp. 393–403.
- Müller-Crepon, Carl. 2020. "Building Tribes: How Administrative Units Shaped Ethnic Groups in Africa." Working Paper.
- Müller-Crepon, Carl, Yannick I Pengl and Nils-Christian Bormann. 2020. "Linking Ethnic Data from Africa." Working Paper.
- Murdock, George Peter. 1959. *Africa: Its Peoples and Their Culture History*. New York: McGraw-Hill.
- Murdock, George Peter. 1967. *Ethnographic Atlas: A Summary*. Pittsburgh, PA: University of Pittsburgh Press.

- Nettle, Daniel. 1998. "Explaining Global Patterns of Language Diversity." *Journal of Anthropological Archaeology* 17(4):354–374.
- Nnoli, Okwudiba. 1978. *Ethnic Politics in Nigeria*. Enugu: Fourth Dimension Publishers.
- Nnoli, Okwudiba. 1998. Ethnic Politics in Africa: a Comparative Analysis. In *Ethnic Conflicts in Africa*, ed. Okwudiba Nnoli. Dakar.: CODESRIA pp. 1–26.
- Nunn, Nathan. 2008. "The Long-term Effects of Africa's Slave Trades." *The Quarterly Journal of Economics* 123(1):139–176.
- Nunn, Nathan and Leonard Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review* 101(7):3221–3252.
- Okoth, Assa. 2006. *A History of Africa: African Nationalism and the De-Colonisation Process*. Vol. II. Nairobi: East African Publishers.
- Oucho, John O. 2002. *Undercurrents of Ethnic Conflicts in Kenya*. Leiden: Brill.
- Paine, Jack. 2019. "Ethnic Violence in Africa: Destructive Legacies of Pre-Colonial States." *International Organization* 73(3):645–683.
- Parkin, Frank. 1974. Strategies of Social Closure in Class Formation. In *The Social Analysis of Class Structure*, ed. Frank Parkin. London: Tavistock Publications pp. 1–18.
- Peel, John David Yeadon. 2003. *Religious Encounter and the Making of the Yoruba*. Bloomington, IN: Indiana University Press.
- Pengl, Yannick I, Philip Roessler and Valeria Rueda. 2021. "Replication Data for: Cash Crops, Print Technologies and the Politicization of Ethnicity in Africa." Harvard Dataverse. Dataset.
URL: <https://doi.org/10.7910/DVN/WQEQPN>
- Posner, Daniel N. 2003. "The Colonial Origins of Ethnic Cleavages: The Case of Linguistic Divisions in Zambia." *Comparative Politics* 35(2):127–146.
- Posner, Daniel N. 2004a. "Measuring Ethnic Fractionalization in Africa." *American Journal of Political Science* 48(4):849–863.
- Posner, Daniel N. 2004b. "The Political Salience of Cultural Difference: Why Chewas and Tumbukas are Allies in Zambia and Adversaries in Malawi." *American Political Science Review* 98(4):529–545.
- Posner, Daniel N. 2005. *Institutions and Ethnic Politics in Africa*. New York: Cambridge University Press.
- Posner, Daniel N. 2017. "When and Why do some Social Cleavages Become Politically Salient rather than Others?" *Ethnic and Racial Studies* 40(12):2001–2019.
- Ranger, Terrence. 1989. Missionaries, Migrants, and the Manyika: The Invention of Ethnicity in Zimbabwe. In *The Creation of Tribalism in Southern Africa*. Berkeley and Los Angeles: University of California Press chapter 4.

- Robinson, Amanda Lea. 2014. "National versus Ethnic Identification in Africa: Modernization, Colonial Legacy, and the Origins of Territorial Nationalism." *World Politics* 66(4):709–746.
- Rodney, Walter. 1972. *How Europe Underdeveloped Africa*. London: Bogle-L'Ouverture Publications.
- Roessler, Philip. 2016. *Ethnic Politics and State Power in Africa*. New York: Cambridge University Press.
- Roessler, Philip, Yannick I Pengl, Robert Marty, Kyle Sorlie Titlow, Nicolas van de Walle et al. 2020. "The Cash Crop Revolution, Colonialism and Legacies of Spatial Inequality: Evidence from Africa." *CSAE Working Paper WPS/2020-12*.
- Rothchild, Donald S. 1997. *Managing Ethnic Conflict in Africa: Pressures and Incentives for Cooperation*. Washington, DC: Brookings Institution Press.
- Rowling, Frank and Charles-Edward Wilson. 1923. *Bibliography of African Christian Literature*. London: Conference of Missionary Societies of Great Britain and Ireland.
- Sasaki, Yu. 2017. "Publishing Nations: Technology Acquisition and Language Standardization for European Ethnic Groups." *The Journal of Economic History* 77(4):1007–1047.
- Simons, Gary F and Charles D Fennig. 2017. *Ethnologue: Languages of the World*. Dallas, TX: SIL.
- Spear, Thomas T. 1997. *Mountain Farmers: Moral Economies of Land & Agricultural Development in Arusha & Meru*. Berkeley and Los Angeles: University of California Press.
- Strommer, Martina A. 2015. Imagined Communities, Invented Tribe? Early Missionary Language Documentation and the Creation of the Herero. In *Colonialism and Missionary Linguistics*, ed. Klaus Zimmerman and Birte Kellermeier-Rehbein. Berlin: De Gruyter pp. 102–128.
- Stryker, Sheldon. 1980. *Symbolic Interactionism: A Social Structural Version*. San Francisco, CA: Benjamin-Cummings Publishing Company.
- Usman, Aribidesi and Toyin Falola. 2019. *The Yoruba from Prehistory to the Present*. New York: Cambridge University Press.
- Vail, Leroy. 1989a. Introduction: Ethnicity in Southern African History. In *The Creation of Tribalism in Southern Africa*, ed. Leroy Vail. Berkeley: University of California Press chapter 1.
- Vail, Leroy. 1989b. *The Creation of Tribalism in Southern Africa*. Berkeley: Univ of California Press.
- van den Berghe, Pierre L. 1981. *The Ethnic Phenomenon*. Amsterdam: Elsevier.
- Vogt, Manuel, Nils-Christian Bormann, Seraina Rügger, Lars-Erik Cederman, Philipp Hunziker and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." *Journal of Conflict Resolution* 59(7):1327–1342.

- Weber, Eugen. 1976. *Peasants into Frenchmen: The Modernization of Rural France, 1870-1914*. Palo Alto, CA: Stanford University Press.
- Weber, Max. 1978. *Economy and Society: An Outline of Interpretive Sociology*. Vol. I. Berkeley, CA: University of California Press.
- Wimmer, Andreas. 2013. *Ethnic Boundary Making: Institutions, Power, Networks*. Oxford: Oxford University Press.
- Woodberry, Robert D. 2012. "The Missionary Roots of Liberal Democracy." *American Political Science Review* 106(02):244–274.