The Seeds of State Capture: Merit and Patrimonialism in the Colonial Bureaucracy*

Sarah Brierley† Noah L. Nathan‡ George Kwaku Ofosu§ Tingxuan Zhu¶

This version: March 2026

Abstract

Bureaucracies vary in the extent to which hiring is meritocratic versus patrimonial. Such outcomes are typically seen as the result of politicians' behavior. However, politicians often delegate hiring to existing bureaucrats. When political leaders are socially distant from communities they seek to govern and these areas are peripheral to politicians' core interests, bureaucrats' private preferences can dominate hiring decisions instead. Colonial and other forms of external state-building typify cases where patrimonialism created by bureaucrats, not politicians, can emerge. Digitizing individual-level records across six decades of British rule in Ghana, we find evidence of patrimonialism rather than meritocracy in the hiring of African staff, with large inequities across ethnic groups. Our results demonstrate the underexplored agency of local bureaucrats in state-building and highlight conditions under which bureaucrats' preferences, rather than only political principals', shape state institutions. They also challenge conventional wisdom about why, and when, newly-independent African states became "neopatrimonial."

Word count: 14,021

*We thank Kweku Darko Ankrah, Benjamin Ayamga, Ruth Okumko, Manasi Rao, Katharin Tai, Michael Thompson-Brusstar, Erik Williams, Patrick Wu, and, especially, Naomi Tilles for research assistance. Jeremy Bowles, Mai Hassan, Nicholas Kuipers, Carl Mueller-Crepon, the Editor and anonymous reviewers, and participants at EPSA, IE University, Universidad Carlos III de Madrid, the London School of Economics and Political Science, MIT Global Diversity Lab, and WGAPE at Ohio State University provided helpful comments on earlier versions. Funding for initial digitization was provided by a University of Michigan LSA Social Science Research Award.

†Associate Professor, Department of Government, London School of Economics and Political Science. Email: s.brierley@lse.ac.uk.

‡Associate Professor, Department of Political Science, MIT. Email: nlnathan@mit.edu.

§Assistant Professor, Department of Government, London School of Economics and Political Science. Email: g.ofosu@lse.ac.uk.

¶PhD Candidate, Department of Political Science, MIT. Email: txzhu@mit.edu.

1 Introduction

Whether the bureaucracy is staffed on merit or through patrimonialism has major implications for governance (Weber, 1949[1919], Rauch and Evans, 2000, Pepinsky, Pierskalla, and Sacks, 2017). A vast literature assumes politicians choose between these equilibria, weighing the benefits of a more competent state against the rents available from raiding state institutions (Shefter, 1977, 1994; Geddes, 1994; Grzymala-Busse, 2007; Egorov and Sonin, 2011; Brierley, 2021). Once merit-based institutions have been formally adopted, scholars typically focus on the strategies that *politicians* adopt to regain control over hiring (Peters and Pierre, 2004; Colonnelli, Prem and Teso, 2020; Pierskalla and Sacks, 2020). But in practice *bureaucrats* are intimately involved in hiring decisions, either as a result of formal delegation or via informal processes. Bureaucrats can use these opportunities to hire staff based on their own preferences, which may deviate from politicians' (Hassan, Larreguy and Russell, 2024). Accordingly, hiring represents a principal-agent problem.

This principal–agent problem is magnified when politicians govern (i) societies from which they are socially distant and (ii) territories beyond their core interests. In such contexts, bureaucrats' preferences over hiring are likely to predominate because they are better informed about applicants and have stronger preferences over whom to hire than principals. These two conditions hold across a broad set of colonial (Herbst, 2000; Garfias and Sellars, 2024), external (Barkey, 1996; Matsuzaki, 2019; Lee, 2022), and subnational (Boone, 2003; Nathan, 2023) state-building scenarios in which existing theories may misattribute bureaucrats' agency in producing patrimonial versus merit-based bureaucracies to their political superiors.

Colonial-era development of central state bureaucracies in Africa often exemplified such dynamics. These central bureaucracies were distinct from the local governments ("Native Authorities") associated with indirect rule (Afigbo, 1972; Mamdani, 1996; Lange, 2009).[1] While initially tiny, central ministries oversaw increasingly complex service delivery and infrastructure portfolios in the later decades of colonial rule (Young, 2004; Lange, 2009; Cogneau and Mesplé-

[1]The separate local bureaucracies built for indirect rule are outside our analysis.

Somps, 2018). As incipient civil services grew, colonial leaders looked to fill positions with locals rather than Europeans. “Africanization" policies resulted in the hiring of the African bureaucrats who then served in state institutions inherited upon independence (Johnson-Kanu, 2021). European leaders hired agents from societies about which they lacked deep knowledge, while operating under monitoring constraints. In colonies that were neither vital to the colonial project nor in sustained revolt, European principals may have lacked strong preferences over which specific Africans to hire. In contrast, existing African bureaucrats possessed deep knowledge of their own societies and likely had strong recruitment preferences given both the high value of state employment and social pressures to reward family or co-ethnics with scarce opportunities (Ekeh, 1975).

Applying our theory suggests a reassessment of bureaucratic development in post-colonial states. Canonical literature argues that the rapid onset of mass politics can incentivize patrimonialism (Huntington, 1968; Shefter, 1977, 1994). In Africa, colonial hiring into central state bureaucracies is often argued to have been meritocratic – with hiring predicted by access to education (Mazrui, 1978; Ricart-Huguet, 2021; Lange, 2025); (neo)patrimonial practices within the central state are often said to have only emerged once post-independence governments faced new pressure to build winning coalitions (Zolberg, 1966; Collier, 1982; Young and Turner, 1985; Chabal and Daloz, 1999; van de Walle, 2001; Lange, 2009). However, most of these claims lack hard evidence about the alleged meritocracy of colonial hiring against which to compare post-independence developments. We instead propose that patrimonial hiring was already commonplace *during* colonial rule – driven by the actions of initial African bureaucrats – despite colonial leaders’ claims to have established meritocratic practices.[2] Consequently, in some cases, independence-era politicians may be better described as responding to *existing* (neo)patrimonialism than creating it.

To assess our theory of bureaucrat-led hiring, we turn to data from Ghana. We use ethnically-biased hiring as a proxy for patrimonial practices. Similar to Johnson-Kanu (2021) for Nigeria, we digitize bureaucrat-level records of the colonial civil service from the British National

[2]Interestingly, the parallel hiring of European officers was also at times non-meritocratic, shaped by political ties in the metropole (Xu, 2018).

Archives. Our main data span 1897 to 1955 – the eve of Ghana's independence. We identify 11,702 unique bureaucrats, including 4,805 Europeans and 6,879 Africans (Ghanaians). Leveraging the universe of names of all 13.8 million contemporary Ghanaian adults (registered voters as of 2015) geocoded against highly-localized ethnicity data from the 2010 census, we distinguish European from African (Ghanaian) bureaucrats and code the ethnicity of every Ghanaian bureaucrat based on their names.[3]

We document persistent hiring inequality throughout the colonial period. From the outset of African employment, two small ethnic groups - the Ga (3% of the population) and Fante (11%) - dominated the bureaucracy, consistently overrepresented and occupying over 70% of public sector positions until independence. We then focus our main attention on evaluating competing explanations for why this inequality emerged. Consistent with a logic of bureaucrat-led insider capture, we show that once any group gained a foothold within senior ranks in any state department (ministry), the chances others from that group were hired in that department rose sharply. Rather than meritocracy, the formal policy asserted by British principals, our evidence suggests patrimonial practices by ethnic groups who held first-mover advantages within the bureaucracy. The Ga and Fante could capitalize the most because they were the first to reach senior bureaucratic positions.

We use additional data to further demonstrate the insider capture mechanism and rule out three alternatives more commonly used to explain ethnic imbalances in incorporation into colonial states. First, we geolocate all schools starting from the 1890s, generating granular measures of both access to and the quality of education per ethnic group over time. Even though early advantages in education help explain the very initial hiring of small numbers of Ga and Fante bureaucrats in the 19th century, we demonstrate that the sustained dominance of these groups during Ghana's main Africanization period in the 1930s through 1950s cannot be explained by education: even many decades after other ethnic groups had caught up in educational opportunities, Ga and Fante bureaucrats continued to dominate new hiring. Employing individual-level results on civil service

[3]While similar hiring records exist for other British colonies, only in Ghana do we have access to the additional data needed to classify bureaucrats' ethnicity at scale.

examinations to test directly for meritocracy, we also show that Ga and Fante applicants were favored far beyond their supply of qualified candidates.

Second, we rule out that Ga and Fante dominance could be explained by their geographic proximity to state offices, showing their advantages extended to both desk jobs in the capital and field positions far from their home regions, and also that Ga and Fante bureaucrats were hired well out of proportion with the presence of state employment opportunities in their home areas. Third, we cast doubt on the presence of strategic British preferences to favor these groups, demonstrating instead that the patterns we document most likely resulted from hiring pressures internal to the bureaucracy itself.

We make several contributions. Theoretically, we advance insights about why bureaucrats are a crucial, but often-overlooked, actor in the development of merit-based or patrimonial bureaucracies (Anzia and Trounstine, 2024; Hassan, Larreguy and Russell, 2024) and suggest how bureaucrats' personal – as opposed to political – connections to other bureaucrats play a role in hiring (Harris et al., 2023). While we focus on a single case, we suggest that bureaucrats' preferences should be similarly important for understanding bureaucratic development across a broader range of examples of external (Herbst, 2000; Matsuzaki, 2019; Lee, 2022) or peripheral (Boone, 2003; Nathan, 2023) state-building. Our theoretical approach also advances an emerging literature exploring how the micro-dynamics of state-building under colonial rule shaped initial distributions of power in the post-colonial period (Harkness, 2018; Johnson-Kanu, 2021; Ricart-Huguet, 2021, 2022; Nathan, 2023; Kuipers, 2025; Lange, 2025), and implies that African employees may have had more agency in aspects of colonial state-building than typically understood.

Methodologically, by linking the universe of adult names in a country to geocoded census data, we introduce a new approach to code ethnicity from names – a common task for scholars of bureaucratic politics (e.g., Hassan, 2020, Kuipers, 2025).[4] Our method offers advantages over ex-

[4]There is already a vibrant methodological literature on how best to derive race from individual-level voter address data in the United States (Enos, 2016; Imai and Khanna, 2016; Imai, Olivella and Rosenman, 2022; Lee and Velez, 2025). We add to this literature in a non-US context, and develop an approach more suited to a broader range of settings where individual location data for the target population of names is absent.

isting dictionary-based, machine learning, or hand-coded approaches that have been used in other contexts in the Global South (Kasara, 2013; Harris, 2015; Bird and Michuda, 2025; Chaturvedi and Chaturvedi, 2024). Importantly, our approach (or one similar to it) can be used in any context in which researchers have access to local ethnic census data and the names of large numbers of individuals linked to approximate locations.

Our rich data also open a long-standing "black box" in the study of African states, with political scientists and historians united by the near-total absence of hard data on the hiring of local staff into colonial central bureaucracies.[5] Most studies have been confined to data on only a single state agency (e.g., Ray 2012) or instead to making broad generalizations about how indigenous populations were incorporated into states (e.g., Apter, 1963[1955]; Shaloff, 1974; Tignor, 1993; Rathbone, 2000*b*; Lange 2025). Opening this black box also creates opportunities to re-interpret the origins of the political cleavages that continue to shape African politics today. Consistent with Ray (2019), Kuipers (2025), and Lange (2025), we show how colonial-era inequities in hiring fed post-colonial demands on politicians to engage in ethnic favoritism to correct them.

2 Theory: bureaucratic hiring in the periphery

Most literature assumes that politicians' incentives and actions determine variation between meritocratic vs. (neo)patrimonial (patronage-based) bureaucracies. Politicians are commonly theorized as facing a "politicians' dilemma" or "loyalty vs. competency trade-off," in which they weigh the temptation to use state jobs as clientelist spoils to build support against the benefits of improving state capacity from hiring on merit (e.g., Shefter 1977, Geddes 1994, Grzymala-Busse 2007, Egorov and Sonin 2011, Hassan 2020, Brierley 2021).

But even after merit-based institutions have been adopted, politicians largely delegate the mechanics of hiring to bureaucrats. Despite the practical role bureaucrats play, they remain an under-theorized actor in state hiring. Scholars recognize that *de jure* merit-based procedures can

[5] Johnson-Kanu (2021) is the notable exception, but focuses on a different research question – path dependence between aggregate colonial-era hiring and Nigeria's modern bureaucracy – rather than variation *within* colonial rule.

be evaded, but almost exclusively focus on *politicians'* rather than bureaucrats' efforts to evade them (Peters and Pierre, 2004; Pierskalla and Sacks, 2020; Brierley, 2021). For example, literature on patronage typically focuses only on connections between politicians and potential state employees, with little attention to relationships between existing bureaucrats and new recruits.[6] Yet public administration scholars and economists have found that non-political connections to other bureaucrats can be more important in hiring and career advancement in some contemporary bureaucracies than connections to politicians (Harris et al., 2023; Cardoso et al., 2025). Nonetheless, Harris et al. (2023, p. 956) explain that non-political connections are a topic most "existing studies of connection-based hiring disregard."

Bureaucrats have hiring preferences that differ from politicians' (Hassan, Larreguy and Russell, 2024). Independent of politicians' incentives, existing bureaucrats sometimes prefer meritocracy. When faced with widespread political interference, bureaucrats are incentivized to seek an insulated, meritocractic public sector to protect job security (Shefter, 1977; Anzia and Trounstine, 2024). Moreover, bureaucrats may be individually "mission-driven" (Honig, 2024) or possess an *esprit de corps* in support of meritocratic principles (Weber, 1978; Evans and Rauch, 1999).

Alternatively, bureaucrats may prefer to hire co-workers and subordinates on the basis of social connections, with little regard for merit (Hassan, Larreguy and Russell, 2024).[7] This is often motivated by social obligations. Ekeh (1975) famously describes extensive social pressure to divert resources to family members and co-ethnics that many African state employees faced in the independence era. Such behavior is rational for bureaucrats, as failure to recognize one's corporatist responsibilities can result in social, psychological, and/or financial distress, as Price (1975) shows in post-independence Ghana. In many settings, social obligations are defined in terms of kinship and ethnicity.

[6]Indeed, common use of the term "patronage" in political science has narrowed over time from its original focus on discretionary hiring by a broad range of potential patrons, including bureaucrats (Grindle, 2012), to a focus only on the behavior of politicians (Kitschelt and Wilkinson, 2007; Stokes, Dunning and Nazareno, 2013; Oliveros, 2021).

[7]Relatedly, Ashraf and Bandiera (2018) summarize an extensive organizational economics literature on this behavior among existing employees who influence hiring in firms.

Beyond an explicit sense of obligation, preferences for group-based hiring can also be a function of social networks: individuals are more likely to interact and reciprocate friendship with people with whom they share identities, such as ethnicity (Wimmer, 2013*a*); if they then favor their own social ties in hiring as a result, ethnically-segregated social networks will produce ethnically-biased hiring, even in the absence of explicit discriminatory preferences. Beyond ethnicity, social networks may be defined by, for example, shared partisan or religious identities and similarly define bureaucratic hiring preferences (Wirsching, 2025).[8] Of course, some bureaucrats may also be motivated by explicit taste-based discrimination, which would produce observationally equivalent homophily in hiring (Wimmer, 2013*a*).

Whatever bureaucrats' motivations, unconstrained hiring based on social obligations or personal connections leads to the overrepresentation of recruits who share salient social identities with the incumbent bureaucrats who influence hiring decisions. These existing bureaucrats can become patrons to new recruits who attain positions through their largesse, a form of patrimonialism.

Divergence in preferences between politicians and bureaucrats results in a principal-agent problem. Bureaucrats can use privileged information they gain over vacancies and roles on hiring and interview boards to gain influence. Hiring outcomes depend on how much politicians are willing to invest to overcome this principal-agent problem. At one end of the scale, where politicians care deeply about who gets hired (or how; e.g., via merit), they are incentivized to develop costly processes to monitor bureaucratic agents' compliance. But in other situations, politicians may prefer a particular type of bureaucracy and invest some resources to promote compliance, but not enough to overcome bureaucrats' own incentives to deviate from politicians' instructions. In these cases, rarely explored in the existing literature, even if politicians do not cede *de jure* control of hiring to bureaucrats, bureaucrats' preferences can still take *de facto* precedence.[9]

[8]Bureaucratic preferences to hire in-group members may also stem from expectations of perceived gains to personal productivity from working with socially-close colleagues (Bozcaga, 2026; Hjort, 2024). This could also mean expectations of being better able to collude (corruptly) within homogeneous teams (Neggers, 2018).

[9]Alternatively, Hassan, Larreguy and Russell (2024) instead provide an example of institutions that explicitly delegate *de jure* independence to bureaucrats in some hiring decisions.

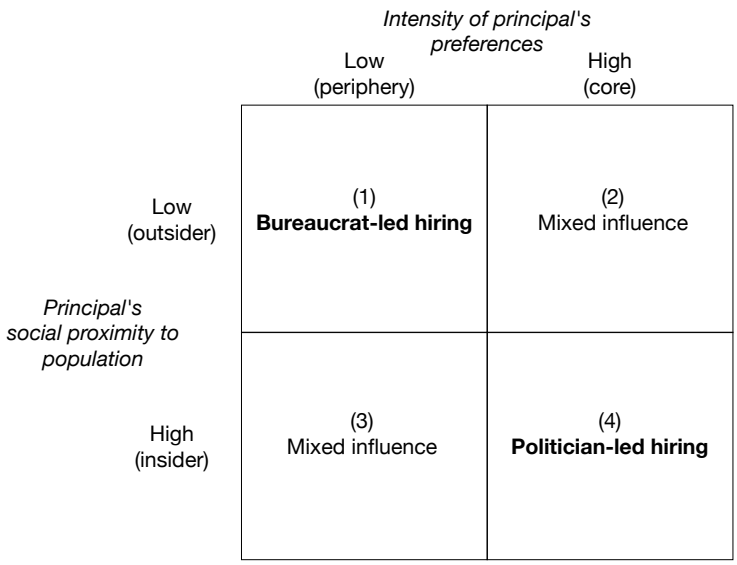


Figure 1: *Scope conditions for de facto hiring influence of politicians versus bureaucrats.*

Figure 1 conceptualizes two dimensions that magnify this problem. Bureaucrat-led hiring is most likely in state-building scenarios exemplified by the top-left cell (#1), where principals' costs of monitoring compliance are highest and incentives to invest in monitoring are lowest.

The vertical axis captures the degree to which political principals overseeing new bureaucracies are ethnically distinct from the populations from which they recruit agents. Principals' *social proximity* to bureaucrats (e.g., speaking the same language, knowing the context, having close relationships to local elites) makes internal hiring processes more legible and less costly to monitor. Proximity is high when principals are from the same society from which they hire agents (cells #3 and #4). Cells #1 and #2 instead capture situations in which principals are outsiders, and a costly monitoring problem emerges in which locally-embedded bureaucrats know much more about potential recruits than principals. Cells #1 and #2 may apply uniformly at the national-level in cases of colonial rule or "external state-building", as foreign powers (re)build state institutions in conquered territories (Migdal, 1988; Matsuzaki, 2019; Lee, 2022). Elsewhere, cells #1 and #2 arise subnationally as central leaders, social insiders in their home regions, instead expand the state out into socially-distinct peripheries (Barkey, 1996; Boone, 2003).

The horizontal axis of Figure 1 instead indicates the *intensity* of principals' preferences to construct a particular type of bureaucracy: how much principals care about seeing through a particular hiring outcome, both in an absolute sense and, especially, in comparison to the intensity of bureaucratic agents' own preferences. High preference intensity (cells #2 and #4) is often taken as given, especially when democratic elections incentivize politicians to use public sector jobs to increase state capacity or instead buy off local support to ensure their own political survival (Shefter, 1977; Geddes, 1994; Grzymala-Busse, 2007). In non-electoral contexts, preference intensity can also be high in territories with high levels of unrest that threaten to depose or eject the principal.

In contrast, low preference intensity (cells #1 and #3) should be especially common when principals seek to expand state institutions into territories that are not essential to their own political or economic interests – and thus that pose little accountability threat to principals' hold on power. This is magnified in the absence of electoral competition, as well as where either social unrest or economic value are low. With lower stakes, principals have less reason to invest extensively in monitoring bureaucrats' hiring decisions.

Many cases of Africanization of central colonial bureaucracies, especially in non-settler colonies, exemplify cases where European principals' high social distance from the communities they governed should have created high costs to monitoring agents' compliance with hiring directives, and in which principals' preferences to hire specific local agents were potentially weak relative to existing bureaucrats' (cell #1). European state-builders were constructing bureaucracies within local societies about which they often had limited knowledge. Meanwhile, colonial subjects could exert little accountability pressure, except through exceptional unrest. And while European firms certainly profited from extractive trade, few colonies were particularly essential to European economies.[10] As long as unrest was low and minimum levels of bureaucratic competence to extract were maintained, European principals may have seen it unnecessary to invest scarce resources in carefully monitoring compliance.

[10]European governments invested correspondingly limited resources in local governance (Herbst, 2000); before WWII British and French colonies were expected to be economically autonomous (Cogneau and Mesplé-Somps, 2018).

Meanwhile, initial African bureaucrats likely had strong incentives to prefer hiring based on their own social obligations and connections, rather than merit. Most African employees of colonial states sought these jobs because they were among the only, most secure, and best remunerated formal sector options in fledgling local economies. There is little reason to expect most would be especially "mission-driven," or committed to the success of the broader (racist and repressive) colonial project. Under these conditions, social obligations and connections should be expected to drive bureaucrats' private hiring preferences, consistent with Ekeh (1975). *De facto* bureaucratic-led hiring shaped by existing bureaucrats' private imperatives could then have quietly become rampant, in spite of colonial leaders' public pronouncements declaring adherence to meritocratic principles (Price, 1975; Cooper, 1996).

By contrast, in settler colonies, the presence of European residents increased direct political accountability of colonial leaders to (segments of) the population, and also increased risks (to Europeans) from mismanaging African discontent. This should have increased the intensity of principals' preference for merit-based criteria to improve performance (even if only performance at repression). Moreover, settlers reduced social distance between colonial leaders and potential employees, and likely created a further interest among European leaders in merit-based procedures that would disproportionately benefit local Europeans, given their advantaged education.

Even though they lacked elections, settler colonies may thus better correspond to more standard models of politicians' preferences over hiring, best captured by cell #4 of Figure 1, where politicians are better-informed about and have strong political incentives to control hiring. Off-diagonal cases (cells #2–#3) instead capture situations in which hiring outcomes reflect mixed influence of both politicians and bureaucrats. These include external state-building in territories vital to the center's economic interests, or prone to very threatening unrest; although monitoring challenges persist, politicians have clearer incentives to invest in costly oversight (cell #2). Similarly, autocratic state-building in contexts where locals share ethnic ties with the principal – reducing social distance – but exert little to no accountability pressure may also generate mixed patterns of influence (cell #3).

In this paper, we seek to establish that cell #1 exists as well by focusing deeply on a single case where new data sources provide a window into bureaucrat-led hiring that previous studies have missed. While the lack of similar data in other cases means it is not possible for us to examine variation across all cells of Figure 1, the figure offers a broader theoretical map of how the social proximity and strategic commitment of political principals should change the structure of bureaucratic hiring across cases.

3 The colonial state and Africanization

Our argument also suggests reinterpreting classic claims about African states. By the end of Africa's post-colonial era ($\sim$the 1990s; Young, 2004) a broad consensus characterized central state bureaucracies as "neopatrimonial," rife with ethnic patronage and limited merit in hiring (Young and Turner, 1985; Bayart, 1993; Chabal and Daloz, 1999; Englebert, 2000; van de Walle, 2001).[11]

This stands in contrast to common claims about the central bureaucracies of colonial states. In many countries, central bureaucracies inherited at independence had already begun Africanization. While many scholars describe indirect rule, a core strategy of *local-level* colonial governance, as inherently patrimonial (Mamdani, 1996; Lange, 2009), scholars frequently suggest that management of *national-level* state agencies in the colonial period instead more tightly fits Weber's bureaucratic archetype (Weber, 1978, p. 220). Young and Turner (1985, p. 27), for example, write that "colonial authoritarianism [in Africa] was exercised through *an autonomous, impersonal bureaucracy*."[12] Lange (2009, p. 4) describes that "central legal-administrative institutions in indirectly ruled colonies were relatively bureaucratic," explicitly referencing Weber's "bureaucratic" type (p. 35).[13] van de Walle (2001, p. 134) observes that "meritocratic hiring and promotion policies... were directly undermined by governments *in the years after independence*", implying their greater presence under colonialism.

[11] This consensus was not without critics (Pitcher, Moran and Johnston, 2009; Mkandawire, 2015).

[12] Emphasis added here and to subsequent quotes in this section.

[13] Lange (2009, 31) explicitly contrasts "patrimonial institutions in peripheral regions" created via indirect rule to "central state institutions... organized bureaucratically."

Existing accounts mostly portray (neo)patrimonialism in central state ministries as something post-colonial leaders did, not inherited.[14] For example, Young and Turner (1985) describe a "distinctive" shift towards patronage after independence (p. 27). Chabal and Daloz (1999) date the emergence of neopatrimonialism to "after independence" (p. 51). van de Walle (2001) writes that, "[t]he *first blows* against state capacity came *early after independence*" (p. 133) dating "the *first step* in the politicization and loss of professionalism of the civil service" to post-colonial presidents (p. 134). Lange (2009, p. 108) describes Sierra Leone similarly.

Colonial central bureaucracies often had *de jure* trappings of meritocracy. However, assertions that hiring was *de facto* meritocratic still remain unsubstantiated – tantamount to simply taking colonial officials' word for it. Beyond the senior-most ranks, relevant data simply have not been available (e.g., Shaloff, 1974). Lacking the data we analyze, Rathbone (2000) observes, for example, that analysis of African hiring beyond the most elite strata of bureaucrats is difficult because, "the exact figures are hard to come by" (p. 71). Others, such as Lange (2009, p. 39), reference that hiring into central bureaucracies was not fully "inclusive" of societal groups, but lack any data with which to empirically assess why or how much. And, critically, such a lack of inclusivity could still be consistent with meritocratic hiring as long as there were inter-ethnic inequalities in access to education (Kuipers, 2025).

Our argument suggests instead that despite adopting impartial institutions, colonial leaders likely had little incentive to ensure African hiring was truly meritocratic. Exploiting information asymmetries and the well-documented unwillingness of colonial principals to make costly investments in colonial states, initially-hired African bureaucrats could have had wide latitude to engage in patrimonial practices that created limited inclusiveness. Wherever Africanization predated independence, the well-documented behaviors of Africa's post-colonial leaders vis-a-vis their central states may then have been a response to patrimonialism and ethnic biases *already* pervasive within state institutions they inherited, not a new bureaucratic politics emerging after independence.

[14]This narrative dates as far back as Zolberg's (1966) account of immediate independence-era politics (Pitcher, Moran and Johnston, 2009).

4 The public sector in the Gold Coast

Ghana's pre-independence state, the Gold Coast, provides an ideal case for understanding hiring of African bureaucrats for several reasons. First, similar to many other African colonies, the Gold Coast had a "bifurcated colonial [state] based on two radically different organizational principles" (Lange, 2009, p. 4), combining indirect rule through local elites as the main method of local governance in the countryside with a nominally-Weberian national-level civil service, our focus here.

Second, the central bureaucracy experienced significant colonial-era Africanization, with over half of senior and mid-level positions already filled by Ghanaians by independence. This is roughly comparable to other major African colonies, such as Nigeria, where 61% of senior civil service posts had been Africanized by independence (Lungu, 1980), or Botswana, where Africans occupied 39% of senior and middle-level posts in 1964, two years before independence (Goldsmith, 1999, 538). Similarly, Africans filled between one quarter and one third of middle- and high-level public service positions in Kenya, Tanganyika, and Uganda (Simson, 2017, 28). In some other settings, however, there were much lower shares of African recruits.[15] Third, similar to most other colonies (Young, 1994), central bureaucracies built under colonialism continued largely uninterrupted, becoming the post-colonial central state. And fourth, significant ethnic competition over positions within these bureaucracies, consistent with neopatrimonialism, occurred after independence (Asante and Gyimah-Boadi, 2004).

Ghana had contact with Europeans at trading forts along its southern coast since the 15th century. In 1821, the British took control of these forts, setting the stage for establishment of a larger colony in 1874, which continued until 1957. The colony only comprised coastal areas until the end of the 19th century, when the British finally conquered Asante and expanded throughout the interior. While the Gold Coast was fiscally successful, especially relative to other African colonies (Austin, 2008, 1011), the climate ensured it remained a non-settler colony. The British

[15] At the senior ranks, Africans occupied only 4 percent of positions in Zambia (Lungu, 1980). In the Democratic Republic of Congo, there were only *three* Africans in the top three civil service categories (compared to 4,642 Europeans) (Young, 1966, 34). Both colonies nonetheless employed large numbers of Africans at lower ranks.

initially invested relatively little in administration and sought to keep expenditures and European staffing low (Sederberg, 1971), but the central bureaucracy expanded greatly in size in the final decades of the colonial period, as the British belatedly took on a more “developmentalist" outlook, especially after WWII (Young, 2004). As the central state grew, locals with Western education eventually filled the gap in staffing.

4.1 Supply of African bureaucrats

On the supply-side, the pool of potential African recruits was initially produced by efforts of European merchants and then Christian missionaries, typical of many British African colonies (Frankema, 2012). Merchants established schools as early as the 1600s within the trading forts. With few students until the 19th century, fort schools mainly educated children born to European merchants and coastal women (Graham, 1971). Enrollment only began to climb after the British takeover in 1821, when the Crown began funding some coastal schools and missionary schools also began to rapidly proliferate beyond zones of British control (Graham, 1971).

The early concentration of schools around Cape Coast meant that the Fante – a sub-group within Ghana’s plurality Akan – had a head start in education before colonial rule. The presence of European firms around the forts increased jobs available to locals and spurred enterprise, and with it class differentiation (Arhin, 1983). By the 1830s, a “genuine mercantile class” (Arhin, 1983, 16) of Fante elites had emerged (Kelly, 1959). Many members of this emerging Fante intelligentsia bore European surnames, owing to their descent from early European traders. The Ga in Accra, another site of trading forts, also enjoyed precolonial education access. Similar to the Fante, a new Westernized elite of clerks and merchants had emerged by the 1820s, educated in schools at or near the forts (Parker, 2000, 32-33).

For the very earliest periods of British state-building in the 19th century, Fante and Ga communities thus enjoyed an advantage in supply of potential job candidates. But by the early 20th century, before significant African hiring into the state had occurred, these advantages had already begun to recede, and many ethnic groups had large numbers of school graduates eligible for state employment. The 19th century expansion of mission schools meant access to education

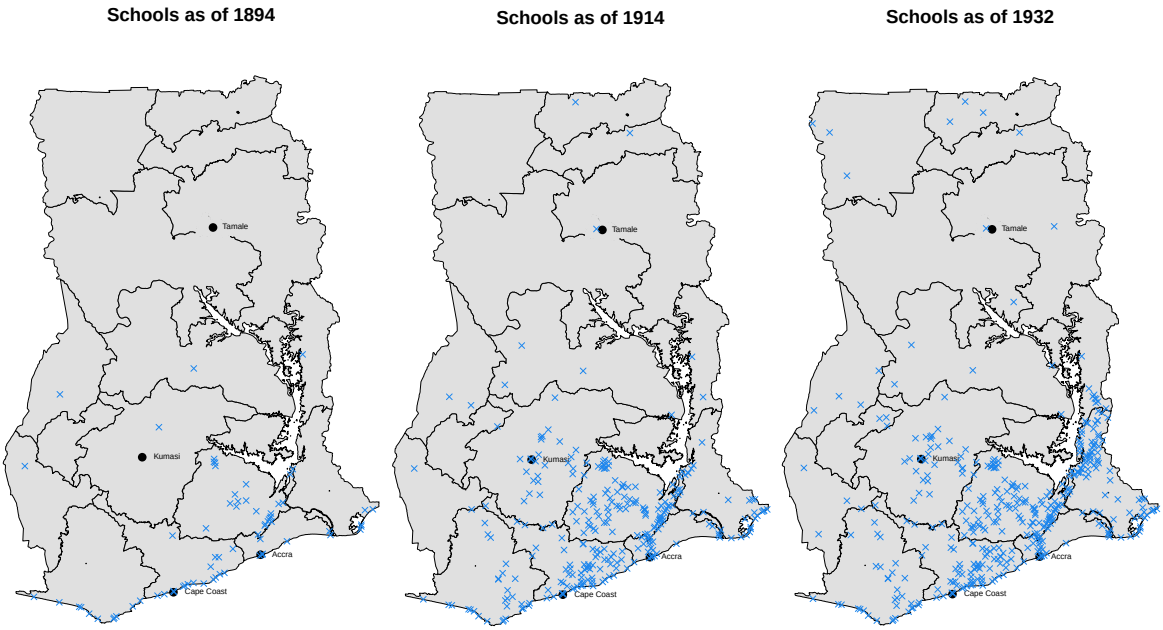


Figure 2: *Location of all schools, primary and secondary (1894-1932).*

already began to equalize throughout Southern Ghana around the turn of the century. School access then expanded rapidly into the interior in the early 20th century, generating an "extremely broad primary base" (Austin, 1964, 13). This is shown in Figure 2 based on geolocating annual lists of schools published through the early 1930s. Only the Northern Territories – modern-day Northern Ghana – was persistently neglected (Nathan, 2023). Crucially, expansion of the education system meant that access was already relatively equal across many groups *decades before* Africanization of the central bureaucracy occurred (1930s-1950s).[16]

4.2 Demand for African bureaucrats

At the apex of the state was the (British) Governor, under whom sat advisory Executive and Legislative Councils with some elite African participation (Price, 1967; Ward, 1966). Under these councils sat the civil service, the core national bureaucracy, divided into departments (ministries or agencies). The number of departments and civil servants increased over time as the state became more complex, with especially steep growth after WWII, similar to many other colonies (Lange,

[16]Moreover, we show in Figure 9 below that school quality was also relatively equal.

2009, p. 42). Separately, local governance occurred via indirect rule, eventually formalized through the creation of much smaller bureaucracies of local clerks to provide support to district-level local governments headed by chiefs. Our focus throughout is instead on the central civil service.

Initial state institutions in the 19th century relied heavily on African staff. von Hesse (2024) documents how Euro-African merchants among the precolonial Ga and Fante elite were in fact "co-architects" (p. 5) of initial colonial institutions, helping to personally fund and physically build the first outposts of the colonial government in the mid-19th century. In 1850, James Bannerman – an educated Ga merchant – even served as the interim Governor of the Gold Coast, "the only British colonial governor of known African ancestry" across the British empire (von Hesse, 2024, 3). Later on in the 19th century, educated Fantes began to dominate the lower ranks of the very nascent civil service, serving as aides to British officers (Kelly, 1959, 33), while a small set of educated Euro-African Fante and Ga elites rose to senior state positions alongside British counterparts. However, by the turn of 20th century, racist British policies against promoting Africans solidified and the vast majority of top positions were reserved for expatriates (Shaloff, 1974).

It was only under Governor Guggisberg (1919-1927), in 1926, that an official African-ization scheme was first adopted, justified in terms of ethical need for local representation, and especially for saving costs because Africans were paid less. Yet despite the new scheme, there continued to be little recruitment of Africans into senior positions. Foreshadowing the agency loss we examine below, European bureaucrats resisted their principal's order to Africanize in the interwar years for "equal parts of racism and self-interest" (Shaloff, 1974, 494), seeking to protect their own jobs.[17] However, as we show below, outside the frame of reference of the existing historical scholarship (e.g., Shaloff, 1974), Africanization became extensive in the 1930s at more junior ranks. During the 1930s and early 1940s, British civil servants pleaded with their superiors that hiring Africans into the senior-most positions was impossible because an insufficient number of well-trained candidates was available (Shaloff, 1974; Rathbone, 2000), but there was already

[17]Note that this dynamic itself already suggests the presence of bureaucrat-led hiring against principals' wishes.

a large *oversupply* of eligible African school graduates for mid- and lower-tier positions. Civil service jobs were scarce and highly-valued among the local population (Hutchful, 1985).

Africanization of the top ranks finally took off after WWII. Between 1949 and 1954, the share of Africans in senior positions rose from 14 to 38 percent (Cooper, 1996, 448). Over the 1950s the rate of Africanization accelerated, such that just after the Gold Coast gained independence, 69% of the 2,864 senior posts were held by Africans (Rathbone, 2000, 80). Importantly, this final wave of hiring occurred after Kwame Nkrumah, Ghana's independence leader, won an election in 1951 and gained control of internal affairs before independence in 1957.

5 Data sources

We unpack potential ethnic hiring in Ghana's colonial bureaucracy using data from multiple sources. First, we digitize data on the colonial civil service between 1897 and 1955 from the British National Archives. These staff lists provide names, positions, salaries, and appointment dates for central state employees, including headquarters and field staff of each ministry or agency.[18] Figure 3 displays an example extract.

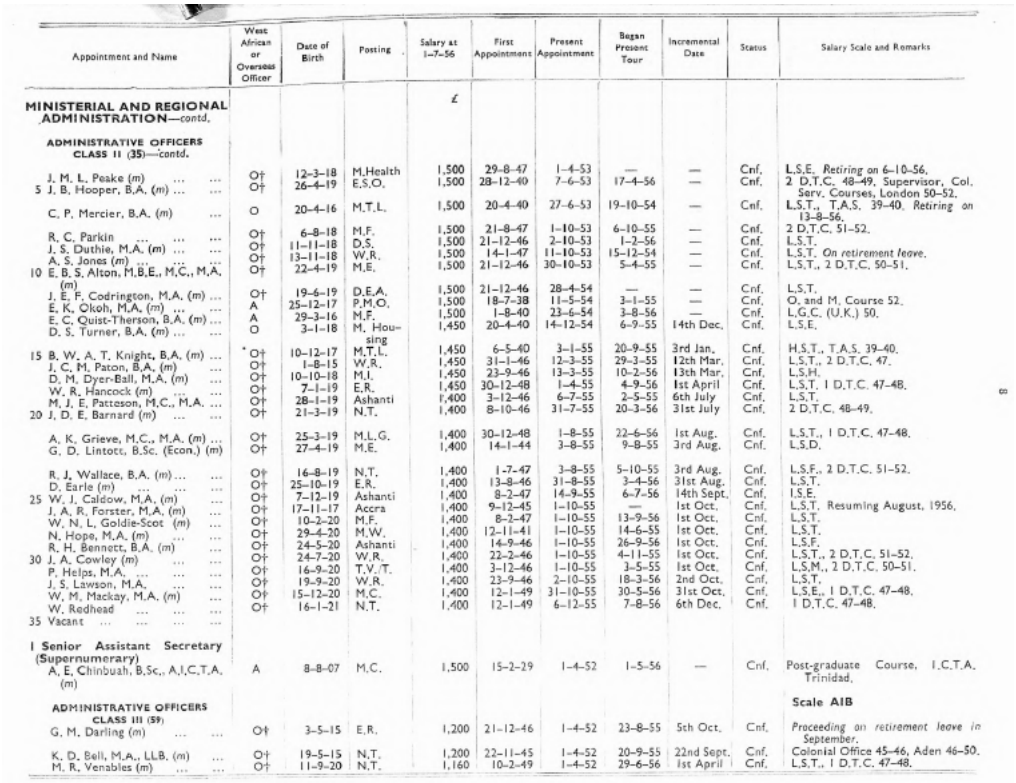
The composition of staff lists changed over time and availability is inconsistent across years. The lists fall into three types: (1) Senior Staff Lists, (2) African Staff Lists, and (3) Hybrid Staff Lists.[19] The *Senior Staff Lists (SSLs)* provide information about the higher ranks of the civil service (e.g., department and section heads and deputies), both European and African, and are available 1897 to 1946. Between 1936 and 1945, we also have *African Staff Lists (ASLs)* — listing only Africans, but spanning all ranks. ASLs offer a detailed snapshot of African representation across various bureaucratic roles, but do not allow for tracking trends across the full colonial period. After 1946, the government altered its reporting format, effectively creating a third, middle-tier list we refer to as the *Hybrid Staff Lists (HSLs)*.[20] The range of positions included in SSLs

[18]These lists exclude employees of Native Authorities, the colonial army, and low-wage manual laborers, such as hired in large numbers by the Railways and Harbor Administration (Jeffries, 1978).

[19]Section S1 in the Supporting Information shows coverage of each type.

[20]These were still labeled "Senior Staff Lists," but we prefer "Hybrid" to reflect the positions included.

Figure 3: Example extract from the 1956 Hybrid Staff List



Appointment and Name	West African or Overseas Officer	Date of Birth	Posting	Salary at 1-7-56	First Appointment	Present Appointment	Began Present Tour	Incremental Date	Status	Salary Scale and Remarks
MINISTERIAL AND REGIONAL ADMINISTRATION—*contd.*				£						
ADMINISTRATIVE OFFICERS CLASS II (35)—*contd.*										
J. M. L. Peake (*m*)	O†	12-3-18	M.Health	1,500	29-8-47	1-4-53	—	—	Cnf.	L.S.E. *Retiring on 6-10-56.*
5 J. B. Hooper, B.A. (*m*)	O†	26-4-19	E.S.O.	1,500	28-12-40	7-6-53	17-4-56	—	Cnf.	2 D.T.C. 48-49, Supervisor, Col. Serv. Courses, London 50-52.
C. P. Mercier, B.A. (*m*) ...	O	20-4-16	M.T.L.	1,500	20-4-40	27-6-53	19-10-54	—	Cnf.	L.S.T., T.A.S. 39-40. *Retiring on 13-8-56.*
R. C. Parkin	O†	6-8-18	M.F.	1,500	21-8-47	1-10-53	6-10-55	—	Cnf.	2 D.T.C. 51-52.
J. S. Duthie, M.A. (*m*)	O†	11-11-18	D.S.	1,500	21-12-46	2-10-53	1-2-56	—	Cnf.	L.S.T.
A. S. Jones (*m*)	O†	13-11-18	W.R.	1,500	14-1-47	11-10-53	15-12-54	—	Cnf.	L.S.T. *On retirement leave.*
10 E. B. S. Alton, M.B.E., M.C., M.A. (*m*)	O†	22-4-19	M.E.	1,500	21-12-46	30-10-53	5-4-55	—	Cnf.	L.S.T., 2 D.T.C. 50-51.
J. E. F. Codrington, M.A. (*m*) ...	O†	19-6-19	D.E.A.	1,500	21-12-46	28-4-54	—	—	Cnf.	L.S.T.
E. K. Okoh, M.A. (*m*)	A	25-12-17	P.M.O.	1,500	18-7-38	11-5-54	3-1-55	—	Cnf.	O. and M. Course 52.
E. C. Quist-Therson, B.A. (*m*) ...	A	29-3-16	M.F.	1,500	1-8-40	23-6-54	3-8-56	—	Cnf.	L.G.C. (U.K.) 50.
D. S. Turner, B.A. (*m*)	O	3-1-18	M. Housing	1,450	20-4-40	14-12-54	6-9-55	14th Dec.	Cnf.	L.S.E.
15 B. W. A. T. Knight, B.A. (*m*) ...	O†	10-12-17	M.T.L.	1,450	6-5-40	3-1-55	20-9-55	3rd Jan.	Cnf.	H.S.T., T.A.S. 39-40.
J. C. M. Paton, B.A. (*m*) ...	O†	1-8-15	W.R.	1,450	31-1-46	12-3-55	29-3-55	12th Mar.	Cnf.	L.S.T., 2 D.T.C. 47.
D. M. Dyer-Ball, M.A. (*m*) ...	O†	10-10-18	M.I.	1,450	23-9-46	13-3-55	10-2-56	13th Mar.	Cnf.	L.S.H.
W. R. Hancock (*m*)	O†	7-1-19	E.R.	1,450	30-12-48	1-4-55	4-9-56	1st April	Cnf.	L.S.T. 1 D.T.C. 47-48.
M. J. E. Patteson, M.C., M.A. ...	O†	28-1-19	Ashanti	1,400	3-12-46	6-7-55	2-5-55	6th July	Cnf.	L.S.T.
20 J. D. E. Barnard (*m*)	O†	21-3-19	N.T.	1,400	8-10-46	31-7-55	20-3-56	31st July	Cnf.	2 D.T.C. 48-49.
A. K. Grieve, M.C., M.A. (*m*) ...	O†	25-3-19	M.L.G.	1,400	30-12-48	1-8-55	22-6-56	1st Aug.	Cnf.	L.S.T., 1 D.T.C. 47-48.
G. D. Lintott, B.Sc. (Econ.) (*m*)	O†	27-4-19	M.E.	1,400	14-1-44	3-8-55	9-8-55	3rd Aug.	Cnf.	L.S.D.
R. J. Wallace, B.A. (*m*)	O†	16-8-19	N.T.	1,400	1-7-47	3-8-55	5-10-55	3rd Aug.	Cnf.	L.S.F., 2 D.T.C. 51-52.
D. Earle (*m*)	O†	25-10-19	E.R.	1,400	13-8-46	31-8-55	3-4-56	31st Aug.	Cnf.	L.S.T.
25 W. J. Caldow, M.A. (*m*) ...	O†	7-12-19	Ashanti	1,400	8-2-47	14-9-55	6-7-56	14th Sept.	Cnf.	I.S.E.
J. A. R. Forster, M.A. (*m*) ...	O†	17-11-17	Accra	1,400	9-12-45	1-10-55	—	1st Oct.	Cnf.	L.S.T. *Resuming August, 1956.*
W. N. L. Goldie-Scot (*m*) ...	O†	10-2-20	M.F.	1,400	8-2-47	1-10-55	13-9-56	1st Oct.	Cnf.	L.S.T.
N. Hope, M.A. (*m*)	O†	29-4-20	M.W.	1,400	12-11-41	1-10-55	14-6-55	1st Oct.	Cnf.	L.S.T.
R. H. Bennett, B.A. (*m*) ...	O†	24-5-20	Ashanti	1,400	14-9-46	1-10-55	26-9-56	1st Oct.	Cnf.	L.S.F.
30 J. A. Cowley (*m*)	O†	24-7-20	W.R.	1,400	22-2-46	1-10-55	4-11-55	1st Oct.	Cnf.	L.S.T., 2 D.T.C. 51-52.
P. Helps, M.A.	O†	16-9-20	T.V./T.	1,400	3-12-46	1-10-55	3-5-55	1st Oct.	Cnf.	L.S.M., 2 D.T.C. 50-51.
J. S. Lawson, M.A.	O†	19-9-20	W.R.	1,400	23-9-46	2-10-55	18-3-56	2nd Oct.	Cnf.	L.S.T.
W. M. Mackay, M.A. (*m*) ...	O†	15-12-20	M.C.	1,400	12-1-49	31-10-55	30-5-56	31st Oct.	Cnf.	L.S.E., 1 D.T.C. 47-48.
W. Redhead	O†	16-1-21	N.T.	1,400	12-1-49	6-12-55	7-8-56	6th Dec.	Cnf.	1 D.T.C. 47-48.
35 Vacant										
1 Senior Assistant Secretary (Supernumerary)										
A. E. Chinbuah, B.Sc., A.I.C.T.A. (*m*)	A	8-8-07	M.C.	1,500	15-2-29	1-4-52	1-5-56	—	Cnf.	Post-graduate Course, I.C.T.A. Trinidad.
ADMINISTRATIVE OFFICERS CLASS III (59)										**Scale A1B**
G. M. Darling (*m*)	O†	3-5-15	E.R.	1,200	21-12-46	1-4-52	23-8-55	5th Oct.	Cnf.	*Proceeding on retirement leave in September.*
K. D. Bell, M.A., LL.B. (*m*) ...	O†	19-5-15	N.T.	1,200	22-11-45	1-4-52	20-9-55	22nd Sept.	Cnf.	Colonial Office 45-46, Aden 46-50.
M. R. Venables (*m*)	O†	11-9-20	N.T.	1,160	10-2-49	1-4-52	29-6-56	1st April	Cnf.	L.S.T., 1 D.T.C. 47-48.

expanded to encompass a broad cross-section of mid-ranking officials. Many positions previously found only in the ASL appear in the HSLs, although the lowest-ranking employees were excluded.

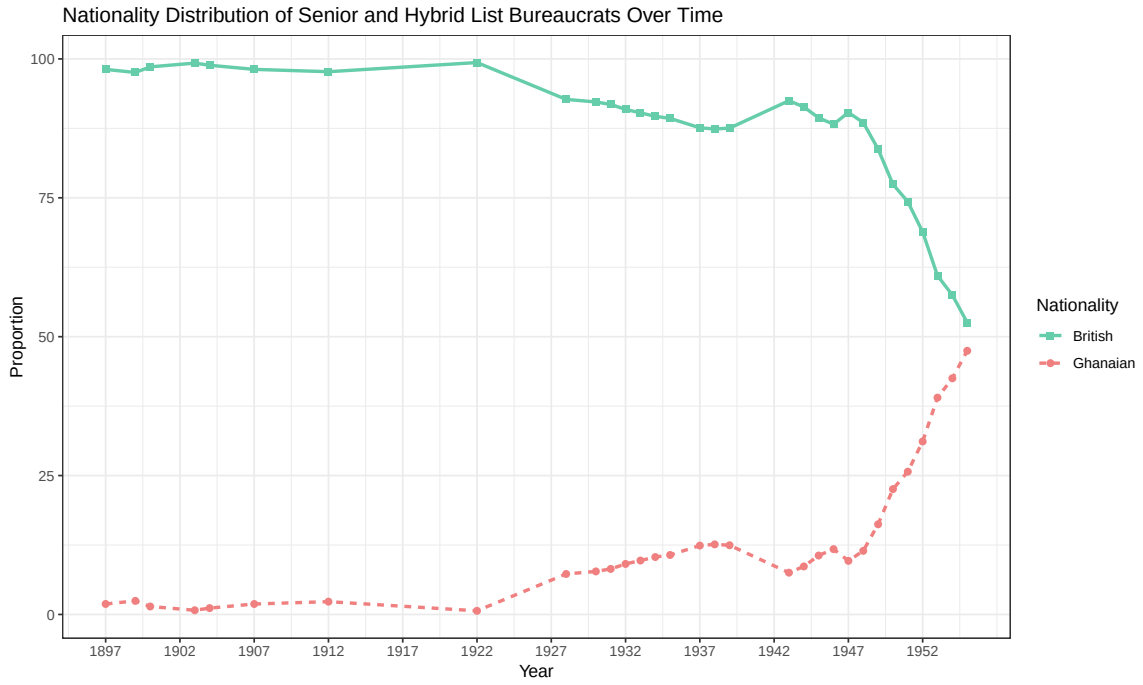
Using these lists, we identify 11,702 unique bureaucrats: 4,805 Europeans and 6,879 Africans (Ghanaians).[21] Figure 4 shows nationalities among senior bureaucrats in the SSLs and HSLs between 1897 and 1955, illustrating the Africanization that began in the late 1920s and rapidly intensified from 1947.

5.1 Classifying the *most likely* ethnicity of each bureaucrat

To establish the ethnic identity of each African bureaucrat, we combine census data containing the ethnic composition of small communities with voter register data that lists names of individuals within these communities. We classify each bureaucrat into one of ten categories, which repre-

[21] We code race based on either a name appearing in the ASL, explicit classifications of race included in the HSL, or school information provided in the SSLs. Race is only ambiguous for 18 bureaucrats.

Figure 4: Bureaucrat Nationality in the Senior and Hybrid Staff Lists (1897-1955)



sent important national-level ethnic cleavages (Asante and Gyimah-Boadi, 2004; Nathan, 2019).[22] We validate our ethnicity coding below. However, we also stress that no classification process is perfect: some names are simply not very ethnically distinct. Similarly, we cannot systematically identify ethnically-mixed individuals.[23] Yet without being able to interview colonial-era bureaucrats to ascertain their ethnicity directly, we believe our approach provides a best-case scenario for extracting ethnic information from names. We outline the general procedure below; a formal definition is in Section S3.

We use data from Ghana's 2015 voter register, which contains a near-universe of names of contemporary adults (13.8 million individuals). Each voter is assigned to a polling station. We combine these names with Ghana's 2010 census, measuring ethnic composition within Enumeration Areas (EAs, or tracts). Geolocating each polling station to these EAs, we split each name

[22] We distinguish among Akan sub-groups the Asante, Akyem, Akuapem, Fante, and Other Akans, and also classify the Dangme, Ewe, Ga, Guan, and a category for all Northern groups. See Section S2 for justification of this classification.

[23] Yet, based on data in Section S4, we expect only a tiny proportion of colonial-era bureaucrats had mixed heritage.

registered at each station into constituent fragments (e.g., "Samuel Okudzeto" becomes "Samuel" and "Okudzeto") and assign each fragment the ethnic population proportions of that station. Summing the frequency of fragment appearances across stations, we compute an average vector of ethnic proportions for each of Ghana's 1.1 million unique name fragments.

We then calculate each fragment's ethnic *distinctiveness*, defined as the distance between the fragment's average group proportions and the base rate of ethnic group proportions in Ghana's overall voter population. Name fragments that carry meaningful ethnic information (high distinctiveness) will exhibit average ethnic distributions that deviate from the national base rates, concentrating more heavily on a particular group. In contrast, names that are widely used across groups – "John" or "Mary" – do not deviate much from the base rate of ethnic group proportions (low distinctiveness). Without accounting for distinctiveness, such widely-used names would disproportionately (and erroneously) be assigned to the most populous ethnic group nationwide. We use distinctiveness scores as weights to ensure that more informative fragments contribute more heavily to classification of each individual; in the example of "Samuel Okudzeto," "Samuel" is common and ethnically uninformative, while "Okudzeto" strongly signals Ewe ethnicity. Our method accordingly upweights the surname's information and classifies him as most likely Ewe.

Our approach builds on methods from Kasara (2013) and Harris (2015) that infer ethnicity from names in Kenya. These methods construct group-specific dictionaries from similar voter registration data by focusing on homogeneous polling stations — e.g., those with over 90% of voters from a single group (Kasara, 2013) — assuming names there are broadly representative of that group. This assumption easily may not hold. Our strategy instead enables classification of individuals even with names used by multiple groups, while also incorporating classification uncertainty in the form of our distinctiveness score.

5.2 Validating our measure

Validating whether our approach correctly identifies ethnicity requires access to a large set of names linked to self-identified ethnicities. Fortunately, a survey from Brierley and Nathan (2021) includes both full names and self-reported, sub-group-level ethnicities for a large sample of Ghana-

ian respondents (N=1,151). Our method performs well in recovering self-identified ethnicities when name distinctiveness is moderate to high. For example, for names in the top quartile of name distinctiveness in the survey, we recover self-reported ethnicity 90% of the time. However, as can be expected, our performance declines for names with low distinctiveness (Section S4).[24]

Rather than indicating a flaw, our difficulty classifying low-distinctiveness names highlights a key feature of our task: names with low distinctiveness convey limited information, as each fragment is widely used across multiple groups. Accordingly, no alternative method — whether algorithmic like ours, machine learning-based, or hand-coded — could reliably extract ethnicity from such names. Indeed, we show that the common alternative of hand-coding names based on local research assistants' knowledge introduces even worse measurement error (Section S5).

Moreover, recent machine learning approaches such as Chaturvedi and Chaturvedi (2024) cannot be applied in our or many other contexts due to the lack of a sufficiently large and statistically-representative training dataset. Machine learning approaches also introduce statistical uncertainty when estimating ethnic labels for names not in the training data. But because we start with the universe of Ghanaian names, over 98% of bureaucrats have a name fragment that matches to the voter register exactly. We thus do not need to infer ethnicity based on similarly-spelled name fragments.

Given the inherent limitations classifying low distinctiveness names, the analyses below drop all bureaucrats scoring in the bottom quartile on our distinctiveness score, removing observations among whom it is especially difficult to ascertain ethnicity. To demonstrate that this specific cutoff does not bias our conclusions, we repeat all tables and figures instead using other cutoffs, as well as when not dropping any observations (Section S7). All key patterns are robust to these alternatives. In fact, Figure S7.1 shows that the central pattern in our results persists impervious to any cutoff for how many low-distinctiveness names to drop all the way up through dropping the full bottom half of our data and is thus very unlikely to be due to measurement error.[25]

[24]This tests against contemporary names, but there could also be concern naming conventions have evolved over time. For further validation, see Section S4.

[25]Section S3 also confirms that distinctiveness is not correlated with ethnicity in ways that could bias our results.

6 Ethnic imbalance in the colonial civil service

Two patterns are evident when disaggregating the bureaucracy by ethnicity. First, two groups – the Ga and Fante – dominated African hiring at both senior and junior levels of the civil service. Second, this dominance persisted over time, both before formal policies of Africanization began in the late 1920s, and as African hiring ratcheted up from the late 1940s. Accordingly, post-independence leaders inherited a civil service that had long been deeply ethnically-imbalanced.

Figure 5 shows ethnicities of African bureaucrats in the Senior and Hybrid Staff Lists over time. The senior bureaucracy expanded after 1900, increasing from 106 posts in 1897 to 906 senior bureaucrats in 1928. Between 1928 and 1948, buffeted by the Great Depression and then war, senior-level hiring stalled for two decades; total senior bureaucrats remained below 1,000 and only 5.1 new Ghanaian officers were hired to senior positions on average annually across departments – a small, but steady trickle into top ranks. Across this period, the Ga and Fante continued dominated, accounting for an overwhelming 82.7% of senior Ghanaian bureaucrats by 1948 (Ga 55.6% and Fante 27.2%), far in excess of their respective 3% and 11% of Ghana's population.[26] After WWII, African hiring into the senior-most ranks picked up dramatically; from 1949 to 1955, there was an average increase of 166 new Ghanaian officers into the senior and mid-level ranks ever year. As this happened, Ga and Fante dominance remained.

Figure 5 also shows from the late 1930s a small increase in representation of some other ethnic groups. For example, the Ewe went from 7.4% of senior bureaucrats between 1928 and 1945 to plateauing around 11% between 1946 and 1955. The Asante began having some representation in the late 1940s, but still reached just 5.8% by 1955. But these groups were still underrepresented relative to their size: 13% and 16% of Ghana's population, respectively. Northern groups, representing 29% of the population, were most severely underrepresented, never accounting for more than 2.3% of senior or mid-level bureaucrats any year.

[26]Lacking historical data on sub-group sizes, we use 2010 census figures as a proxy. Usefully, de Graft-Johnson (1969) reports disaggregated figures from the 1960 census for a few sub-groups, and these closely align with the 2010 proportions, supporting the validity of using 2010 data to measure historical group size.

Figure 5: *Bureaucrat ethnicity among Ghanaians in Senior and Hybrid Staff Lists:* Top figure shows the raw ethnic distribution, and bottom figure shows the deviation from respective population shares. 2010 population proportions were: Asante (15.5%), Fante (11.2%), Akyem (3.1%), Akuapem (2.5%), Other Akan (13.3%), Ga (3.0%), Dangme (4.1%), Ewe (13.4%), Northern (29.1%), and Guan (1.2%). The figure drops all observations below the bottom 25th percentile in name distinctiveness.

Although we only have complete data on junior positions for a more limited period – 1936-1945 – a similar pattern of Ga and Fante dominance emerges (Section S6). Across 1936 to 1945, Ga and Fante bureaucrats held the large majority of these positions, with Gas on average 42.8% and the Fante 28.3%. There is virtually no change in any group's percentages year on year. While the Ewe hold steady around their share of the national population, the Asante, other Akan sub-groups, and Northerners are all heavily underrepresented.

We also identify African bureaucrats holding top positions within their departments. We use salary information to classify top bureaucrats, focusing on salaries in the top 10th percentile in each department (Section S6).[27] The vast majority (over 80%) of high-salaried African bureaucrats are either Ga or Fante throughout the entire period in which Africanization was occurring. As we describe further below, this likely provided officials from these groups with the ability to influence hiring within their departments. Other groups, such as the more populous Asante and Ewe, had few similarly well-placed high-ranked officials.

6.1 Ethnic "lock-in" within departments

Many studies note that some ethnic groups gained disproportionate access to state employment in former colonies (e.g., Ricart-Huguet 2021, Lange 2025). Conventional explanations for this dynamic are examined in Section 7 below, and do help explain the very *initial* dominance of Ga and Fante bureaucrats. But we argue that *persistent* Ga and Fante overrepresentation was instead sustained through to independence via a previously unexplored process of patrimonial bureaucrat-led hiring, resulting from existing African bureaucrats' private preferences to hire co-ethnics. We show that this is the most likely explanation by triangulating across a series of analyses.

Our first piece of evidence is the self-replication of ethnic hires within departments (ministries). Disaggregating by ethnicity, we analyze who gets hired each new year (time t) compared to the ethnic composition of existing staff (time t-1) in that same department. Figure 6 displays a strong positive association between the number of existing senior or mid-level staff each year

[27] Job titles are too inconsistent between departments and across years to use them to classify ranks.

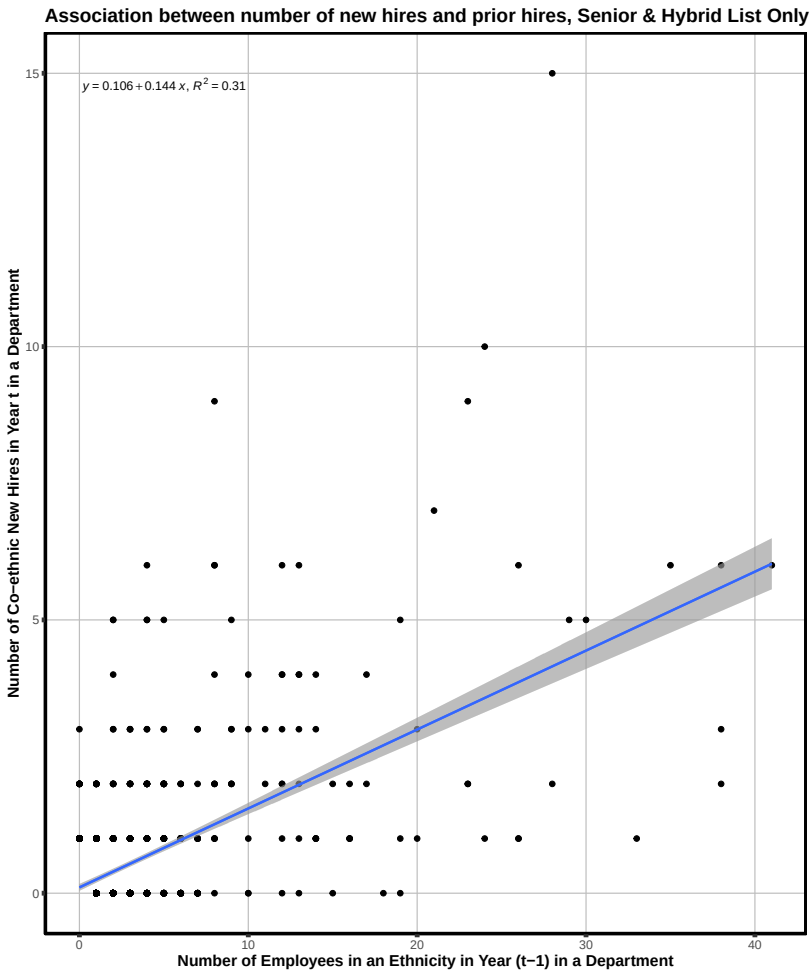


Figure 6: *Current Employees and Coethnic Hires in Each Department-Year:* the X-axis is the number of bureaucrats of a given ethnicity employed in a given department in the preceding year; the Y-axis is the number of new hires from that group in the same department in the following year. Each dot represents a department-year-ethnicity observation. Subset to the Senior and Hybrid Staff Lists. Drops all observations below the bottom 25th percentile on the name distinctiveness score.

from a given group and the entry of new senior or mid-level staff from the same group in the same department the following year. This implies that wherever an ethnic group gained a foothold in senior ranks of a department, it used its position to hire (or promote) other co-ethnic applicants.

Figure 7 repeats Figure 6 broken out by ethnic group. Crucially, it shows a similar positive correlation both among the Ga and Fante *and* among most other groups too.[28] In fact, the serial correlation between prior and new hires within departments is even stronger among some other

[28]Except for the Dangme, all other groups deviating from the general pattern in Figure 6 had such negligible numbers of senior employees to begin with that it is difficult to even observe whether such a correlation exists.

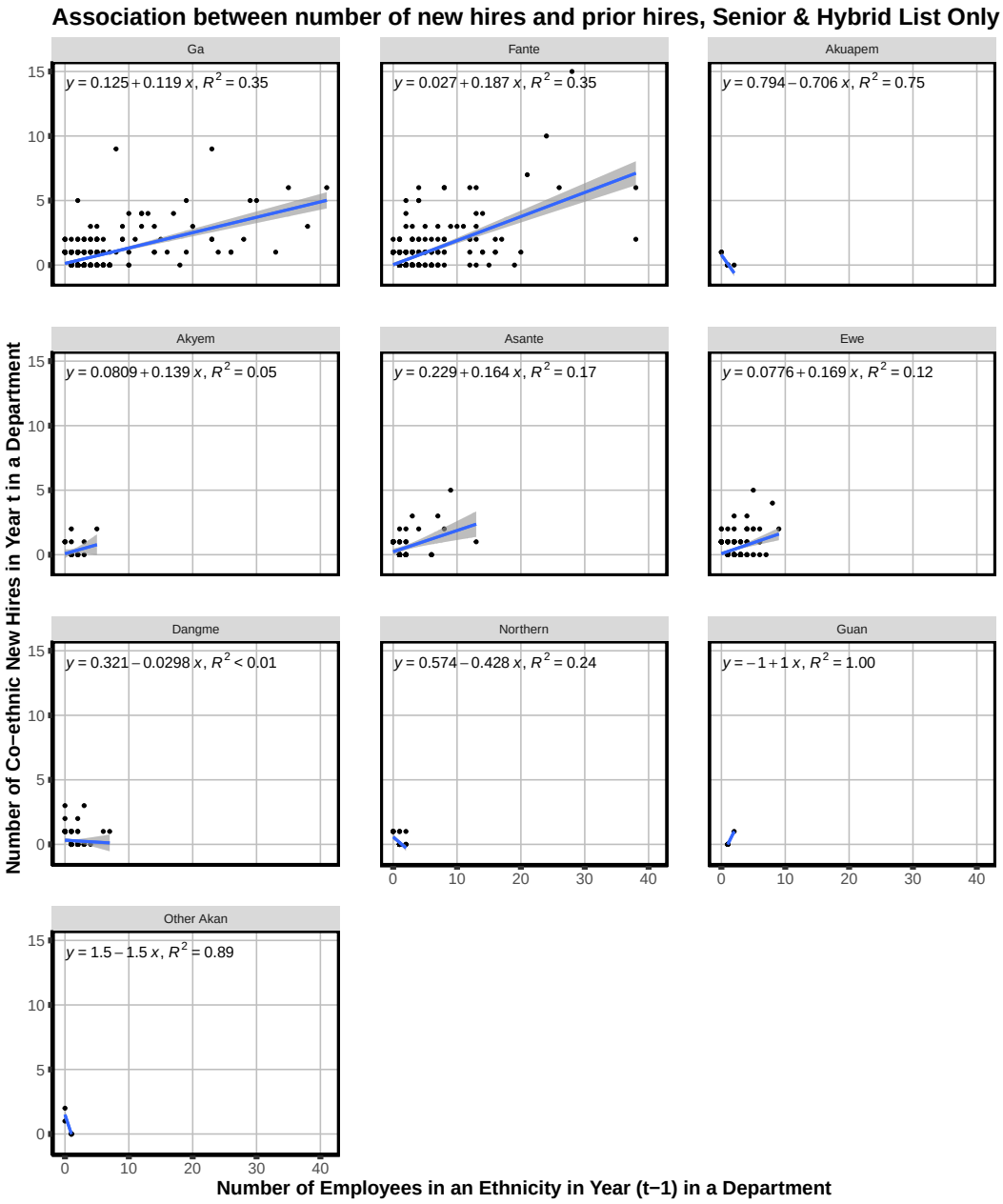


Figure 7: *Current Employee and Coethnic Hires in each Department-Year by Group*: drops all observations below the bottom 25th percentile on the name distinctiveness score.

Table 1: New hires against existing senior staff and lagged school access

Outcome: New hires from group in year t	1	2	3	4	5	6
Bureaucrats from group among senior staff (year $t-1$)	0.079**	0.080**	0.061†			0.241*
	(0.019)	(0.018)	(0.029)			(0.086)
Avg. distance (km) to school for group (year $t-25$)	−0.008		−0.080	−0.003		−0.013
	(0.010)		(0.083)	(0.005)		(0.013)
Avg. distance (km) to school for group (year $t-30$)		−0.008			−0.007	
		(0.006)			(0.005)	
Years included	1928-1955	1928-1955	1936-1955	1928-1955	1928-1955	1928-1955
Ethnic group FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*N*
Department FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
Year FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
N	5929	5929	4301	7161	7161	5929
adj. R^2	0.197	0.197	0.202	0.144	0.144	0.176

† significant at $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$. The unit is the ethnic group-department-year. OLS regressions with standard errors clustered by ethnic group, department, and year. Results with other SE clusterings are similar (and less conservative). Column 3 restricts only to years for which both Senior and African staff lists (1936-1945) and/or the Hybrid Staff list (1946-) are available. All bureaucrats in bottom 25th percentile of name distinctiveness are dropped.

groups, such as the Asante and Ewe, than for the Ga. The ethnic lock-in pattern in Figure 6 is thus not unique to the Ga and Fante: many groups replicated themselves within departments once they had an initial foothold of senior staff. However, such behavior disproportionately favored the Ga and Fante because they had footholds in most departments the earliest.

While the lock-in pattern is already evident from these descriptive statistics, we also demonstrate path dependence using a regression pooling across list types. In Table 1 the dependent variable is the count of new hires at any rank in year t for group g in department d and the independent variable is the count of existing employees at the senior or mid-level ranks of the same department from that same group, but in year $t-1$. We also include a measure of the population-weighted average distance of group g to colonial schools as of 25 or 30 years earlier, to proxy for the group's education access as of the time a new hire in year t would have been starting school.[29] By controlling for this, our results are independent of potential selection effects that result from inequality in education access. We also include fixed effects by ethnic group, department, and year, controlling for group-specific, department-specific, and year-specific patterns, to examine how marginal changes within groups and departments in the presence of senior bureaucrats and/or in (lagged) education access reflect in marginal changes in subsequent hiring. Standard errors

[29] We introduce this measure in more detail in the next section, as well as in Section S10.

are clustered by group, year, and department – the most conservative specification. We restrict to 1928-1955, the only years for which we have a large number of African hires.[30]

The pattern in Table 1 is again clear: each additional bureaucrat among the senior or mid-level staff from a given ethnic group in a given department in a given year predicts an additional 0.06-0.24 members of that group were hired into that department (at any rank) the subsequent year. This pattern is suggestive of relatively senior African bureaucrats influencing their agencies to hire group members. Importantly, this result does not depend on the group fixed effects (column 1 vs. column 6), suggesting it is not specific to the Ga and Fante alone, but a more general pattern. Moreover, changes in group-specific exposure to education at the time bureaucrats likely began schooling are consistently uncorrelated with marginal changes in future hiring, including when put in alone as a predictor (columns 4 and 5). This suggests that shifts in hiring patterns are not explained by shits in education access – discussed further below.

6.2 Archival evidence of bureaucrat-led hiring

Ethnic “lock-in" within departments was likely facilitated by how hiring was structured. We examine the extensive correspondence among British officials about African hiring in Ghana’s National Archives, spanning the 1930s through early 1950s. These materials establish three important facts. First, many hiring and promotion decisions for Africans were decentralized to departments. Second, as a result, incumbent bureaucrats had points of discretion to shape hiring. Third, these bureaucrats sometimes abused these powers – and the British were aware of the problem.

Evidence of decentralization comes from materials that show that prospective employees were instructed to apply directly to departments, without a centralized screening process.[31] The archives include a large volume of handwritten letters from African applicants to top British officials asking about vacancies, with attached originals of exam certificates and (again often handwrit-

[30] In column 3, as a robustness test we instead restrict to 1936-1955, the periods during which we either have both the Senior and African Staff lists (1936-1945) and/or have the Hybrid Staff List that includes mid-ranked positions (1947-1955).

[31] See GH/PRAAD/CSO.2/7/130, file 1680/30/S33 (“Candidates for... appointments to the African pensionable staff, other than clerical... are selected by Heads of Departments"); and also, for example, folios GH/PRAAD/CSO.2/8/184, GH/PRAAD/CSO.2/8/193, and GH/PRAAD/CSO.2/8/195.

ten) reference letters. While unqualified applicants were rejected outright, most instead received a curt letter returning their documents and stating that they must write to each state department separately to inquire about jobs.[32] An exception applied to African clerical staff (First and Second Division Clerks), a narrow stratum of the total civil service, appointed, promoted, and rotated among departments by a central Selection Board (or Promotion Board) staffed by three senior British officials.[33] But memoranda from this committee demonstrate that these British officers had little independent information about most prospective African clerks, appearing to mostly act on recommendations of department heads, recommendations in turn based on qualitative evaluations relayed upwards from conversations with subordinates within each department.[34]

This largely decentralized process should have provided discretion to existing Africans within departments to influence future hiring and promotions by helping to identify and screen potential recruits, as well as in subjective advice and performance evaluations about candidates provided to department heads. As one example, a 1947 memo to customs collectors explicitly instructs that "*they should obtain [African] recruits themselves*" [emphasis added] for their individual customs offices, and were provided with a list of job applicants who had written to the department head and told to make contact with them "when vacancies occurred" without first working through the head office.[35]

The need to apply to individual departments should have facilitated existing bureaucrats' discretion by ensuring that information on vacancies was fragmented. These bureaucrats could presumably provide privileged information about vacancies in their departments to preferred recruits. Indeed, that so many applicants painstakingly compiled their documents by hand and wrote directly to central British officials to ask about vacancies, only to be immediately informed this was an invalid way to apply for jobs in the first place, suggests that full information about vacancies did not circulate widely, even among these educated Ghanaians. And the need to then apply

[32] For example, folios GH/PRAAD/CSO.2/8/184, GH/PRAAD/CSO.2/8/193, and GH/PRAAD/CSO.2/8/195.
[33] The Colonial Secretary, the head of the Education Department, and the head of the Department of Native Affairs.
[34] See Section S13 for more details.
[35] GH/PRAAD/CSO.2/8/640/671.

across several dozen departments separately, attaching originals of your references and school and exam certificates, in the hope there might be a vacancy erected time-consuming barriers to entry for applicants who lacked inside information.

Finally, the archival materials show that bureaucrats were known to exploit this discretion to achieve hiring outcomes at odds with British intentions. The Colonial Secretary observed in a 1938 bulletin: "Cases have been known to occur in the subordinate ranks of the service in which an African officer's colleagues required or received from him on his first appointment or promotion a sum of money, gift, or entertainment."[36] Although the reason attributed to bureaucrats' actions in this quote is a distinct, but related motivation for corruption than what we propose (rent extraction vs. social favoritism), the statement exemplifies the same principal-agent dilemma: a top principal explicitly admitting that African agents within departments were using their influence to deviate from merit-based procedures when selecting and promoting other African staff.

6.3 Evidence of ethnic bias conditional on merit

Further evidence consistent with bureaucrat-led ethnic favoritism is that Ga and Fante applicants appear to have been favored over applicants from other groups, even when they performed similarly on formal examinations for entry into the bureaucracy.

Between 1906 and 1930, Gold Coast colonial yearbooks list names of individuals who passed civil service examinations to qualify as teachers and agricultural trainers and (often) their rank scores relative to others who passed.[37] We digitize these lists to obtain data on 111 unique exam sittings by 2,423 unique qualifiers. We assign each qualifier a most likely ethnicity and match against our staff lists, observing which qualifiers were subsequently hired into the central bureaucracy. Unsurprisingly, the most common state posts were as teachers (52%) or agriculture officers (8%), but we also find that 40% of those subsequently hired were employed across a

[36] GH/PRAAD/CSO.2/7/130/19. The statement continued that this was forbidden, but did not state what, if anything, might be done to try to stop it.

[37] The lists of exam candidates unfortunately do not include the names of individuals who did not pass. These yearbooks also mention other civil service exams were held for other positions, but we could not locate similar individualized results lists.

wide range of additional departments. Successful examinees were eligible for many positions, as a teaching exam certification was listed among the formal educational qualifications for broader civil service employment.[38] We take individual performance on these exams as a general indicator of merit – an indicator the British themselves created for this purpose – and explore subsequent hiring.[39]

Figure 8 visualizes the pool of exam candidates, disaggregated by ethnic group, indicating both exam ranks (scores) and whether each examinee subsequently appeared in staff lists (highlighted by ▲). The horizontal bars indicate the average exam rank of hired bureaucrats from each group. This data can help draw conclusions about supply and demand-side explanations for hiring.

On the supply side, Figure 8 indicates a clear over-representation of Ga and Fante among successful examinees (total number of dots by group). Among those who passed exams across these years, 35.0% were Fante and 17.8% were Ga, significantly higher than their 11% and 3% of the population. It is unlikely that this over-representation is merely the result of greater access to schooling; while the Ga and Fante initially had greater education access, other groups who later experienced rapid increases in school access never caught up in success on these exams. Gas and Fantes comprise 58.1% of successful examinees in the 1900s, 55.4% in the 1910s, and still 50.5% in the 1920s. Over these decades, the Ga and Fante had stopped being any nearer to schools than, for example, Akyems and Akuapems. However, the proportion of Akyem and Akuapem among successful examinees instead fell, from 4.2% in the 1900s to just 1.7% in the 1920s, even as school access increased. Similarly, the proportion of Asante examinees declined from 6.7% in the 1900s to 5.7% by the 1920s, despite steadily increasing school access.[40]

[38]For example, see GH/PRAAD/CSO.2/7/129/18a. Until the creation of Ghana's first university in 1948, teacher training colleges were essentially the only tertiary institutions; teacher trainees (the main pool of examinees in this data) pursued many state positions.

[39]Successful examinees were also qualified to teach at mission schools. However mission teachers were typically paid less than government teachers (or other civil servants), such that government employment was more desirable and should still correlate with exam rank if state hiring were meritocratic.

[40]This over-supply of Gas and Fantes was also not simply because training colleges and testing centers were only located in Ga- or Fante-dominant areas. See Appendix S8 for further details. We unfortunately do not know the hometowns of exam takers.

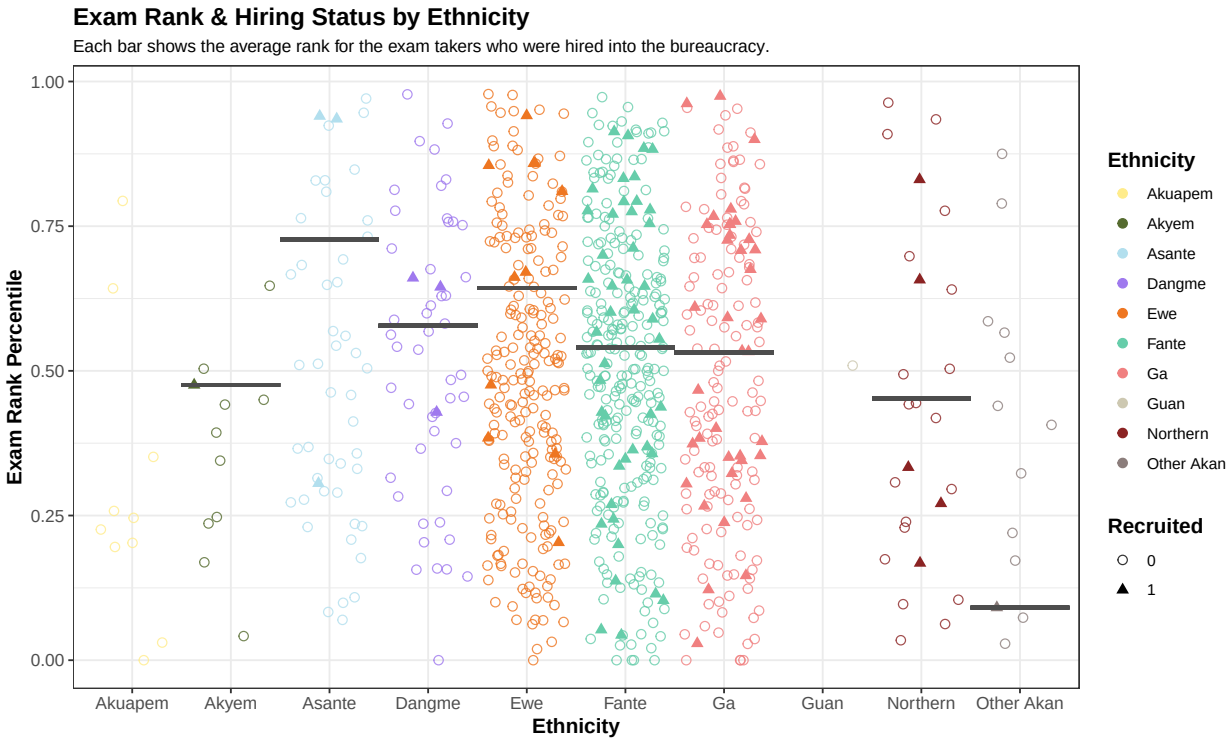


Figure 8: *Distribution of exam ranks (1906-1930) and subsequent state hiring status by group*: each dot is an exam candidate; triangles are those matched to a subsequent staff list. Horizontal bars indicate the average rank of hired candidates. Note that these averages are calculated from very few observations in groups that experienced little hiring.

One potential explanation for this oversupply is the development of cultural norms such that members of some groups particularly aspired to gain state employment. Such norms could have been culturally-reinforced by a history of early access to schooling or state jobs. Alternatively, they may have been due to differences in outside economic opportunities. For example, the Ewe may have been particularly overrepresented among successful examinees, while the Asante were particularly underrepresented, because Ewe areas largely lacked cash crop agriculture, while educated Asante (and Akyem and Akuapem), living in the heart of colonial Ghana's cocoa economy, may have had less overall need to seek state employment.[41]

But regardless of why some groups became oversupplied among test takers, evidence on hiring demand is where support for ethnic favoritism within the bureaucracy becomes most apparent. The exam data show that demand for Ga and Fante to enter the public sector is disproportionally high *even conditional on exam performance.* The Ga and Fante represent a startling 31.0% and 43.3% of successful examinees eventually hired, an additional 13.1 and 8.3 p.p. above their already over-represented share of those passing the civil service exams. By contrast, and consistent with significant hiring bias on the demand side, despite comprising 25.9% of successful examinees between 1906 and 1930, the Ewe are just 10.5% of hired examinees, a drop of 15.4 p.p and apparent evidence of significant hiring discrimination. The Asante too had 3.2 p.p. fewer bureaucrats hired, even relative to their small 5.5% of passing examinees. This bias conditional on exam performance persists equally among examinees hired as teachers and agricultural workers, and across those hired into various other state departments. It also cannot be explained by the locations of jobs, such as teaching positions at government-run schools (Section S12).

Analyzing each exam candidate's performance relative to their cohortmates at the same exam sitting further emphasizes demand-side bias. The median Ewe hired after passing an exam ranked at the 67th percentile in their exam sitting (horizontal bar in Figure 8); the median Asante ranked at the 74th percentile. By contrast, the median Fante hired ranked at the 58th percentile

[41] Arhin (1983) links other differences in the underlying structure of local economies to Asante versus Fante norms around Western education in the precolonial period as well.

and the median Ga hired ranked at just the 53rd percentile, showing that more mediocre Ga and Fante candidates received jobs over more qualified Ewe and Asante candidates. Most other groups in Figure 8 had so few (or none) of their successful examinees even hired (e.g., just one Akyem, and zero Akuapem) that a comparison of exam ranks among hired bureaucrats is uninformative.

Overall, these data suggest exam performance did not purely guide selection, supporting the idea that nominally-meritocratic institutions were exactly that: examinations gave an appearance of meritocracy, consistent with the *de jure* priorities of British principals, but in fact, selection was based on other attributes.

6.4 Nepotistic hiring among Euro-African families

Another piece of evidence for insider capture comes from the further over-representation of particular families within the Ga and Fante. As noted above, the earliest opportunities for Western education in the pre-colonial Gold Coast (17th and 18th centuries) were for Africans with mixed European descent, creating high-status families within Ga and Fante society who maintained European surnames into subsequent generations.[42] Historical sources make clear that these Euro-African families were the very first to be hired into the colonial state in the 19th century (Parker, 2000; von Hesse, 2024). To the extent that any African officials had the ability to use an initial foothold in the colonial state to then employ their own relatives or other close social ties in further positions through nepotistic insider hiring, it was these families.

In Section S9 we document just this: each additional Ga or Fante bureaucrat from a Euro-African family at a senior or mid-level position in the past year is a strong correlate of an additional new hire from a Euro-African family within that same state department in the next year, consistent with socially-driven hiring among job seekers from the same narrow social stratum.

[42] Unlike other colonies with culturally-separate mixed-race populations, having a European surname as a legacy of (potentially centuries-old) patrilineal descent from a European did not on its own indicate being phenotypically distinct from (or a lack of integration into) Ga or Fante society.

7 Alternative arguments

Support for our interpretation that ethnic "lock-in" results from bureaucrat-led hiring also emerges from ruling out three major alternative explanations. First, unequal education could explain hiring. Second, these patterns may reflect proximity to physical offices of the state. Third, these patterns could reflect principals' strategic preferences to prioritize certain groups. However, all three possibilities cannot account for the above patterns.

7.1 Ethnic inequality in education

Many scholars argue ethnic inequality in colonial hiring was a consequence of inequality in Western education (Mazrui, 1978; Ray, 2018; Johnson-Kanu, 2021; Ricart-Huguet, 2021; Lange, 2025).[43] Education reached groups unevenly because of differences in proximity to the coast and other zones of investment (Ricart-Huguet, 2022), as well as differences in the spread of missionaries (Frankema, 2012; Jedwab, Meier zu Selhausen and Moradi, 2022). With ethnic differences in the supply of qualified candidates, inequality in employment can still be consistent with meritocracy (Kuipers, 2025).

This surely accounts for at least one pattern in Figure 5: the enormous under-representation of Northerners cannot be separated from the near-total absence of schools throughout the Northern Territories until the 1950s (Figure 2 above; Grischow, 2006; Nathan, 2023). Evidence that education accounts for other groups' hiring is less compelling. As detailed above, the Ga and Fante enjoyed the greatest *initial* education access in the 19th century and before, which partly explains why they were the very first to be hired at these earliest moments of state-building. But these groups' early advantages are causally overdetermined. The Ga and Fante also initially benefited from their especially fortuitous locations, proximate to the earliest zones of formal British control. For example, Euro-Ga elites already were "in on the ground floor" of the colonial state in Accra from the 1850s in a quite literal sense: they were the British government's landlords, leasing their

[43] For example, Ricart-Huguet (2021, p. 2459) argues that "districts [within African colonies] with more primary education became more represented in the civil service" as a "byproduct of education-based recruitment into the colonial state."

Table 2: Approximate average distance (km) to the nearest primary school by ethnic group

Ethnic group	1894	1906	1909	1914	1921	1927	1932
Akyem	17.01	4.05	3.47	3.47	3.29	2.87	2.77
Akuapem	15.73	6.19	5.56	5.55	4.85	4.75	4.53
Asante	45.32	12.88	9.57	9.37	8.78	8.56	8.12
Dangme	12.73	8.34	7.52	7.16	6.76	6.60	6.34
Ewe	25.19	13.78	10.92	10.8	9.15	7.53	6.46
Fante	11.02	3.66	3.28	3.28	3.24	3.20	3.17
Ga	6.04	3.24	3.19	3.19	3.19	3.19	3.18
Guan	56.42	21.34	16.47	16.47	10.19	6.39	3.70
Northern	248.95	248.95	82.14	71.23	38.06	33.43	31.68
Other Akan	36.07	21.34	16.76	16.08	13.72	12.29	11.64

homes for state offices and residences (von Hesse, 2024, p. 4). Meanwhile, regardless of education, at the start of our data in 1897, groups like the Asante mechanically could not have had similar access to jobs because they were not formally incorporated into the colony for another few years.

But the *continued* dominance of the Ga and Fante deep into the 20th century remains more puzzling. The exam data above already clearly shows that ethnic differences in hiring cannot be fully accounted for by the supply of qualified candidates per group. Moreover, combining the schools data in Figure 2 with geolocated ethnic census data, Table 2 displays the population-weighted average distance of each group to schools between 1894 and 1932.[44] Using birth dates, we calculate that the typical civil servant was hired in their mid-to-late 20s. Accordingly, we use the availability of primary schools 25 years before a hiring decision to proxy for the supply of educated candidates. Multiple other ethnic groups in Southern Ghana caught up in education access in the early 20th century. But state hiring never updated, even as the vast majority of African hiring occurred between the 1930s and 1950s.

Table 2 shows that while the Ga and Fante had the best school access in 1894, this advantage soon dissipated. By 1909 the Akyem and Akuapem had become almost as close to schools. By 1921, the Akyem had pulled even, and by 1927, when many bureaucrats hired in the 1950s would have begun school, schools were more numerous in Akyem territory than Ga or Fante areas. The

[44] We discuss how these figures are constructed in Section S10.

typical member of every other group except Northerners and Other Akans was, on average, now within walking distance of a school. And yet, there is little evidence in Figures 5 or Section S6 that future hiring kept pace, a descriptive pattern we already confirmed above via multivariate regression (Table 1). For example, by 1954, 45 years *after* the Akyem had reached virtual parity with the Ga in proximity to schools, just 1.4% of Africans in the Hybrid Staff List were Akyem, compared to 38.5% Ga, even as the groups have identical population sizes (both 3% nationally).

In addition, in Figure 9 we also rule out that differences in hiring could be due to differences in school quality. Digitizing annual Education Department reports available from 1906 to 1932 that provide exact enrollment numbers and total operating budgets for primary schools, including the mission schools,[45] we calculate *expenditure per pupil* across the schools in each ethnic group's home region to proxy for school quality. There is no evidence that Ga and Fante areas had systematically better schools, even very early in the 20th century. School quality instead appears quite similar across groups, and, if anything, the Fante had among the colony's worst-funded schools despite dominating civil service hiring.

7.2 Geographic proximity to state offices

Another possible explanation concerns distances to state offices. A challenge in explicitly assessing a proximity mechanism is that the staff lists we digitize list job titles, not locations. However, we can cast doubt on this possibility in several ways.

First, the capital shifted from Cape Coast to Accra in 1877, along with most state jobs (Parker, 2000, 121). As indigines of Accra, it may seem natural the Ga would have become overrepresented among state employees. However, the same is not true of the Fante, whose main population mass (around Cape Coast) lies further from Accra than those of several other groups

[45]This budget includes both the portions contributed by the mission society and by the state. Mission schools operated independently, but were heavily state subsidized (Frankema, 2012). Figure 9 focuses on primary schools. Secondary schools were largely boarding institutions; only among primary schools can we interpret expenditures at schools located in a group's home area as investments in the education of that specific group.

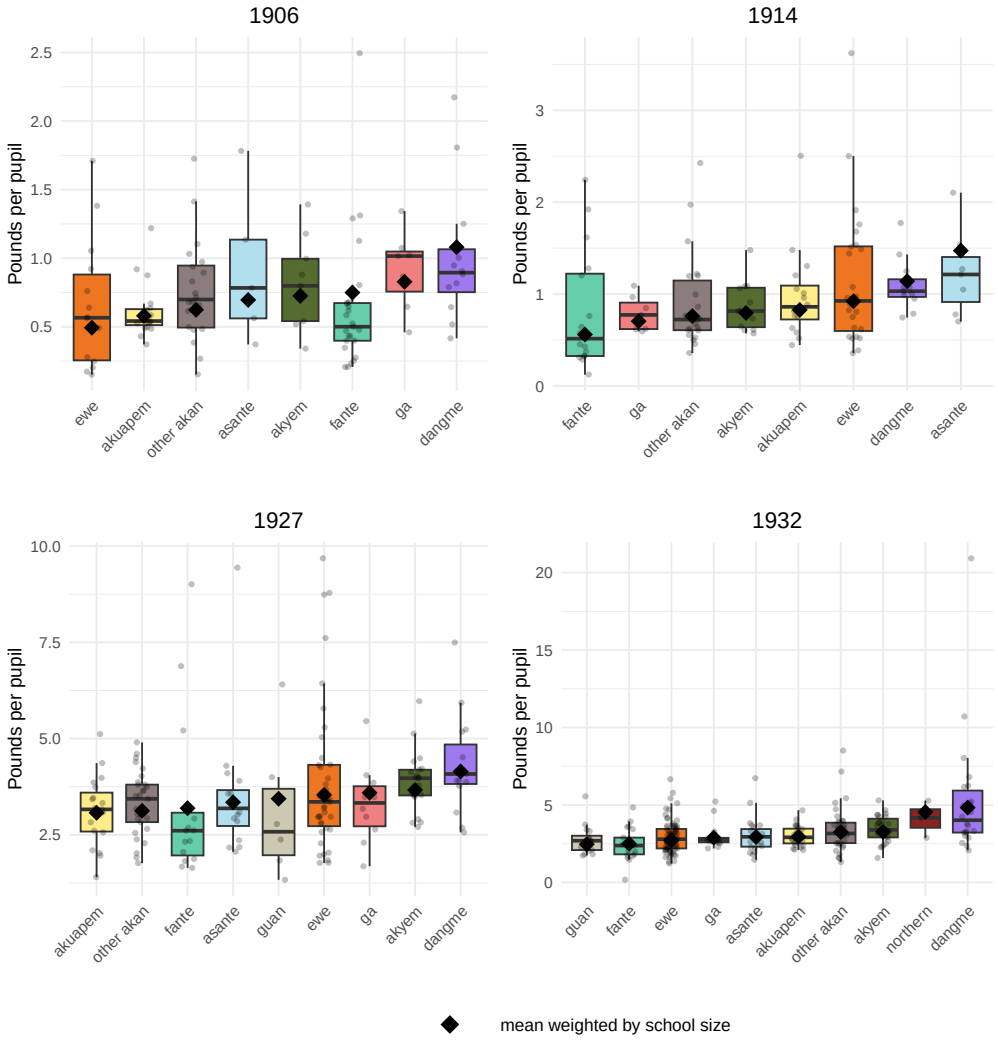


Figure 9: *Expenditure (quality) by school location*: British pounds per enrolled student at primary schools at four snapshots in time. Each individual school (grey dot) is geolocated and categorized by the local ethnic majority in that community. The central bar in each box plot shows mean expenditure *per school* per ethnic group; the black diamond instead weights the mean by school size, to show mean expenditure *per pupil* per group.

such as the Akyem, Dangme, and Akuapem that already had similar education access prior to the main Africanization period.[46] These groups are more proximate, yet remained underrepresented.

Second, Ga and Fante bureaucrats were also vastly overrepresented in *both* desk-based jobs in Accra and field-based jobs that would have been spread elsewhere throughout the colony. Without location information in the staff lists, we classify job titles in terms of whether they were likely desk-based or field-based jobs.[47] Examples of field-based jobs include teachers, postal workers, and nurses. On average, there are more Gas and Fantes in desk-based than field-based jobs, which suggests proximity influenced hiring at least to some extent: 74.8% of desk-job holders are Ga or Fante compared to 56.9% of field-job holders. Yet the extremely large share of Ga and Fante workers in field jobs relative to their combined population size (14%) still strongly implies that a proximity-based explanation remains insufficient.

Third, in Section S12 we can directly compare the geographic location of jobs to the ethnicity of state employees for the hiring of government teachers in a way that we cannot for other types of civil servants. We confirm that Ga and Fante government teachers were hired in excess of the proportion of government schools located in Ga and Fante communities, providing clear evidence that hiring did not simply reflect the ethnic breakdown of workplace locations.

Finally, the colonial state expanded significantly in complexity and geographic footprint over time, with new departments coming online and new state institutions (e.g., clinics, post offices) opening continually further into the interior. If jobs were a function of location, we should expect a strong upward trend in employment of non-coastal groups. But Ga and Fante dominance remains virtually constant (Figure 5 and Section S6).

7.3 British ethnic preferences

A final possibility is that British principals had strategic preferences for Ga and Fante employees, potentially perceiving them as more loyal or competent. For example, Ray (2018, 2019) and

[46] Cape Coast (the biggest Fante town) is 125km from central Accra, further than Kibi and Koforidua (the main Akyem towns, 60-75km away), the Akuapem ridge (35km away), and major Dangme settlements ringing the city (20-40km away).

[47] Coding rules are in Section S11.

Lange (2025) suggest that British colonial leaders often favored historically weak, especially pre-colonially non-centralized, ethnic groups for key posts to offset the existing strength of groups that had pre-colonial states.

We cast doubt on this possibility in several ways. Most directly, Ray (2018, 2019) and Lange's (2025) theory of British strategy would explicitly *not* predict the Ga and Fante dominance we observe: unlike some other Ghanaian groups (Nathan, 2023), both the Ga and Fante were centralized in the pre-colonial period; the Fante, in particular, came from a well-established state. To our knowledge, historical literature on Africanization in the Gold Coast also does not document such a preference (e.g., Shaloff, 1974, Rathbone, 2000). We also found no evidence of it in the materials on hiring in Ghana's National Archives: it is notable that across the many hundreds of documents we reviewed discussing African hiring and promotions, we found *zero* mentions of anyone's ethnicity. If the British did have an explicit and sustained preference for these groups, it is reasonable to expect they would have discussed it internally.[48]

More indirectly, we assess observable implications of such a preference. First, the archival documents described above demonstrate that senior British officials took an active role in hiring and promotion decisions for the minority of African employees who were First or Second Division Clerks, while leaving other decisions decentralized to individual departments. But in Section S13 we show that the Ga and Fante are still heavily overrepresented across positions for which hiring was not handled centrally and also that the overrepresentaion of Fante bureaucrats, in particular, is *higher* for jobs not directly selected by top British leadership, seemingly inconsistent with this bias only occurring due to these same officials' top-down directive.

Second, if a top-down directive existed, hiring should have been sensitive to changes in the principals overseeing the process. Until 1951, Africanization was conducted under the ultimate command of British officials. Starting from 1952, before independence in 1957, senior authority passed to Kwame Nkrumah's new CPP government. The CPP leadership primarily came from other ethnic groups, and Nkrumah (an Nzema; "Other Akan") viewed existing African civil ser-

[48]These documents are mostly private memoranda, with little reason to have hidden such a preference if it existed.

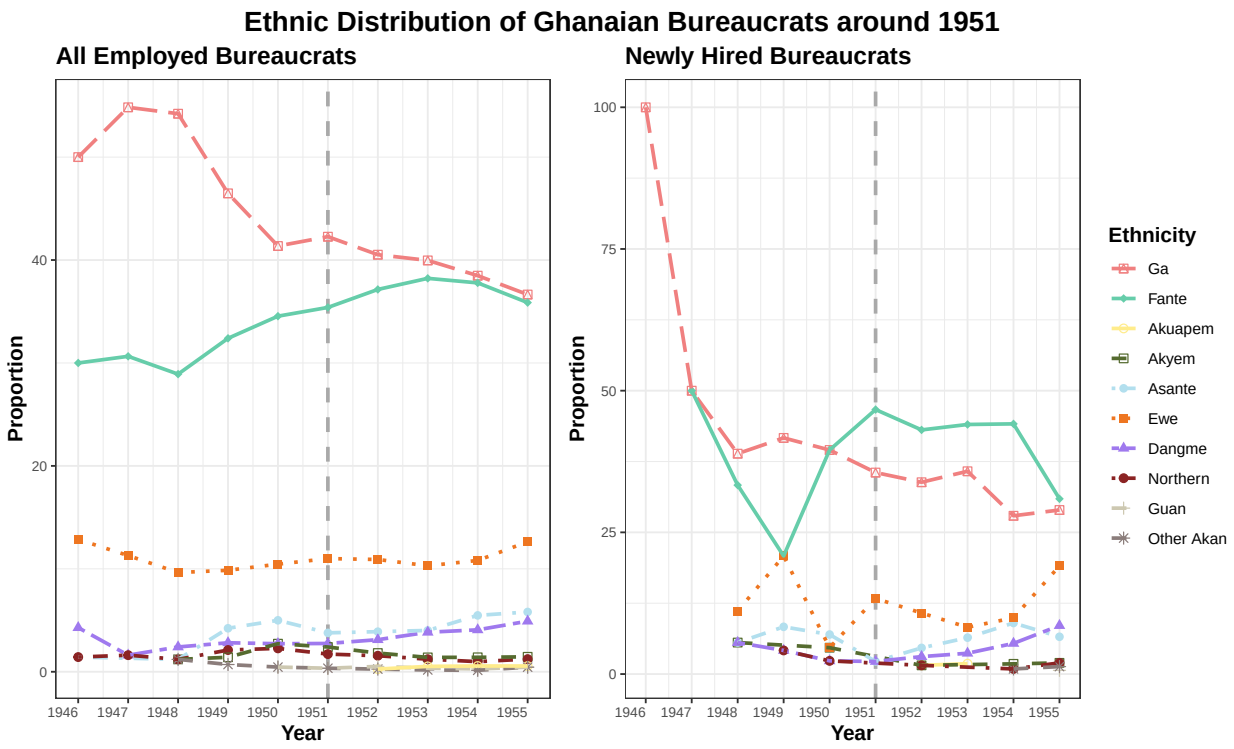


Figure 10: *Distribution of ethnicity among Ghanaian bureaucrats around 1951*: drops all observations below the bottom 25th percentile on the name distinctiveness score.

vants (the large majority of whom were Ga or Fante) as part of his opposition (Zolberg, 1966, pg. 71), such that he had little incentive to continue favoring Ga or Fante hires. Yet in Figure 10, we find no sharp discontinuity in ethnic hiring across the 1951-1952 transfer in power: the right panel shows that Ga and Fante were similarly over-represented as a proportion of new hires regardless of which principal led the process.[49] This relative stasis in new hiring meant little meaningful aggregate shift in the years after Nkrumah took internal power in the overall proportion of Ga and Fante in the central bureaucracy (left panel), inconsistent with their overrepresentation having been due to top-down British orders to favor them.

Third, other accounts of strategic efforts to favor specific groups in hiring suggest such preferences should be visible in variation across positions, departments, or time. For example,

[49]Fisher's exact tests comparing proportions of new hires from each ethnic group just before versus just after the change in power fail to reject a null hypothesis of no significant change in proportions of new hires from each group when comparing any of the 1, 2, 3, or 4 years before vs. after Nkrumah's election to each other.

groups perceived as inherently more loyal should be privileged for key security or administrative posts (Ray, 2012; Harkness, 2018; Hassan, 2017, 2020), or assigned to higher-rank jobs. Similarly, assuming British leaders' most pressing local threats must have evolved across our long time series, we would expect the groups favored in hiring to have changed as well (Hassan, 2020).

Yet in Figure 11 we see no systematic pattern in other groups' presence across departments of varying strategic importance, while the Ga and Fante dominate across the board.[50] Ga and Fante bureaucrats are not differentially concentrated in core security or administrative departments most vital to law and order or political governance. They were also actually *less* overrepresented in departments concerned with agricultural production and customs – the primary revenue streams sustaining British rule – than in other less vital parts of the state.[51] Ga and Fante dominance also holds across both senior (Figure 5) and junior (African staff list; Section S6) roles, and at all tiers of the salary distribution — a pattern inconsistent again with British prioritization of groups thought most competent for higher-rank positions.

8 Implications for post-independence politics

Similar to many other post-colonial settings in which ethnic grievances sparked by uneven colonial state-building led to post-colonial competition and conflict (Wimmer, 2013*b*; Ray, 2018, 2019; Kuipers, 2025; Lange, 2025), the already ethnically-captured, and non-meritocratic, central state inherited at independence created challenges for Ghana's immediate post-independence leaders, who struggled to dismantle existing hiring practices and create more diverse institutions.

Ethno-regional opposition parties quickly formed partly in response to their exclusion from state institutions, centered around three disadvantaged groups. Anticipating marginalization under

[50] We code strategic importance in three areas: agriculture, governance, and law and order. For agriculture, Agricultural and Customs were coded "high," others "low." For governance we classified as high: accountant general, audit, administrative, commerce, customs, governor, income tax, judiciary, political administration, and supreme court; for law and order: police, prisons, and regiment. Results by department appear in Section S14, where we likewise find no coherent pattern for other ethnic groups.

[51] As discussed below, the colonial army (not in our data) had relatively more Northerners at low ranks (Harkness, 2018) and Ewes at high ranks (Hutchful, 1985) than the civil service. To the extent that literally arming one's subjects reveals preferences about perceived loyalty (Ray, 2012; Roessler, 2016), it is notable the British did not favor the Ga and Fante in this separate institution.

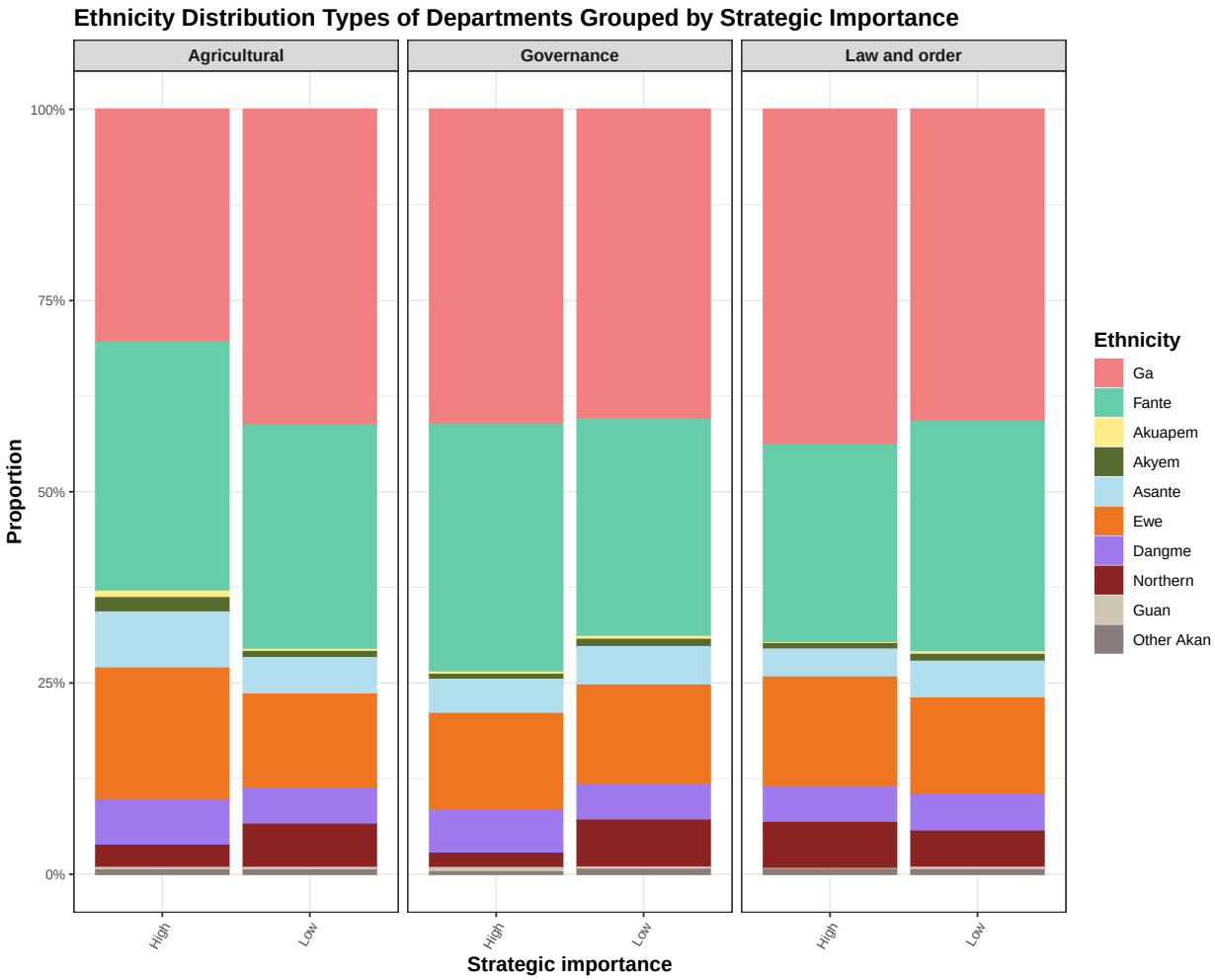


Figure 11: *Distribution of ethnicity among Ghanaian bureaucrats by strategic importance of department*: drops all observations below the bottom 25th percentile on the name distinctiveness score.

a strong post-independence government, the Asante-led, pro-federalist National Liberation Movement (NLM) emerged in 1954, also drawing support from other major Akan groups (e.g., the Akyem). The NLM was sustained at the grassroots by educated, but underemployed Asante who were frustrated by their exclusion from both traditional power structures and the civil service (Allman, 1990, 1993). Similarly, the Northern People's Party (NPP) pushed for greater Northern inclusion in national institutions, while Ewe elites established the Togoland Congress and Anlo Youth Association, mobilizing around a (ultimately failed) 1956 UN-sponsored referendum on secession of Togoland, the core Ewe area (Austin, 1964; Skinner, 2015).

Confronted with these widespread ethnic grievances, Ghana's new leaders attempted to address imbalances in state institutions. In the early 1960s, Nkrumah's government pivoted towards explicitly recruiting more personnel from then-marginalized groups (Adei, 2008), especially the Akan, into both the bureaucracy and military (Harkness, 2018) – efforts at rebalancing that future scholars would instead (erroneously in our view) cite as examples of *new* ethnic patrimonialism (e.g., Zolberg, 1966, Adei, 2008).

These explicit top-down measures to reshape the bureaucracy reduced Ga and Fante dominance somewhat, but given the large accumulated stock of bureaucrats from these two groups to that point, their advantages proved fairly durable. Digitizing additional, individual-level bureaucratic data for junior bureaucrats in 1968 – more than a decade after independence – Fantes and Gas still represented 24% and 26%, respectively. Ewe representation rose to 20%, followed by Asantes at 14%.[52] These patterns were consistent across nearly all ministries and departments.[53]

Ethnic imbalances in the central state also had implications for political stability, as disgruntled, educated youths unable to find central state employment instead joined the army, an institution where Ga and Fante advantages were less entrenched. While the army's rank-and-file was predominantly Northern (Ray, 2012, p. 565), a consequence of British assumptions that uneducated Northerners made better soldiers (Harkness, 2018, p. 146), the officer corps was more

[52]This data digitizes archival lists of all bureaucrats as of July 1, 1968, the only post-colonial staff list for Ghana that we could locate. Section S15.

[53]Section S16.

diverse. Educated Ewe, in particular, had joined the army in large numbers in the 1950s. Consistent with our exam data above, Hutchful (1985, p. 168) notes that "an important factor that made the Army appealing to Ewe... was the limited opportunities for educated youth," as other "government departments" could not "offer employment that met the aspirations of [Ewe] school leavers." In response to the resulting ethnic imbalances, Nkrumah attempted to diversify the officer corps by also incorporating more Akan (including Asante) officers (Harkness, 2018). Yet this strategy backfired, as the new Asante and Ewe officers ultimately (briefly) united to depose him (Asante and Gyimah-Boadi, 2004). Ghana then cycled through multiple further military coups and purges from the state.

Thus, despite Nkrumah's efforts at ethnic inclusion, entrenched patterns of dominance contributed to instability, and sparked further struggles for ethnic rebalancing. After the Asante and Ewe coup plotters fell out and Asante elites consolidated control in the late 1960s, new attempts to increase Akan representation sparked fears of exclusion among other groups. Under the Busia government (1969) more than 560 civil servants – predominantly Ewe and Ga – were dismissed in an effort to reverse their bureaucratic power, another case cited as evidence of *new* patrimonial practices in the bureaucracy (Asante and Gyimah-Boadi, 2004).

The exact manner in which legacies of patrimonialism in colonial-era hiring shaped post-independence politics is no doubt a function of a wider set of factors that varied across African cases. Among others, this includes the overall extent of Africanization prior to independence, the resulting level of ethnic (im)balance in hiring, whether the initial post-colonial leader was from the same group(s) favored or excluded from the colonial bureaucracy, and how these distributions of power corresponded to overall demographic and economic conditions. But the Ghanaian experience elucidates the broader insight that the manner in which the nascent bureaucracy was staffed in the colonial period cannot be separated from our understanding of a country's immediate post-independence political contestation over who controlled the state.

9 Conclusion

This paper theorizes public sector recruitment and argues that bureaucrats' preferences in hiring will outweigh those of politicians when politicians lack social connections to the communities they govern and have relatively low intensity preferences about who to hire. Both social distance and weak preferences exacerbate principal-agent problems in recruitment by making it more difficult for politicians to overcome information asymmetries with existing bureaucrats. Digitizing colonial records of bureaucrats in Ghana under British colonial rule and using a novel method to link individuals to ethnic groups, we reveal the overwhelming dominance of bureaucrats from two small ethnic groups. We argue that this dominance was the result of first-mover advantages combined with bureaucrat-led insider capture, with groups with more existing bureaucrats at high ranks within state departments able to hire more of their own regardless of formal policies of meritocracy.

Our work is not the first to observe ethnic disparities in colonial hiring (Ray, 2012; Johnson-Kanu, 2021; Ricart-Huguet, 2021; Kuipers, 2025; Lange, 2025). But our key contribution is to propose a novel explanation for *why* this imbalance occurred. In addition to arguing that imbalances resulted from strategic British preferences (e.g., Harkness 2018, Lange 2025), the most common explanation in existing literature is that colonial inequality in access to education drove colonial hiring (Mazrui, 1978; Ray, 2018; Johnson-Kanu, 2021; Ricart-Huguet, 2021; Kuipers, 2025; Lange, 2025). We suggest, however, that when using richer data than previously available to unpack dynamics in hiring within the colonial period against dynamics in the development of colonial education systems, this common explanation may not always retain the power its proponents claim. Even in settings with stronger ethnic educational inequalities than Ghana, especially those that were not settler colonies, our theory suggests bureaucrat-led ethnic lock-in may have played a previously-unexplored role under the surface in explaining these ethnic imbalances. We leave it to future work to investigate evidence for bureaucrat-led ethnic lock-in across other cases.

Our results also suggest the need to update conventional wisdom about post-colonial leaders as the chief orchestrators of ethnic packing within the central state after independence. Post-colonial central bureaucracies were not blank slates upon which new leaders built. Instead, many

leaders inherited packed bureaucracies. They then faced the challenge of both building the public sector and dismantling existing biases. As they did so, post-colonial bureaucracies sometimes became less ethnically homogeneous over time (Hassan, 2020), challenging standard accounts of rampant post-colonial favoritism.

More generally, fruitful avenues for future work also include developing a deeper understanding of the strategies that bureaucrats take to informally influence hiring outcomes. We also hope that our analysis stimulates further research using contemporary data to classify the relevant social identities of bureaucrats in prior periods, thereby developing a deeper understanding of how historical events may continue to shape modern states.

References

Adei, Stephen. 2008. Building a Capable Public Administration. In *Ghana: Governance in the Fourth Republic*, ed. Baffour Agyeman-Duah. Ghana Center for Democratic Development.

Afigbo, Adiele Eberechukwu. 1972. *The Warrant Chiefs: Indirect Rule in Southeastern Nigeria, 1891-1929*. London: Longman Publishers.

Allman, Jean M. 1990. "The Youngmen and the Porcupine: Class, Nationalism and Asante's Struggle for Self-Determination, 1954-57." *The Journal of African History* 31(2):263–279.

Allman, Jean M. 1993. *The Quills of the Porcupine: Asante Nationalism in an Emergent Ghana*. University of Wisconsin Press.

Anzia, Sarah F. and Jessica Trounstine. 2024. "Civil Service Adoption in America: The Political Influence of City Employees." *American Political Science Review* pp. 1–17.

Apter, David E. 1963(1955). *Ghana in Transition (2nd edition)*. Atheneum Press.

Arhin, Kwame. 1983. "Rank and Class among the Asante and Fante in the Nineteenth Century." *Africa* 53(1):2–22.

Asante, Richard and Emmanuel Gyimah-Boadi. 2004. "Ethnic Structure, Inequality and Governance of the Public Sector in Ghana." UNRISD Working Paper.

Ashraf, Nava and Oriana Bandiera. 2018. "Social Incentives in Organizations." *Annual Review of Economics* 10(1):439–463.

Austin, Dennis. 1964. "Politics in Ghana, 1946-1960.".

Austin, Gareth. 2008. "The 'reversal of fortune'thesis and the compression of history: Perspectives from African and comparative economic history." *Journal of International Development: The Journal of the Development Studies Association* 20(8):996–1027.

Barkey, Karen. 1996. *Bandits and Bureaucrats: The Ottoman Route to State Centralization*. Ithaca, NY: Cornell University Press.

Bayart, Jean-Francois. 1993. *The State in Africa: The Politics of the Belly*. Longman Publishers.

Bird, Samuel S and Aleksandr Michuda. 2025. "Ethnic diversity and voting behavior at scale: Evidence from Uganda." *World Development* 196:107181.

Boone, Catherine. 2003. *Political Topographies of the African State: Territorial Authority and Institutional Choice*. Cambridge University Press.

Bozcaga, Tugba. 2026. "The Social Bureaucrat: How Social Proximity among Bureaucrats Affects Local Governance." *American Journal of Political Science, Forthcoming* .

Brierley, Sarah. 2021. "Combining Patronage and Merit in Public Sector Recruitment." *Journal of Politics* 83(1):182–197.

Brierley, Sarah and Noah L. Nathan. 2021. "The Connections of Party Brokers: Which Brokers Do Parties Select?" *Journal of Politics* 83(3):884–901.

Cardoso, Danilo, Flavio Cireno, Julien Labonne, Pedro Masson, Pedro Palotti, Flavio de Vitoria and Martin Williams. 2025. "Interpersonal Connections and Political Ties in Bureaucratic Networks: Evidence from Brazil.".

Chabal, Patrick and Jean-Pascal Daloz. 1999. *Africa Works: Disorder as Political Instrument*. Bloomington, IN: Indiana University Press.

Chaturvedi, Rochana and Sugat Chaturvedi. 2024. "It's All in the Name: A Character-Based Approach to Infer Religion." *Political Analysis* 32(1):34–49.

Cogneau, Denis, Yannick Dupraz and Sandrine Mesplé-Somps. 2018. "African states and development in historical perspective: Colonial public finances in British and French West." *Paris School of Economics, Working Paper 29* .

Collier, Ruth Berins. 1982. *Regimes in Tropical Africa: Changing Forms of Supremacy, 1945-1975*. University of California Press.

Colonnelli, Emanuele, Mounu Prem and Edoardo Teso. 2020. "Patronage and selection in public sector organizations." *American Economic Review* 110(10):3071–3099.

Cooper, Frederick. 1996. *Decolonization and African society: The labor question in French and British Africa*. Number 89 Cambridge University Press.

de Graft-Johnson, J.C. 1969. "The Population of Ghana 1846-1967: A Digest and Discussion of the Data in the Official Counts and Censuses." *Transactions of the Historical Society of Ghana* 10(1):1–12.

Egorov, Georgy and Konstantin Sonin. 2011. "Dictators and their Viziers: Endogenizing the Loyalty-Competence Trade-off." *Journal of the European Economic Association* 9(5):903–930.

Ekeh, Peter P. 1975. "Colonialism and the Two Publics in Africa: A Theoretical Statement." *Comparative Studies in Society and History* 17(1):91–112.

Englebert, Pierre. 2000. *State Legitimacy and Development in Africa*. Boulder, CO: Lynne Riener Publishers.

Enos, Ryan D. 2016. "What the demolition of public housing teaches us about the impact of racial threat on political behavior." *American Journal of Political Science* 60(1):123–142.

Evans, Peter and James E Rauch. 1999. "Bureaucracy and growth: A cross-national analysis of the effects of "Weberian" state structures on economic growth." *American sociological review* 64(5):748–765.

Frankema, Ewout H.P. 2012. "The Origins of Formal Education in Sub-Saharan Africa: Was British Rule More Benign?" *European Review of Economic History* 16:335–355.

Garfias, Fracisco and Emily A. Sellars. 2024. "Fiscal Legibility and State Development: Theory and Evidence from Colonial Mexico." *American Journal of Political Science* pp. 1–16.

Geddes, Barbara. 1994. *Politician's Dilemma: Building State Capacity in Latin America*. Berkeley, CA: University of California Press.

Goldsmith, Arthur A. 1999. "Africa's overgrown state reconsidered: Bureaucracy and economic growth." *World Politics* 51(4):520–546.

Graham, C.K. 1971. *The History of Education in Ghana: From the Earliest Times to the Declaration of Independence*. London: Frank Cass and Co.

Grindle, Merilee S. 2012. *Jobs for the Boys*. Harvard University Press.

Grischow, Jeff D. 2006. *Shaping Tradition: Civil Society, Community and Development in Colonial Northern Ghana, 1899-1957*. Boston, MA: Brill.

Grzymala-Busse, Anna. 2007. *Rebuilding Leviathan: Party competition and state exploitation in post-communist democracies*. Cambridge University Press.

Harkness, Kristen A. 2018. *When Soldiers Rebel: Ethnic Armies and Political Instability in Africa*. Cambridge University Press.

Harris, Adam S, Jan-Hinrik Meyer-Sahling, Kim Sass Mikkelsen, Christian Schuster, Brigitte Seim and Rachel Sigman. 2023. "Varieties of connections, varieties of corruption: Evidence from bureaucrats in five countries." *Governance* 36(3):953–972.

Harris, J. Andrew. 2015. "What's in a Name? A Method for Extracting Information about Ethnicity from Names." *Political Analysis* 23:212–224.

Hassan, Mai. 2017. "The Strategic Shuffle: Ethnic Geography, the Internal Security Apparatus, and Elections in Kenya." *American Journal of Political Science* 61(2):382–395.

Hassan, Mai. 2020. *Regime Threats and State Solutions: Bureaucratic Loyalty and Embeddedness in Kenya*. New York: Cambridge University Press.

Hassan, Mai, Horacio Larreguy and Stuart Russell. 2024. "Who Gets Hired? Political Patronage and Bureaucratic Favoritism." *American Political Science Review* p. 1–18.

Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton, NJ: Princeton University Press.

Hjort, Jonas. 2024. "Ethnic Divisions and Production in Firms." *Quarterly Journal of Economics* 129(4):1899–1946.

Honig, Dan. 2024. *Mission Driven Bureaucrats: Empowering People to Help Government Do Better*. Oxford University Press.

Huntington, Samuel P. 1968. *Political Order in Changing Societies*. New Haven: Yale University Press.

Hutchful, Eboe. 1985. "The Development of the Army Officer Corps in Ghana, 1956-1966." *Journal of African Studies* 12(3):163–173.

Imai, Kosuke and Kabir Khanna. 2016. "Improving ecological inference by predicting individual ethnicity from voter registration records." *Political Analysis* 24(2):263–272.

Imai, Kosuke, Santiago Olivella and Evan TR Rosenman. 2022. "Addressing census data problems in race imputation via fully Bayesian Improved Surname Geocoding and name supplements." *Science Advances* 8(49):eadc9824.

Jedwab, Remi, Felix Meier zu Selhausen and Alexander Moradi. 2022. "The economics of missionary expansion: evidence from Africa and implications for development." *Journal of Economic Growth* 27:149–192.

Jeffries, Richard. 1978. *Class, Power, and Ideology in Ghana: The Railwaymen of Sekondi*. Cambridge University Press.

Johnson-Kanu, Ada. 2021. "Colonial Legacies in State Building: Ethnicity and Bureaucratic Representation in Nigeria." Working Paper.

Kasara, Kimuli. 2013. "Separate and Suspicious: Local Social and Political Context and Ethnic Tolerance in Kenya." *Journal of Politics* 75(4):921–936.

Kelly, Gail Margaret. 1959. The Ghanaian Intelligentsia PhD thesis The University of Chicago.

Kitschelt, Herbert and Steven I Wilkinson. 2007. *Patrons, clients and policies: Patterns of democratic accountability and political competition*. Cambridge University Press.

Kuipers, Nicholas. 2025. *States Against Nations: Meritocracy, Patronage, and the Challenges of Bureaucratic Selection*. Cambridge University Press.

Lange, Matthew. 2009. *Lineages of Despotism and Development: British Colonialism and State Power*. University of Chicago Press.

Lange, Matthew. 2025. *Legacies of British Rule: Colonialism, Statehood, and Nationalist Civil War*. Princeton University Press.

Lee, Diana Da In and Yamil Ricardo Velez. 2025. "Measuring descriptive representation at scale: Methods for predicting the race and ethnicity of public officials." *British Journal of Political Science* 55:e110.

Lee, Melissa M. 2022. "International Statebuilding and the Domestic Politics of State Development." *Annual Review of Political Science* 25(1):261–281.

Lungu, Gatian F. 1980. "Africanization and the merit principle in the Zambian public service." *Journal of Administration Overseas* 19(2):88–99.

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

Matsuzaki, Reo. 2019. *Statebuilding by Imposition: Resistance and Control in Colonial Taiwan and the Philippines*. Ithaca, NY: Cornell University Press.

Mazrui, Ali A. 1978. *Political values and the educated class in Africa*. University of California Press.

Migdal, Joel S. 1988. *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*. Princeton, NJ: Princeton University Press.

Mkandawire, Thandika. 2015. "Neopatrimonialism and the Political Economy of Economic Performance in Africa: Critical Reflections." *World Politics* 67(3):563–612.

Nathan, Noah L. 2019. *Electoral Politics and Africa's Urban Transition: Class and Ethnicity in Ghana*. Cambridge University Press.

Nathan, Noah L. 2023. *The Scarce State: Inequality and Political Power in the Hinterland*. Cambridge University Press.

Neggers, Yusuf. 2018. "Enfranchising Your Own? Experimental Evidence on Bureaucrat Diversity and Election Bias in India." *American Economic Review* 108(6):1288–1321.

Oliveros, Virginia. 2021. *Patronage at work: Public jobs and political services in Argentina*. Cambridge University Press.

Parker, John. 2000. *Making the Town: Ga State and Society in Early Colonial Accra*. Oxford, UK: James Currey Press.

Pepinsky, Thomas B., Jan H. Pierskalla and Audrey Sacks. 2017. "Bureaucracy and Service Delivery." *Annual Review of Political Science* 20(1):249–268.

Peters, B Guy and Jon Pierre. 2004. *Politicization of the civil service in comparative perspective: The quest for control*. London: Routledge.

Pierskalla, Jan H and Audrey Sacks. 2020. "Personnel politics: Elections, clientelistic competition and teacher hiring in indonesia." *British Journal of Political Science* 50(4):1283–1305.

Pitcher, Anne, Mary H. Moran and Michael Johnston. 2009. "Rethinking Patrimonialism and Neopatrimonialism in Africa." *African Studies Review* 52(1):125–156.

Price, Joseph Henry. 1967. *Political Institutions of West Africa*. Taylor & Francis.

Price, Robert M. 1975. *Society and Bureaucracy in Contemporary Ghana*. Univ of California Press.

Rathbone, Richard. 2000. "The transfer of power and colonial civil servants in Ghana." *The Journal of Imperial and Commonwealth History* 28(2):67–84.

Rauch, James E and Peter B Evans. 2000. "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries." *Journal of Public Economics* 75(1):49–71.

Ray, Subhasish. 2012. "The Nonmartial Origins of the 'Martial Races': Ethnicity and Military Service in Ex-British Colonies." *Armed Forces & Society* 39(3):560–575.

Ray, Subhasish. 2018. "Beyond Divide and Rule: Explaining the Link between British Colonialism and Ethnic Violence." *Nationalism and Ethnic Politics* 24(4):367–388.

Ray, Subhasish. 2019. "History and Ethnic Conflict: Does Precolonial Centralization Matter?" *International Studies Quarterly* 63:417–431.

Ricart-Huguet, Joan. 2021. "Colonial Education, Political Elites, and Regional Political Inequality in Africa." *Comparative Political Studies* 54(14):2546–2580.

Ricart-Huguet, Joan. 2022. "The Origins of Colonial Investments in Former British and French Africa." *British Journal of Political Science* 52(2):736–757.

Roessler, Philip. 2016. *Ethnic Politics and State Power in Africa: The Logic of the Coup-Civil War Trap*. Cambridge University Press.

Sederberg, Peter C. 1971. "The Gold Coast under colonial rule: an expenditure analysis." *African Studies Review* 14(2):179–204.

Shaloff, Stanley. 1974. "The Africanization Controversy in the Gold Coast, 1926-1946." *African Studies Review* 17(3):493–504.

Shefter, Martin. 1977. "Party and Patronage: Germany, England, and Italy." *Politics and Society* 7:403–451.

Shefter, Martin. 1994. *Political Parties and the State: the American Historical Experience*. Princeton, NJ: Princeton University Press.

Simson, Rebecca. 2017. (Under) privileged bureaucrats? The changing fortunes of public servants in Kenya, Tanzania and Uganda, 1960–2010 PhD thesis London School of Economics and Political Science.

Skinner, Kate. 2015. *The fruits of freedom in British Togoland : literacy, politics and nationalism, 1914-2014*. Cambridge University Press.

Stokes, Susan C, Thad Dunning and Marcelo Nazareno. 2013. *Brokers, voters, and clientelism*. Cambridge University Press.

Tignor, Robert L. 1993. "Political Corruption in Nigeria before Independence." *Journal of Modern African Studies* 31(2):175–202.

van de Walle, Nicolas. 2001. *African Economies and the Politics of the Permanent Crisis, 1979-1999*. New York: Cambridge University Press.

von Hesse, Hermann W. 2024. "More Than an Intermediary: James Bannerman and Colonial Space-Making on the Nineteenth-Century Gold Coast." *African Studies Review* pp. 1–20.

Ward, William Ernest Frank. 1966. *A history of Ghana*. Routledge.

Weber, Max. 1949(1919). Politics as a Vocation. In *Max Weber: Essays in Sociology*, ed. Hans H. Gerth and C. Wright Mills. Oxford University Press.

Weber, Max. 1978. *Economy and society: An outline of interpretive sociology*. Vol. 1 University of California Press.

Wimmer, Andreas. 2013*a*. *Ethnic Boundary Making: Institutions, Power, and Networks*. Oxford University Press.

Wimmer, Andreas. 2013*b*. *Waves of War: Nationalism, State Formation, and Ethnic Exclusion in the Modern World*. Cambridge University Press.

Wirsching, Elisa Maria. 2025. "Barriers to Representation: Selection Processes and Political Diversity in US Urban Bureaucracy." Working paper, Center for the Study of Democratic Politics (CSDP), Princeton University.

Xu, Guo. 2018. "The Costs of Patronage: Evidence from the British Empire." *American Economic Review* 108(11):3170–3198.

Young, Crawford. 1994. *The African Colonial State in Comparative Perspective*. New Haven, CT: Yale University Press.

Young, Crawford. 2004. "The End of the Post-Colonial State in Africa? Reflections on Changing African Political Dynamics." *African Affairs* 103:23–49.

Young, Crawford and Thomas E. Turner. 1985. *The Rise and Decline of the Zairian State*. Madison: University of Wisconsin Press.

Young, M. Crawford. 1966. "Post-Independence Politics in the Congo." *Transition* (26):34–41.

Zolberg, Aristide R. 1966. *Creating Political Order: The Party-States of West Africa*. Rand McNally.

Supporting Information (SI)

Table of Contents

S1 Coverage of colonial bureaucrats staff list (pg. 17)

Table SI1: Data Coverage in Senior and Hybrid Staff Lists

Year	Total #	Ghanaian #	British #	Ghanaian %	British %
1897	106	2	104	1.9	98.1
1899	123	3	120	2.4	97.6
1900	140	2	138	1.4	98.6
1903	265	2	263	0.8	99.2
1904	263	3	260	1.1	98.9
1907	266	5	261	1.9	98.1
1912	391	9	382	2.3	97.7
1922	613	4	609	0.7	99.3
1928	906	66	840	7.3	92.7
1930	1007	78	929	7.7	92.3
1931	903	74	829	8.2	91.8
1932	748	68	680	9.1	90.9
1933	711	69	642	9.7	90.3
1934	716	74	642	10.3	89.7
1935	710	76	634	10.7	89.3
1937	734	91	643	12.4	87.6
1938	777	98	679	12.6	87.4
1939	859	107	752	12.5	87.5
1943	798	60	738	7.5	92.5
1944	706	61	645	8.6	91.4
1945	773	82	691	10.6	89.4
1946	791	93	698	11.8	88.2
1947	881	85	796	9.6	90.4
1948	1057	121	936	11.4	88.6
1949	1220	198	1022	16.2	83.8
1950	1374	310	1064	22.6	77.4
1951	1599	411	1188	25.7	74.3
1952	1792	558	1234	31.1	68.9
1953	2078	811	1267	39.0	61.0
1954	2416	1027	1389	42.5	57.5
1955	2697	1280	1417	47.5	52.5

Years in **bold (1947-1955)**: "hybrid" lists that also include more mid-tier positions.

Table SI2: Data Coverage in African Staff List

Year	Total
1936	2973
1937	2956
1938	3208
1940	3769
1941	3848
1942	3876
1943	3957
1944	4090
1945	4263

S2 Breakdown of Northern and Guan ethnicity classifications (pg. 19)

Our analyses lump all bureaucrats from Northern Ghana (comprising a small number of observations overall) into a single “Northern" category. Ethnic distinctions among Northerners are very salient for local politics in the North. However, both the colonial and post-independence states were overwhelmingly dominated by Southerners, with Northerners treated as a bloc in independence-era politics (Nathan, 2023). We can thus collapse the over 30 separate Northern groups into a single category for analytical simplicity.

All sub-groups within the upper-level Mole-Dagbon, Grusi, Gurma, and Mande categories on Ghana’s census are categorized together in this Northern category. Groups originally from other Sahelian countries to Ghana’s north that fall within the census’s "Other" upper-level category, such as the Mosi, Hausa, Fulani, and Zabrama, are also categorized as Northern, as most of their populations in Ghana are in the North. One tiny Akan sub-group, the Chokosi, that is primarily settled in Northern Region is also categorized in the Northern set.

Separately, the Guan upper-level category on the census includes small linguistically-similar sub-groups scattered throughout the country. The largest of these – the Gonja – is one of the main ethnic groups in Northern Ghana, concentrated in what is now the Savannah Region (formerly Northern Region). Gonjas are included in the Northern category. Most other Guan sub-groups live in the south (primarily in the Volta, Oti, and Central Regions) and are grouped together as a residual category in our analysis.

S3 Ethnicity coding procedure (pg. 19)

We take the 13.8 million names of registered voters in the 2015 Ghana voters register and split them into each of 1.1 million unique name fragments (e.g., “John Mensah" becomes “John" and “Mensah").[54] Each name fragment, f, is linked to the polling station, p, at which that name is registered. Polling stations are matched by geolocation to their surrounding 2010 census Enumeration Areas (EAs, or tracts) and we calculate the ethnic group proportions of the population residing at EAs falling within 500m of the polling station’s location, producing a vector $\mathbf{E}_p$ of group proportions by polling station, with G elements, indexing a mutually-exclusive set of ethnic categories. Let $g \in \{1,\ldots,G\}$ index ethnic groups, and let E_{pg} denote the proportion of group g at polling station p.

For each unique name fragment in the voters register (f), we then calculate the average of ethnic group population proportions across all polling stations ($\mathbf{E}_f$) at which that name appears, weighting by the frequency of appearances of that name fragment at each polling station (n_{fp}):

$$\mathbf{E}_f = \frac{\sum_{p=1}^{P} n_{fp}\mathbf{E}_p}{\sum_{p=1}^{P} n_{fp}} \tag{1}$$

[54] Note that the frequency distribution of name fragments is *very* right-skewed: a relatively small proportion of the total fragments appear a very large number of times, representing all commonly used names in the country, while 52% of the unique fragments appear exactly once. From visual inspection, many of these cases appear to represent unique misspellings of more common names. The median name fragment thus appears once, but the mean fragment appears 39.8 times. The most common name in Ghana is the Akan surname “Mensah", used by 437,425 different voters.

$\mathbf{E}_f$ thus represents the ethnic composition of the local communities in which name fragment f appears. Each element, E_{fg}, is the average proportion of group g in these communities.

Separately, we calculate the base rate $\mathbf{B}$ of ethnic group proportions across all polling stations, to approximate the group proportions among all registered voters in Ghana:

$$\mathbf{B} = \frac{1}{P} \sum_{p=1}^{P} \mathbf{E}_p \tag{2}$$

where each element, B_g, is the nationwide proportion of group g across all polling stations.

For each name fragment (f), we then calculate the Euclidean distance (D_f) between the vector of ethnic group population proportions of the locations at which the name fragment appears ($\mathbf{E}_f$) and the vector of base rate population proportions nationwide ($\mathbf{B}$):

$$D_f = \left\| \mathbf{E}_f - \mathbf{B} \right\|_2 = \sqrt{\sum_{g=1}^{G} \left(E_{fg} - B_g \right)^2} \tag{3}$$

where g indexes ethnic groups (e.g., Asante, Ewe, Fante, Ga, Dangme, etc.). Euclidean distance is a measure of how "atypical" the ethnic distribution of fragment f is compared to the overall population $\mathbf{B}$. For ease of interpretation, we then transform the raw Euclidean distance into its percentile score across all name fragments, producing our *distinctiveness score* for each fragment:

$$\text{Distinctiveness}_f = \frac{1}{F} \sum_{f'=1}^{F} \mathbb{I}(D_{f'} \leq D_f) \tag{4}$$

Distinctiveness_f ranges from 0 to 1 with higher values corresponding to more distinctive name fragments.

For each African bureaucrat i in our data, we similarly split their names into all component name fragments, j, dropping any single letter initials.[55] Because spellings of each bureaucrat's name sometimes differ slightly across years in our time-series dataset, the name fragments j associated with each bureaucrat i include all versions of the name. We exact match each fragment (j) from the bureaucrat data to the name fragments from the register (f); 98.1% of the fragments match.[56]

We then calculate $\mathbf{E}_i$ the average ethnic group population proportions of the name fragments j associated with bureaucrat i, by weighting the $\mathbf{E}_j$ vectors by the distinctiveness scores associated with each of their component name fragments (Distinctiveness_j):

$$\mathbf{E}_i = \frac{\sum_{j \in J_i} \text{Distinctiveness}_j \cdot \mathbf{E}_j}{\sum_{j \in J_i} \text{Distinctiveness}_j} \tag{5}$$

This is explicitly designed such that more distinctive name fragments count more, as they provide more information about i's likely ethnicity.

[55] "John Mensah" becomes "John" and "Mensah" but "J. Mensah" only becomes "Mensah". The staff lists we digitize, unfortunately, varied widely year by year – and even across departments within years – in whether they included given names or only first initials.

[56] The tiny set of bureaucrats with no matching fragments have ethnicity coded as NA throughout our analysis.

For each bureaucrat, we then subtract the base rate ethnic proportions nationwide $\mathbf{B}$ from $\mathbf{E}_i$, to identify the ethnic group g^* from which E_{ig} is most positively deviant from the base rate for that group (B_g):

$$g_i^* = \arg \max_{g \in \{1,\ldots,G\} : E_{ig} > B_g} (E_{ig} - B_g) \tag{6}$$

We label this group (g^*) as that bureaucrat's "most likely" ethnicity. As our measure of uncertainty in this classification we also assign each bureaucrat the average distinctiveness score of its component name fragments.

We also deviate from this general procedure for an exceptional situation, based on case knowledge of naming conventions used in a small subset of our data. Because Western education came very late to Northern Ghana, the use of standardized surnames was not yet widespread in Northern Ghana in the colonial period. Instead, upon inspecting the colonial staff lists, we realized that the British engaged in a practice of inventing surnames on the spot for some of (their small number of) uneducated Northern employees (e.g., who often worked as guards or in other low-wage positions) in which they often simply assigned the bureaucrat his ethnic group's name as a placeholder surname. Names such as "Musah Dagomba", "Issa Grunshi" appear in the staff lists and quite literally signal ethnicity, with no ambiguity.[57] This antiquated naming practice is no longer common today and surnames like "Dagomba" are not widely used in the contemporary voters register. In cases where a bureaucrat's name fragment exactly matches to the name of a Northern ethnic group in the census, we automatically assign the bureaucrat that group as their "most likely" ethnicity and set their distinctiveness score to 1, to reflect our certainty in their ethnicity coding.

Finally, to ease numeric interpretation of the distinctiveness scores for each bureaucrat, we note that in some of the data plots below we scale these scores by their percentile within our dataset. For example, when we drop low distinctiveness names in our main specifications in the main text, this is executed as dropping names falling below the 25th percentile that appears in our data, not based on a specific raw numeric value of the distinctiveness score.

We show in S3.1 and S3.2 below that, except for the Northern names that have minimal ambiguity as detailed above, there is no substantial variation that signals an alarming bias across ethnicity groups or salary ranks in the bureaucrats' name distinctiveness in both the full data and the subset that dropped the lowest 25th percentile. Akuapem, Akyem and Asante names do display relatively low distinctiveness scores, on average. This is because names from these closely-related Akan sub-groups often get cross-labeled for *each other* (not for Gas or Fantes), but such a pattern should not raise a serious concern for our main conclusions, because there is no clear directional bias in this cross-identification and these three groups all appear underrepresented relative to their population sizes in our bureaucrats dataset. Scrambling the Akuapem, Akyem, and Asante labels among each other would change nothing about our core conclusions in the main text.

[57] These Muslim (Arabic) names literally translate to Moses the Dagomba or Jesus the Grunshi, for example.

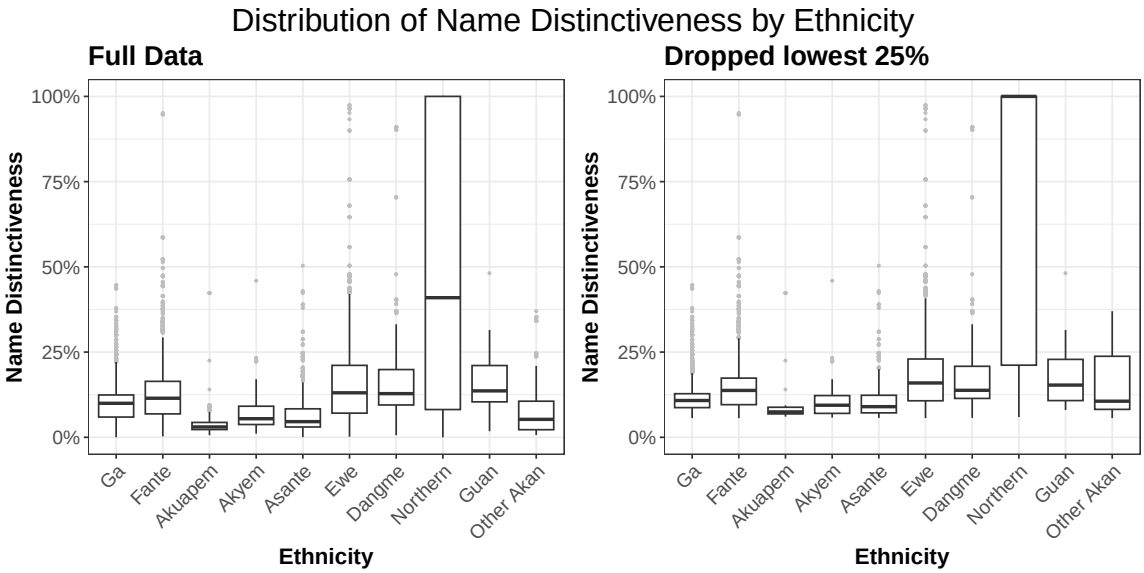


Figure S3.1: *Name Distinctiveness by Ethnicity*

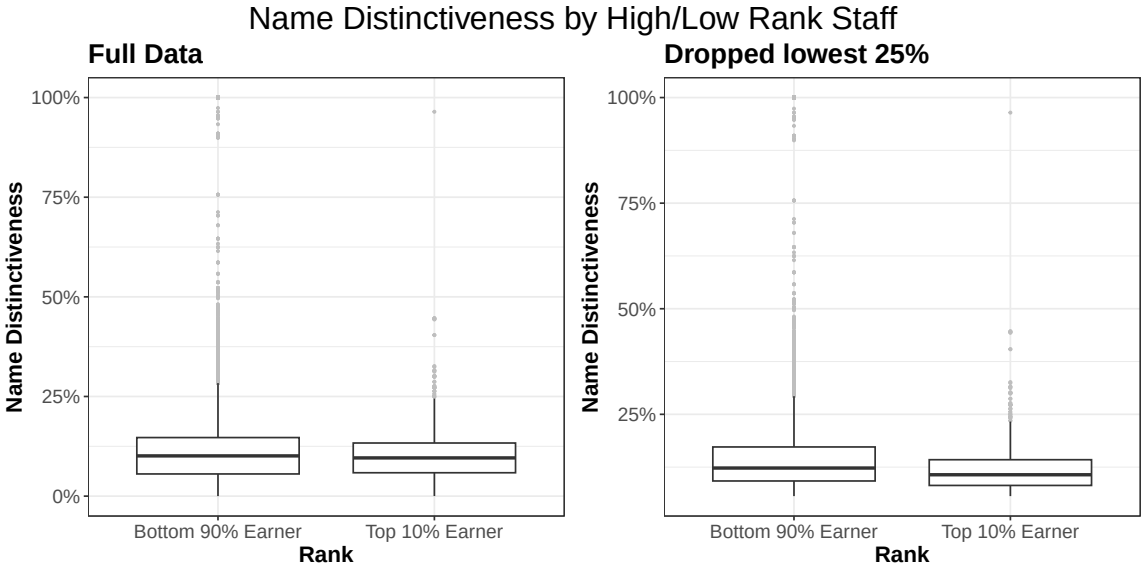


Figure S3.2: *Name Distinctiveness by Salary Rank*

S4 Validating the ethnicity coding approach (pg. 21)

The best "ground truth" dataset we could locate to validate our ethnicity coding approach is from Brierley and Nathan (2021).[58] Their 2018 survey of N=1,151 grassroots political activists in Ghana's ruling party helpfully contains both respondents' full names and their self-reported ethnicities at the sub-group level, allowing us to aggregate their ethnicities into the same 10 middle-tier ethnicity categories used in our analyses.[59] We break the name for each respondent into its fragments and apply the same coding procedure as used for our bureaucrats data. We then compare the ethnicity assigned by our coding to the respondents' self-reports.

Using this survey for our validation dataset is not without limitations. First, Brierley and Nathan (2021) is not a representative sample of the full Ghanaian population, but a purposive sample of the local party agents working in a randomly-selected set of parliamentary constituencies in Southern Ghana only.[60] This means that ethnic groups in the survey data do not appear in equal frequency as they do in the general population,[61] and there could be some possibility names used by local politicians from a particular group are not fully representative of the group overall. Yet, we note that the medians and ranges of the name distinctiveness score in the survey data and in our colonial bureaucrats dataset are virtually identical, suggesting the survey data captures a similar variety of names as in our main dataset.

Second, there could be survey administration errors that produced measurement error in the survey data, such that it would not be reasonable to expect perfect performance; in particular, the spellings of names recorded in Brierley and Nathan (2021) relied on enumerators[62] and self-reports of ethnicities were recorded by the enumerators in a long drop-down menu of options on a small touch screen, such that there easily could have been a few cases of "fat thumbs" in which enumerators ticked a wrong, adjacent option. Third, mixed-ethnicity respondents were allowed to assign themselves multiple identities; 54 (4.7%) respondents reported two choices and 1 reported three choices.[63] We consider our coding as successfully recovering the true ethnicity if it matches any of those that respondent listed.

With this data, we find an overall success rate of our coding scheme recovering respondents' self-reported ethnicity 62% of the time. But this masks enormous variation in performance

[58] This source is still much too small to use as a training dataset for a machine learning model, as in Chaturvedi and Chaturvedi (2024).

[59] Other surveys in Ghana that we are aware of either don't collect respondent names, only ask ethnicity at a more aggregated level (e.g., only recording Akan, rather than Asante vs. Fante), or ask ethnicity at inconsistent levels of aggregation (e.g., some respondents label themselves Akans, others label themselves Asantes vs. Fantes); any of these make the data unsuitable for validating our coding.

[60] The survey still has a meaningful sub-sample of Northerners; 75 respondents (6.5%) self-report as members of Northern ethnic groups, reflecting the fact that there are now Northern migrant populations throughout much of Southern Ghana (see Section S10).

[61] However, if we re-run our validation tests in this section applying survey weights that re-weight respondents back to the ethnic group proportions in the overall population, the statistics in Figure S4.1 are effectively identical.

[62] Interviews occurred orally, with enumerators transcribing responses, rather than respondents spelling their names directly.

[63] Notably, that there are still so few self-declared mixed ethnicity individuals in a modern sample collected after decades of sustained rural-urban migration and intermarriage suggests that the proportion of colonial-era bureaucrats with mixed ethnicity – born long before much of the social and spatial interethnic contact that could produce interethnic marriages had yet happened – was likely quite small.

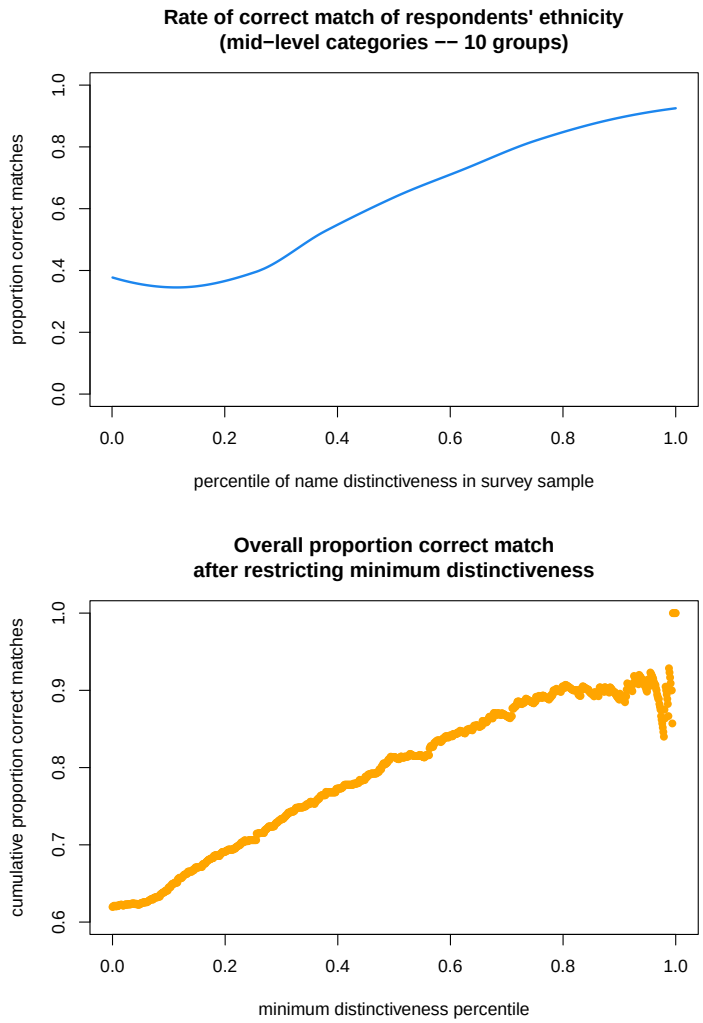


Figure S4.1: *Correct matches against Brierley and Nathan (2021) survey data*. Top panel: loess curve of proportion of correct ethnicity matches from our coding approach compared to survey respondents' self-reported ethnicity against the percentile score of each respondent's name's ethnic distinctiveness within the survey sample. Bottom panel: cumulative proportion of correct ethnicity matches when dropping all respondents below the name distinctiveness percentile on the x-axis. Both panels aggregate ethnicity to the mid-level categories (10 groups) used in our main analyses.

based on the name distinctiveness score – our measure of how much ethnic information is contained in each name.[64]

In the top quartile of name distinctiveness in the survey, we recover the correct ethnicity 90% of the time. The top panel of Figure S4.1 shows how our success rate rapidly increases in the percentile of each name's distinctiveness within the survey. The bottom panel of Figure S4.1 shows how our overall proportion of correct matches rises as we drop progressively larger subsets of the data below a minimum distinctiveness score.

In the bottom quartile of name distinctiveness in the survey – among the most uninformative names – our coding only recovers the correct ethnicity 34% of the time.[65]

Based on these validation statistics, we believe it is best to drop the lowest confidence observations from our bureaucrat data in our main analyses, as described in the main text. These are names that signal little ethnic information. We show robustness to different choices of cutoff under which to drop low distinctiveness observations below.

Validation using immediate post-independence ministers' names

Our "ground truth" validation relies on contemporary Ghanaian names. There could be concern that the method performs better with these names because they are matched against recent names contained in the voters' list. Meanwhile, our main analyses apply our modern dictionary to classify bureaucrats from the colonial era, raising concerns about potential changes in naming conventions over time. Unfortunately, we lack a large dataset of names from the colonial period with self-identified ethnicities to conduct a direct validation in the same time period.

As an alternative, we compiled the names of ministers who served in Kwame Nkrumah's two governments (1957–1960 and 1960–1966) in the immediate post-independence era. These represent a sample of individuals all born during colonial rule – and whose names thus should reflect naming conventions of that era – who were sufficiently prominent figures that there is enough detailed secondary source information available to reasonably deduce their ethnicity.

Drawing on biographical information from the Oxford Reference Dictionary of African Biography, local media sources, and Wikipedia entries, we coded each official's hometown and, based on that, assigned their ethnicity. We then applied the same dictionary-based classification procedure described above to predict each official's likely ethnicity and compared these predictions to our hometown-based coding.

Table SI1 reports the accuracy rates for the full sample and the subset above our preferred distinctiveness ethnicity threshold (25th percentile) as described in the main text. In the full sample, the accuracy rate is 62.5 percent, while in the restricted sample dropping low distinctiveness names it is 75 percent. These figures are virtually identical to those from the modern survey data above

[64] If we instead analyze performance at the upper-level, ethnolinguistic group tier of ethnic categorization, our overall accuracy is instead 83%. But an analysis of our bureaucrat data at this higher tier aggregation would miss substantively-important variation uncovered in the main text, such as between Fantes and Asantes (both Akans) or Gas and Dangmes (both Ga-Dangmes).

[65] This rate of 34% is still substantially better accuracy than a naive guess, showing that our approach still is able to leverage some information about each name even in cases of low distinctiveness. In simulations, randomly assigning ethnic groups to names with equal probability (a completely naive guess approach), only recovers the correct ethnicity of these respondents 8.7% of the time on average. Instead, randomly assigning ethnic groups to names in proportion to groups' share of the general population instead only recovers the correct ethnicity 14.1% of the time, on average.

Table SI1: Accuracy of the ethnic classification of ministers in Nkrumah's governments (1957–1960 and 1960–1966)

Accurate	Full list	Restricted list
Yes	62.5% (35)	75% (30)
No	37.5% (21)	25% (10)
Total	56	40

and lend further credibility to our ethnic classification approach, suggesting that measurement error from time-varying naming conventions is minimal.

S5 Comparison to hand-coding of ethnicity by local RAs (pg. 21)

A common alternative method of coding ethnicity from names is to hire local RAs with deep case knowledge and have them manually assign ethnicity. For comparison, we also hired three experienced RAs in Ghana and sent them 300 randomly-selected bureaucrats from the African Staff List. Each RA is a graduate student at the University of Ghana-Legon, the leading university in the country, with scholarly expertise in Ghana's colonial history and/or ethnic politics. Each personally comes from a different ethnic group and region of the country, ensuring a diversity of perspectives. We see this set of RAs as a simulation of the best-case scenario for hiring highly-knowledgeable coders familiar with ethnicity in a context like ours. These RAs were asked to give up to 3 guesses (ranked in order of confidence) of the ethnicity of each sampled bureaucrat's name.

The RAs struggled to agree on codings for most names. Overall, the three coders unanimously assigned the same mid-level ethnic category to only 27% of names. Two of three coders agreed in their first guess for another 41% of names, suggesting moderate confidence about the bureaucrat's ethnicity. In another 8% of cases, there was no agreement in first guesses, but the same option came up in one coder's first and two coders' second guesses, indicating a possible ethnicity with more limited confidence. In the remaining 24% of cases, there was no agreement at all in the coding, indicating that these knowledgeable RAs had no idea what the ethnicity might be based on the bureaucrat's name. Importantly, this means that if we had relied on manual coding of ethnicity by RAs instead, we would have still had to drop roughly a quarter of our data because of NAs on the ethnicity variable.

Moreover, the degree to which the RAs agreed with each other is correlated with the name distinctiveness score of each bureaucrat's name, as calculated using our voter register and census data approach. Among randomly-sampled names in the bottom quartile of name distinctiveness for the full bureaucrat dataset, the RAs only reached unanimous agreement about ethnicity in 25% of cases. In the top quartile – among the names our approach flags as most ethnically distinctive – the rate of unanimous agreement among RAs was instead 38%. But while our other validation test above shows that we code ethnicity with 90% accuracy in this same quartile, the RAs still couldn't agree at all on their coding for 20% of these most informative names. This implies significantly more measurement error from a hand-coding approach.

S6 Bureaucrats' ethnicity in the junior ranks and among high-earners (pg. 6

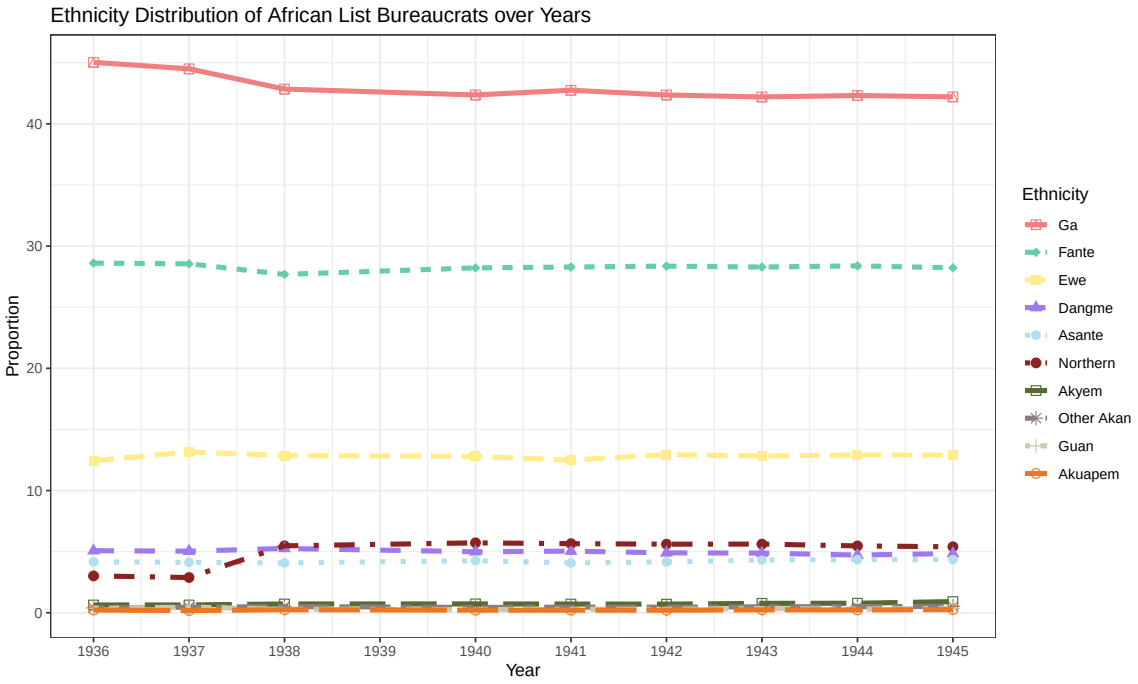


Figure S6.1: *Bureaucrat ethnicity in African Staff List*: drops all observations below the bottom 25th percentile on the name distinctiveness score.

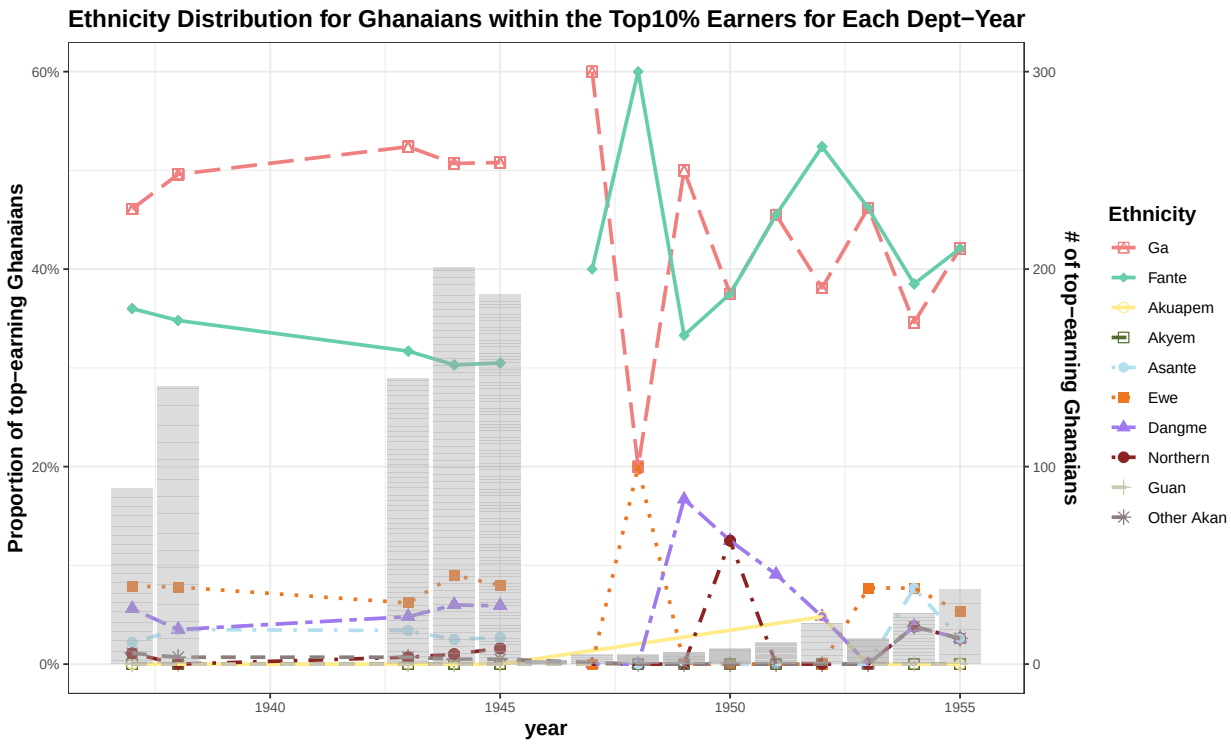


Figure S6.2: *Distribution of ethnicity among top-earning Ghanaians*: drops all observations below the bottom 25th percentile on the name distinctiveness score. Covers all years when either both the Senior and African staff lists are available, and/or in which the Senior Staff has become a Hybrid Staff List and includes a larger share of posts (from 1947 on).

S7 Robustness to ethnicity confidence thresholds (pg. 21)

To weed out bureaucrats with names that do not convey clear ethnicity information, our main specifications drop all bureaucrats below the bottom 25th percentile in our data in distinctiveness score. However, our results do not depend on this specific cutoff. For the most concise demonstration of this point, Figure S7.1 recalculates the overall proportion of African bureaucrats who are Ga or Fante – our paper's headline descriptive statistic – while iteratively dropping increasingly large portions of the data by name distinctiveness. The magnitude of Ga and Fante dominance in the civil service holds virtually identically even up through dropping the entire bottom half of our dataset. Clearly, this finding is not an artifact of measurement error in our ethnicity coding.

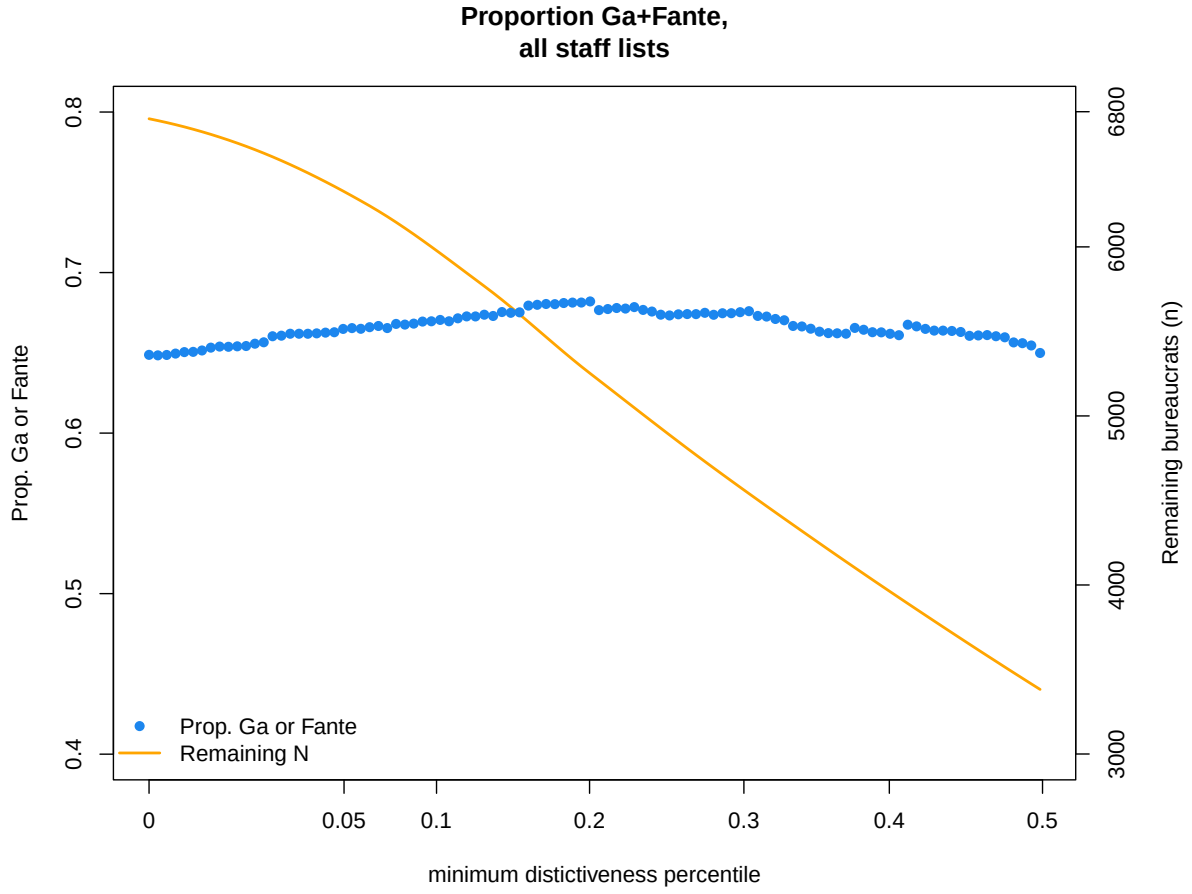
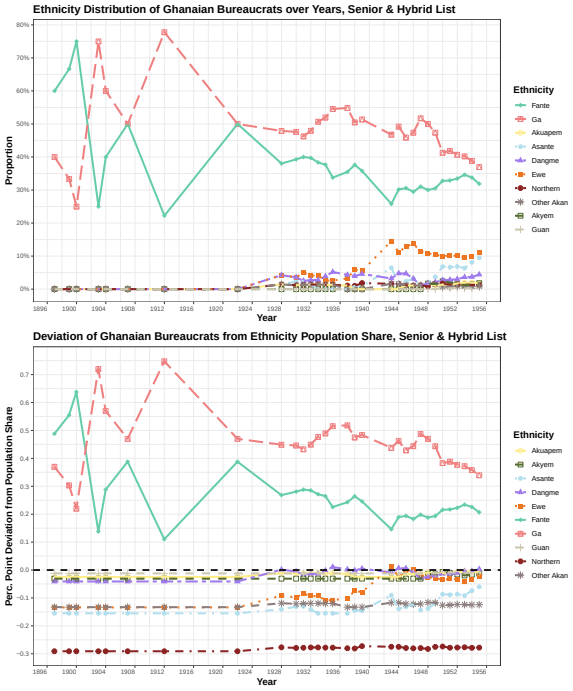


Figure S7.1: Overall proportion of all unique African bureaucrats (across combined Senior, Hybrid, and African staff lists) who are Ga or Fante (left Y axis) after dropping successively higher percentages of observations by minimum name distinctiveness percentile (x axis), reducing the remaining number of bureaucrat observations to the number on the right Y axis.

To provide more detail on robustness to our specific choice of minimum distinctiveness score, we also reproduce each figure and table in the main text at other equally-plausible cutoffs below (20th percentile) and above (30th percentile) our main choice, as well as repeat these same figures and tables without dropping any observations. Specific point estimates of course jump

Figure S7.2: Figure 5 with full data (all names)



around as additional observations enter or exit in each permutation, but the core patterns in the main text remain substantively robust across every alternative specification in each figure and table.

Figure S7.3: Figure 5 dropping bottom 20th percentile on distinctiveness

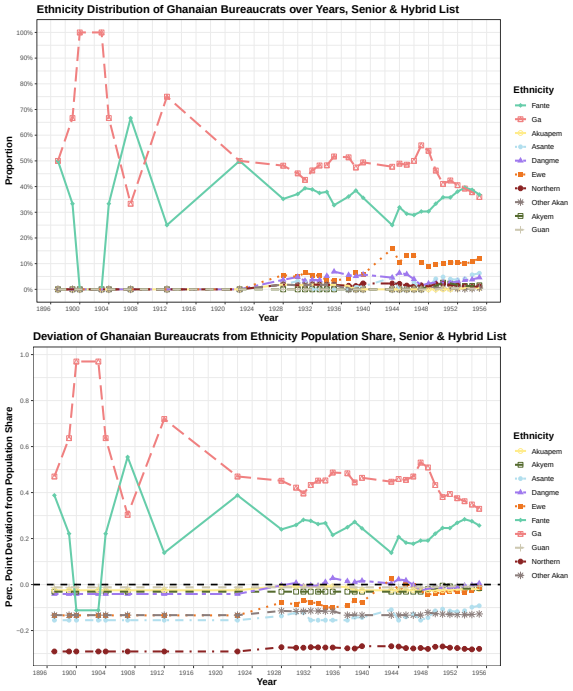


Figure S7.4: Figure 5 dropping bottom 30th percentile on distinctiveness

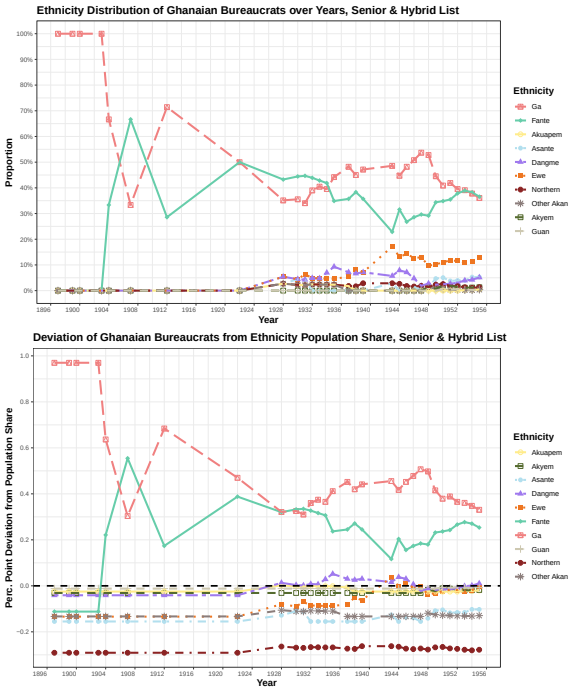


Figure S7.5: Figure S6.1 with full data (all names)

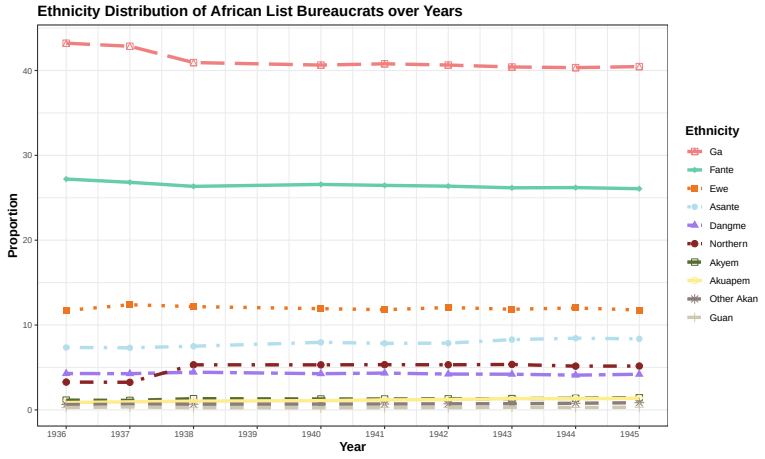


Figure S7.6: Figure S6.1 dropping bottom 20th percentile on distinctiveness

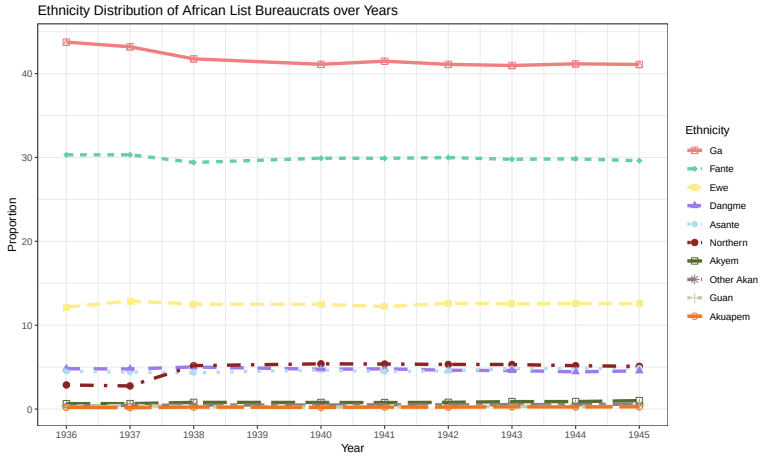


Figure S7.7: Figure S6.1 dropping bottom 30th percentile on distinctiveness

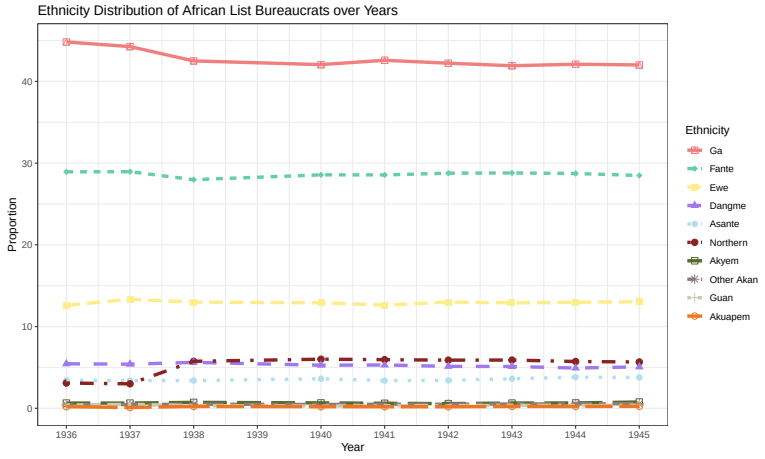


Figure S7.8: Figure S6.2 with full data (all names)

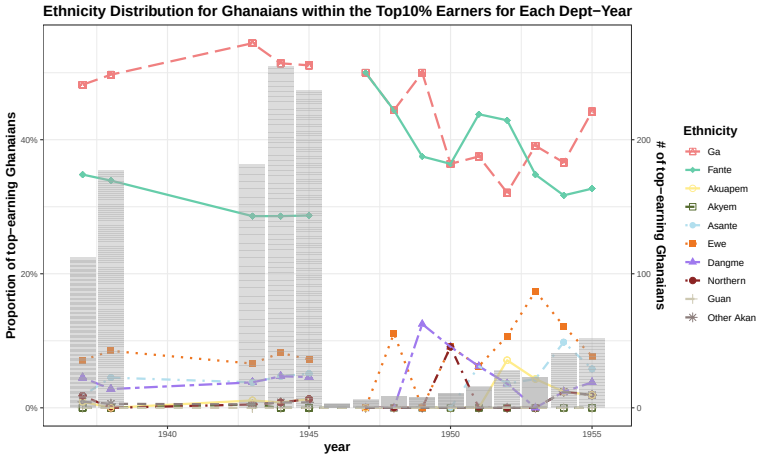


Figure S7.9: Figure S6.2 dropping bottom 20th percentile on distinctiveness

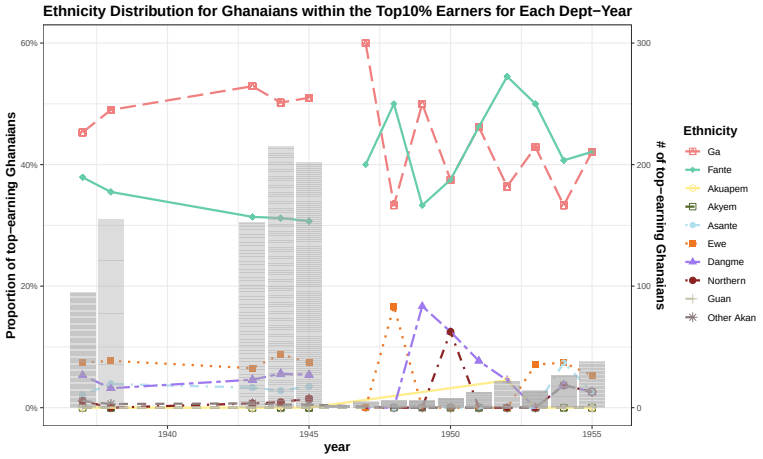


Figure S7.10: Figure S6.2 dropping bottom 30th percentile on distinctiveness

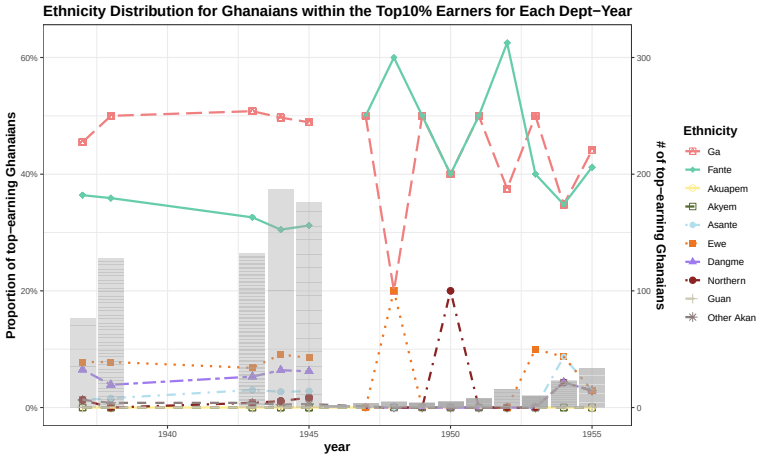


Figure S7.11: Figure 10 with full data (all names)

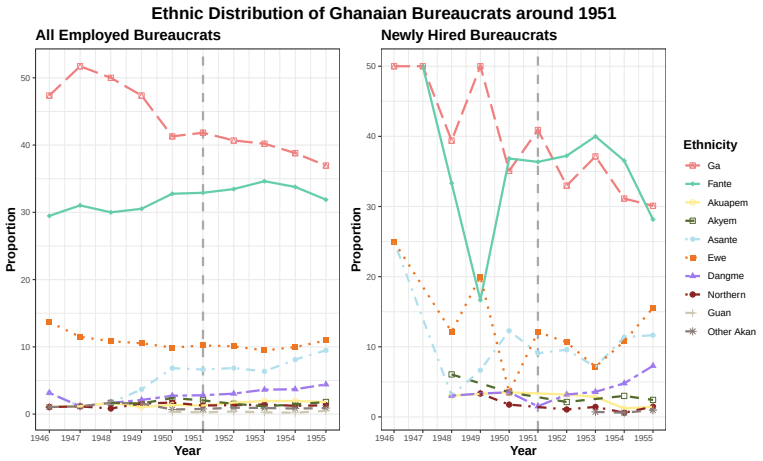


Figure S7.12: Figure 10 dropping bottom 20th percentile on distinctiveness

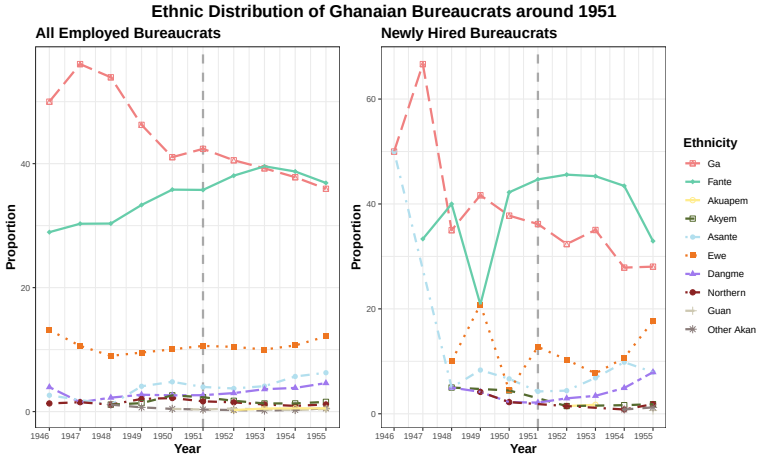
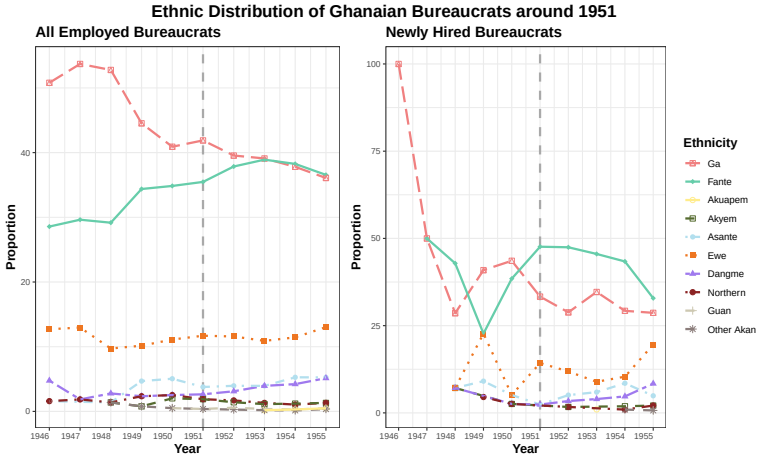


Figure S7.13: Figure 10 dropping bottom 30th percentile on distinctiveness



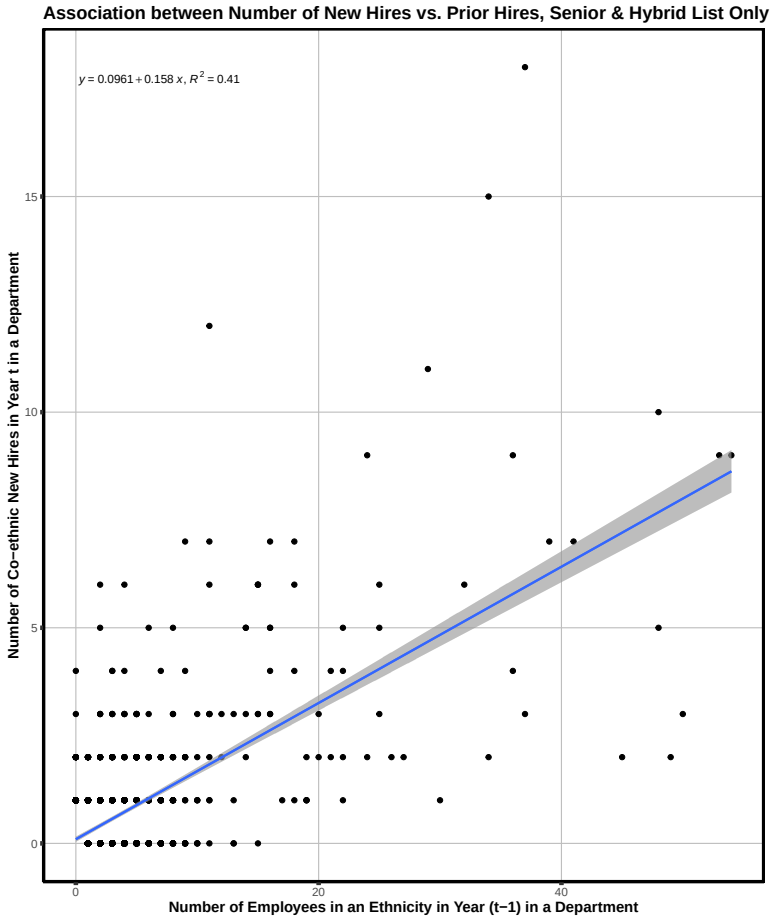


Figure S7.14: Figure 6 with full data (all names)

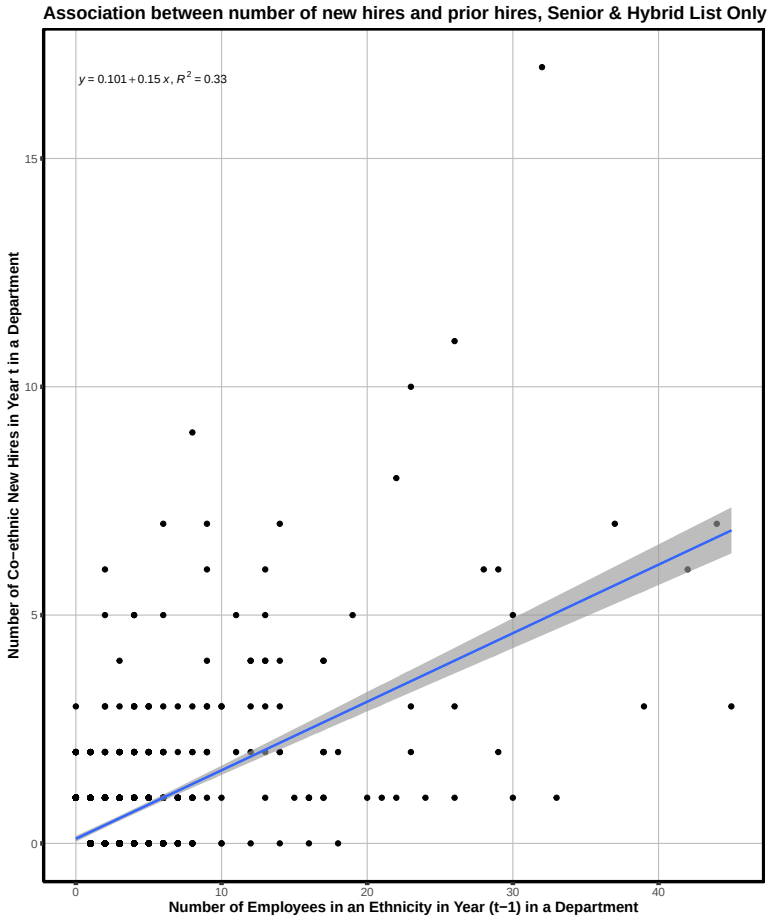


Figure S7.15: Figure 6 dropping bottom 20th percentile on distinctiveness

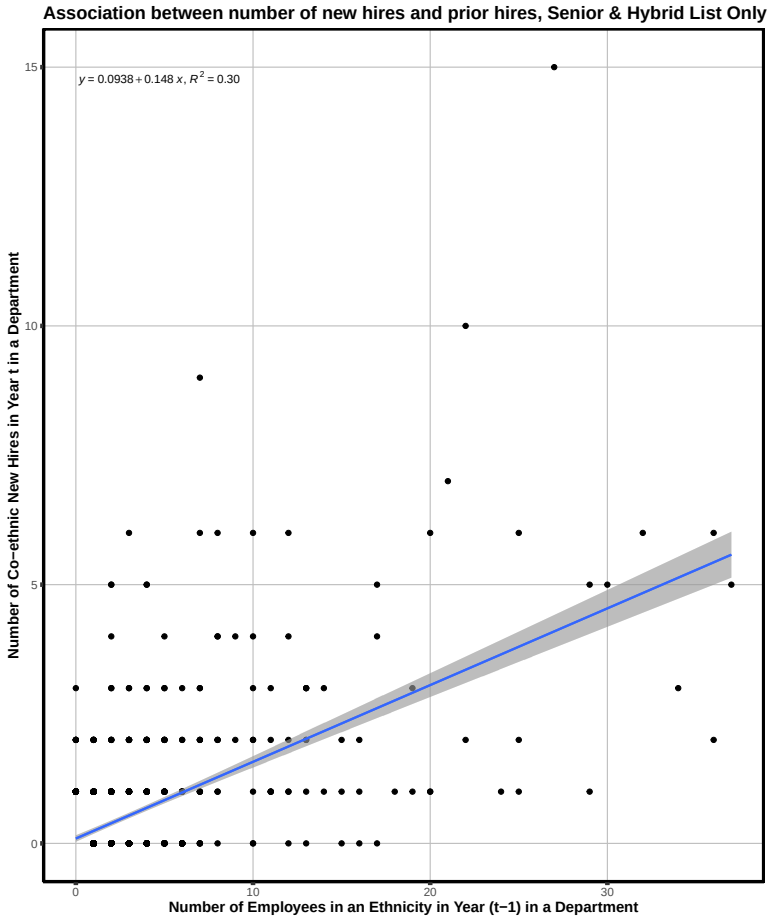


Figure S7.16: Figure 6 dropping bottom 30th percentile on distinctiveness

Figure S7.17: Figure 7 with full data (all names)

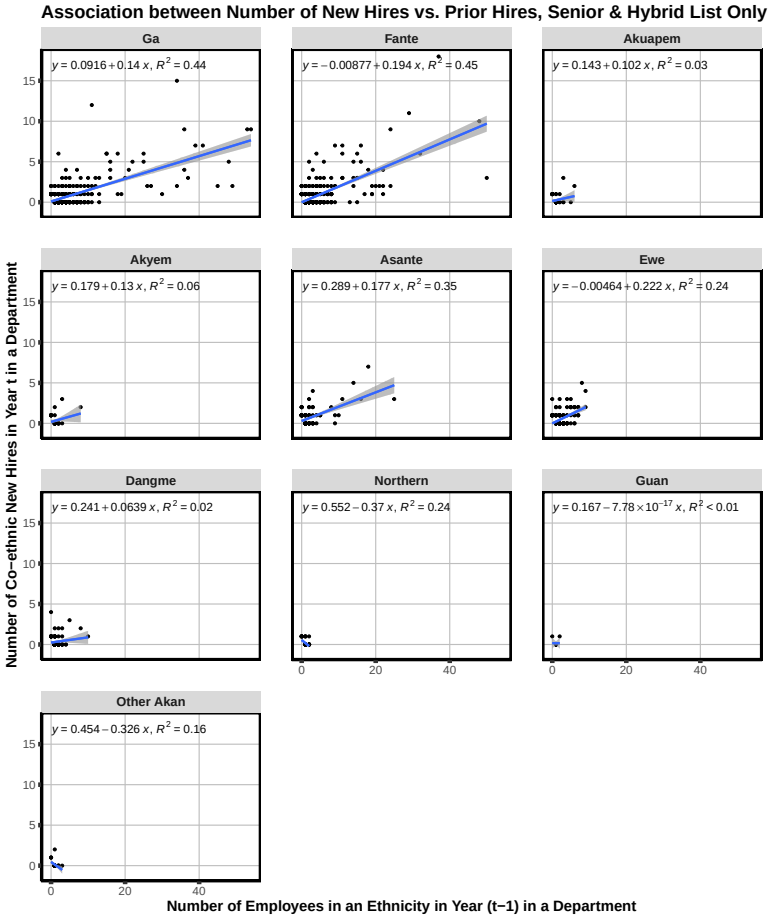


Figure S7.18: Figure 7 dropping bottom 20th percentile on distinctiveness

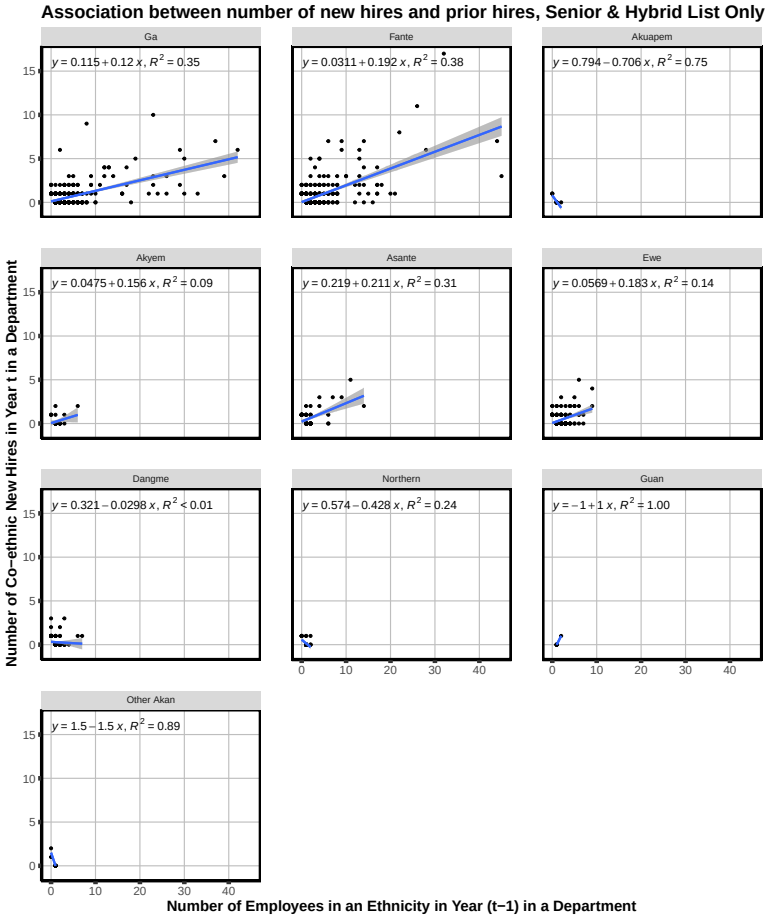


Figure S7.19: Figure 7 dropping bottom 30th percentile on distinctiveness

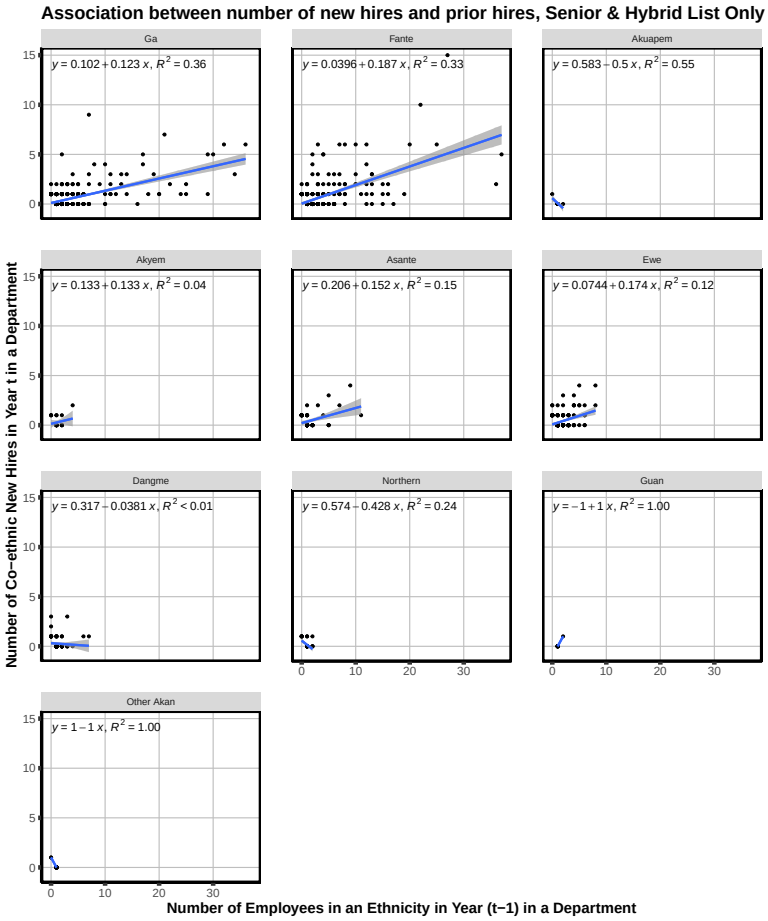


Figure S7.20: Figure 8 with full data (all names)

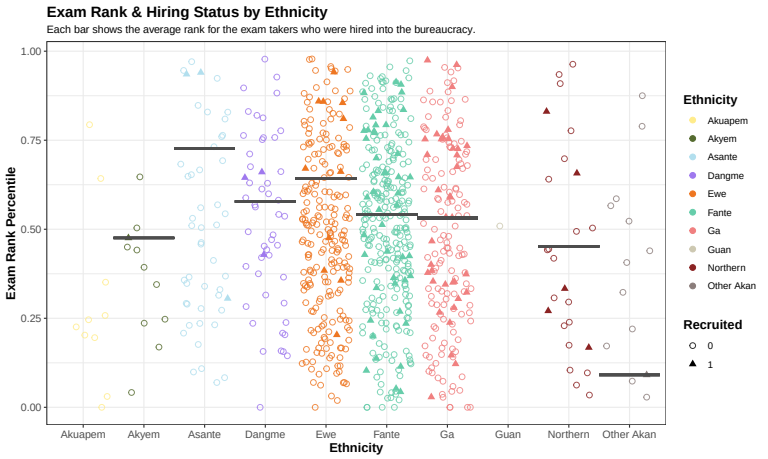


Figure S7.21: Figure 8 dropping bottom 20th percentile on distinctiveness

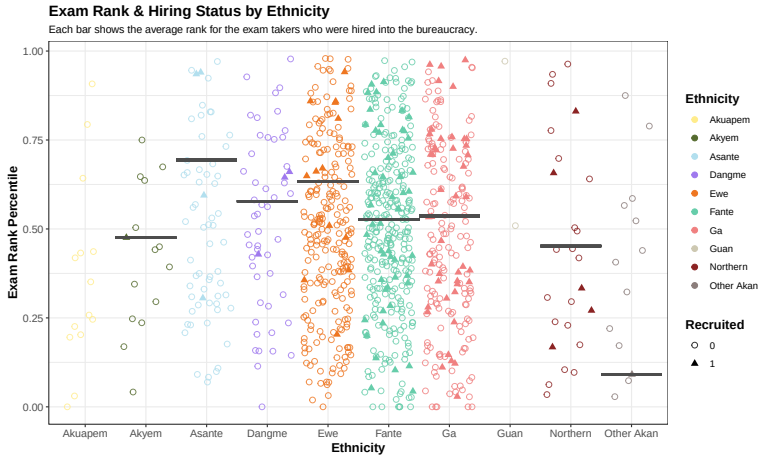


Figure S7.22: Figure 8 dropping bottom 30th percentile on distinctiveness

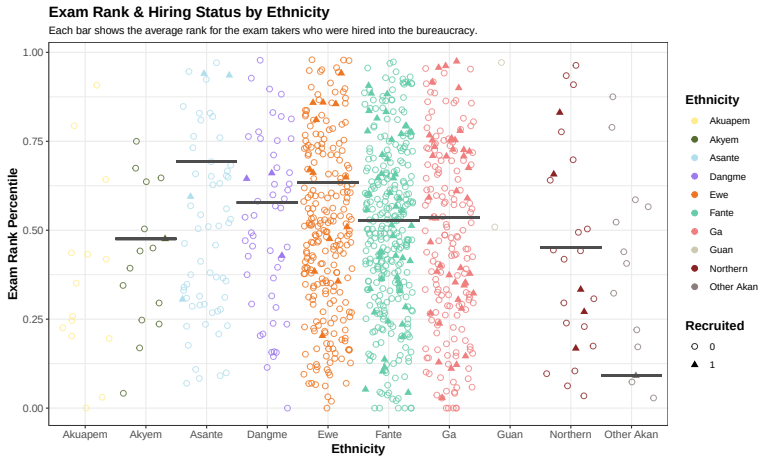


Figure S7.23: Figure 11 with full data (all names)

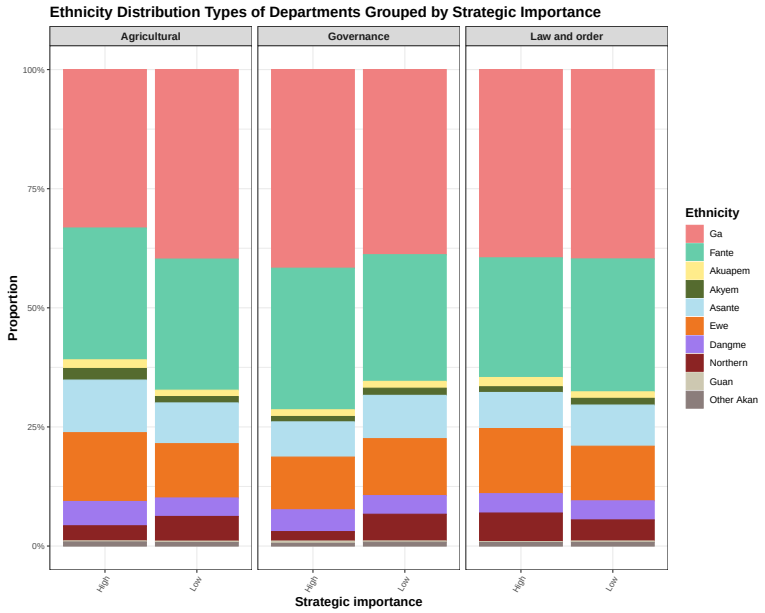


Figure S7.24: Figure 11 dropping bottom 20th percentile on distinctiveness

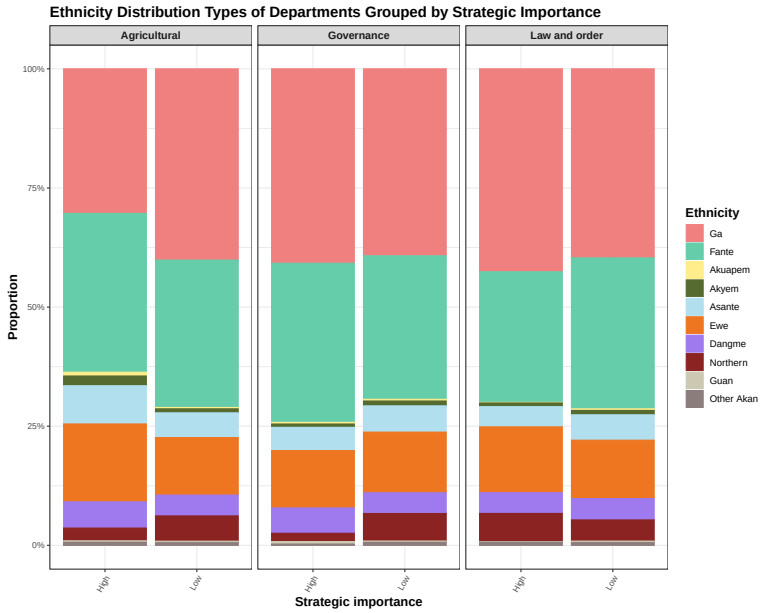


Figure S7.25: Figure 11 dropping bottom 30th percentile on distinctiveness

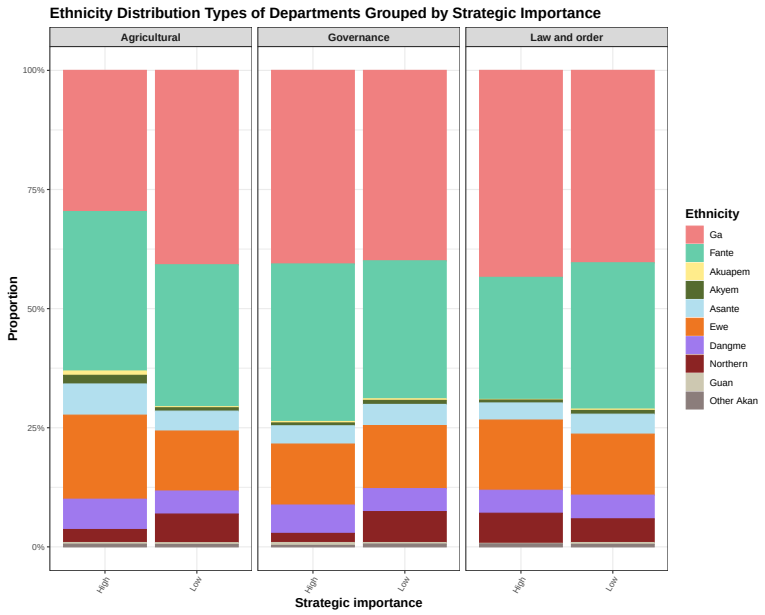


Table SI1: Table 1 with full data (all names)

Outcome: New hires from group in year t	1	2	3	4	5	6
Bureaucrats from group among senior staff (year $t-1$)	0.145**	0.146**	0.139**			0.256*
	(0.038)	(0.038)	(0.038)			(0.091)
Avg. distance (km) to school for group (year $t-25$)	−0.008		−0.093	−0.003		−0.012
	(0.009)		(0.090)	(0.004)		(0.012)
Avg. distance (km) to school for group (year $t-30$)		−0.007			−0.006	
		(0.006)			(0.004)	
Years included	1928-1955	1928-1955	1936-1955	1928-1955	1928-1955	1928-1955
Ethnic group FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*N*
Department FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
Year FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
N	8217	8217	6116	9284	9284	8217
adj. R^2	0.169	0.170	0.175	0.155	0.155	0.153

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. The unit is the ethnic group-department-year. OLS regressions with standard errors clustered by ethnic group, department, and year. Results with other SE clusterings are similar (and less conservative). Column 3 restricts only to years for which both Senior and African staff lists (1936-1945) and/or the expanded Senior Staff list (1946-) are available. Includes full bureaucrat data, regardless of name distinctiveness score.

Table SI2: Table 1 dropping bottom 20th percentile of distinctiveness

Outcome: New hires from group in year t	1	2	3	4	5	6
Bureaucrats from group among senior staff (year $t-1$)	0.095**	0.097**	0.084**			0.257*
	(0.024)	(0.024)	(0.025)			(0.096)
Avg. distance (km) to school for group (year $t-25$)	−0.009		−0.090	−0.004		−0.014
	(0.011)		(0.089)	(0.005)		(0.014)
Avg. distance (km) to school for group (year $t-30$)		−0.009			−0.007	
		(0.007)			(0.005)	
Years included	1928-1955	1928-1955	1936-1955	1928-1955	1928-1955	1928-1955
Ethnic group FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*N*
Department FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
Year FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
N	5929	5929	4301	7161	7161	5929
adj. R^2	0.201	0.202	0.207	0.148	0.148	0.181

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. The unit is the ethnic group-department-year. OLS regressions with standard errors clustered by ethnic group, department, and year. Results with other SE clusterings are similar (and less conservative). Column 3 restricts only to years for which both Senior and African staff lists (1936-1945) and/or the expanded Senior Staff list (1946-) are available. All bureaucrats in bottom 20th percentile of name distinctiveness are dropped.

Table SI3: Table 1 dropping bottom 30th percentile of distinctiveness

Outcome: New hires from group in year t	1	2	3	4	5	6
Bureaucrats from group among senior staff (year $t-1$)	0.055†	0.057†	0.030			0.219*
	(0.027)	(0.026)	(0.046)			(0.069)
Avg. distance (km) to school for group (year $t-25$)	−0.007		−0.072	−0.003		−0.012
	(0.009)		(0.076)	(0.004)		(0.012)
Avg. distance (km) to school for group (year $t-30$)		−0.007			−0.006	
		(0.005)			(0.004)	
Years included	1928-1955	1928-1955	1936-1955	1928-1955	1928-1955	1928-1955
Ethnic group FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*N*
Department FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
Year FEs	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
N	5885	5885	4257	7095	7095	5885
adj. R^2	0.193	0.194	0.198	0.141	0.142	0.172

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. The unit is the ethnic group-department-year. OLS regressions with standard errors clustered by ethnic group, department, and year. Results with other SE clusterings are similar (and less conservative). Column 3 restricts only to years for which both Senior and African staff lists (1936-1945) and/or the expanded Senior Staff list (1946-) are available. All bureaucrats in bottom 30th percentile of name distinctiveness are dropped.

S8 Location of training colleges and test centers (pg. 31)

From 1906, annual exam sittings for teachers occurred simultaneously in at least six different locations spread around the colony, including at testing centers in Akyem (Abetifi), Akuapem (Akropong), Ewe (Keta), and Nzema (Other Akan; Axim) territories. By the early 1920s, teacher exams were also being held in Kumasi, the Asante capital. By that point, many sat for these exams only after completing studies at one of the colony's three teacher training colleges, two of which were located in Akuapem (Akropong) and Asante (Kumasi) communities. Sitting for the exam in agriculture required first completing an agricultural training course; these courses were also held annually outside Ga and Fante home areas, such as at demonstration farms in Aburi (Akuapem), Kibi (Akyem), and Kumasi (Asante). It is clearly not the case that Ga and Fante job-seekers were advantaged in the exams simply because they had easier access to test centers than other Southern Ghanaian ethnic groups.

S9 Lock-in pattern subset to Euro-African families (pg. 34)

To further investigate the lock-in effect, we explored the employment patterns among Euro-African descendants. We code each Ga and Fante bureaucrat for whether they have a European (British, Dutch, or Danish) surname.[66] We compare the frequency of these surnames in the general Ga and Fante population using our contemporary universe of voter registrations against their frequency among the population of Ga and Fante bureaucrats in the colonial period.

Subsetting to all name fragments in the voter register data coded as most likely being either Ga or Fante, and then weighting name fragments by their empirical frequency of use in the general population, the European surnames that appear among our bureaucrats represent 14% of the Ga population and 11% of the Fante population. However, these names correspond to 31% of Ga and 37% of Fante bureaucrats in our data, a substantial degree of overrepresentation.

Figure S9.1 plots the hiring dynamics of senior bureaucrats among Euro-Africans within the Ga and Fante ethnic groups. It shows a clear positive relationship between the number of existing senior Euro-African bureaucrats in a department and the number of new hires in the subsequent year. This pattern again demonstrates the lock-in effect in Ghanaian bureaucracy, as the early presence of a certain population subgroup influences subsequent hiring and thus the future demographic composition of the bureaucracy.

[66] Common Euro-Ga surnames include Bruce, Bannerman, Brown, Barnor, Crabbe, Schandorf, Vanderpuye (or Vanderpuije), and Hesse (or von Hesse). Common Euro-Fante surnames include Thompson, Arthur, Fynn, Hagan, Brew, and Hayford.

Figure S9.1: Lock-in Effect in Euro-African Bureaucrat Hiring Dynamics

S10 Calculating average distance to colonial schools (pg. 36)

We calculate a proxy for the average distance a member of each ethnic group would have had to travel to a colonial school at different points in time by aggregating several data sources. First, we digitize lists of all government-supported/sanctioned schools, by town/village name, provided in the annual education department reports of the Gold Coast available in the British National Archives. Full lists are only available in some years' reports. After 1932, they are also no longer broken down to individual schools (only aggregating counts by province thereafter), so we focus on 7 snapshots of the set of official schools as of 1894, 1906, 1909, 1914, 1921, 1927, and 1932. By matching place names, we manually geolocate these schools to contemporary census enumeration areas, as visualized in Figure 2.[67]

Second, we then produce a granular measure of the geographic distribution of each ethnic group's likely population in the early 20th century. This is complicated by the fact that no detailed census data with ethnicity information exists from before the contemporary period. We instead leverage access to Enumeration Area (census tract) level data from the 2010 census, which records ethnicity at the sub-group level, combined with the historical fact that the vast majority of Ghanaian ethnic groups' (rural) population centers have remained very geographically consistent over time, largely due to the incentives created by persistent neocustomary property rights institutions (e.g., Boone 2014) in which most Ghanaians continue to only be able to securely own/inherit agricultural land in ancestral home communities.

In Figure S10.1 we aggregate the ethnic sub-groups up into the mid-level ethnicity categories used in our main analyses above and shade each census EA in 2010 by its majority population. Importantly, however, we first restrict only to EAs in which the local majority group is known to have been historically from that same administrative region (province) as of the pre-colonial period. This corrects for subsequent migration in the 20th century, in which some groups' populations have shifted to other regions of the country in which they did not originally live.[68] The remaining EAs provide a good proxy for the rural home communities of each group and the distribution in Figure S10.1 very closely matches with expert case knowledge of where these groups historically settled.

By adding an additional assumption that the relative ranking of rural community size as of 2010 is roughly correlated with the ranking of community sizes in the past (e.g., the same

[67] These data do not include "unassisted" schools, sometimes called "bush schools": private primary schools created by Ghanaian graduates in an informal capacity, outside the control of the state or missionary societies (Graham, 1971; Nathan 2023; p. 286-287). Location data for these schools – which existed only in Southern Ghana – is not included in any government reports. But our understanding is that employment in the civil service typically required graduation from a formal, government-recognized school, such that this omission should not speak to the supply of qualified candidates for the civil service.

[68] There are two forms of migration for which it is particularly important to correct. First, there has been widespread rural-urban migration over time, with Ghana's major cities all now having some EAs with majority populations from ethnic groups originally from other parts of the country. But most urban growth and rural-urban migration long post-dates our schools data. Second, in the mid-to-late 20th century, many members of the Northern ethnic groups engaged in widespread rural-rural migration to the South to become the agricultural (i.e., share-cropping) labor force in the South's more profitable rural cash crop economy (e.g., especially in cocoa), such that there are now scattered EAs throughout rural Southern Ghana with majority Northern populations that would clearly not have existed in those locations c. 1900.

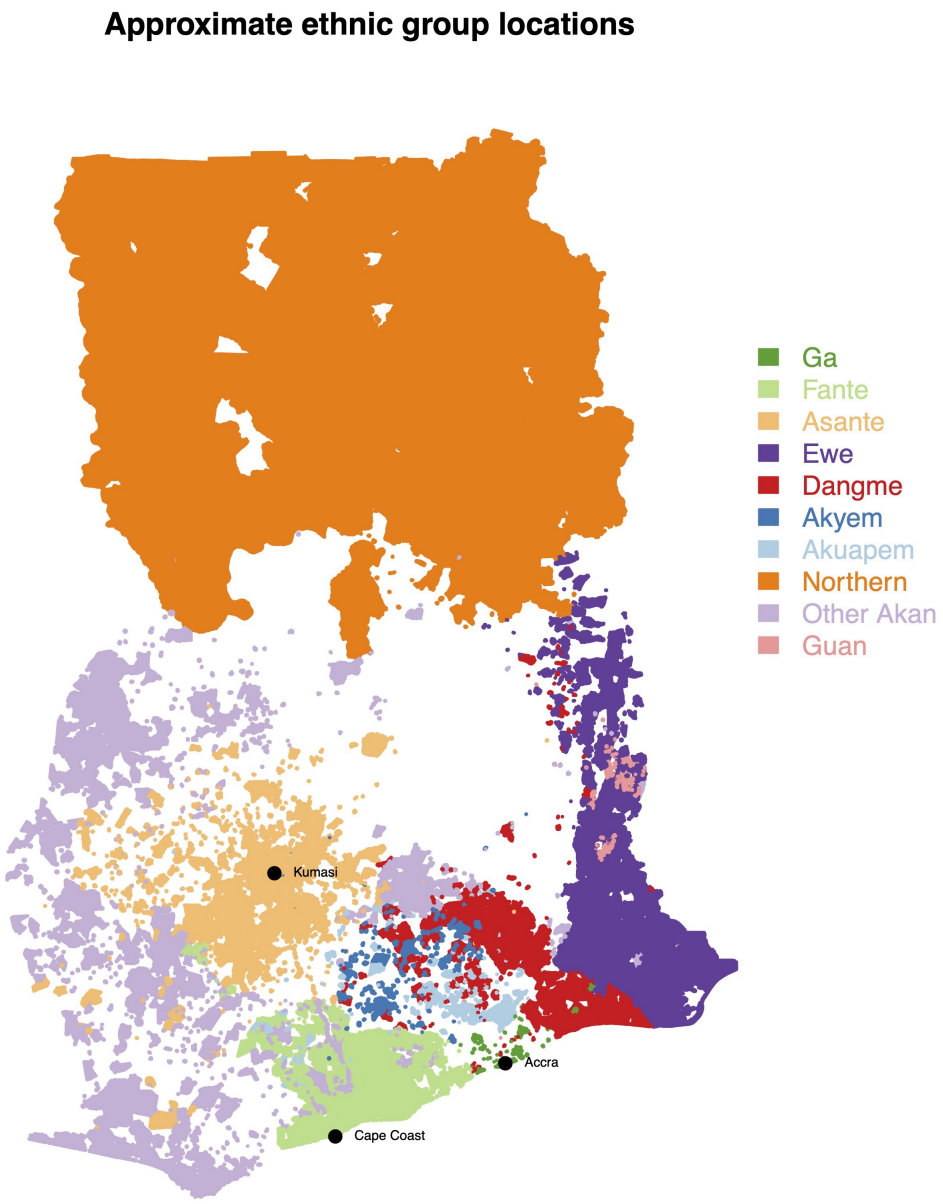


Figure S10.1: *Approximate locations of ethnic group populations:* 2010 census enumeration areas (EAs) shaded with majority populations from each ethnic group, after restricting to EAs within the historical home regions of each group to adjust for post-colonial migration. The blank areas in this map have: (a) no ethnic majority population; (b) an ethnic majority from a group historically from a different region; or (c) are Lake Volta or other uninhabited areas.

towns with the relatively largest populations within a rural district now were similarly the largest population towns in that same area historically),[69] we then use the 2010 count of members of each group in each EA to calculate a population-weighted average distance (as the crow flies) to the nearest school as of each year in our schools data. This proxies for the average distance the typical group member living in their rural home region would have had to travel to reach a school as of that year.

S11 Classifying field- vs. desk-based job titles (pg. 39)

Unfortunately we do not have the precise posting location for each job, so we classify the field- vs. desk-based jobs based on the type of work jobs generally entail. We classify jobs as field-based when the jobs' location is more likely to have been outside the capital city than inside. This is not to say that there were not teachers, nurses or postal workers in Accra. But there is a greater chance of an individual teacher working outside than inside the capital. Table SI1 lists the jobs we code as field-based.

[69] We believe this is a reasonable assumption given that most large rural towns in Ghana have grown up around the seats of higher-ranking (e.g., divisional or paramount) chieftaincies, which date back into the early colonial period or before. District boundaries and the administrative seats of local district governments (which typically comprise the largest rural towns in each area) today are also extremely path dependent to the administrative seats chosen in the colonial period, and were often chosen specifically because the locally-dominant chieftaincy was already sited there.

Table SI1: Job titles coded as field-based

Job Title
aidedecamp
agricultural officer
agricultural survey officer
agriculture officer
cadet
chaplain
collector
community development officer
cooperative officer
dental surgeon
deputy headteacher
field assistant
fisherman
forest guard
forest ranger
foreman
headteacher
health officer
instructor
leprosy survey control mo
letter carrier
lieutenant
major
marine branch lighthouse
matron
medical officer
medical specialist
meter reader
meter tender
midwife
mistress
nurse
pay sergeant
postman
postmaster
rector
regiona officer
sergeant
sheriff
soil survey officer
surgeon
teacher
tsetse control officer
vaccinator
veterinary
warder

S12 Teacher hiring vs. school locations (pg. 39)

We are unable to geolocate the specific physical state offices at which most of the bureaucrats in our data might have been employed, precluding an analysis for most state agencies of how employment patterns track with specific job locations.

Helpfully, one exception is for government teachers. Up until 1932, Education Department annual reports include the full list of every government (and mission) school by town/village, at all levels of education (primary and secondary), as well as their exact enrollment figures.[70] Starting from 1936, when the African Staff List first becomes available, we also have the complete set of Africans employed as government teachers. As noted in the main text, the majority of schools in colonial Ghana were mission schools, run independently from the state by mission societies, even as they received state funding. The remainder of schools were instead government run. Teachers at mission schools were private employees of the mission societies. But teachers at the government-run schools were civil servants and appear in the staff lists we digitize.

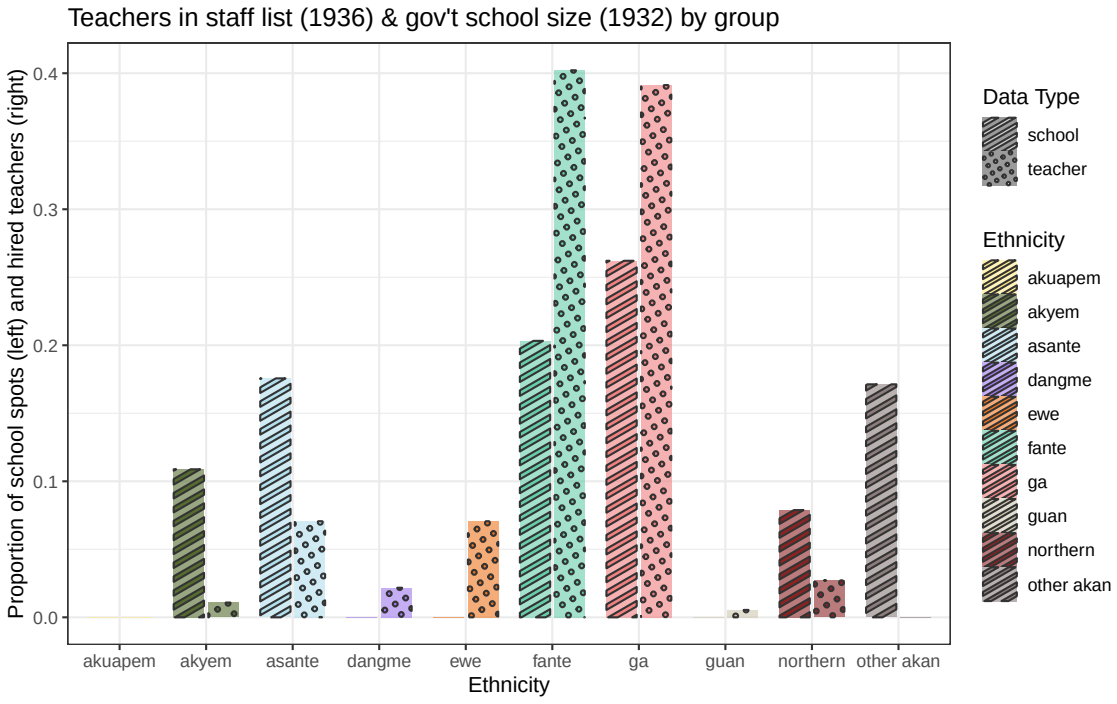


Figure S12.1: Government teachers by group vs. government school seats by group

Figure S12.1 uses these data sources to compare the geographic location of jobs in government schools in 1932 to the ethnic composition of government employees in 1936.[71] For each

[70] Unfortunately, this data series stopped with the 1933 annual report; after that point they no longer reported these statistics individually by school, only summarizing by province (mostly too broad a categorization for us to tie schools to specific ethnic groups' homelands).

[71] Note that our interpretation of Figure S12.1 operates under the assumption that there was no massive change in the geographic distribution of government schools in the four years in between these two dates. We are aware of no discussion in secondary sources suggesting anything like that happened.

ethnic group along the x-axis, it plots first (left side bar of each pair) the proportion of all students in government schools that were enrolled (as of 1932) in government schools located in towns/villages in which that group is the local majority. This measures government staffing need by group location. It then plots (right side bar of each pair) the proportion of all government teachers (in 1936) employed from that group.

If state hiring simply tracked the ethnic composition of the communities in which state jobs were located, the left and right side bars for each ethnic group in Figure S12.1 should be equal. But they are not. Both Fante and Ga teachers are employed by the state very far in excess of the government's teaching staff needs at schools located in Fante and Ga areas. Meanwhile, many other groups – such as the Akyem, Asante, Northerners, and Other Akans – have large numbers of students in government schools located in their territories, but very few or no teachers employed to teach at them. This pattern is completely inconsistent with state hiring simply being a function of the geographic distribution of jobs.

Separately, we recognize that the data point for the Ewe here might raise a question for readers about how Figure S12.1 is consistent with our separate analysis of the teacher exam data in the main text. The main text shows that successful Ewe candidates in the teacher certification exams were disproportionately passed over and underrepresented in subsequent hiring into the civil service. Meanwhile, Figure S12.1 shows that although there were no government schools located in Ewe communities (note: there were many mission schools located in Ewe communities) there were still some Ewe employed as government teachers.

The overall pattern of the underrepresentation of Ewe relative to their supply of qualified candidates emphasized in the main text cannot simply be explained by the lack of government schools in Ewe communities. First, the share of Ewe government teachers in Figure S12.1 (roughly 7%) is still dramatically lower than the proportion of successful teaching examinees in the 1920s and 1930s who were Ewe (29%). And because government teaching jobs were generally better paid and more desirable than mission positions, this huge underrepresentation is unlikely to be explained by Ewe teachers all opting to work in mission schools instead.

Second, the analysis in the main text focuses on subsequent hiring of successful examinees into *all* civil service positions, both as teachers and in other positions throughout the bureaucracy (both within and beyond the Education ministry), while the analysis here focuses more narrowly only on civil servant teachers. Importantly, the anti-Ewe bias we document in the main text among successful teaching examinees persists among *both* those subsequently employed as government teachers and those employed in all other state positions. If this anti-Ewe bias were only explained by the lack of government schools in Ewe territory, evidence for it should appear in the teacher data alone.

S13 Clerks vs. other jobs (pg. 40)

A formal Board of Selection (and Promotion) staffed by three senior British officials (at least nominally) handled all hiring of African civil servants with the job titles First or Second Division Clerks. The clerical service was handled separately because it was viewed as a subset of the broader civil service with skill sets that were not specific to any particular department, so these officials were regularly rotated and promoted across departments. Meanwhile, all other hiring of African civil servants was handled separately by each department, and these other hires typically instead remained with the same departments over the course of their careers (e.g., an African hired as a nurse remained in the Health Department, and was selected internally by other officials within the Health Department).

The result of this variation in hiring procedure is that the senior-most British officials in Accra had more latitude to directly decide clerk hiring decisions compared to all other job descriptions. If the bias in favor of Ga and Fante in the hiring process was primarily due to top down British orders to favor these groups, we should see it most clearly among this sub-population of African bureaucrats.

To be clear, archival folios on the internal operation of the Selection and Promotion Board do not suggest that any clear imperative to explicitly favor Ga and Fante hires was actually in effect. Instead, the board's memoranda make clear that the senior British officers sitting on this board likely had little independent information or coherent performance metrics about candidates to work with, and instead relied heavily on personal recommendations from subordinates in the various departments.

These folios contain many memos to department heads asking for recommendations about who to promote, and then perfunctory notices that the recommendation was acted upon. They also contain spreadsheets of performance evaluations ostensibly used for meritocratic selection of candidates for promotion (e.g., lists of all Second Division Clerks in a given department being considered for a promotion to First Division Clerk) in which the column that indicated the main performance metric amounted to little more than a summary of department leaders' subjective/qualitative recommendations: e.g., no more than "good" vs. "very good." In such a setting, any applicants who were well liked by their superiors lower down in the bureaucratic hierarchy – by senior and mid-level officials who could themselves be African and would be in a position put in a good word – would seem to have had a leg up in the process. For example, folios GH/PRAAD/CSO.2/4/258, GH/PRAAD/CSO.2/4/841, GH/PRAAD/CSO.2/7/129, GH/PRAAD/CSO.2/7/130, GH/PRAAD/CSO.2/7/134.

We nonetheless test to see if there is evidence that senior British leaders engaged in very different hiring patterns when they had explicit oversight over the process (among the clerks). Limiting to the years covered by the African Staff List, for which we have the entire roster of African civil servants available, we compare the representation of the Ga and Fante among clerks (left column) and all other positions (right column) in Figure S13.1. We show that: (a) Ga and Fante bureaucrats continue to be overwhelmingly overrepresented even among the bureaucrats *not* selected centrally by top British officials (69% combined in the right column, compared to 14% of the general population); (b) while there are more Ga hired among clerks than in other positions, there are instead more Fante hired among other positions (30.1%) than among the clerks (27.6%) – inconsistent with a top-down directive to favor *both* groups in a manner consistent with the patterns in the rest of our analyses; (c) the biggest difference between the columns seems to be due instead

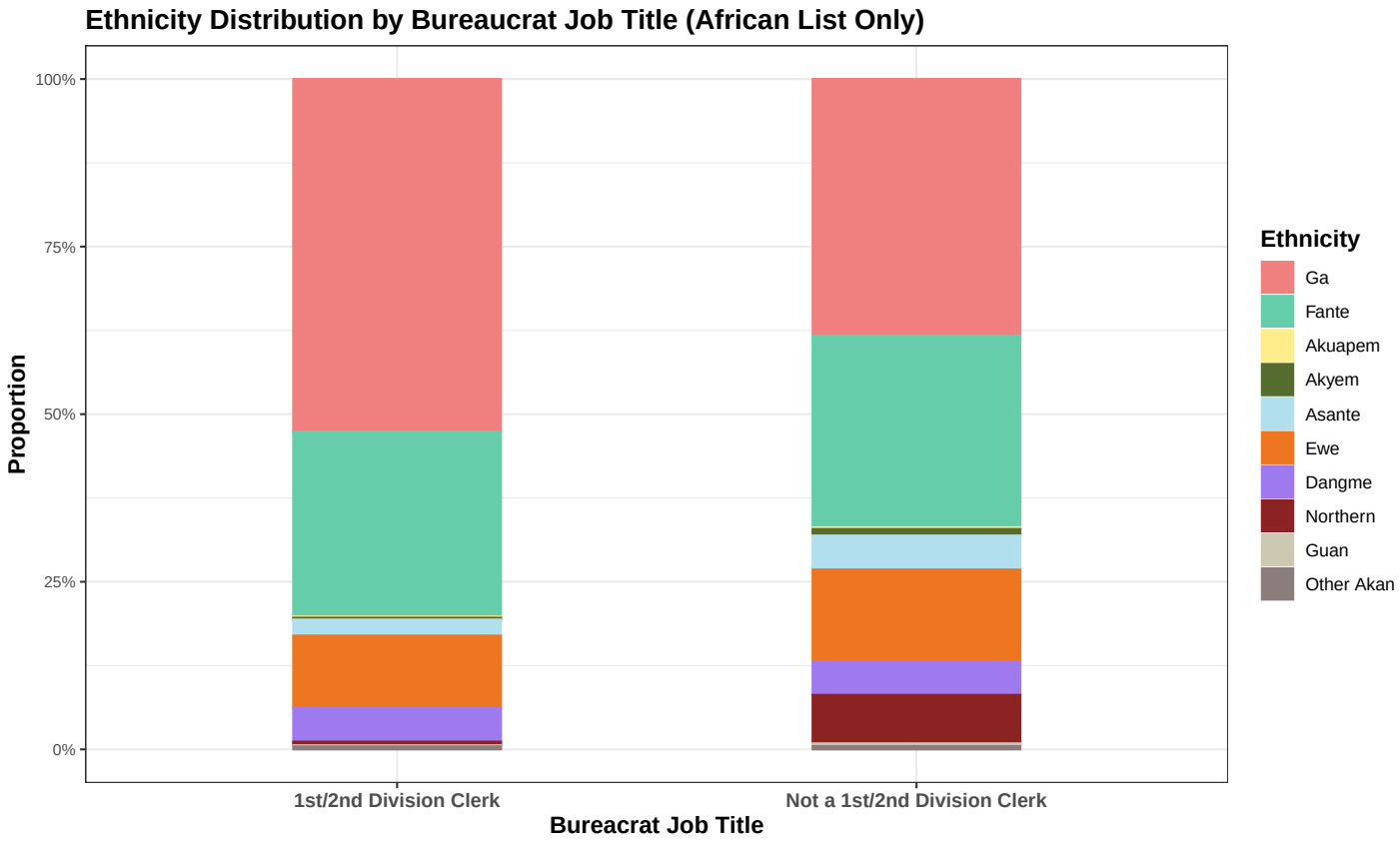
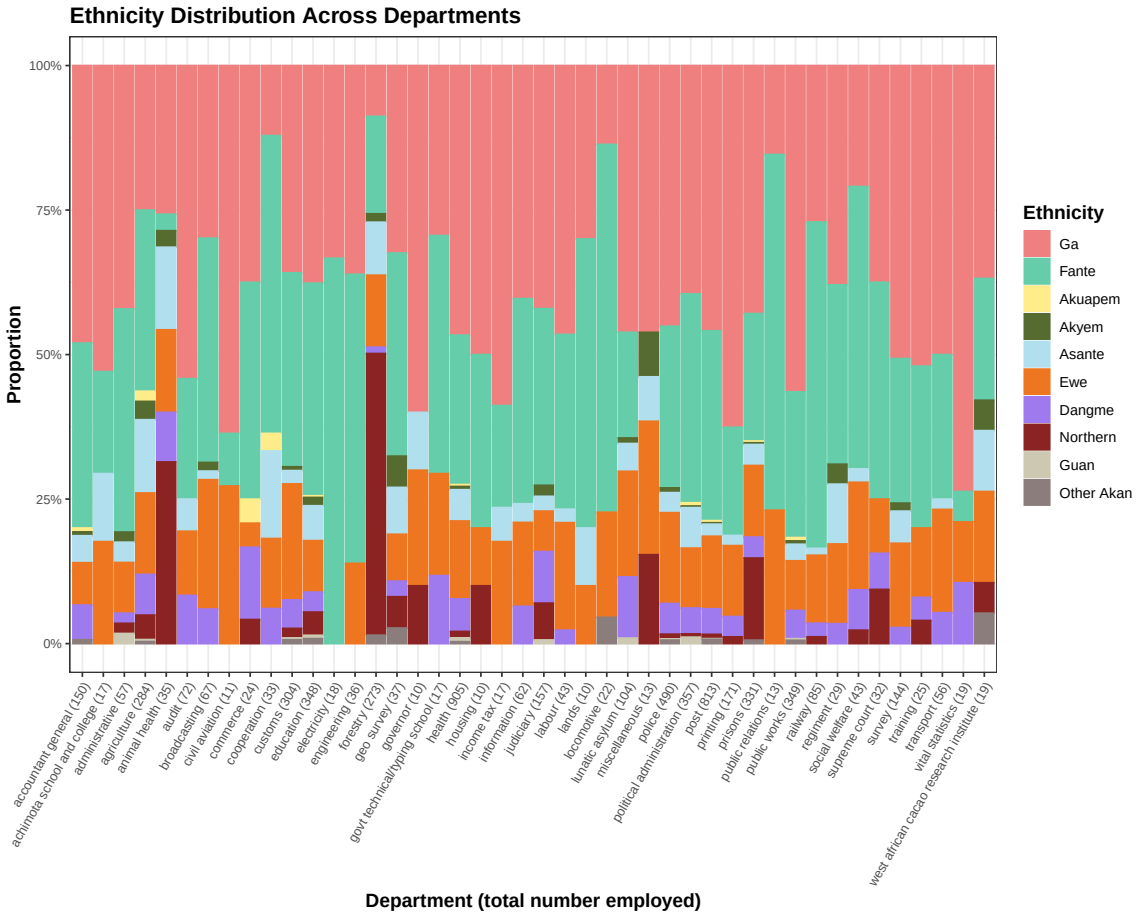


Figure S13.1: Clerks vs. all other job descriptions, African Staff List only

to the greater proportion of Northerners in non-clerk positions. This can be best explained by education, not explicit British preferences: as we describe in the main text, Northerners had by far the least access to education in the colonial period, and had very few educated applicants qualified to serve as clerks; the overwhelming majority of Northerners in the data instead served in very low rank (and low pay) positions, such as "Forest Guard" in the Forestry Department, all of which are grouped in the right column of Figure S13.1.

Figure S14.1: Ethnic Composition Across Departments



S14 Ethnic distribution by department (pg. 42)

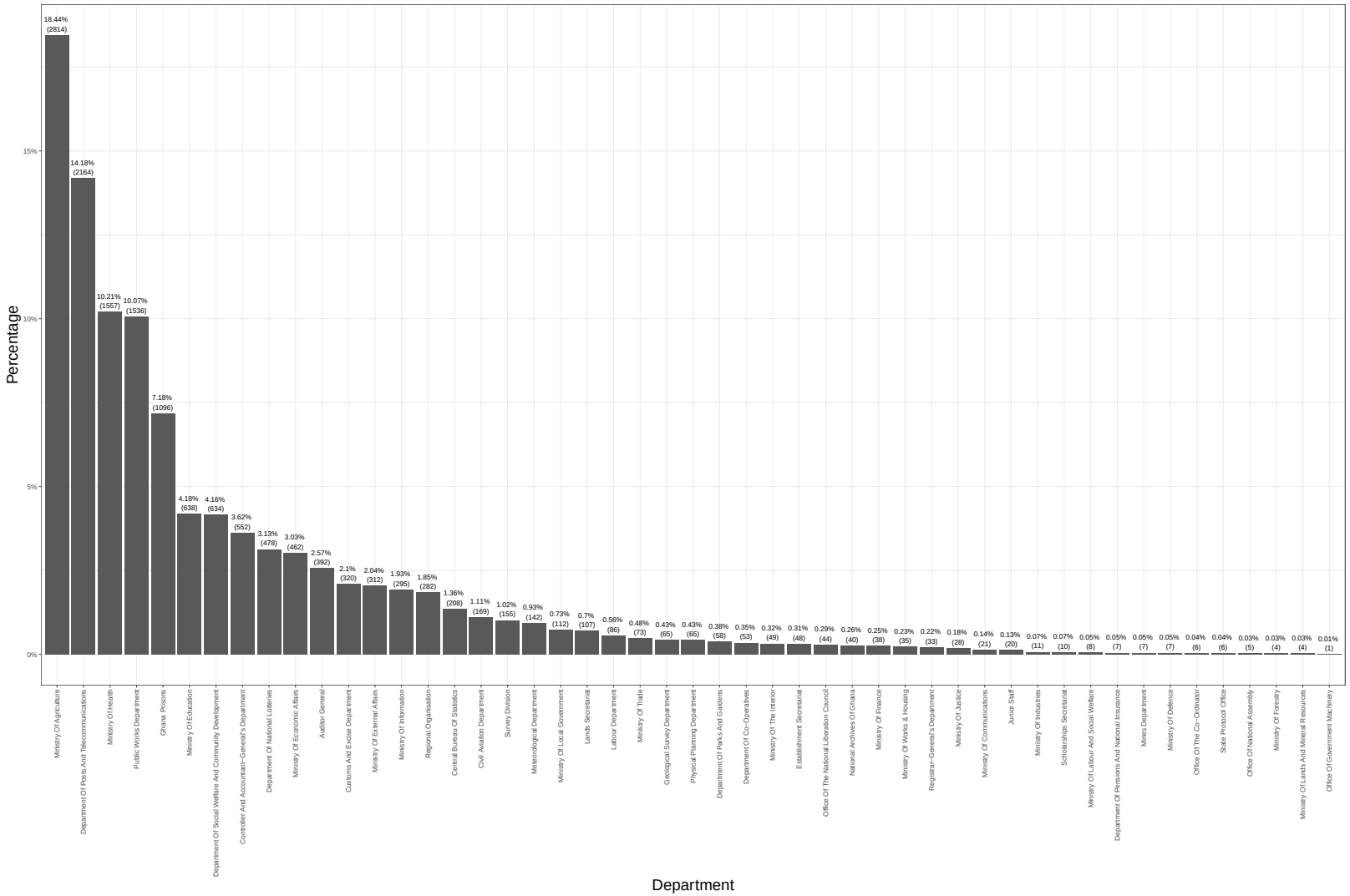
To better understand the pattern of ethnic distribution in the Ghanaian bureaucracy, we plot the overall ethnic distribution of all departments in the combined staff lists, as shown in Figure S14.1. Each vertical bar along the X-axis shows one public department, with the total number of Ghanaian employees by the department recorded in parentheses. Because the structure of Ghana's bureaucracy changed substantially over the temporal coverage of our data, we include only the departments that employed more than 10 Ghanaian bureaucrats to exclude those that are either too small or established for only a short period of time.

The figure shows a diverse distribution of ethnic groups across departments, with 44 out of 45 departments employing bureaucrats from more than two ethnic groups. Moreover, Figure S14.1 shows a consistent domination of Ga and Fante groups across the departments, with the combined Ga and Fante population accounting for over 50% of the bureaucrats employed in 42 of 45 departments. This pattern suggests that the dominance of Ga and Fante is not restricted to a subset of departments due to skill advantage or department preference, but rather a cross-cutting phenomenon throughout the colonial bureaucracy.

S15 Ethnic composition of post-independent Ghana's bureaucracy (as of July 1, 1968) (pg. 44)

We investigate whether the lock-in patterns we observed in Ghana's colonial period persisted ten years after independence. We obtained the junior staff list as of July 1, 1968, from the University of Essex library, digitized and coded by ethnicity using our approach outlined in Section 5.2. The dataset comprises 15,278 bureaucrats – consistent with a vast expansion in the overall size of the state since independence – across 49 ministries and departments, including the Ministry of Agriculture (2,814 employees), Ministry of Health (1,557), Department of Posts and Telecommunications (2,164), and Department of Public Works and Housing (1,536). Figure S15.1 gives the number and percentage of junior staff across all these departments.

Figure S15.1: Departmental Composition of Junior Staff in Post-independent Ghana (last at July 1, 1968)



S16 Ethnic composition of bureaucrats by department (1968) (pg. 44)

Figure S16.1 shows the ethnic composition of junior bureaucrats in 1968. Consistent with our findings of lock-in, the Fantes (26%) and Ga (24%) remained dominant. Ewes representation rose to 20%, followed by Asantes (14%). Other Akan subgroups (Akyem, Bono, Akuapem) remained significantly underrepresented.

We disaggregate our results by the thirteen top departments (that is, the proportion of staff employed in this department). The patterns of Fante and Ga dominance remain in all institutions except in the Ministry of Agriculture and the Ghana Prisons.

Figure S16.1: Ethnic Composition of Junior Staff in Post-independent Ghana (last at July 1, 1968)

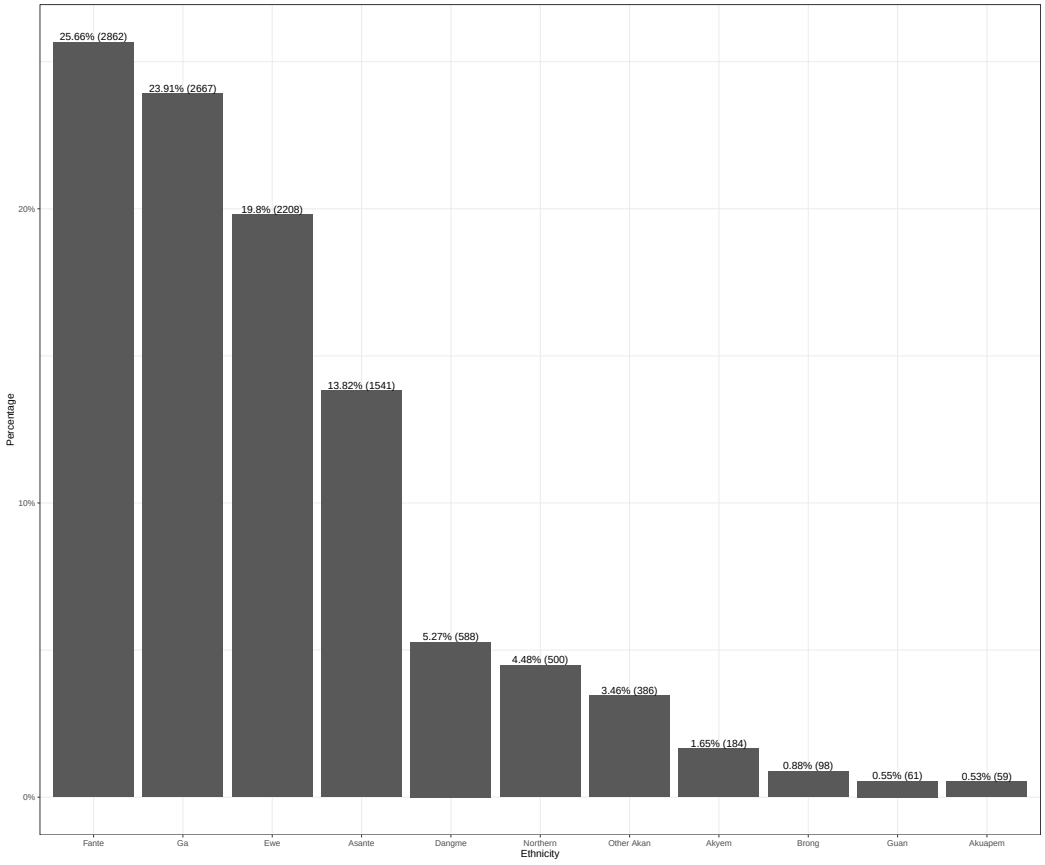
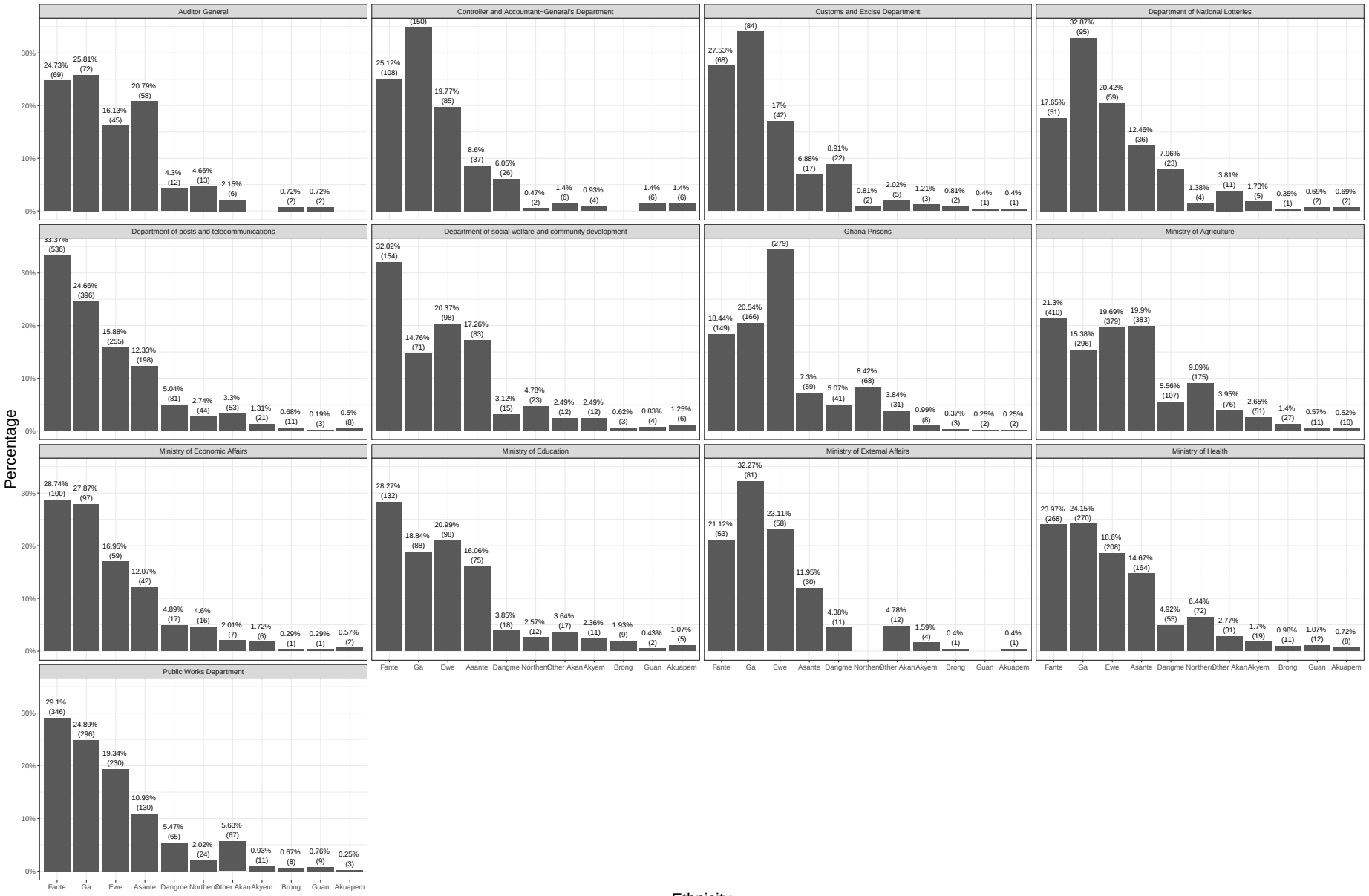


Figure S16.2: Ethnic Composition of Junior Staff in Post-independent Ghana (last at July 1, 1968) by Institution



S17 Data transparency

We will post full reproduction files and code to replicate all analyses, including the raw bureaucrats data used in our main analyses. However, for three of the auxiliary datasets also used in certain portions of our analysis, there are ethical or legal constraints against publicly posting on the open internet a full raw dataset in the uncleaned form it first appears at the very beginning of our code pipeline. We understand that this is what the APSR's reproduction package policy now requires.

For these particular datasets, we would instead have to upload reproduction data that begins instead only after several initial data cleaning steps have already occurred to preserve the anonymity and identities of living persons (in two situations) or to comply with a signed data use agreement with the Ghanaian government (in the final situation).

We will post the full data cleaning code used for these three datasets, even without the uncleaned versions of the completely raw data. We are happy to work with the Editors at the appropriate time on whatever the details of such a process should look like.

The details of these limitations are:

1. Our method of assigning ethnicity to names draws on the complete Ghana voters register from 2015, which is an identifiable list of real (mostly, still living) people located to specific places. Although we are under no specific legal restrictions in using this public record, we are unaware of it ever having been posted in its raw form on the open internet. We believe it would be very ethically inappropriate for us to do so when Ghana's Electoral Commission, the official creator of the data, has always elected not to post it publicly on the internet for personal privacy reasons. In the wrong hands, the raw data could be used to "dox" and locate real people. Our reproduction code would begin after the dataset has already been pre-processed and consolidated into name fragments, with polling station identifiers removed, to preserve privacy of living individuals.

2. We validate our ethnicity coding against survey data from Brierley and Nathan (2021). The raw survey data used for this validation contains names and other personal identifying information for each respondent. To protect their anonymity, we would again instead only post pre-processed data already consolidated into de-identified name fragments. The public reproduction code would continue from there and still can be used to replicate all analyses using the data. Note also that an anonymized version of this dataset is already publicly available as part of the replication package for Brierley and Nathan (2021).

3. The creation of the "distance to school" covariate (Figure S10.1) draws on a geo-coded (GIS) Enumeration Area map that corresponds to Ghana's 2010 census, obtained from the Ghana Statistical Service (the Government of Ghana) under terms of a signed data use agreement that precludes posting the full raw GIS map on the public internet. (The underlying census data itself is instead publicly available and can be posted in its raw form.) However, we have faced no problems in the past posting cleaned, pre-processed excerpts of specific portions of this map as part of the reproduction packages for other published articles and plan to do so again here so that the creation of this variable can still be reproduced as part of our replication package, even as the full EA map is not released to the public in contravention of our agreement with the GSS.