Historical Origins of Modern Ethnic Violence: Pre-Colonial States and Separatist Civil Wars

Jack Paine[*]

September 4, 2019

Abstract

This paper explains how precolonial statehood has triggered postcolonial ethnic violence. Groups organized as a pre-colonial state (PCS groups) often leveraged their historical privileges to control the postcolonial state while also excluding other ethnic groups from power, creating motives for rebellion. The size of the PCS group determined other groups' opportunities for either gaining a separate state or overthrowing the government at the center. Regression evidence based on a novel global dataset of historical statehood demonstrates a strong positive correlation between stateless groups in countries with a PCS group and separatist civil war onset. Although the typical PCS group is large enough to deter center-seeking rebellions, in countries where the PCS group is small, stateless groups in their countries fight center-seeking rebellions at high rates. By contrast, particularly large PCS groups disable any rebellion prospects. These findings also explain cross-regional patterns in ethnic civil war onset and aims.

1 INTRODUCTION

Large-scale ethnic conflict is strikingly and tragically common in the postcolonial world. Numerous states outside of Western Europe fit the categorization of weakly institutionalized polities in which armed rebellion provides a viable avenue for aggrieved groups to achieve political goals. Many scholars analyze prospects for powersharing coalitions in these countries and consistently find that ethnic groups that lack access to power in the central government more frequently fight civil wars (Cederman, Gleditsch and Buhaug 2013; Roessler 2016). But experiences prior to independence undoubtedly affected which groups commanded power, their ability to commit to powersharing deals, and other opportunities for excluded groups to rebel. Which historical factors matter for explaining contemporary ethnic violence?

Prior to the period of European colonialism, the political and territorial organization of ethnic groups across the world differed considerably. Nomadic and other stateless or segmentary groups occupied some areas, such as the Maasai in Kenya and various groups in the Zomia region of southeast Asia (Scott 2010). By contrast, states with standing armies and centralized tax collection formed elsewhere. In some cases, pre-colonial states mapped quite closely to a modern country. For example, various Chinese dynasties over several millennia engendered a distinct Han Chinese ethnic group that still governs China today. In other cases, such as the Buganda kingdom in modern-day Uganda, European colonizers combined members of the historical kingdom into a colony (and, later, sovereign country) containing numerous other ethnic groups. This paper argues that these two related historical factors—pre-colonial statehood, and the size of ethnic groups organized as a state—provide considerable insight into modern ethnic violence. Groups organized as a pre-colonial state (PCS groups) often leveraged their historical privileges to control the postcolonial state while also excluding other ethnic groups from power, creating motives for rebellion. The size of the PCS group determined other groups' opportunities for either gaining a separate state (separatist civil war) or overthrowing the government at the center (center-seeking civil war).

I first develop a general theoretical framework that explains civil war onset and aims. Existing theories of domestic conflict explain why governments' inability to commit to proposed powersharing deals causes fighting, which I build upon by additionally analyzing how the size of the ruling group affects incentives for different types of civil war. A small ruling group is vulnerable to either center-seeking or separatist rebellions, but a large ruling group can deter either type of rebellion. The distinction between civil war

aims has bite if the ruling group is medium-sized. Such groups are large enough to deter attacks against the center—for which numerical size more directly affects the probability of winning—but cannot deter guerrilla conflicts in the periphery that aim to secede. But regardless of the ruling group's size, only if the government exhibits low ability to commit to powersharing deals will any conflict occur in equilibrium.

The main contribution is to apply this logic to explain the long-term conflict legacies of historical statehood. Empirically, precolonial statehood affected three parameters in the theory of civil war aims: the identity of the ruling group, its ability to commit to powersharing deals, and its numerical size. Ethnic groups with a precolonial state were, on average, distinguished from non-PCS groups through diverse historical channels: precolonial warfare and slaving, privileges in colonial governance (indirect rule), and incentives to create regionally rather than nationally oriented policies during the post-World War II decolonization era or to continue governing ethnically exclusive monarchies. These factors not only enabled many PCS groups to control the government at independence, but also undermined the ruling group's ability to commit to powersharing deals with other ethnic groups in the postcolonial country, given the divisive interethnic relationships created by these precolonial and colonial interactions.[1] Consequently, ruling PCS groups faced incentives to exclude from power other ethnic groups within their country, which in turn created motives for groups that lacked a state before colonialism to organize a violent rebellion. However, variance in the size of the PCS group affected the opportunities that disfavored groups faced to launching a rebellion.

The main unconditional hypothesis is that members of *stateless* ethnic groups in countries with a PCS group should more frequently fight separatist civil wars than ethnic groups in countries without a PCS group, which tended to face lower incentives to violently control the political arena. Members of the PCS group itself should not rebel at elevated rates because their tendency to control the government should obviate the need to organize a rebellion around ethnic aims and recruitment; and the hypothesis is limited to center-seeking civil wars because, typically, PCS groups were large enough to deter fights for the center. But variability in the size of PCS groups yields conditional hypotheses: in countries with a small PCS group, other ethnic groups in their country should exhibit elevated rates of center-seeking rebellions in addition to separatist wars; and in countries with a particularly large PCS group, other ethnic groups in their country should not frequently fight either type of civil war. Finally, conditional on facing exclusion from power, PCS

---

[1]Throughout, statements such as these refer to individuals who belong to broader identity groups, rather than suggesting that ethnic groups act monolithically.

groups should fight center-seeking civil wars at elevated rates because their numerical size usually makes it feasible to take the center.

To test these hypotheses, I collected an original dataset on historical states that codes ethnic groups in over one hundred modern countries by whether or not, prior to European colonial rule, members of the group governed a polity that exhibited hierarchical organization across the territory inhabited by the modern ethnic group. Across a sample of ethnic groups between 1946 and 2013, stateless groups in countries with a PCS group participated in a separatist rebellion more than eight times as frequently as groups in countries without a PCS group. This correlation is robust to controlling for region fixed effects, standard civil war covariates, geographic and demographic traits of groups that facilitate separatism, precolonial development and conflict, colonial traits, and the artificiality of international borders. Regressions that condition either on the size of the PCS group or on ethnopolitical exclusion also support the theoretical predictions.

These findings relate to a growing literature on the long-term legacies of historical statehood, including several recent studies that analyze postcolonial civil war and reach mixed conclusions. Depetris-Chauvin (2015) and Wig (2016) find that precolonial statehood covaries with less frequent civil war onset in Africa, although Paine (2019) shows the opposite result when using precolonial state data with less measurement error and imposing other theoretically appropriate changes to the statistical models.[2] Ray (2019) analyzes ex-British colonies and also finds that precolonial statehood positively covaries with postcolonial civil war onset. However, he does not address groups' civil war aims, and concludes that *precolonial state* groups rather than stateless groups in their countries tend to fight civil wars—the opposite of my conclusion. The differences in findings likely arise from how we measure precolonial state groups: in his list of postcolonial ethnic war onsets (his Table 2), very few were fought by groups coded as PCS for the present paper. Section 3.2 and Appendix A.1 motivate how my PCS measure improves on Murdock's (1967) measure of precolonial statehood, which Ray (2019) and almost all existing articles use. Other related studies find that precolonial wars positively covary with post-independence civil war onset (Besley and Reynal-Querol 2014; Fearon and Laitin 2014; Dincecco, Fenske and Onorato 2016), study the impact of colonial policies on ethnopolitical

---

[2]These earlier studies control for country fixed effects in every model. However, both here and in Paine (2019), the theory predicts that PCS groups caused civil war through their effect on excluding other ethnic groups *within their postcolonial country*, implying that controlling for country fixed effects generates uninformative counterfactual comparisons.

access and civil war (Cederman et al. 2015; Wucherpfennig, Hunziker and Cederman 2016), or show that ethnic groups partitioned by European powers frequently fight civil wars (Englebert, Tarango and Carter 2002; Michalopoulos and Papaioannou 2016).

The final component of the empirical analysis shows that differences in precolonial states can explain variance in types of ethnic civil wars across world regions. Center-seeking civil wars are particularly prominent in Sub-Saharan Africa (SSA), which corresponds with the relatively high percentage of countries in the region with a PCS group but also the relatively small size of PCS groups in SSA. Although Asian ethnic groups participate in separatist civil wars at similar rates as those in SSA, the overwhelming majority of ethnic civil wars in Asia are separatist. Despite even higher prevalence of historical states in Asia, these groups tend to be large as a percent of the population, which deters center-seeking and sometimes separatist civil wars as well. By contrast, there were few precolonial states in Latin America, which has experienced almost no major ethnic wars since 1946. Many studies provide insight into dynamics of civil war or specifically of separatism within particular regions, but without elaboration cannot explain these cross-regional patterns.[3] Others provide more general theories of separatist civil war onset (Walter 2009; Lacina 2015), but the causal factors that they analyze do not provide obvious explanations for these cross-regional discrepancies nor how rebels choose civil war aims. More broadly, we lack general theoretical frameworks that explain civil war aims, and present analysis makes progress on this front.

2 THEORY

2.1 THE STRATEGIC LOGIC OF REBELLION AIMS

This section explains the theoretical logic connecting the size of a country's ruling group and the credibility of its powersharing offers to a challenger's incentives to fight a center-seeking or separatist civil war. I explain the intuition verbally here and formalize the logic in Appendix D.1.

The first theoretical building block concerns how the government's ability to commit to powersharing deals affects motives to rebel. Many theories of domestic conflict assume that a government interacts with a challenger group that can fight against the government if not offered sufficient concessions. A central

---

[3]See, for example, Englebert (2009), Roessler and Ohls (2018), and Paine (2019) for SSA; Toft (2005) and Smith (2013) for Eastern Europe; and Ross (2010) for Latin America.

impediment to striking a peaceful bargain is the government's inability to commit to deals made with the opposition. For example, the government may offer certain cabinet positions to members of the challenger group that—if they retained perpetual control over the positions—would facilitate sufficient rents that the challenger would prefer the powersharing agreement over rebelling against the government. Alternatively, the ruler may agree to compete in somewhat competitive executive elections or at least allow the opposition to compete in legislative elections. However, even if the government makes these concessions initially, they may later renege on the deal. Below, I discuss and provide examples of members of ethnic groups with historical states undermining the types of interethnic organizations that often promote commitment ability in otherwise weakly institutionalized polities.

The second theoretical building block addresses how the size of the ruling group affects the challenger's opportunity to fight either a center-seeking or separatist civil war. I assume that although a larger ruling group decreases the challenger's probability of winning either type of rebellion, the effect of a larger ruling group more strongly diminishes the challenger's probability of winning a center-seeking civil war. When the government's forces are numerically superior to the challenger's, it is exceedingly difficult to defeat the military in the capital. However, although sizable government forces can also be used to fight outside the capital, factors such as greater knowledge of terrain and local support help rebels to prolong a peripheral insurgency. Stated differently, the marginal effect of an additional government soldier on diminishing the challenger's probability of winning is larger in magnitude if the government defends the capital than if it fights in the periphery. This logic relates to Buhaug's (2010) empirical finding that regimes with greater coercive strength tend to fight battles farther from the capital. Rebels only stand a chance against strong regimes by fighting in areas that minimize power differential.[4]

Given these premises, Figure 1 summarizes the theoretical logic. It shows how the challenger's expected utility under five different outcomes varies in the size of the governing group. The black lines present the challenger's utility to initiating either type of rebellion, separatist or center-seeking. Both decrease in

---

[4]There are also geographical constraints that affect the viability of secession, which I operationalize below but are not necessary to derive the core theoretical logic. In addition to the size of the ruling group, its geographic dispersal—in particular, whether its territorial location overlaps with that of the challenger group—affects separatist prospects, as does whether or not the challenger resides in a territorially concentrated area that serves as a regional basis for an autonomous region or independent state.

Figure 1: Theoretical Logic
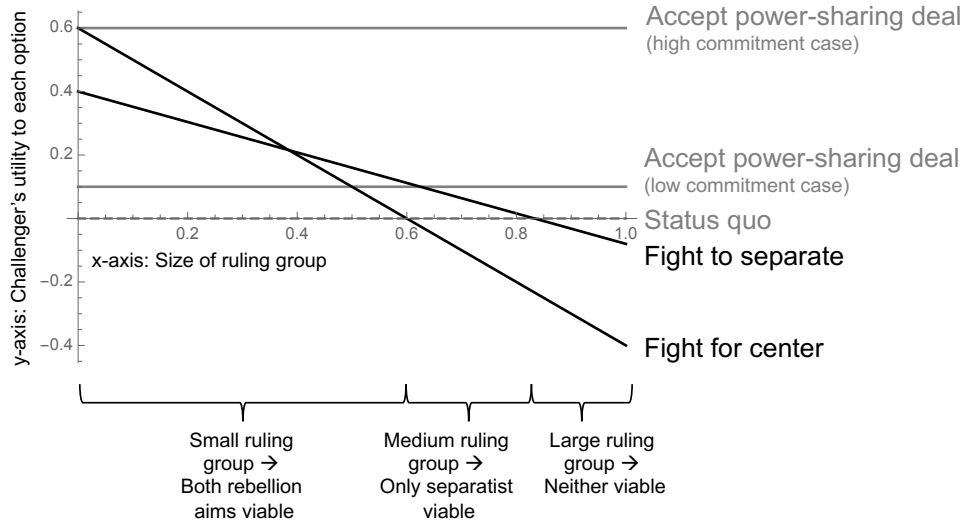
Notes: Appendix D.1 presents the functional form assumptions and assumed parameter values used to generate the expected utility terms in Figure 1.

the size of the ruling group, but the decrease for center-seeking is steeper. Comparing these lines to the dashed gray line for the challenger's utility if it does not fight in reaction to the government offering nothing (i.e., the status quo) highlights its equilibrium opportunity for rebelling. If the ruling group is small, then the challenger prefers either type of rebellion to the status quo. If the ruling group is medium-sized, then the challenger can still viably threaten to secede, but its expected utility to center-seeking fighting drops below its utility from accepting no government concessions. Finally, if the ruling group is very large, then separatism also loses its viability.

The government can offer a powersharing deal to attempt to prevent fighting. If the government's ability to commit to deals is high (top solid gray line), then it can always offer a powersharing deal that the challenger prefers to either type of rebellion, that is, the motives for rebellion are low. However, a government with limited commitment ability (lower solid gray line) creates high motives for rebellion and the logic is largely the same as comparing the two rebellion options to the status quo: the challenger prefers either type of rebellion to accepting the powersharing deal if the governing group is small, the group will fight to separate but not for the center if the governing group is medium-sized, and a large ruling group deters war.[5]

---

[5]The theoretical framework assumes that commitment ability is orthogonal to group size, which enables distinguishing motives from opportunity from rebellion. Although in reality the two may correlate, below I provide examples to distinguish the two. Furthermore, the theoretical logic is identical as long as the

2.2 Acronyms

Before explaining how precolonial statehood affected key parameters in the model, I introduce the key acronyms used to distinguish countries and ethnic groups. Countries that contain at least one PCS group are "PCS countries" and those that do not are "non-PCS countries." PCS countries contain precolonial state ethnic groups ("PCS groups") and stateless groups ("SLPCS groups"). All ethnic groups in non-PCS countries are stateless ("SL groups").

Figure 2 illustrates the abbreviations. The Baganda in Uganda were organized under the Buganda kingdom before colonization and are a PCS group (black). Therefore, stateless ethnic groups in Uganda, such as northern groups, are SLPCS (light gray), and Uganda is a PCS country. By contrast, Kenya contains no PCS groups, making it a non-PCS country. All its groups, including the Kikuyu, are SL (dark gray).

Figure 2: Acronym Examples
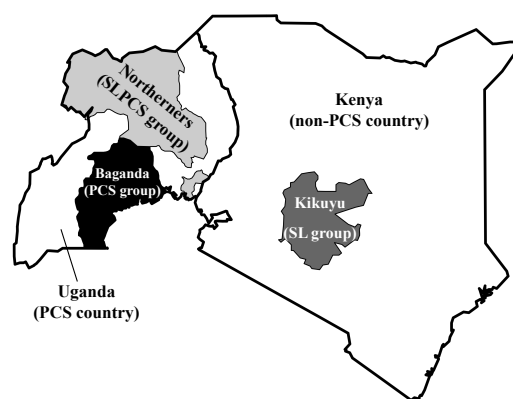
2.3 Controlling the State at Independence

PCS groups affected three main parameters in the theoretical framework: (1) which group governed, (2) the government's ability to commit to powersharing, and (3) size of the ruling group. PCS groups contributed to intergroup inequality before, during, and at the end of colonial rule, which gave them advantages for controlling the state at independence while also diminishing their ability to commit to share power. In the precolonial era, they frequently fought wars and raided their stateless neighbors for slaves. In every non-colonized PCS country, the PCS group controlled an ethnically exclusive monarchy into the twentieth century and, in many cases, through 1945. For PCS countries that experienced European colonization, the colonizer often favored the PCS group through indirect rule, which led many PCS groups to create regionally rather than nationally oriented policies during the post-World War II decolonization era.

Precolonial polities and ethnic groups differed in ways that affected the political salience of ethnic identities. Centralized states often participated in violent activities posited to promote intergroup inequality and domi-

expected utility of rebellion lines more steeply decrease in ruling group size than does commitment ability.

7

nation (Cederman, Gleditsch and Buhaug 2013, 33), including frequent wars with neighboring peoples. To demonstrate this point, I coded precolonial internal wars at the level of modern-day countries, using Brecke's (1999) list of conflicts between 1400 and the present. I counted only wars fought between groups within the same modern-day country boundaries—hence "internal" wars—as opposed to any war fought across modern-day borders. For example, the entry "Funj-Musabaat Arabs (Sudan), 1747" constitutes an internal war in Sudan in 1747, but "Ethiopia-Funj (Sudan), 1744" is not an internal war for either Sudan or Ethiopia. These data show that for the sample of countries described below, between 1400 and either the last year before colonization or 1900 (whichever occurred earlier), PCS countries experienced internal wars in 4.8% of precolonial years compared to 0.4% of years in non-PCS countries.[6] During this period, 65% of PCS countries experienced at least one war compared to 24% of non-PCS countries.[7] More precisely geocoded data for Africa between 1400 and 1700 reveals a similar discrepancy when analyzing territories in which modern PCS groups reside (29% experienced at least one precolonial war) versus territories without a PCS group (14%), and many of these PCS groups raided stateless neighbors for slaves (Paine 2019, 11).

The onset of colonial rule or, in some cases, the perpetuation of a historical state (e.g., China, Saudi Arabia) solidified the historical advantages of state-like groups. In 1945, across non-colonies and colonies, 41% of PCS countries (using modern-day borders) had either a ruling monarch or a monarch with an elevated position in the government, compared to only 6% in non-PCS countries.[8] In all cases, the ethnic group coded as PCS for the present paper controlled the monarchy. In some countries, a PCS group continued to govern an ethnically exclusive monarchy after 1945. For example, in Iran, the Soviet Union promoted Kurdish nationalism during its occupation of Iran in World War II, leading to a brief Kurdish state (the Republic of Mahabad) in 1946. However, the reassertion of the Persian monarchy after the war enabled the government to forcibly end the Republic, and over the next three decades Mohammad Reza Pahlavi consolidated his personal rule while excluding Kurds and members of all other ethnic groups (except Azeri) from power during his reign (Katouzian 1998).

---

[6]The p-value for this difference in a two-sided difference-in-means test is 0.0037. China is a vertical outlier, and excluding China reduces average precolonial years for PCS countries to 3.6% but also decreases the p-value for the difference to 0.0000 by decreasing variance in the outcome.

[7]The p-value for this difference in a two-sided difference-in-means test is 0.0000.

[8]Author's coding. The three exceptions in non-PCS countries were British-governed territories with a colonially created monarchy: Iraq, Jordan, and Libya.

PCS groups that experienced European colonization also tended to perpetuate their privileges, even if they lost their monarchy, through elevated positions in the colonial governance hierarchy. British indirect rule was best-suited for hierarchically organized groups and many PCS groups gained considerable autonomy: Arabs in the Persian Gulf protectorates, Asante in Ghana, Bhutanese in Bhutan, Baganda in Uganda, Hausa and Fulani in Nigeria, Sotho in Lesotho, Swazi in Swaziland, and various princely states in India. However, indirect rule through PCS groups was not limited to the British empire. France pursued a similar policy with large monarchies in its protectorates for whom they allowed legislative power, as in Tunis (Tunisia), Morocco, Annam (Vietnam), and Cambodia (Savary 1952); and also practiced indirect governance for various ethnic groups for whom they had previously deposed the monarch, such as Fon in Benin, Muslim Sahelian groups in Chad, and Wolof in Senegal. Germany, and later Belgium, allowed considerable autonomy for ruling monarchs in Ruanda-Urundi (Rwanda and Burundi).

Even in cases where the colonizer did not favor the PCS group, group members sometimes leveraged their history of statehood to reverse their fortunes. For example, in Burma (Myanmar), Britain fought a series of wars against the historical Burman state and excluded Burmese from the armed forces because Britain feared that Burmese could join an anticolonial nationalist movement.[9] In 1931, despite composing 75% of the population, only 12% of the indigenous soldiers in Burma's army were Burmese. Instead, members of the minority groups Karen, Chin, and Kachin provided the majority of troops. However, during World War II, British authority collapsed and Burmese individuals created the Burma Independence Army (BIA), which helped to consolidate a sense of common citizenship. The BIA was almost exclusively Burman and they used their force against minority groups in the colony during World War II, which many argue "irrevocably sealed a split between Karens and Burmans." The BIA preceded postcolonial Burma's ethnically stacked and repressive military, despite a decade-long interlude of civilian rule and tenuous ethnic powersharing immediately after independence.[10]

---

[9]This paragraph draws from Callahan (2005, 35-6, 53, 56, 75).

[10]Before their post-World War II resurgence, the Burmese instead more closely fit the pattern discussed in Ray (2019) and many other accounts of colonial security forces: British officials underrepresented ethnic groups with precolonial states, instead relying on "martial" races to provide security. This possibility is not mutually exclusive from the present discussion because precolonial state groups could have been disadvantaged in the colonial security forces yet still tended to gain control of the state at independence. However, Ray (2019) collected novel data on the composition of British colonial police forces that enables assessing

In the two decades following 1945, European colonizers introduced elections in most colonies, which engendered political parties. The historical privileges and distinction of PCS groups created the organizational ability to form ethnically oriented parties that engendered a fractured political arena at independence, as opposed to participating in broad interethnic governance coalitions. Examples include the Baganda's KY ("King only") party in Uganda, northern Nigeria's Northern People's Congress, and the Burmese-controlled Anti-Fascist People's Freedom League that organized to displace Karen organizations for control of Burma. Analyzing SSA countries, Paine (2019) provides statistical evidence that each of PCS and SLPCS groups were considerably more likely to be represented by ethnic parties than SL groups in the final elections before independence, and ruling monarchies in many Asian countries created similar conditions of ethnic distinction and dominance elsewhere.

The Uganda case is illustrative, especially because the relatively small size of its PCS group (Baganda were 16% of the population at independence) creates a harder case for finding evidence for disruption from a PCS group. When Britain colonized Uganda, it bestowed the powerful state of Buganda—which frequently warred with and slave-raided from neighboring groups also incorporated into modern Uganda (Reid 2012, 115-16)—with significant self-governance privileges. "The special status of Buganda in Uganda was the most important legacy of the colonial era" and their founding treaty with Britain in 1900 "appeared to the Baganda as in some sense at least an agreement between equals" (Rothchild and Rogin 1966, 341). Therefore, Buganda's later "integration within the rest of Uganda posed serious problems first to colonial officials and subsequently to nationally oriented African politicians [because] Buganda could not be dethroned from its dominant position without seriously compromising the viability of Uganda as a whole" (Doornbos 1977, 241). In response to Britain's attempt to unify colonial administration after World War II, Buganda attempted to secede from the rest of Uganda to "safeguard the traditions, Kabakaship, and the customs of Buganda in an independent Uganda." (Rothchild and Rogin 1966, 348). The king—known as the *kabaka*—cited Buganda's distinct status in the Uganda Agreement of 1900 to promote his claim. Consequently, "the power of traditional groups . . . precluded the success of a centralized, ideological mass party"

this pattern systematically. Appendix Table C.1 shows that when using my PCS variable, there is no association between PCS and ethnic colonial police imbalance. Section B.6 of Paine (2019) provides additional evidence that martial race recruitment cannot explain patterns of violence in PCS countries in Sub-Saharan Africa.

among all Ugandans (389). Supporters of the *kabaka* instead created the Kabaka Yekka party—meaning "king only"—after the *kabaka* led a highly successful boycott of the 1961 Legislative Council elections in which less than 2 percent of eligible Baganda voted. Kabaka Yekka provided "a practical avenue through which Buganda could enter national politics and yet preserve its own autonomy and unity" (358). The ethnically oriented party received 26 percent of parliamentary seats in the final pre-independence elections in 1962 (Schmidt 1999, 934). In response to political deadlock created by the sizable vote share of Buganda's ethnically oriented party, an interethnic ruling coalition formed at independence that composed an "alliance of complete opposites" (Decalo 1990, 152) between Kabaka Yekka and a major party led by a member of an SLPCS group, Milton Obote's UPC party.

Figure 3 systematically demonstrates PCS groups' power access advantages at independence by depicting patterns of ethnopolitical access in the central government for all ethnic groups in 1946 or their country's first year of independence (whichever is later).[11] The figure sorts power access by SL, SLPCS, and PCS groups, and lists the percentage with an ethnopolitical power access score at least as high as the level stated: monopoly, dominant, senior partner, and junior partner. At each level, PCS groups stand out for high political power status at independence, especially relative to stateless groups in their country—15.4 times more likely to be monopoly or dominant, 7.0 times for senior partner or higher, and 2.6 times for junior partner or higher—but also relative to SL groups. Appendix Table C.2 shows that these correlations are statistically significant and are robust to controlling for the covariates used in the statistical models below (including ethnic groups' population share) as well as to modeling country fixed effects to create within-country comparisons between PCS and SLPCS groups.
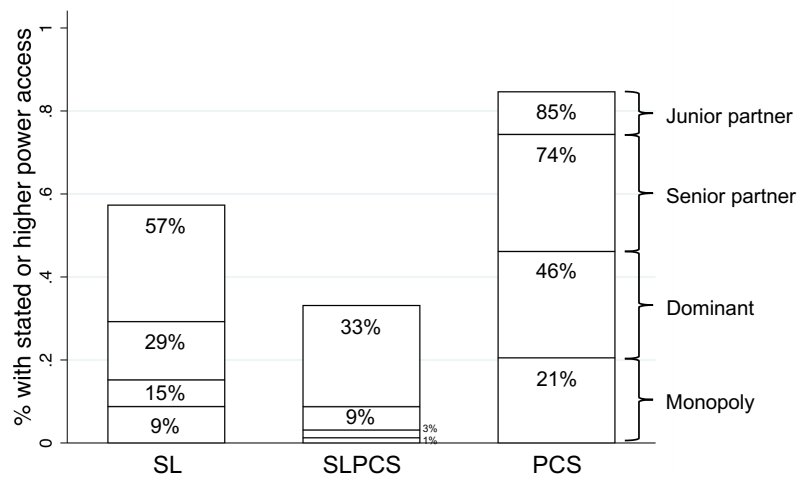
2.4   ETHNOPOLITICAL EXCLUSION

The same historical advantages of PCS groups that enabled controlling the postcolonial state also undermined the ability of governments in their countries to commit to powersharing deals. Ethnically exclusive monarchies or political parties—which reinforced precolonial and early colonial governance differences between PCS and stateless groups—prevented the creation of stable interethnic coalitions, as the Iran, Burma, and Uganda examples highlighted. Typically, PCS groups excluded others from power, as Figure 3 shows with the low prevalence of any cabinet positions for SLPCS groups. Yet even in countries like Uganda and

---

[11]Data from the Ethnic Power Relations database (EPR; Vogt et al. 2015), described below.

Figure 3: Ethnopolitical Power Status at Independence

Nigeria that inherited powersharing arrangements at independence, less than a decade after independence, crises occurred—followed by ethnopolitical exclusion and civil war. In Uganda, "it is hard to determine at what stage Prime Minister Obote made up his mind to confront the Kabaka and the State of Buganda . . . but it is tempting, from the small amount of evidence available and his careful preparing of the ground, to think that he had intended it all along" (Dinwiddy 1981, 514).

By contrast, even in settings where the rule of law and constitutions are not well established, alternative institutions can enable political leaders to commit to bargains. In many non-PCS countries, such as Côte d'Ivoire and Tanzania, rulers created nationalist parties that facilitated interethnic cooperation after independence. Even Lebanon, which later broke into mass conflict, experienced three decades of internal peace through its consociational arrangement that facilitated power access for parties representing the major ethnic groups. More broadly, the authoritarian politics literature shows how party institutions can facilitate credible commitment (Magaloni 2008; Svolik 2012). Using the related metric of cabinet positions, Francois, Rainer and Trebbi (2015) show in a sample of fifteen countries in Africa that rulers tend to allocate cabinet positions in proportion to ethnic group size to balance sectional interests—contrary to notions that everywhere and always rulers attempt to achieve ethnic hegemony. However, highlighting how PCS countries differ from others, Roessler (2016, 68) notes that their sample excludes the six most ethnically dominant regimes in Africa—Angola, Burundi, Ethiopia, Rwanda, Sudan, and South Africa—all PCS countries.

In some cases, the divisive interethnic relationships caused by PCS groups undermined the government's ability to commit to power-sharing deals because the ever-present fear of a coup attempt created incentives

to purge members of the ruling coalition, especially those from other ethnic groups. These fears were most prevalent in African countries given the relatively small size of PCS groups and their inability to monopolize the colonial officer corps (Roessler 2016; Paine 2019). The fear of coup attempts by stateless groups was less acute in Asian countries governed by large PCS groups (and in many cases named after the dominant ethnic group), but a long history ethnically exclusive monarchies more frequently engendered ethnonational principles of political legitimacy in which "the state is ruled in the name of an ethnically defined people and rulers should therefore care for 'their own people'" (Wimmer, Cederman and Min 2009). Reinforcing these tendencies, in many cases where PCS groups succeeded in gaining their own countries, the ethnic geography created precipitous conditions for "sons of the soil" separatist conflicts. In China, India, Bangladesh, Thailand, and Myanmar, the PCS group lives in a densely populated riverine area. This created incentives for the government to encourage poor co-ethnic farmers to migrate to hinterland areas, and the desire to exploit hinterland ethnic groups undermined incentives to share power with them (Fearon and Laitin 2011).
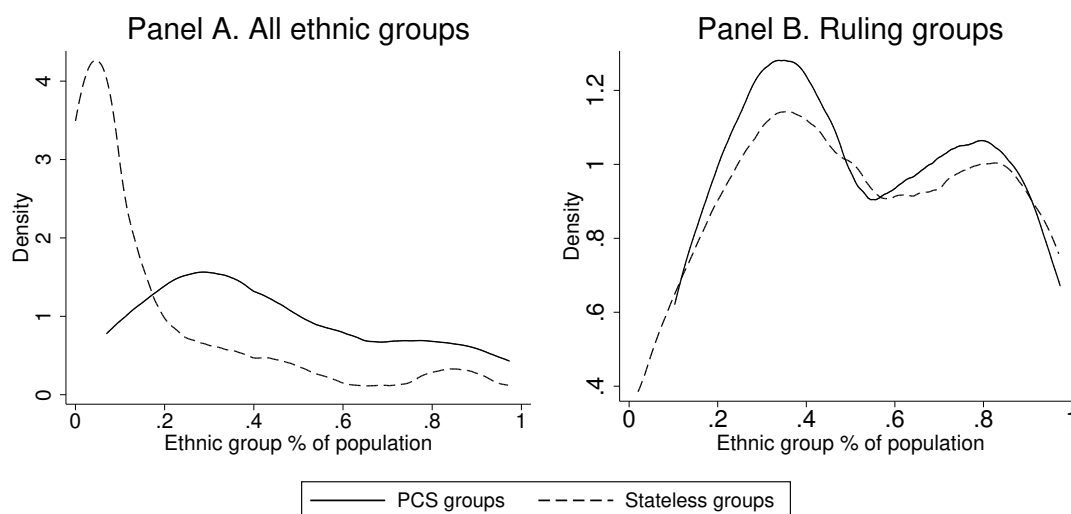
2.5 SIZE OF PCS GROUPS

PCS groups vary considerably in the percentage of their country's population, which resulted from varying precolonial and colonial experiences. In some cases, such as China and Japan, the PCS group governed a large precolonial empire in which members of the PCS group composed the overwhelming majority of the population, and this empire formed the basis of a modern country that experienced either no or minimal colonial rule. By contrast, empires such as Buganda and Dahomey were considerably smaller and fell to European powers, who incorporated the Baganda and Fon into colonies (Uganda and Benin, respectively) that contained numerous other ethnic groups. But even PCS groups in non-colonized countries exhibit considerable variance based on the historical process of creating their state. In the first year of independence, PCS groups averaged 57% of the population in non-colonized countries—including cases such as Amhara in Ethiopia and Zaydis in Yemen in which the PCS group was less than 30% of the population—compared to 49% of the population in colonized countries (this difference is not statistically significant). PCS groups in countries that experienced colonization also exhibit considerable variance. In contrast to patterns in Buganda and Dahomey, in some cases, European colonizers preserved very small kingdoms into their own colonies for idiosyncratic reasons, such as Rwanda, Burundi, Lesotho, and Swaziland. Koreans and North African Arabs were each partitioned into different colonies, but some of these countries were themselves highly

ethnically homogeneous (North Korea, South Korea, Tunisia, Egypt).

Figure 4 examines aggregate patterns. The sample in Panel A is all politically relevant ethnic groups in their country's first year in the sample. This shows that PCS groups tend to be larger than stateless groups as a percentage of their countries' population, with respective averages of 45% and 18%. However, when analyzing the size of ruling groups at independence, PCS countries are very similar to non-PCS countries, and the average for each is 54% of the population. Relating this percentage to the geography of rebellion, it is large enough that few groups besides the ruling group will have a viable chance to overthrow the government in the capital, but low enough that in many countries there will be groups that reside in territories distinct from that of the ruling group, which provides a potential base for seceding. Figure 7, discussed below, provides more evidence of this claim.

Figure 4: Size of Ethnic Groups

Although it is beyond the scope of the present paper to explain the size of ethnic groups before independence and how European border drawing mapped those groups into modern countries, the key observations are: PCS groups tended to compose a relatively large percentage of the population of their postcolonial country, but also exhibited heterogeneity in size that corresponds with the differential opportunities for rebellion highlighted in the theory of civil war aims.[12]

---

[12]There is considerably less quantitative research on the causes of ethnolinguistic diversity than on its effects for conflict and other outcomes. African ethnic groups tend to be smaller because of precolonial factors—higher geographic variability (Michalopoulos 2012) and heavy participation of certain areas in the

2.6 HYPOTHESES

Combining the historical background with the logic of strategic rebellion aims yields the hypotheses. The pernicious effects of PCS groups on stable interethnic powersharing relationships created motives for rebellions in PCS countries, specifically among SLPCS groups given their frequent exclusion from power. However, despite these motives, the relatively large size of the average ruling PCS group often prevented SLPCS group from having a viable opportunity to take the center, leading them instead to pursue separatism. Crucially, the posited within-country spillover effects of PCS groups—in the sense of creating rebellion incentives for *stateless* groups in their countries—implies that the relevant comparison for SLPCS groups is to SL groups. Therefore, theoretical considerations require disaggregating stateless groups by whether or not any PCS groups reside in their country. If the theory is correct, then comparing PCS groups to SLPCS groups would yield the incorrect implication that precolonial statehood exerts a *pacifying* effect by ignoring within-country spillover effects. This logic yields the main hypothesis:

> **Hypothesis 1.** *SLPCS (but not PCS) groups should participate in separatist (but not center-seeking) civil wars more frequently than SL groups.*

Although the typical PCS group was large enough to deter center-seeking but not separatist fighting, variance in the size of PCS groups yields two conditional predictions that follow directly from the posited theory of civil war aims.

> **Hypothesis 2.**
>
> > *a. Among countries in which the PCS group is small, SLPCS groups should participate in center-seeking civil wars more frequently than SL groups.*
> >
> > *b. Among countries in which the PCS group is large, SLPCS groups should not participate in separatist civil wars more frequently than SL groups.*

A more precise mechanism linking large PCS groups to a lack of separatist attempts by SLPCS groups is that when the ruling group is large, fewer groups will have distinct territories from which they can attempt to create an autonomous state. Therefore, considerations about historical statehood relate to more general

transatlantic slave trade (Whatley and Gillezeau 2011)—and Europeans' decisions to create geographically large colonies in areas with low population density among peoples that lacked a shared precolonial history (Green 2012). This is a fruitful area for future research.

geographical and demographic characteristics that affect rebellion prospects.

Hypothesis 3. *Among ethnic groups with high territorial overlap with the ruling group, SLPCS groups should not participate in separatist civil wars more frequently than SL groups.*

Finally, the present theoretical framework can also be productively combined with that from Cederman, Wimmer and Min (2010), which analyzes rebellion incentives for groups that lack access to power at the center. Unconditionally, we should not expect PCS groups to rebel at high rates because they should tend to control the state. However, conditional on ethnopolitical exclusion, PCS groups should face high rebellion incentives. Although some groups excluded from power pose minimal rebellion risk because leaders chose to exclude them specifically because they do not pose a threat (*opportunistic* exclusion), this will not typically be the case for PCS groups. Instead, the threat they pose to others creates incentives to exclude them for *strategic* reasons, perhaps because other groups fear that a powersharing arrangement with a PCS group will enable staging a coup (Roessler 2016, 60-81). And, since the ruling group in PCS countries will tend to be small if the PCS group is excluded from power, conditional on fighting they should usually seek the center. A similar logic applies to SLPCS groups conditional on exclusion, although since these groups tend to be smaller, conditional on fighting their rebellion aims should be separatist.

Hypothesis 4. *Conditional on ethnopolitical exclusion:*

> ***a.*** *PCS groups should participate in center-seeking civil wars more frequently than SL groups.*

> ***b.*** *SLPCS groups should participate in separatist civil wars more frequently than SL groups.*

3   DATA

3.1   SAMPLE

The unit of analysis in the main regression specifications is ethnic group-years. The sample includes politically relevant ethnic groups from most countries in Latin America, North Africa and the Middle East, Sub-Saharan Africa, and Asia between the later of 1946 and independence, and 2013.[13] The baseline speci-

---

[13]Future drafts of the paper will add Eastern European and former Soviet countries. I do not include countries that violently seceded from PCS countries but lacked any PCS groups after secession (East Timor, Eritrea, South Sudan) or country-years following the loss of a country's only PCS group (Pakistan after

fication contains 580 ethnic groups in 101 countries. I draw ethnic groups from the Ethnic Power Relations database (EPR; Vogt et al. 2015), which provides panel data on politically relevant ethnic groups and their access to power in the central government. Cederman, Wimmer and Min (2010, 99) "classify an ethnic group as politically relevant if at least one political organization claims to represent it in national politics or if its members are subjected to state-led political discrimination."

3.2 PRECOLONIAL STATEHOOD

Coding the variable. The operational definition for coding an EPR ethnic group as governing a precolonial state is that co-ethnics governed a substantial percentage of members of the EPR ethnic group through a single or small number of political organizations that exhibited some degree of centralized rule on the eve of colonization. Although the operational definition of a state is minimal, the paucity of reliable historical information for many precolonial political organizations across the global sample makes it difficult or perhaps impossible to operationalize a conceptual definition that requires more nuanced information about the degree or origins of centralization for each group. I restrict attention to states that existed at the eve of colonization because the theory posits that precolonial states' influence on colonial policies is a key persistence mechanism, and for noncolonized countries I examined conditions between 1800 and 1945.

To construct the dataset, I consulted various general sources to generate a list of possible precolonial states: *Encyclopaedia Britannica* (2019), historical dictionaries, and Library of Congress country studies.[14] In many cases, these sources were sufficient to determine either an EPR ethnic that governed a precolonial state, or that the country contained no centralized ethnic groups before independence. If there was ambiguity regarding whether a precolonial polity qualified as centralized or with which EPR ethnic group to

Bangladesh seceded in 1971) because it is unclear how to code groups in these countries. On the one hand, there was no longer a PCS group in the country after secession, which suggests that coding the groups as SLPCS is incorrect. On the other hand, in all these cases, previous interaction with the PCS group contributed considerably to the country's fragile institutions, and therefore recoding the groups as SL would be inappropriate given the posited theoretical mechanisms.

[14]This builds on an earlier coding project by Paine (2019) to code precolonial states in Africa. Although he consulted a wider range of sources to generate a list of candidate precolonial states, I verified that every precolonial state he associated with an EPR ethnic group that he coded as PCS was mentioned (although often not discussed in detail) in its country's *Encyclopaedia Britannica* article.

match the state, I consulted additional sources listed in Appendix A.1. I used the following specific criteria, which correspond with the operational definition of precolonial states: (1) co-ethnic governance requires the state was independent rather than tributary; (2) some evidence of central authority such as acknowledged hierarchy of authority in regions outside the capital and centralized tax collection, as opposed to nomadic confederations or trading centers; (3) one (or a small number of) states governed a substantial percentage of members of the EPR ethnic group, as opposed to groups such as Yoruba in Nigeria or Bamileke in Cameroon that were fractured into dozens or hundreds of mini-states; and (4) these conditions held on the eve of colonization. Appendix Tables A.1 through A.4 list every PCS group and PCS country, and Appendix A.1 provides country-by-country coding justifications for the variable.

I coded only the largest ethnic group in each country with a precolonial state as a PCS group. This primarily reflects theoretical considerations, as smaller groups in a country with a larger PCS group would likely suffer many of the same injustices as stateless groups (e.g., Fur in Sudan). Regarding measurement error concerns, the existence of at least one PCS group in a country implies that no groups in that country are coded as SL, and extending the coding to classify smaller ethnic groups as PCS would not change that important distinction.[15] This decision also reflects practical considerations given the large number of countries and ethnic groups in the sample and the higher degree of difficulty to finding reliable sources on smaller ethnic groups, especially because the general sources inconsistently address smaller groups. This coding decision also implies a caveat for SLPCS groups: not every one was stateless, although none were the *largest* state in their modern-day country.

Comparison to existing measures. Any attempt to measure historical characteristics of modern ethnic groups is inherently imperfect. Neither ethnic identity itself nor its political relevance is static. Ethnic identification underwent particularly important changes during the colonial period as European administrators attempted to demarcate all persons as members of a specific ethnic group for taxation and land allocation purposes. However, given the prevalence of research on postcolonial ethnic conflict, the present project attempts to extend our knowledge of historical factors that matter—given the implausibility that the causes

---

[15]In other words, given the crucial theoretical distinction between ethnic groups in PCS countries and those in non-PCS countries, the greater potential source of measurement error arises from miscoding a country as containing *no* PCS groups when in fact it does contain at least one; but deciding how to code groups smaller than the largest PCS group does not raise this concern.

of modern ethnic violence are entirely postcolonial.

Despite numerous important research questions to which existing measures of historical political centralization contribute, properly assessing the present hypotheses requires a new measure. Bockstette, Chanda and Putterman (2002) code a territory's history of state-like institutions dating back over two millennia to the year 0 CE. Their data set uses modern country boundaries as the unit of analysis, which precludes assessing ethnic group behavior, although a robustness check for the main regression tables uses their state antiquity variable measured in 1500.

Murdock's (1967) *Ethnographic Atlas* and Murdock and White's (1969) Standard Cross-Cultural Sample code an ordinal political "jurisdictional hierarchy" variable at the ethnic group level on the eve of colonization, used widely in the literature to measure precolonial statehood. However, even with subsequent updates, the Standard Cross-Cultural Sample covers only 186 ethnic groups across the world. Even in Sub-Saharan Africa, where the *Ethnographic Atlas* has comprehensive coverage, Paine (2019) shows that the jurisdictional hierarchy variable is questionably coded for many cases and Murdock's list of ethnic groups is extremely difficult to merge with EPR ethnic groups or any other dataset with extensive coverage of postcolonial civil wars.

3.3 Dependent Variable

The dependent variable in most regressions is the onset of either a major separatist or center-seeking civil war. I coded ethnic group-level civil war onset by assigning Fearon and Laitin's (2003) major civil wars (at least 1,000 battle deaths) to EPR ethnic groups primarily using ACD2EPR (Vogt et al. 2015), which codes ethnic wars as involving ethnic-specific recruitment and war aims. As Appendix B.1 details, this civil war onset variable has advantages over UCDP/PRIO conflict data by using rigorous criteria for coding civil war "onset" as well as excludes minor conflicts. Regarding war aims, empirically, almost all post-1945 civil wars enable relatively unambiguous codings about center-seeking versus separatist goals. For the present civil war variables, I combined information from Fearon and Laitin (2003) and other conflict datasets to code war aims. Only two cases yielded codings of multiple war aims for the same rebel group: the SPLM/A in Sudan, and the EPRDF and constituent groups in Ethiopia. More frequently, center-seeking and separatist civil wars occurred simultaneously within the same *country*—including Angola, India, and Myanmar—although each *rebel group* in these conflicts pursued either center-seeking or separatist aims (but not both). The sample of

19

post-independence years implies that I do not include wars fought to gain independence.

3.4 STATISTICAL MODELS

The main regression tables estimate logistic regressions:

$$\ln\left(\frac{Y_{it}}{1 - Y_{it}}\right) = \beta_0 + \beta_P \cdot P_i + \beta_S \cdot S_i + \boldsymbol{X}'_{it}\beta_X + \boldsymbol{T}'_{it}\beta_T + \epsilon_{it}, \tag{1}$$

where Y_{it} is an indicator variable for either separatist or center-seeking civil war onset (with years of ongoing civil war dropped), P_i indicates PCS groups, S_i indicates SLPCS groups, β_P and β_S are the main parameters of interest, and $\boldsymbol{X}_{it}$ is a vector of covariates that differs by column. The vector $\boldsymbol{T}_{it}$ contains standard event history controls for civil wars—years since the last with an ongoing civil war of the specified type, and cubic splines—plus lagged country-level civil war incidence (either center-seeking or separatist). I cluster standard errors at the ethnic group level given the repeated time observations for each group.

Importantly, the main specifications do not include country fixed effects. A key premise of the theory is that PCS groups caused within-country spillover effects that should raise the civil war propensity of *stateless* groups in their countries, which necessitates comparisons between SLPCS and SL groups—as the hypotheses posit—rather than comparisons among groups within PCS countries.

4 STATISTICAL RESULTS

4.1 PRECOLONIAL STATEHOOD AND SEPARATIST CIVIL WAR ONSET

Figure 5 depicts the core pattern, which supports Hypothesis 1: stateless ethnic groups in countries with at least one PCS group (SLPCS groups) are 8.8 times more likely than stateless ethnic groups in countries without any PCS groups (SL groups) to participate in separatist civil wars, whereas PCS groups are slightly less likely than SL groups to attempt violent secession. In fact, the sample contains only two instances of separatism by PCS groups, Luba Kasai in the Democratic Republic of the Congo in 1960 and Bengali in Pakistan in 1971. By contrast, for center-seeking civil wars, SLPCS groups exhibit only a marginally higher frequency than SL groups, and PCS groups exhibit the highest relative frequency (although only 1.7 times greater than SL groups).

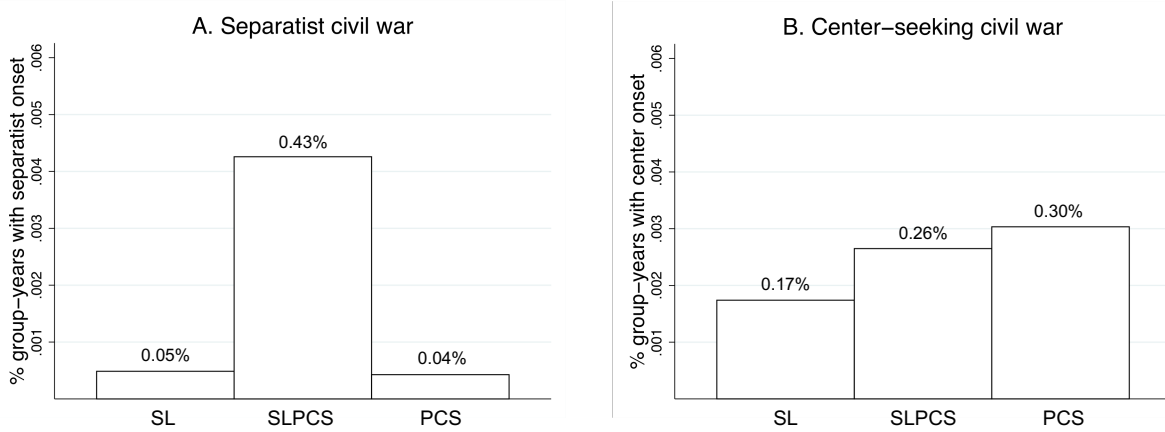**Figure 5: Civil Wars by Aims and PCS Indicators**

Table 1 examines the separatist relationship in more depth by assessing numerous alternative explanations. The first column runs a logit model with the PCS and SLPCS indicators—leaving SL groups as the omitted basis category—and event history controls but no substantive covariates, and shows that the correlation is statistically significant. Appendix Table C.4 contains the same set of specifications using center-seeking civil war onset as the dependent variable, and demonstrates a null estimated effect for PCS groups and SLPCS groups in almost every specification.

The remaining columns add substantive controls to account for alternative hypotheses.[16] The coefficient estimates for SLPCS remain relatively stable across the various specifications and retain statistical significance. Column 2 adds fixed effects for world regions (previewed above and analyzed in more depth below; Asia is the omitted basis category). Although one goal of studying historical statehood is to assess whether in can explain differences across world regions, it is important to show that precolonial statehood can also explain variance within regions (this specification drops every Latin American country because the fixed effect perfectly predicts the absence of separatist civil war onset). Column 3 adds four standard civil war co-variates, measured annually: logged income per capita, logged population, democracy level, and an indicator for ethnic groups that reside in territory with a giant oil field. Column 4 controls for several geographical and demographic characteristics that affect the feasibility of separating: the ethnic group's size as a percentage of the population, an indicator for whether the group is geographically concentrated in a particular territory, an indicator for whether the group's location is noncontiguous from the region of the country containing

---

[16]Appendix Section B.2 provides additional information and references for each covariate. Table C.3 provides summary statistics.

Table 1: Pre-Colonial Statehood and Separatist Civil War Onset

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	DV: Major separatist civil war onset						
SLPCS group	2.246***	2.061***	2.311***	1.803***	2.599***	2.214***	2.325***
	(0.638)	(0.623)	(0.661)	(0.665)	(0.763)	(0.683)	(0.776)
PCS group	0.842	0.609	-0.0576	0.785	1.166	0.837	0.400
	(0.932)	(0.924)	(1.200)	(0.931)	(1.012)	(0.958)	(1.270)
SSA		0.401					
		(0.317)					
MENA		0.471					
		(0.478)					
Latin America		-					
ln(GDP/pop)			-0.347**				
			(0.155)				
ln(Pop.)			-0.0833				
			(0.0639)				
Democracy			0.000151				
			(0.0207)				
Giant oil field			1.213***				
			(0.348)				
Group % pop.				-0.303			
				(1.012)			
Geo. concentrated				-			
Noncontiguous				1.248***			
				(0.396)			
Dist. from capital				0.0921			
				(0.0693)			
Neolithic transition					0.0730		
					(0.0621)		
Precolonial wars					-2.653***		
					(0.949)		
British colony						0.567	
						(0.386)	
Other colony						0.185	
						(0.440)	
Settler colonialism						-0.436	
						(0.590)	
Squiggly borders							-22.19
							(13.57)
Partitioned group							-0.348
							(0.315)
Group-years	25,646	21,769	24,006	20,383	24,016	25,646	21,156
Ethnic groups	580	506	553	458	533	580	476
Countries	101	78	95	86	91	101	79
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES

Notes: Table 1 summarizes a series of logistic regressions by presenting coefficient estimates for the variables, and ethnic group-clustered standard errors in parentheses. Column 2 drops every country in Latin America because this fixed effect perfectly predicts no separatist civil war. ***$p < 0.01$,** $p < 0.05$,* $p < 0.1$.

the capital city (either because it is an island or an enclave), and the logged distance between the group's location polygon and the capital city.

The remaining specifications address the origins of precolonial states and alternative historical explanations. Column 5 controls for two variables measured for modern-day countries: thousands of years since a territory experienced a Neolithic transition, and precolonial wars (percentage of years between 1400 and the onset of colonial rule in which a territory experienced an internal war). Previous research shows that each helps to explain the emergence of states,[17] although precolonial wars may themselves be endogenous to the earlier emergence of states.

Column 6 contains indicators for British colonialism and other colonies (leaving uncolonized territories as the reference category) and for settler colonialism. Recent research studies how colonial policies affected ethnopolitical power status and civil war (Cederman et al. 2015; Wucherpfennig, Hunziker and Cederman 2016). Following earlier historical research, they focus on differences between British and French colonialism. This distinction is relevant if British officials more frequently perpetuated the authority of precolonial rulers, although above I provided examples of other European colonial officials also ruling indirectly through PCS groups—and, in uncolonized territories, precolonial rulers enjoyed an even stronger position. Regarding settler colonialism, precolonial states may have exerted less effect on postcolonial outcomes in cases where the colonial period transformed the population, an especially acute concern for New World colonies in which natives lacked immunity to Eurasian diseases. The settler variable indicates whether Europeans ever exceeded 5% of the total colonial population, which includes every New World country and several in Africa.

Column 7 controls for other effects of European colonialism that could have affected separatism. Alesina, Easterly and Matuszeski (2011) measure two aspects of a country's artificiality: the extent to which a country's international borders are squiggly rather than straight (with straight corresponding with higher artificiality); and ethnic groups partitioned within a country, which other scholars link to civil war prospects (Englebert, Tarango and Carter 2002; Michalopoulos and Papaioannou 2016). For partition, I use an ethnic group-level measure coded by EPR that indicates whether the group is partitioned across international borders.

---

[17]Putterman (2008) provides empirical evidence for Neolithic transition. Many scholars associate warfare with state formation, most prominently Tilly (1992).

A broader consideration about confounders is that although precolonial statehood undoubtedly affected which ethnic groups were placed into which modern-day countries, it is unclear exactly how that would create bias in favor of a positive relationship between SLPCS and separatist civil wars. Not only do the aforementioned covariates directly control for the most obvious concerns, a distinct story would be required to explain away the relationship shown for *stateless* groups in countries with a precolonial state. These groups tended to exert less influence on the borders drawn for modern-day countries. Furthermore, Appendix Table C.5 shows that only a large amount of bias from unobserved covariates could explain away the correlation, given its relative immunity to controlling for various observable characteristics. Nor are the patterns from Table 1 altered when controlling either for every statistically significant covariate in Columns 2 through 7 or for every covariate across the specifications, although the latter drops many observations due to different patterns of missingness across the covariates (results available upon request).

Appendix C.1 presents additional robustness checks that yield similar findings. Tables C.6 and C.7 present results using a cross-section of politically relevant ethnic groups in the first year of the sample (for many countries, the year of independence) and either a binary indicator for any civil wars between then and 2013, or a count of civil war onsets. Table C.8 replaces the PCS group and SLPCS group indicators with Putterman's (2012) state antiquity index, measured in 1500. This is a weighted sum of number of years with government above the local level between 0 CE and 1500, measured using modern country boundaries. Table C.9 replaces the present civil war measure with the ACD2EPR measure that uses a lower death threshold and short lapse rules for coding onsets. I also assessed the robustness of the estimates to jackknife sample alterations. For each column in Table 1, I iteratively dropped all ethnic groups from each country in the sample. In each of these 631 regressions, the SLPCS indicator is statistically significant at 1% (results available upon request).

4.2 SIZE OF PCS GROUPS AND CIVIL WAR AIMS

Although, on average, SLPCS groups fought separatist civil wars more frequently than SL groups, PCS countries vary considerably. Table 2 lists every country in which a SLPCS group participated in an ethnic rebellion, disaggregated by size of the largest PCS group in the country (using thresholds described below) and rebellion aims. Eighteen countries are positive-positive cases that support Hypothesis 1 because a SLPCS group violently seceded, and eight experienced at least two onsets. Yet in 25 PCS countries no

Table 2: Civil War Onsets by SLPCS Groups

|	Small PCS group	Medium PCS group	Large PCS group
Separatist CW	Angola* Congo, DRC Ethiopia* India* Nigeria Sudan*	Indonesia* Iran* Morocco Myanmar* Senegal Sri Lanka Turkey Yemen	Bangladesh China* Mali Thailand
Center-seeking CW	Burundi* Chad* Congo, DRC* Ethiopia* Rwanda Sudan* Uganda*	Afghanistan*	

*Multiple onsets of the specified type of war.

SLPCS group initiated a separatist civil war, and 17 PCS countries experienced no ethnic conflict.

To explain this heterogeneity, this section assesses conditional effects for the size of PCS groups (Hypothesis 2). Figure 6 summarizes the main findings. The panels disaggregate by civil war aims, and each contains four bars. The first is for SL groups (black bar). The next three are for SLPCS groups, disaggregated by the size of their country's PCS group (white bars). The figure yields two main takeaways.[18] First, SLPCS groups exhibit their largest discrepancy from SL groups if the PCS group is less than 30% of the population. The difference for separatist civil wars (nearly eighteen-fold difference) is even larger than the discrepancy in Figure 5 between SLPCS and SL groups. There is also a large-magnitude difference when analyzing center-seeking civil wars: SLPCS groups in small PCS countries experienced new center-seeking civil wars 3.9 times more frequently than SL groups. This difference is consistent with Hypothesis 2a.

Second, the rate of civil war onset drops precipitously as the size of the PCS group increases, although the decrease is sharper for center-seeking civil wars. Separatist civil war onsets occur almost twice as frequently in small PCS countries than in medium-sized ones, and nearly six times more frequently than in large PCS countries. By comparison, center-seeking civil war onsets occurred 11 times more frequently in small PCS countries than in medium-sized ones, and no ethnic center-seeking civil wars occurred in large PCS countries. These patterns are consistent with Hypothesis 2b. Appendix Table C.10 demonstrates similar findings when using the same set of control variables from Table 1.

[18]Appendix Figure C.1 depicts similar patterns using a continuous measure of PCS group size.

Figure 6: Civil Wars by Stateless Groups: Aims and Size of PCS Group
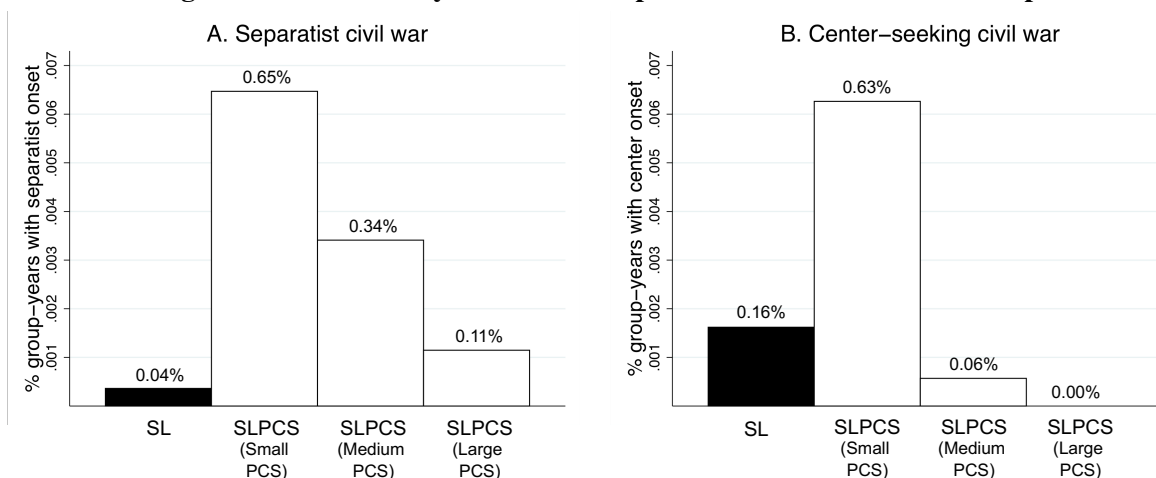
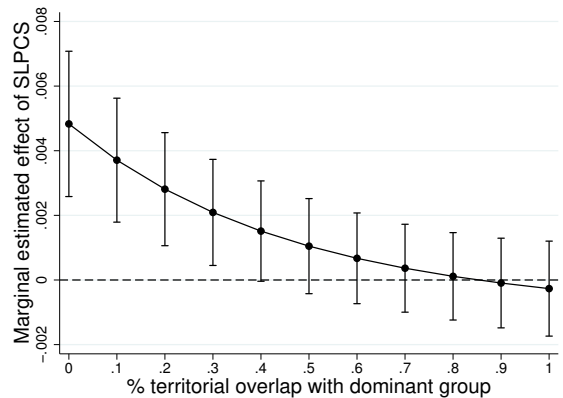## 4.3 PCS GROUPS AND THE GEOGRAPHY OF REBELLION

The theory posits that although large PCS groups provide motives to rebel, their large size eliminates the opportunity for aggrieved groups to rebel (Hypothesis 3). Figure 7 shows that among PCS countries, the number of ethnic groups that could feasibly attempt to separate decreases in the size of the PCS group. The sample for this figure includes six countries for which I code a PCS group but EPR codes ethnicity as not politically relevant: Lesotho, North Korea, Oman, South Korea, Swaziland, and Tunisia. I count an ethnic group as a separatist threat if it has a geographically concentrated settlement area and less than 50% of its location polygon overlaps with the location polygon with the ethnic group that controls the government in that year, which is usually the PCS group. By construction, this measure drops the group that controls the government from the sample. I measure both variables in the country's first year in the sample, which for most is the year of independence. Appendix Table C.12 shows that this correlation is statistically significant, even when accounting for outlying countries with numerous potential separatist groups.

Figure 7: PCS Group Size and Potential Separatist Challengers
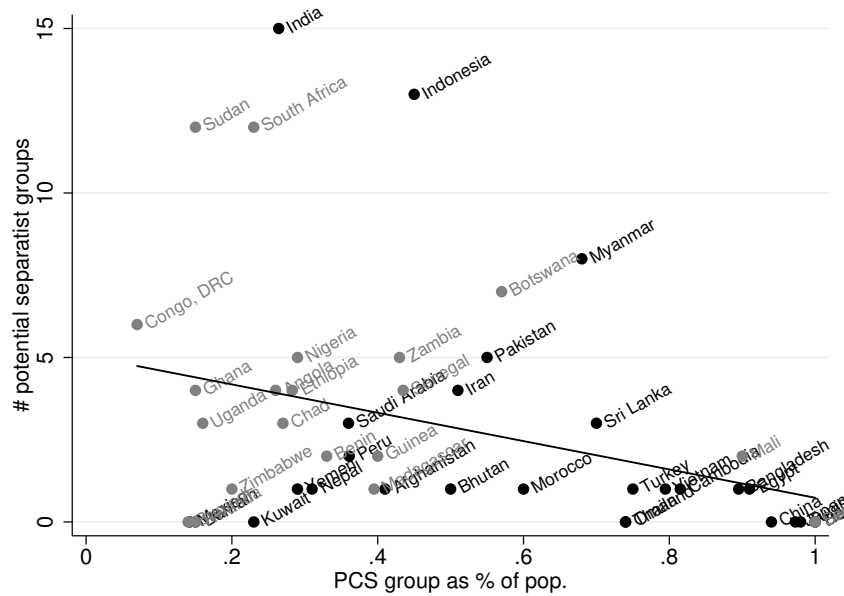
Appendix Table C.13 shows that higher territorial overlap with the dominant group mitigates the correlation between SLPCS groups and separatist civil wars. It runs the same specifications as in Table 1 except it controls for overlap with the dominant group and interacts this variable with the SLPCS and PCS indicators. Figure 8 plots the marginal estimated effect of SLPCS on separatist civil war onset for different values of dominant overlap with 95% confidence intervals, although the estimated

Figure 8: Margins Plot: SLPCS and Sep. CW

effects are similar across the different specifications. The estimated marginal effect decreases in percent territorial overlap with the dominant group. Among ethnic groups for which this overlap is less than 50%, the positive correlation between SLPCS groups and separatist civil war onsets is statistically significant, but it loses significance among groups with higher overlap. This provides evidence of a particular mechanism through which large PCS groups can deter separatism.

4.4 Ethnopolitical Exclusion and Rebellions by PCS Groups

The theory posits that PCS groups indirectly caused civil war by controlling the state and excluding SLPCS groups from power, leading the latter to rebel. However, members of PCS groups occasionally lost their grip on power and faced exclusion, such as the Baganda in Uganda following Milton Obote's coup and purge in 1966. Across the sample, PCS groups faced exclusion in 15% of group-years (specifically, years in which EPR codes the group as politically relevant but not monopoly, dominant, senior partner, or junior partner). Only relatively small PCS groups ever faced exclusion: 0% of group-years for PCS groups above the average size for all PCS groups, and 23% of group-years for smaller PCS groups.

Hypothesis 4 predicts that, conditional on ethnopolitical exclusion, PCS groups are more likely than SL groups to fight center-seeking civil wars, and SLPCS groups are more likely to separate than SL groups. Figure 9 supports these predictions. This figure is identical to Figure 5 except the sample contains only years in which an ethnic group is excluded from power. SLPCS groups are nearly five times more likely than SL groups to participate in a separatist civil war, and PCS groups are nearly six times more likely than SL groups to participate in a center-seeking civil war. These findings are consistent with the theoretical motivation that wherever SLPCS groups reside in a country with a reasonably large PCS group, it is not a viable option to take the center. However, PCS groups that are excluded from power do not face the same hurdle. The disruptive political conditions created by PCS groups make groups in those countries more likely to fight, and since center-seeking is a viable option for many excluded PCS groups, they choose this path. Appendix Table C.14 provides the accompanying regression table.

Figure 9: Conditioning on Ethnopolitical Exclusion
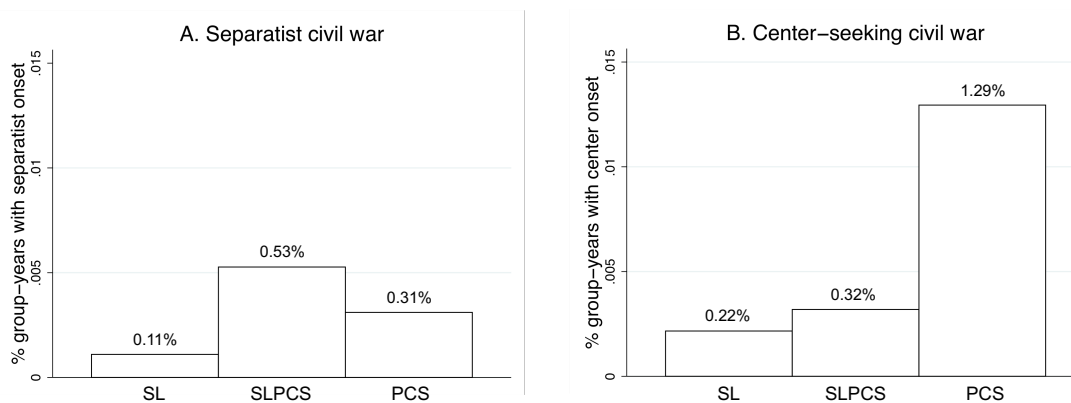
5   IMPLICATIONS FOR CROSS-REGIONAL DIFFERENCES

Studying the existence and size of historical states also helps to explain cross-regional patterns of ethnic civil war. Panel A of Figure 10 plots the frequency of major ethnic civil war onsets, disaggregated by civil war aims, between 1945 and 2013 among ethnic groups in four major world regions: Sub-Saharan Africa (SSA), Asia (excluding the Middle East), the Middle East and North Africa (MENA), and Latin America. The remaining panel plots different characteristics of PCS groups across regions: prevalence (Panel B); average percentage of the population (Panel C); and among groups in PCS countries, the percentage for which at least half of their territorial settlement overlaps with the dominant group in the country. The following relates these differences to the cross-regional civil war patterns.

Figure 10: Cross-Regional Patterns
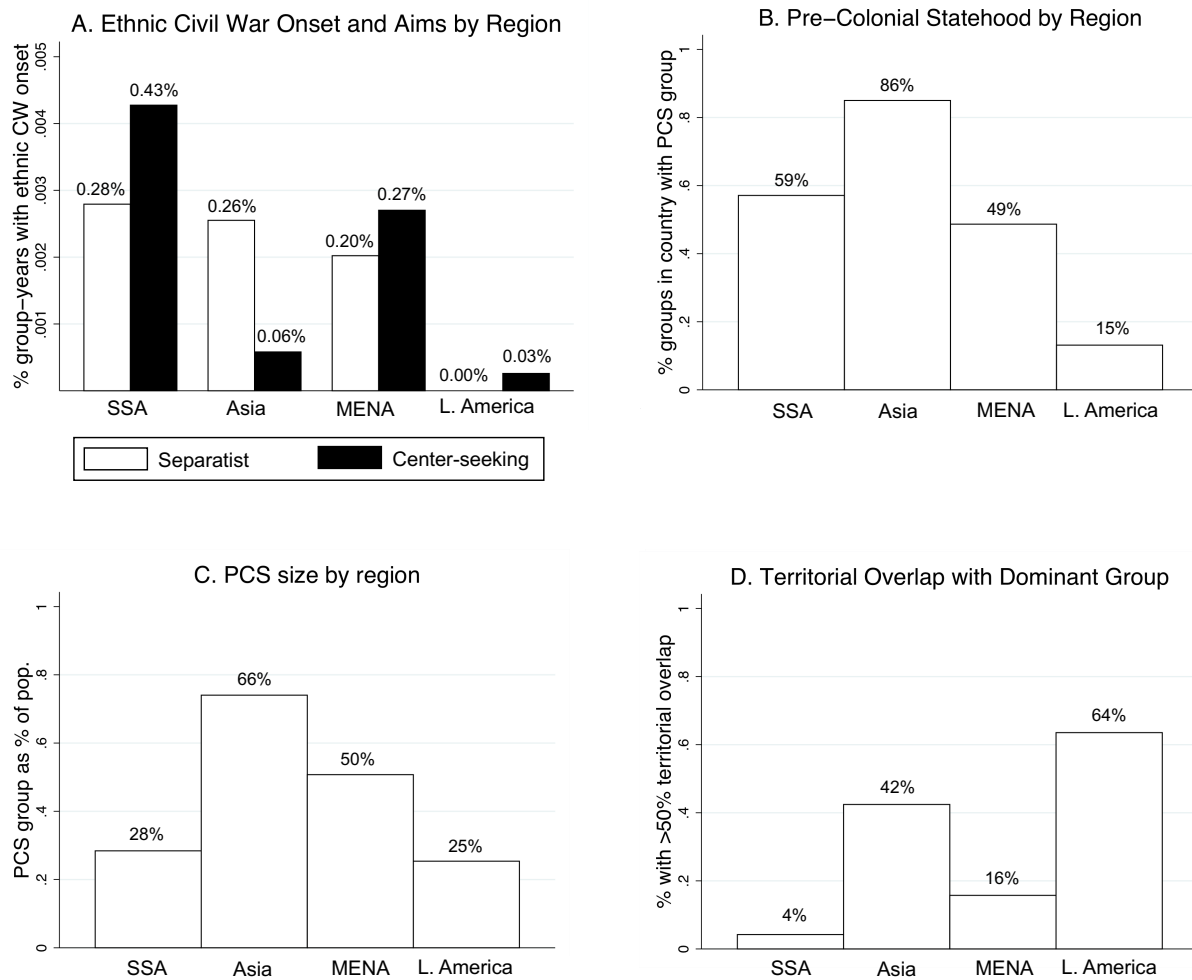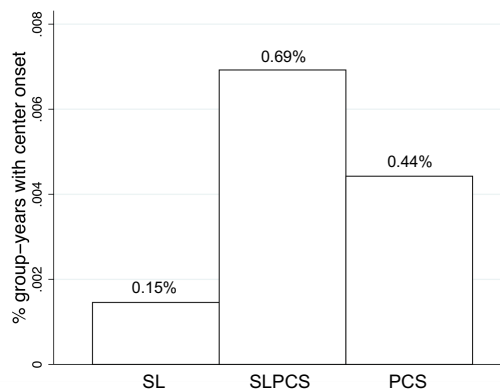
5.1 SUB-SAHARAN AFRICA

Many SSA ethnic groups reside in countries with a PCS group (59%) and PCS groups in SSA tend to be small (they average 28% of the population). The theory anticipates that the presence of PCS groups should trigger rebellions, and their relatively small size implies that center-seeking rebellions should often be viable (Hypothesis 2a). Consistent with this implication, center-seeking rebellions are more likely in SSA than in any other region, whether measured absolutely or as a percentage of all civil wars. Within-region evidence further supports Hypothesis 2a. Figure 11 replicates Panel B of Figure 5 while truncating the sample to SSA countries. SLPCS groups were 4.6 times more likely to initiate center-seeking civil wars than SL groups, and PCS groups were 2.9 times more likely. Appendix Table C.11 shows that the difference for SLPCS groups is statistically significant when controlling for the different covariates from Table 1.

SSA ethnic groups are also distinct for the very low percentage of groups (4%) for which at least half their territory overlaps with that of the dominant group in the country, compared to 42% of ethnic groups in Asia. In SSA, this corresponds with conditions under which either center-seeking or separatist rebellions are feasible. Tabulating conditions in individual cases suggests that if PCS groups in SSA were medium rather than large—which should deter center-seeking but not separatist attempts—

Figure 11: Center-Seeking CWs in SSA

then the percentage of separatist civil wars in SSA would be higher. Among the 31 center-seeking civil war onsets by SLPCS groups, in 26 cases (84%) the group faced favorable separatist conditions (geographically concentrated and less than 50% territorial overlap with the dominant group). Twenty-eight of these cases occurred in SSA, and in each case the country's largest ethnic group was below the population threshold used for medium sized PCS group (30% of the country's population). By contrast, in Asia, even groups that face favorable conditions for separating usually face impediments to taking the center because of the larger PCS group.

Finally, this analysis of civil war aims in Africa differs from Englebert's (2009) characterization of Africa's separatist deficit. In my ethnic-group level data with major civil wars, the *prevalence* of center-seeking

civil wars as opposed to the *rarity* of separatist conflict distinguishes the region—in fact, SSA ethnic groups are slightly more likely than Asian ethnic groups to participate in a separatist rebellion. Instead, Africa's separatist deficit only appears *relative to center-seeking wars*. Furthermore, observing that many African ethnic groups face favorable prospects for separatism yet have instead fought for the center suggests that explanations for this relative separatist gap should instead focus on civil war substitution, which has received minimal attention in existing research.

5.2 ASIA

Asian countries (excluding the Middle East) are distinct because a very high percentage of their wars are fought to secede rather than for the center. Although separatist civil wars occur in Asia roughly as frequently as in SSA, center-seeking conflicts are quite rare—only 19% of all ethnic wars compared to 55% in SSA. An even higher percentage of ethnic groups in Asia reside in a country with a PCS group (86%) compared to SSA, and the average size of Asian PCS groups is more than twice as large (66% of the population). Although PCS groups create grievances and motives for armed conflict, the large size of the ruling group implies that many stateless groups in Asia lack a viable opportunity to separate (Hypothesis 2b). This is particularly true when the PCS group resides throughout the country, leaving aggrieved ethnic minorities without a distinct territorial base from which to launch secession (Hypothesis 3). However, when conflicts occur, they should usually be separatist given the even greater extent to which large ruling groups deter center-seeking rebellions.

This explains why many Asian countries—in some cases, the country is named after the PCS ethnic group— that seemingly have the preconditions for separatism because they contain a PCS group instead have not faced secession attempts. In Japan, the Japanese compose almost the entire population (97%) and are located across the country's entire territory. Consequently, none of the small minority groups can viably secede.[19] Less trivially, in Cambodia, the PCS group Khmer composes 82% of the population and the country contains three territorially concentrated minority groups: Vietnamese (5% of the population), Cham and Malays (3.5%), and Thai-Lao (1%). However, although Khmers compose a lower share of the population

---

[19]Although it is theoretically trivial that separatism is unlikely if one group composes almost the entire population, note that cases with large PCS groups mitigate against finding the correlation shown in Table 1, which assesses the effect of historical statehood without taking into account the size of the group.

in Cambodia than Japanese in Japan, their residential territory includes the entirety of Cham/Malays' and Vietnamese's settlement areas, and the majority of Thai-Lao's (61%). Therefore, the large size of the PCS group makes separatism quite difficult.

By contrast, although in Myanmar the Burmese (PCS group) compose 68% of the population, their lack of territorial overlap with many minority groups in the country (see the outlying point for Myanmar in Figure 7) has caused the country to be wracked by separatist attempts since independence. Figure 12 plots the location polygon of the Burmese in black, and for eight minority groups with favorable separatist geography in gray. Three of these have initiated major separatist rebellions—Karens, Kachins, and Shan—and this number is higher when counting conflicts with less than 1,000 battle deaths. In this sense, despite the *numerical* preponderance

Figure 12: Burmese in Myanmar

of the Burmese, their relative lack of geographical dispersion makes Myanmar more similar to a typical SSA country in which minority groups almost always enjoy settlement areas unimpeded by the PCS group.

5.3 MIDDLE EAST AND NORTH AFRICA

In MENA, overall conflict propensity and center-seeking onset each lie in between the corresponding figures for SSA and Asia. PCS groups in MENA countries exhibit greater variance, ranging from quite small PCS groups in Bahrain and Yemen to hegemonic Arabs in Egypt. Some cases in the region are typical of the cross-regional patterns, such as Yemen (small PCS group, both center-seeking and separatist conflicts) and Iran, Morocco, and Turkey (medium PCS groups, one or multiple separatist wars). However, other cases flip a traditional question among scholars regarding why the region is so conflict-prone. Instead, countries like Kuwait, Bahrain, and Saudi Arabia—all with relatively small PCS groups—are outliers in a global comparison for their *lack* of conflict, and other factors such as extreme oil wealth may account for this discrepancy (Paine 2016).

5.4 LATIN AMERICA

The rarity of ethnic conflict in Latin America—and the complete absence of separatist wars—corresponds with the dearth of countries in the region with PCS groups (only Mexico and Peru). Studies of Latin American colonialism typically focus on Mexico and Peru, which Spain made the centers of its American empire so that Spanish settlers could exploit the native labor supply. However, despite the Spanish preference for densely settled regions with an organized labor force, precolonial states were exceptional rather than typical in the region.[20] This factor combined with natives' lack of immunity to European diseases to engender a demographic transformation during the colonial era in which natives died in genocidal proportions while millions of Europeans migrated overseas. Although microbes would seem to alone account for the entire outcome here, thinking about counterfactual scenarios regarding precolonial statehood suggests otherwise. Had historical states with densely settled populations pervaded the region, it is unlikely that European colonial rule would have so dramatically changed the demographic landscape. Scholars analyzing global samples consistently show that territories with higher historical population density and a longer history of statehood tended to experience less European migration and more indirect rule (Acemoglu, Johnson and Robinson 2002; Gerring et al. 2011; Hariri 2012), even if within Spanish America this correlation was reversed (Mahoney 2010). Furthermore, even within Latin America, there is evidence that precolonial institutions have persisted to affect modern-day socioeconomic outcomes, suggesting their relevance despite intensive European colonization (Angeles and Elizalde 2017).

The demographic shift in Latin America altered the region from one set of conditions that in the global sample correspond with a lack of ethnic wars—the near-absence of PCS groups—to another that also predicts a lack of ethnic conflict: the rise of a white/mestizo class that in most countries came to compose the overwhelming majority of the population and acted as a de facto hegemonic PCS group. The creation of a distinct class of non-indigenous peoples is unique in Latin America compared to the other regions studied here, although Latin America's lack of both historical states and separatism corresponds with the general pattern.[21] This argument departs from existing research that specifically examines Latin America's separatist deficit,

---

[20]The coding notes in Appendix A.1 shows that this is true even without trying to map historical states in the region to EPR ethnic groups.

[21]Also recall that the main results are robust to (1) controlling for region fixed effects, which drops every Latin American country and (2) controlling for settler colonialism.

such as Ross (2010) who analyzes the length of time since independence. Analyzing precolonial states and the colonial transformation can explain the absence of separatism over Latin America's entire postindependence period—even in the nineteenth century, the entire region experienced only four separatist civil wars and two of these occurred in Mexico, a PCS country[22]—as opposed to only the post-1945 period.

Overall, this paper explains why precolonial statehood (PCS) has triggered modern ethnic violence, focusing primarily on why stateless groups in countries with a PCS group (SLPCS groups) fight separatist civil wars at elevated rates. I also explain why small PCS groups facilitate center-seeking challenges whereas particularly large PCS groups deter any type of civil war. The theory, the data, and implications for explaining differences across regions offer various avenues for constructive future research.

REFERENCES

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117(4):1231–1294.

Alesina, Alberto, William Easterly and Janina Matuszeski. 2011. "Artificial States." *Journal of the European Economic Association* 9(2):246–277.

Angeles, Luis and Aldo Elizalde. 2017. "Pre-colonial Institutions and Socioeconomic Development: The Case of Latin America." *Journal of Development Economics* 124:22–40.

Besley, Timothy and Marta Reynal-Querol. 2014. "The Legacy of Historical Conflict: Evidence from Africa." *American Political Science Review* 108(2):319–336.

Bockstette, Valerie, Areendam Chanda and Louis Putterman. 2002. "States and Markets: The Advantage of an Early Start." *Journal of Economic Growth* 7(4):347–369.

Brecke, Peter. 1999. "Violent Conflicts 1400 AD to the Present in Different Regions of the World." 1999 Meeting of the Peace Science Society.

---

[22]Data from Wimmer and Min (2009). This count excludes wars of independence fought against Spain or Portugal, which mirrors the exclusion of decolonization wars in my main sample.

Buhaug, Halvard. 2010. "Dude, Where's My Conflict? LSG, Relative Strength, and the Location of Civil War." *Conflict Management and Peace Science* 27(2):107–128.

Callahan, Mary Patricia. 2005. *Making Enemies: War and State Building in Burma*. Cornell University Press.

Cederman, Lars-Erik, Andreas Wimmer and Brian Min. 2010. "Why Do Ethnic Groups Rebel?" *World Politics* 62(1):87–119.

Cederman, Lars-Erik, Kristian Skrede Gleditsch and Halvard Buhaug. 2013. *Inequality, Grievances, and Civil War*. Cambridge University Press.

Cederman, Lars-Erik, Simon Hug, Andreas Schädel and Julian Wucherpfennig. 2015. "Territorial Autonomy in the Shadow of Conflict: Too Little, Too Late?" *American Political Science Review* 109(2):354–370.

Decalo, Samuel. 1990. *Coups and Army Rule in Africa*. Second ed. Yale University Press.

Depetris-Chauvin, Emilio. 2015. "State History and Contemporary Conflict: Evidence from Sub-Saharan Africa.". Working paper.

Dincecco, Mark, James E. Fenske and Massimiliano Gaetano Onorato. 2016. "Is Africa Different? Historical Conflict and State Development.". Working paper.

Dinwiddy, Hugh. 1981. "The Search for Unity in Uganda: Early Days to 1966." *African Affairs* pp. 501–518.

Doornbos, Martin R. 1977. Ankole. In *African Kingships in Perspective*, ed. Rene Lemarchand. Frank Cass & Co. Ltd. pp. 236–259.

Encyclopædia Britannica. 2019. "Encyclopædia Britannica Online.".

Englebert, Pierre. 2009. *Africa: Unity, Sovereignty, and Sorrow*. Lynne Rienner Publishers Boulder, CO.

Englebert, Pierre, Stacy Tarango and Matthew Carter. 2002. "Dismemberment and Suffocation." *Comparative Political Studies* 35:1093–1118.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–90.

Fearon, James D. and David D. Laitin. 2011. "Sons of the Soil, Migrants, and Civil War." *World Development* 39(2):199–211.

Fearon, James D. and David D. Laitin. 2014. "Does Contemporary Armed Conflict Have 'Deep Historical Roots'?". Prepared for the 2014 Annual Meeting of the American Political Science Association, August 28-31. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1922249. Accessed 6/6/17.

Francois, Patrick, Ilia Rainer and Francesco Trebbi. 2015. "How is Power Shared in Africa?" *Econometrica* 83(2):465–503.

Gerring, John, Daniel Ziblatt, Johan Van Gorp and Julian Arevalo. 2011. "An Institutional Theory of Direct and Indirect Rule." *World Politics* 63(3):377–433.

Green, Elliott. 2012. "On the Size and Shape of African States." *International Studies Quarterly* 56(2):229–244.

Hariri, Jacob Gerner. 2012. "The Autocratic Legacy of Early Statehood." *American Political Science Review* 106(3):471–494.

Katouzian, Homa. 1998. The Pahlavi Regime in Iran. In *Sultanistic Regimes*. Johns Hopkins University Press pp. 182–205.

Lacina, Bethany. 2015. "Periphery Versus Periphery: The Stakes of Separatist War." *Journal of Politics* 77(3):692–706.

Magaloni, Beatriz. 2008. "Credible Power-Sharing and the Longevity of Authoritarian Rule." *Comparative Political Studies* 41(4-5):715–741.

Mahoney, James. 2010. *Colonialism and Postcolonial Development: Spanish America in Comparative Perspective*. Cambridge University Press.

Michalopoulos, Stelios. 2012. "The Origins of Ethnolinguistic Diversity." *American Economic Review* 102(4):1508–39.

Michalopoulos, Stelios and Elias Papaioannou. 2016. "The Long-Run Effects of the Scramble for Africa." *American Economic Review* 106(7):1802–1848.

Murdock, George P. and Douglas R. White. 1969. "Standard Cross-Cultural Sample." *Ethnology* 8(4):329–369.

Murdock, George Peter. 1967. *Ethnographic Atlas*. University of Pittsburgh Press.

Paine, Jack. 2016. "Rethinking the Conflict "Resource Curse": How Oil Wealth Prevents Center-Seeking Civil Wars." *International Organization* 70(4):727–761.

Paine, Jack. 2019. "Ethnic Violence in Africa: Destructive Legacies of Pre-Colonial States." *International Organization* 73(3):645–683.

Putterman, L. 2012. "State Antiquity Index, Version 3.1." *Brown University* . Available at http://www.brown.edu. Accessed 5/16/16.

Putterman, Louis. 2008. "Agriculture, Diffusion and Development." *Economica* 75(300):729–748.

Ray, Subhasish. 2019. "History and Ethnic Conflict: Does Precolonial Centralization Matter?" *International Studies Quarterly* 63(2):417–431.

Reid, Richard J. 2012. *Warfare in African History*. Cambridge University Press.

Roessler, Philip. 2016. *Ethnic Politics and State Power in Africa*. Cambridge University Press.

Roessler, Philip and David Ohls. 2018. "Self-Enforcing Power Sharing in Weak States." *International Organization* 72(2):423–454.

Ross, Michael. 2010. "Latin America's Missing Oil Wars." Working paper, UCLA.

Rothchild, Donald and Michael Rogin. 1966. Uganda. In *National Unity and Regionalism in Eight African States*, ed. Gwendolen M. Carter. Cornell University Press pp. 337–440.

Savary, Alan. 1952. "The French Union: Centralism or Federalism?" *International Journal* 7(4):258–264.

Schmidt, Siegmar. 1999. Uganda. In *Elections in Africa*, ed. Dieter Nohlen, Bernard Thibaut and Michael Krennerich. Oxford University Press.

Scott, James C. 2010. *The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia*. New Haven, CT: Yale University Press.

Smith, Benjamin. 2013. "Separatist Conflict in the Former Soviet Union and Beyond: How Different was Communism?" *World Politics* 65(2):350–381.

Svolik, Milan W. 2012. *The Politics of Authoritarian Rule*. Cambridge University Press.

Tilly, Charles. 1992. *Coercion, Capital, and European States, AD 990-1992*. Blackwell Oxford.

Toft, Monica Duffy. 2005. *The Geography of Ethnic Violence*. Princeton University Press.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." *Journal of Conflict Resolution* 59(7):1327–1342.

Walter, Barbara F. 2009. *Reputation and Civil War*. Cambridge University Press.

Whatley, Warren and Rob Gillezeau. 2011. "The Impact of the Transatlantic Slave Trade on Ethnic Stratification in Africa." *American Economic Review* 101(3):571–76.

Wig, Tore. 2016. "Peace from the Past: Pre-Colonial Political Institutions and Civil Wars in Africa." *Journal of Peace Research* 53(4):509–524.

Wimmer, Andreas and Brian Min. 2009. "The Location and Purpose of Wars Around the World: A New Global Dataset, 1816–2001." *International Interactions* 35(4):390–417.

Wimmer, Andreas, Lars-Erik Cederman and Brian Min. 2009. "Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Data Set." *American Sociological Review* 74(2):316–337.

Wucherpfennig, Julian, Philipp Hunziker and Lars-Erik Cederman. 2016. "Who Inherits the State? Colonial Rule and Post-Colonial Conflict." *American Journal of Political Science* 60(4):882–898.

Online Appendix

CONTENTS

LIST OF FIGURES

LIST OF TABLES

A.1   CODING PRE-COLONIAL STATES

Throughout, I refer to pages from Encyclopædia Britannica (2019) as "EB [page title]."

A.1   SUB-SAHARAN AFRICA

Paine (2019*b*) provides extensive coding notes for most Sub-Saharan African countries. I use his coding for all cases except for groups that he codes as PCS but are not the largest PCS group in the country. That source, however, does not have data for the following African countries, although some of these are not in my core sample because EPR codes ethnicity as politically irrelevant in the country.

Table A.1: List of PCS Groups in SSA

Country	EPR ethnic group	Historical state(s)
Angola	Mbundu-Mestico	Kasanje/Matamba
Benin	South/Central (Fon)	Dahomey
Botswana	Tswana	Tswana chiefdoms
Burundi	Tutsi	Burundi
Chad	Muslim Sahel groups	Ouaddai
DRC	Luba Kasai	Luba
Ethiopia	Amhara	Ethiopia
Ghana	Asante (Akan)	Asante
Guinea	Peul	Futa Jalon
Lesotho*	Sotho	Sotho
Madagascar	Highlanders	Merina
Mali	Blacks (Mande, Peul, Voltaic etc.)	Tukulor
Nigeria	Hausa-Fulani and Muslim Middle Belt	Sokoto
Rwanda	Tutsi	Rwanda
Senegal	Wolof	Walo/Kajor/Bawol/Jolof
South Africa	Zulu	Zulu
Sudan	Shaygiyya, Ja'aliyyin and Danagla	Mahdist
Swaziland*	Swazi	Dlamini
Uganda	Baganda	Buganda
Zambia	Bemba speakers	Kazembe/Bemba
Zimbabwe	Ndebele-Kalanga-(Tonga)	Ndebele

*EPR codes ethnicity as not politically relevant in the country.

A.1   Burkina Faso

EB mentions Mossi and Gurma kingdoms (EB Burkina Faso). According to many ethnicity datasets (e.g., Alesina et al. 2003, Fearon 2003), Mossi compose half the population of Burkina Faso. However, the Mossi were fractured among many precolonial states. EB Mossi states describes: "Mossi states, complex of independent West African kingdoms (fl. c. 1500-1895) around the headwaters of the Volta River (within the modern republics of Burkina Faso [Upper Volta] and Ghana) including in the south Mamprusi, Dagomba, and Nanumba, and in the north Tenkodogo, Wagadugu (Ouagadougou), Yatenga, and Fada-n-Gurma (Fada Ngourma)." McFarland and Rupley (1998, 86-7) also refer to distinct groupings of southern and northern Mossi kingdoms. Englebert (1996, 11) states that "Ouagadougou should not be regarded as the capital of an alleged Mossi 'empire,' as there was considerable autonomy, and even infighting, among the different kingdoms and principalities." He mentions four main kingdoms and 19 additional autonomous minor kingdoms (12).

A.2 Lesotho

EPR lists one ethnic group, Sotho, but does not provide population estimates because ethnicity is coded as politically irrelevant in Lesotho. The largest ethnic group according to standard ethnicity datasets (e.g., Alesina et al. 2003, Fearon 2003), **Sotho**, were organized under a single state in 1824, the Sotho kingdom. British colonial rule began in 1868 and the monarchy lasted through the colonial era.

A.3 Mauritius

Uninhabited prior to European colonial rule.

A.4 Somalia

The only EPR ethnic group, Somali, is fractured into numerous tribes with different political histories. Scarritt and Mozaffar's (1999) coding of ethnic groups in African countries, for example, list six main tribes plus additional sub-tribe divisions. Fearon (2003, 197-8) discusses the difficulty of deciding whether Somalia has a single ethnic group, Somalis, or numerous ethnic groups corresponding to the major clans. Scarritt and Mozaffar's (1999) and Fearon (2003) use clans, whereas EPR uses the broader grouping.

Historically, most Somalis have been pastoral nomads (EB Somalia) and have lacked unified political organizations. EB Somalia and Mukhtar (2003, 3-4) discuss many Somali local sultanates (some of which correspond to the different tribes) that did not exhibit unified political organization among the EPR ethnic group.

A.5 Swaziland

EPR lists one ethnic group, Swazi, but does not provide population estimates because it codes ethnicity as politically irrelevant in Swaziland. The largest ethnic group according to Fearon (2003), **Swazi**, were organized under a single state in 1770, Swaziland (also known as the Dlamini kingdom). Colonial rule began in 1890 and the monarchy has lasted through the colonial era and into the present.

A.6 Tanzania

Modern Tanzania consists of two former colonies: Tanganyika, the mainland, and Zanzibar, an island. Neither EB Tanzania nor Ofcansky and Yeager's (1997, 1-8) historical overview mentions any indigenous states in mainland Tanzania. Iliffe's (1979) more detailed discussion of political organization in precolonial Tanganyika asserts: "The political systems existing in Tanganyika in 1800 ranged from complete statelessness to chiefdoms administered by appointed officers" (21). He disaggregates Tanganyika into three regions. "Statelessness was most common in the sparsely-populated, tsetse-fly infested south-east" (21). In the northeast, he mentions the small Shambaa kingdom (capital contained 3,000 people) (22). The west tended to feature more state-like forms of political organization because it bordered major centers of African political development, such as the Luba (in modern DRC) and Lake Victoria (in modern Uganda). He mentions "several small kingdoms," including Haya states, Karagwe, Zinza, and Ukerewe Island. But EPR's ethnic group "mainland Africans" (96% of Tanzania's population) was clearly not unified into a single or a small number of states.

According to EB Tanzania and the relevant entries from Ofcansky and Yeager ([1997](#)), the island of Zanzibar has a longer history of statehood due to its incorporation into the Omani empire, and in fact was the capital of Oman from 1832 to 1861. The Omani ruler who moved the capital to Zanzibar, Said Ibn Seyyid, died in 1856. One son governed Oman and the other governed Zanzibar, and through British assistance, Zanzibar became independent in 1861 after threats of re-colonization by Oman. Zanzibar retained this status until becoming a British protectorate in 1890, despite losing almost all of its claimed mainland territory during this period. Therefore, local control over Zanzibar was achieved only in the context of European interference during a period of increasing European control over the area, and therefore I do not code the corresponding EPR ethnic group, Zanzibar Arabs (0.5% of Tanzania's population), as PCS. Zanzibar's precolonial history contrasts with many cases in which a precolonial state survived, precariously, despite concerted European attempts to undermine the state before the territory officially became a colony.

A.2 ASIA

Table A.2: List of PCS Groups in Asia

Country	EPR ethnic group	Historical state(s)
Afghanistan	Pashtuns	Durrani, Barakzay
Bangladesh	Bengali Muslims	Bengali province of Mughal empire
Bhutan	Bhutanese	Bhutan
Cambodia	Khmer	Khmer
China	Chinese (Han)	Qing
India	Hindi (Non SC/ST/OBCs)	Maratha
Indonesia	Javanese	Mataram
Japan	Japanese	Tokugawa, Meiji
Myanmar	Bamar (Burman)	Alaungpayo
Nepal	Caste Hill Hindu Elite	Nepal
North Korea*	Koreans	Choson
Pakistan (pre-1971)	Bengali	Bengali province of Mughal empire
South Korea*	Koreans	Choson
Sri Lanka	Sinhalese	Kotte, Sitawaka, Kandyan
Thailand	Thai	Rattanakosin
Vietnam	Kinh (Vietnamese)	Nguyen

*EPR codes ethnicity as not politically relevant in the country.

A.1 Afghanistan

The largest ethnic group, **Pashtuns**, were organized into various large states for hundreds of years prior to the 20th century (Afghanistan was never formally colonized by a European power), including the Durrani dynasty until 1826 and the Barakzay dynasty until 1973 (EB Afghanistan).

A.2 Bangladesh

The entry for Pakistan explains why **Bengali Muslims** are coded as PCS.

A.3   Bhutan

The largest ethnic group, **Bhutanese**, were organized under a monarchy at least as far back as 1616. The first theocratic king "united the leaders of powerful Bhutanese families in a land called Drukyul. He promulgated a code of law and built a network of impregnable dzong, a system that helped bring local lords under centralized control and strengthened the country against Tibetan invasions. Many dzong were extant in the late twentieth century" (Worden 1991, 256). Although the kingdom suffered several invasions from Tibet and long periods of civil war in the 18th century, the kingdom has remained intact through the present (British influence became strong in the 19th century and Bhutan agreed to a treaty with Britain to relinquish its external affairs in 1910).

A.4   Cambodia

Cambodia's first large unified state existed between 802 and 1432, when Thais sacked the capital city of Angkor. Within the next hundred years, Cambodian kings had re-established their court in Phnom Penh and, briefly, Lovek (Corfield and Summers 2003, xxxiii). Into the 19th century, Cambodia was considerably weaker than in its previous period and frequently sought protection from their stronger neighbors, Siam/Thailand and Vietnam (EB Cambodia). Vietnam occupied most of the Cambodian kingdom between 1835 and 1841, after which "Cambodians [were] able to exercise a small degree of independence" (EB Cambodia). This political status persisted until the Cambodian king requested and became a French protectorate in 1863.

This is a borderline case because of Cambodia's varying and unclear sovereign status relative to neighboring kingdoms. However, because the **Khmer** exhibited unified rule (as opposed to fractured rule in Laos) with some degree of independence at different times within reasonable temporal proximity to its European colonization, it appears appropriate to code Khmers as having a precolonial state.

A.5   China

The largest ethnic group, **Chinese (Han)**, were organized into various large states for millennia prior to the 20th century (most of China was never colonized by a European power), including the Qing dynasty from 1644 to 1912 (EB China).

A.6   Fiji

The only indigenous EPR ethnic group in the country, Fijians, were politically fractured for almost their entire history prior to British colonization in 1874, and the brief political unity achieved just prior to colonization was European-sponsored. Historically, every Fijian belonged to a collection of ranked family groups, and these groups occasionally united to form larger confederacies, called *matanitu*. "At the beginning of the 19th century, there were about a dozen *matanitu* in Fiji . . . The struggle for political supremacy among the leading confederacies was the dominant feature of early 19th-century Fijian politics, with Bau family finally becoming ascendant under Ratu Seru Cakobau toward the middle of the century" (Lal 2015, 4). However, Cakobau's ascendancy and the creation of a King of Fiji, which occurred several years prior to colonization, was facilitated by aspiring European colonizers: "In 1867, an allegedly pan-Fiji government with Cakobau as the head was established with the backing of some politically ambitious Europeans in the east" (Parke 2014, 1). Similarly, EB Fiji refers to the 1860s and early 1870s as a period in which

European influence began to structure Fijian politics: "Disputes ensued over land and political power within and between European and Fijian communities, and problems arose with labourers introduced from other Pacific islands. Those factors contributed to violent confrontations, exacerbated the implicit instability of Fijian society, and ensured that no Fijian chief could impose his rule on the whole group. European attempts at government were doomed by the greed and factionalism of their members and by the interference of European governments and consuls."

A.7   India

The Indian subcontinent has a long history of statehood. The largest ethnic group, **Hindi (Non SC/ST OBCs)**, was most recently organized under the Maratha empire prior to colonization, rising during the break-up of the Mughal empire in the first half of the 18th century and ending in 1818 after their third war with the British East India company, after which the company had achieved territorial control over most of modern India. Maratha controlled most of northern India, which is where the EPR Hindi polygon is located, and the Maratha are "famed in history as yeoman warriors and champions of Hinduism" (EB Maratha). The empire exhibited many characteristics of states, for example, "they seem to have consolidated methods of assessment and collection of land revenue and other taxes" (EB India). However, despite providing a unifying identity for Hindus in northern India in which Hindus "fought under the flag of religion for the defense of their country against Muslim rulers [NB: Mughal was a Muslim empire]" (Mansingh 1996, 250), for most of its history, Maratha was a decentralized confederacy composed of five chiefdoms nominally organized under a single *peshwa* (prime minister), rather than a tightly centralized state. In this sense, it resembles states such as the Sokoto caliphate in Nigeria and the Tswana chiefdoms in Botswana (both of which are coded as PCS), and perhaps also the Tutsi kingdom in Burundi (also coded as PCS) in which there was a nominal king but considerable in-fighting among local officials.

A.8   Indonesia

The largest ethnic group, **Javanese**, were organized into various large states for approximately 1,000 years prior to Dutch colonization, including the Mataram kingdom from the late 16th century until displaced by Dutch colonizers in the 18th century (EB Indonesia).

A.9   Japan

The largest ethnic group, **Japanese**, were organized into various large states for over one millennium prior to the 20th century, including the Tokugawa Shogunate between 1603 and 1868 and the Meiji Restoration period until 1947 (EB Japan).

A.10   Laos

This is a borderline case that, due to fractures within the ethnic groups and their tributary status for more than a century before European colonization, is coded as having no states. The EPR group "Lao (incl. Phuan)" does have a history of statehood. The first "extensive Lao kingdom" (Stuart-Fox 2001, 5) arose in 1353 and lasted until 1713 when it split into three rival kingdoms: Champassak, Vien Chan/Vientiane, and Luang Prabang (EB Laos). By 1778, all three had become Siam/Thai tributaries and Siam decided who took the throne. The Vien Chan dynasty ended in 1828 after a failed attack on Siam, and was subsequently

made into a Siamese province. The other two dynasties remained intact but remained under Siamese control, and Siam garrisoned troops in Champassak starting in 1846 and in Luang Prabang starting in 1886. This political situation lasted until France gained control of modern-day Laos through treaties with Siam in 1904 and 1907.

A.11   Malaysia

The majority group Malays composed a large state, the Malacca kingdom, until 1511 when it was defeated by Portugal. However, colonial control over a significant portion of modern-day Malaysia did not occur until 1874 under expanded British penetration, the year that Ertan, Fiszbein and Putterman (2016) use to code the onset of European colonial rule. Therefore, it is relevant that in the centuries between these dates, Malay fractured into numerous states, and "no Malay state rose to match the former commercial strength and cultural splendor of Melaka ... Furthermore, neither the Europeans nor the indigenous states could achieve sustained preeminence over the other" (Kaur 2001, 8). The most prominent successor to Malacca, Johor, had itself broken up into multiple states by the end of the 17th century. British Malaya was a federated collection of nine separate Malay states. Given the numerous Malay states and the lack of even nominal subordination to a single ruler, this ethnic group is too fractured to be coded as composing a precolonial state.

A.12   Myanmar

The largest ethnic group, **Bamar (Barman)**, were organized into various large states for over 1,000 years prior to British colonization, including the Alaungpayo dynasty from 1752 to 1885 (EB Myanmar).

A.13   Nepal

The largest ethnic group, **Caste Hill Hindu Elite**, was organized into various states for over a millennium prior to the 20th century (Nepal was never formally colonized by a European power), including the Kingdom of Nepal/Gorkha from 1679 to 2008 (EB Nepal).

A.14   North Korea

The only EPR ethnic group, **Koreans**, was organized as a large state for a millennium before Japanese annexation in the 20th century, including the Choson dynasty from 1392 to 1910 (EB Korea).

A.15   Pakistan

The largest ethnic group at independence, **Bengali**, organized an autonomous state in the 50 years between the decline of the Mughal empire in Bengal in 1707 and the onset of British colonial rule in 1757 (Baxter and Rahman 1996, 8). Although still officially part of the Mughal Empire, during this period: "The links binding the three provinces to the imperial centre, itself seriously weakened, had become very tenuous. An independent state was in the making ... By 1740 the three provinces had acquired an administrative system that was almost entirely separate from that of the rest of the empire. The Nawabs [provincial governors] had also developed their own military forces" (Marshall 1987, 48, 50-1).

Bengalis successfully seceded in 1971 to create the independent state of Bangladesh. This left the Punjabi as the largest ethnic group in Pakistan. A large, locally ruled state existed in the Punjab (capital was Lahore) between 1799 and 1849, when British colonialism began in current-day Pakistan (known as West Pakistan prior to Bangladesh seceding). However, Pakistan's largely Muslim Punjabi population did not govern the empire, which was instead ruled by Sikhs. Most of the Sikh population was incorporated into India during India and Pakistan's partition in 1947 (Burki 1991, 170).

A.16 Papua New Guinea

EB Papua New Guinea does not mention any precolonial states. Turner (1994, 2) asserts for the precolonial period that "small-scale societies developed largely in isolation."

A.17 Philippines

Despite some evidence of statehood for the EPR ethnic group Moro, who are Muslims in southern Philippines, I did not code any PCS groups in the Philippines. EB Philippines asserts: "The Philippines is the only country in Southeast Asia that was subjected to Western colonization before it had the opportunity to develop either a centralized government ruling over a large territory or a dominant culture." However, it also mentions that the spread of Islam to the southern Philippines in the 15th century led to new forms of political organization, and Guillermo (2011) mentions the Sulu among other Muslim sultanates. Although these sultanates exhibited hierarchical organization, none seems to have encompassed the majority of Philippines' Moro population. "The Sulu Sultanate was a centralized political system whose territorial sovereignty was recognized most strongly at the centre, in the environs of Jolo town, and shaded off into ritual hegemony in distant areas" (Warren 2007, xliv). However, Jolo is located on a separate island south of Mindanao, the largest island with a large Moro population. Other sultanates included Maguindanao and Buayan (Majul 1999, 29), which were located on the island of Mindanao, the largest island containing Moros. The sultanates fought Spain from the 16th through 19th centuries, and collectively prevented Spanish colonization of the southern Philippines until the 19th century (Majul 1999). However, given the lack of available sources about these states (Majul 1999, 1), there is not sufficient evidence to suggest that any of the states ruled a large enough percentage of Moros to code the ethnic group as a precolonial state. [NB: Olsson (2009) codes the Philippines' first year of colonial rule as 1565 because the the first Spanish settlement occurred then. Ertan, Fiszbein and Putterman (2016) code colonial onset as 1600 because around this time Spain met their definition of controlling at least 20% of the territory of the modern country. However, these dates are largely irrelevant for coding precolonial states among Moros because Spain did not colonize their territory until the 19th century.]

A.18 Singapore

EB Singapore mentions sporadic raids of the island by foreign states but no states before British colonial rule began in the 19th century. Corfield (2010, 4-5) reiterates the absence of indigenous centralized states.

A.19 South Korea

The only EPR ethnic group, **Koreans**, was organized as a large state for a millennium before Japanese annexation in the 20th century, including the Choson dynasty from 1392 to 1910 (EB Korea).

A.20 Sri Lanka

This case is difficult to code because European interference with local states began centuries before more permanent colonial rule began, similar to colonies such as Angola. However, Sri Lanka's largest EPR ethnic group, **Sinhalese**, appears to meet the definition of a state.

Portugal first arrived on the island in 1505 and fought wars with native Sinhalese kingdoms. Olsson (2009) codes Sri Lanka's colonization year as 1619, when Portugal annexed a Tamil kingdom in the north, Jaffna. Ertan, Fiszbein and Putterman (2016) use a more stringent definition of colonial onset, requiring the European power to control at least 20% of the modern country's territory, which the Netherlands met in 1658 when they gained more effective control over the country's interior than had the Portuguese. Therefore, the relevant benchmark for "the eve of colonization" is sometime during the 16th or 17th centuries.

According to EB Sri Lanka, Peebles (2015, 2-5), and Blood (1990, 3-27), the Sinhalese have a history of kingdoms dating back to before 0 CE. The predominant Sinhalese kingdom on the eve of Portugal's arrival was Kotte, which Portugal weakened during the 16th century through warfare and trade. Two rival Sinhalese kingdoms arose during this century: Sitawaka and Kandyan. The latter remained independent through the 17th and 18th centuries—when the Netherlands colonized the rest of the island and Kandy was the only remaining indigenous state—until 1815 when it signed a treaty that granted sovereignty to Britain.

Although the Sinhalese were not united under a single state on the eve of colonization, there is clear evidence that a small number of Sinhalese states were vying for control over the island on the eve of colonization. Furthermore, the Sinhalese had a long history of statehood, and fractionalization among the Sinhalese in the 16th and 17th centuries existed to some extent because of colonial penetration—and therefore this fractionalization is not precolonial.

A.21 Taiwan

Neither EB Taiwan nor Copper's (2000, 6-28) historical overview mention any indigenous states prior to the island's alternation between European and Chinese colonial rule during the 17th through 20th centuries.

A.22 Thailand

The largest ethnic group, **Thai**, were organized into various large states for hundreds of years prior to the 20th century (Thailand was never formally colonized by a European power), including the Rattanakosin Kingdom from 1782 until the reform of absolute monarchy in 1932 (EB Thailand).

A.23 Timor Leste

Neither EB East Timor nor Gunn (2010, 6-9) mention any precolonial states. The island attracted attention from European powers vying to monopolize its sandalwood trade.

A.24 Vietnam

The largest ethnic group, **Kinh (Vietnamese)**, were organized into various large states for hundreds of years prior to French colonization in the 19th century, including the Nguyen dynasty from 1802 to 1945 (the

monarchy remained in place throughout the colonial period) (EB Vietnam).

A.3 Middle East and North Africa

Table A.3: List of PCS Groups in MENA

Country	EPR ethnic group	Historical state(s)
Algeria	Arabs	Algerian province of Ottoman empire
Bahrain	Sunni Arabs	al-Khalifa
Egypt	Arab Muslims	Muhammed Ali
Iran	Persians	Qajar, Pahlavi
Kuwait	Kuwaiti Sunni (Arab)	al-Sabah
Morocco	Arabs	Alawite
Oman*	Ibadhi Muslims (Arab)	al-Said
Saudi Arabia	Sunni Wahhabi (Najdi) (Arab)	al-Saud
Tunisia*	Arabs	Husaynid (Tunis province of Ottoman empire)
Turkey	Turkish	Ottoman
Yemen	Zaydis	al-Mutawakkil

*EPR codes ethnicity as not politically relevant in the country.

A.1 Algeria

Algeria was nominally an Ottoman colony between 1518 and 1830, when France colonized Algeria. However, a military revolt by Ottoman officers in 1689 created a state that sources describe as governing Algeria "independently from the Ottoman government" (EB North Africa) and as "in practice an autonomous state" (Naylor 2015, 12). This state was both local—"the Ottomans who settled there identified themselves as 'Algerian'" (Naylor 2015, 12)—and covered the northern part of the territory occupied by the largest EPR ethnic group in Algeria, **Arabs** (Oliver and Atmore 2005, 7-9), which is where the overwhelming majority of Arabs in Algeria live; although EPR's Arab polygon stretches into the Sahara desert, few people live there.

A.2 Bahrain

The minority ethnic group **Sunni Arabs** has ruled the island since the al-Khalifa dynasty began in 1783, including during and after an intervening period of British rule in the 20th century (EB Bahrain).

A.3 Egypt

The largest ethnic group, **Arab Muslims**, had been organized into various large states for over a millennium prior to the late 19th century—when Britain colonized Egypt—albeit with extensive stretches of foreign rule as well. In 1805, despite nominal Ottoman rule, viceroyal Muhammed Ali created an expansive state throughout Egypt, in part by overhauling the tax system. His successors governed the country until British colonization in 1882 (EB Egypt).

A.4 Iran

The largest ethnic group, **Persians**, had been organized into various large states for over a millennium prior to the 20th century (Iran was never formally colonized by a European power), including the Safavid empire from 1501 to 1736, the Qajar dynasty from 1796 to 1925, and the Pahlavi dynasty from 1925 to 1979 (EB Iran).

A.5 Iraq

According to EB Iraq and Ghareeb and Dougherty (2004), following earlier caliphates based in Baghdad, the Ottoman Empire nominally ruled Iraq as three separate vilayets/provices between 1534 and 1918, when British colonial rule began. The Ottoman period featured considerable variation in the extent of foreign rather than local governance in Iraq. Between 1704 and 1831, Ottoman mamlukes ruled a de jure independent regime from Baghdad. However, despite establishing Baghdad-based rule, the mamluk rulers were mostly Georgian-based rather than local Iraqis. Between 1831 and 1918, the Ottomans reasserted more direct rule over Iraq. This period, rather than earlier periods of statehood, shaped the post-independence Iraqi elite: "As secular reforms were implemented and the role of the state expanded in the 19th century, Iraqi religious notables and officeholders—both Shiite and Sunnite—suffered a relative loss of status, influence, and wealth. Meanwhile, Ottoman civil administrators and army officers, virtually all of whom were Sunnites, came to constitute a political elite that carried over into post-1918 Iraq" (EB Iraq). Therefore, there is no period within 500 years of colonial rule in which either Iraqi Sunni, Iraqi Shi'a, or Iraqi Kurds (the three EPR groups) governed their own state.

A.6 Israel

The largest EPR groups at independence are two Jewish groups, composed mainly of recent migrants to Israel. The Ottoman empire ruled native Palestinian Arabs between 1516 and 1918, and other foreign powers ruled in the preceding centuries. Unlike some Ottoman territories, there is no evidence of an autonomous local regime during the Ottoman period (EB Israel), and Nazzal and Nazzal's (1997, 1-3) historical overview only discusses rule by foreign powers.

A.7 Jordan

"Jordan as a distinct geographical or political entity did not exist throughout most of recorded history . . . [it] came into existence as a British mandate following the breakup of the Ottoman Empire at the end of World War I" (Gubser 1991, 2). During Ottoman rule, the modern-day territory of Jordan belonged to the Damascus vilayet. During the second millennium CE, it alternated between foreign imperial rule and local nomadic organization (EB Jordan; Gubser 1991, 2).

A.8 Kuwait

The largest ethnic group, **Kuwaiti Sunni (Arab)**, has ruled Kuwait since the al-Sabah dynasty began in 1756, including during and after an intervening period of British rule in the 20th century (EB Kuwait).

A.9 Lebanon

Ottoman rule began during the 16th century. "Lebanon did not enjoy an independent status in the Ottoman Empire . . . Administratively, Lebanon was split into districts that made no distinction between Lebanon, Syria, and Palestine" (AbuKhalil 1998, 6, 7). Mount Lebanon became a "relatively autonomous local district of the empire" (7) after massacres of Maronite Christians by Druze, but this occurred in response to European intervention rather than from a local state.

A.10 Libya

Libya has historically been divided among three regions, Tripolitania, Cyrenaica, and Fezzan, across which the EPR Arabs polygon spans, the largest ethnic group in Libya. Ottoman rule began in the 16th century, was interrupted by the local Karamanli dynasty from 1711 to 1835, and then resumed until Italian colonization in 1911. Karamanli was an independent state despite nominal Ottoman rule, going as far as to conduct its own diplomatic relations with European powers (Vandewalle 2012, 16). It also extended its rule through most of modern Libya—"At their peak, the Karamanlis' influence reached Cyrenaica and Fezzan, covering most of Libya" (EB Karamanli dynasty)—although it likely exerted little effective influence outside Tripolitania among people that were largely nomadic and where tribal leaders consistently resisted central influence in the vast hinterlands (approximately 90% of Libya' territory is desert) (Vandewalle 2012, 15, 18). However, Libya experienced 76 years of relatively intensive foreign Ottoman rule prior to European colonization, yielding a coding of no indigenous states prior to European colonization.

A.11 Morocco

The largest ethnic group, **Arabs**, were organized into various large states for hundreds of years prior to the 20th century when it was colonized by France, including the Alawite dynasty from the 17th century through the present (EB Morocco).

A.12 Oman

The only ethnic group, **Ibadhi Muslims (Arab)**, was organized into various large states for over a millennium prior to the 20th century (Oman was never formally colonized by a European power), and the al-Said family has ruled from 1749 until the present (EB Oman).

A.13 Qatar

Prior to British colonization, different Arab families lived in modern-day Qatar in small towns. The family with which Britain originally signed a treaty, the al-Thani (which throughout the colonial era and until today has been the ruling dynasty) was one of several families and only became ascendant because of their alliance with Britain. The lack of centralized structures and lack of unity among the different families prior to colonization do not meet the criteria for coding Qatari Arabs as PCS.

A.14   Saudi Arabia

The largest ethnic group, **Sunni Wahhabi (Najdi) (Arab)**, established a state during the 19th and early 20th centuries. Members of the al-Saud family battled with the Ottomans and Egyptians during the 19th century, resulting in alternating control of major Arabian cities, before Abdul Aziz took Riyadh in 1902 and conquered all parts of the modern country by the 1920s. Saudi Arabia was never colonized by a European power (EB Saudi Arabia).

A.15   Syria

According to EB Syria and Commins (1996, 7-9), Syria was ruled by the Ottoman Empire between 1516 and 1831, by Egypt between 1831 and 1840, and again by the Ottomans until 1918, when French colonial rule began. "Although parts of Syria enjoyed some local autonomy, the area as a whole remained for 400 years an integral section of the Ottoman Empire" (EB Syria). Weakened Ottoman rule enabled largely autonomous rule in some parts of Syria in parts of the 18th century, such as Damascus by the al-Azm family and Sidon by a Bosnian governor. These dynasties, however, had ceased ruling their former parts of Syria by the time Egyptian ruler Muhammed Ali conquered Syria. After re-integration with the Ottoman Empire in 1840, the Ottomans imposed administrative and legal reforms during the rest of the 19th century to tie Syria closer to Istanbul. In addition to a period of firm non-local rule for nearly a century prior to European colonization, Syria was also split into multiple provinces (which changed over time), including Damascus, Aleppo, Sidon, and—outside modern Syria—Tripoli and Beirut. Therefore, the largest EPR group, Sunni Arabs, whose EPR polygon covers most of Syria, did not compose an indigenously governed state prior to European colonization.

A.16   Tunisia

Tunisia was officially an Ottoman colony between 1574 and 1881, when France colonized Tunisia. However, a military revolt by Ottoman officers in 1591 created a state governed "independently from the Ottoman government" (EB North Africa). Perkins (1997, 7) adds: "As elsewhere in North Africa, direct Ottoman rule was short-lived. Less than two decades [after 1574], a revolts of the deys of Tunis, junior Ottoman officers, severely weakened Istanbul's hold over the province." Members of the Husaynid family composed the ruling dynasty of Tunisia between 1705 and 1957, shortly after independence (EB North Africa; EB Husaynid dynasty). "Al-Husayn conducted his affairs without Ottoman interference and with a measure of independence that allowed him to maintain separate treaties" with several European powers. Finally, this state covered the territory occupied by the only EPR ethnic group in Tunisia, **Arabs** (Oliver and Atmore 2005, 7-9).

A.17   Turkey

The largest ethnic group, **Turkish**, was organized as a large state for hundreds of years under the Ottoman empire. The Ottoman dynasty began in 1293, expanded across Anatolia and beyond in subsequent centuries, and ended in 1922 after defeat in World War I (EB Turkey).

A.18 United Arab Emirates

The United Arab Emirates is composed of seven separate tribal governments that belonged to the British Trucial States from the 19th century onward. These were separate families that separately signed treaties with Britain throughout the 20th century before becoming British colonies (Hawley 1970, 126-141). Warfare between the different Arab tribes continued throughout the 19th century (133). Despite similarities to cases such as the Tswana in Botswana, the lack of coordination across the different tribes justifies coding that Arabs in the UAE did not meet the criteria for PCS.

A.19 Yemen

EB Yemen mentions one modern state, the Zaydi imamate. This state corresponds to the second-largest EPR ethnic group in Yemen in 1946, the **Zaydis**. According to EB Yemen, EB Yahya, and Burrowes (2010), the Zaydi imamate dates back to the 9th century. More recently, the Ottoman empire occupied Yemen in the 16th century but the Zaydi imamate forcibly expelled the Ottomans in the 17th century. Modern Yemen consists of two countries that were separate before 1990, North Yemen and South Yemen, because of differing colonial histories. South Yemen, based in Aden, was a British colony from 1839 until independence in 1967. The Ottoman empire in 1849 re-colonized North Yemen, based in Sanaa, which lasted until 1918. Although Ottoman rule disrupted local rule in the north, the Zaydi imamate continued to resist Ottoman rule and the imam Yahya Mahmud al-Mutawakkil was the acknowledged ruler of the country after the Ottoman departure, although the country's borders were not defined. During the subsequent decades until the EPR dataset begins in 1946, North Yemen remained under Zaydi imamate rule and independent of European colonial rule (unlike Mesopotamian and North African countries that were colonized by Europe after Ottoman rule) and built the modern Yemeni state by suppressing tribal warfare, building a modern military, and defining the country's borders.

A.4 LATIN AMERICA AND CARIBBEAN

Linking precolonial states with EPR's postcolonial ethnic groups is more difficult in Latin America than in other regions because (a) European colonization occurred much earlier here than in most of the rest of the world and (b) European rule more intensively shaped demographics because natives lacked immunity to European diseases, resulting in genocide-level population declines. Although these distinctions raise the concern that I would code certain Latin American EPR ethnic groups as stateless because colonization destroyed many precolonial states—and therefore their corresponding ethnic groups were not politically relevant at independence—the sources consulted show that there were only two major states in Latin America on the eve of colonization, Aztec and Inca, both of which I link to EPR ethnic groups.

Table A.4: List of PCS Groups in Latin America

Country	EPR ethnic group	Historical state(s)
Mexico	Other indigenous groups	Aztec
Peru	Indigenous peoples of the Andes	Inca

A.1 Argentina

Although the Inca empire (see the Peru entry) expanded into northwest Argentina, no indigenous states arose in Argentina prior to Spanish colonization. EB Argentina characterizes the absence of precolonial states: "The population of the area now called Argentina may have totaled 300,000 before the arrival of

the Europeans. Some of the indigenous peoples were nomadic hunters and fishers, such as those in the Chaco, the Tehuelche of Patagonia, and the Querand and Puelche (Guennakin) of the Pampas, but others, such as the Diaguita of the Northwest, developed sedentary agriculture. The highlands of the Northwest were a part of the Inca empire." Mahoney (2010, 77) asserts: "in the vast territories south and east of the Inca empire—in most of Chile and Argentina and all of Uruguay—complex indigenous societies were *not* present" [emphasis in original].

A.2   Bolivia

The country featured a large precolonial empire, the Tiwanaku, but this fell in the 11th century (EB Bolivia). The empire was likely governed by the Aymara people (EB Tiwanaku), an EPR ethnic group. However, the Aymara were fractured in the 15th century. According to EB Bolivia: "By the 15th century the region was controlled largely by some 12 groups of Aymara-speaking Indians; they, in turn, fell under the control of the expanding Inca empire, which had its capital in Cuzco (now in Peru)." Mahoney (2010, 72) discusses the hierarchical political organization of an unnamed number of "Aymara kingdoms," but suggests considerable fractionalization within the Aymara. The Inca empire colonized these kingdoms in the 15th century.

A.3   Brazil

Neither EB Brazil, McCann (1998, 6-9), nor Levine (1979) mention precolonial states. EB Brazil and McCann (1998, 6-9) mention warfare involving Tupian-speakers, but they did not exhibit hierarchical or otherwise unified political organization (EB Tupian).

A.4   Chile

Although the Inca empire (see the Peru entry) expanded into northern Chile, no indigenous states arose in Chile prior to Spanish colonization. EB Chile characterizes the absence of precolonial states: "At the time of the Spanish conquest of Chile in the mid-16th century, at least 500,000 Indians inhabited the region. Nearly all of the scattered tribes were related in race and language, but they lacked any central governmental organization." Mahoney (2010, 77) asserts: "in the vast territories south and east of the Inca empire—in most of Chile and Argentina and all of Uruguay—complex indigenous societies were *not* present" [emphasis in original]. He then states that although the Mapuche lived in somewhat densely settled areas, they were not fully sedentary and "lived in disconnected and temporary settlements" (79).

A.5   Colombia

Although EB (see its Colombia and Chibcha entries) and Mahoney (2010, 105) each mention nascent Chibcha/Musica states, they also describe considerable fractionalization within the Musica. "The Musicas were really two loosely knit, tribute-collecting kingdoms—the Zaque of Tunja and the Zipa of Bogotá— together with other smaller and mostly independent chiefdoms . . . Musicas might have been on the verge of developing a more unified empire, but at the onset of Spanish colonialism, they were smaller states, more like the Mayas than the Aztecs" (Mahoney 2010, 105). EB Colombia adds: "At the time of the Spanish conquest, the Chibcha were in the process of consolidation by warfare and had not achieved firm union and political institutions." Therefore, even if Musicas were an EPR ethnic group, they would not count as a

state. Related, the EPR group "Indigenous peoples" is not coded as a state because of the many decentralized groups in precolonial Colombia (Musicas composed about one-third of the population).

A.6 Costa Rica

EB Costa Rica does not mention any specific precolonial groups. Mahoney (2010, 92-4) discusses numerous precolonial chiefdoms in Costa Rica, El Salvador, Honduras, and Nicaragua. However, all of them lacked centralized political authority. He describes the Pipil in El Salvador as "organized into state-like polities" (92) but mentions that there were numerous Pipil states without an overarching ruler. The Lenca and Chorotega chiefdoms in El Salvador and Honduras existed "in the absence of overarching political authority" (93). The Nicaraos chiefdom in Nicaragua and Costa Rica was "organized as a 'maximal chiefdom,' not a proto-state" (93). The other regional peoples he discusses exhibited even fewer state-like characteristics.

A.7 Cuba

Country-specific EB pages and the respective Library of Congress country studies for Cuba, Dominican Republic, Haiti, and Jamaica mention one group with some evidence of hierarchical organization: the Tainos. However, their organizational structure more closely resembles either petty or paramount chiefdoms rather than a state. For example, EB Taino asserts they had "a complex social order, with a government of hereditary chiefs and subchiefs and classes of nobles, commoners, and slaves." Various sources (Suchlicki 2002, 8; EB Dominican Republic) emphasize their less advanced governance and civilization than the Aztec, Inca, and Maya (NB: I coded Aztec and Inca as a state but not Maya). This decision is also consistent with Putterman's coding of no government above the local level prior to Spanish colonization.

A.8 Dominican Republic

No precolonial states, see coding notes for Cuba.

A.9 Ecuador

EB Ecuador mentions divisions among warring chiefdoms and asserts that "Ecuador lacked cities and states until after the Inca conquest." Mahoney (2010, 110) states that among Ecuador's highlanders: "Only when the Incas captured the region in the late fifteenth century did the chiefdoms become part of a state empire."

A.10 El Salvador

EB El Salvador does not mention any specific precolonial groups. Mahoney (2010, 92-4) discusses numerous precolonial chiefdoms in Costa Rica, El Salvador, Honduras, and Nicaragua. However, all of them lacked centralized political authority. He describes the Pipil in El Salvador as "organized into state-like polities" (92) but mentions that there were numerous Pipil states without an overarching ruler. (EB also mentions Pipil.) The Lenca and Chorotega chiefdoms in El Salvador and Honduras existed "in the absence

of overarching political authority" (93). The Nicaraos chiefdom in Nicaragua and Costa Rica was "organized as a 'maximal chiefdom,' not a proto-state" (93). The other regional peoples he discusses exhibited even fewer state-like characteristics.

A.11 Guatemala

Although the Maya "were one of the most highly developed peoples of precolonial America" (EB Guatemala), their civilization collapsed around 900. "When Spanish conquerors arrived in the 16th century, they found many cities in ruins and encountered little organized resistance" (EB Guatemala). Mahoney (2010, 101) mentions "several disunited Maya kingdoms" interspersed with various chiefdom Maya societies. Finally, EB Central America asserts: "Although the Maya were the most advanced pre-Columbian civilization in the hemisphere, they were never unified. Unlike the Aztec and Inca empires, their autonomous city-states remained independent, presaging the political fragmentation that would characterize Central America to the present day. What unity existed was cultural rather than political."

A.12 Haiti

No precolonial states, see coding notes for Cuba.

A.13 Honduras

Although EB Honduras mentions the Maya, as my Guatemala entry describes, Maya civilization had collapsed around 900. Mahoney (2010, 92-4) discusses numerous precolonial chiefdoms in Costa Rica, El Salvador, Honduras, and Nicaragua. However, all of them lacked centralized political authority. He describes the Pipil in El Salvador as "organized into state-like polities" (92) but mentions that there were numerous Pipil states without an overarching ruler. The Lenca and Chorotega chiefdoms in El Salvador and Honduras existed "in the absence of overarching political authority" (93). The Nicaraos chiefdom in Nicaragua and Costa Rica was "organized as a 'maximal chiefdom,' not a proto-state" (93). The other regional peoples he discusses exhibited even fewer state-like characteristics.

A.14 Jamaica

No precolonial states, see coding notes for Cuba.

A.15 Mexico

EB Mexico and Mahoney (2010, 54-5) discuss the Aztec empire, which exhibited considerable political hierarchy: "During the reign of Montezuma II, the ninth Aztec king (1502-20), Aztec officials produced codices that recorded the organization of the empire into provinces and the payment of tribute according to the production of each region. A gigantic political, military, and religious bureaucracy was built up, with governors, tax collectors, courts of justice, military garrisons, mail and messenger services, and other civil offices" (EB Mexico). The Aztec empire itself contained many ethnic groups (Mahoney 2010, 55), and considerable Spanish migration and natives' susceptibility to European diseases further ensured that Aztecs were not a politically relevant ethnic group at Mexico's independence (or in the first year of EPR data,

1946). I code the EPR group **Other indigenous groups** as a state to ensure that the PCS variable includes one of Latin America's two major precolonial states. Notably, it is unclear what percentage of this EPR group Aztecs actually ruled, although Other indigenous groups is mostly concentrated in southern Mexico—where the Aztec empire was located and where the preponderance of Mexico's precolonial population lived (Mahoney 2010, 54-6)—and the Aztec capital city, Tenochtitlán, is located right outside an EPR polygon for Other indigenous groups.

A.16   Nicaragua

EB Nicaragua does not mention any specific precolonial groups. Mahoney (2010, 92-4) discusses numerous precolonial chiefdoms in Costa Rica, El Salvador, Honduras, and Nicaragua. However, all of them lacked centralized political authority. He describes the Pipil in El Salvador as "organized into state-like polities" (92) but mentions that there were numerous Pipil states without an overarching ruler. The Lenca and Chorotega chiefdoms in El Salvador and Honduras existed "in the absence of overarching political authority" (93). The Nicaraos chiefdom in Nicaragua and Costa Rica was "organized as a 'maximal chiefdom,' not a proto-state" (93). The other regional peoples he discusses exhibited even fewer state-like characteristics.

A.17   Panama

Neither EB Panama nor Leonard's (2014, 1-19) or Black and Flores's (1989, 6-7) historical overviews mention any precolonial states.

A.18   Paraguay

EB Paraguay mentions the Guaraní, but they were a "warlike seminomadic people." Mahoney (2010, 86-7) does not describe any of Paraguay's precolonial groups as having states.

A.19   Peru

EB Peru and Mahoney (2010, 64-6) discuss the Inca empire, which exhibited considerable political hierarchy despite conquering expansive territories (1460s) not long before Spanish colonialism began. Mahoney describes the empire's tribute and forced labor system. The Inca correspond with the EPR group **Indigenous peoples of the Andes**.

A.20   Trinidad and Tobago

Neither EB Trinidad and Tobago and **?**, 165 mention any precolonial states. Both mention the Arawaks and Caribs, but neither exhibited hierarchical political organization (Anthony 1997, 23-4, 102).

A.21   Uruguay

EB Uruguay does not mention any precolonial states: "[t]he principal groups were the seminomadic Charra, Chan (Chanes), and Guaran Indians." Mahoney (2010, 77) asserts: "in the vast territories south and east of

the Inca empire—in most of Chile and Argentina and all of Uruguay—complex indigenous societies were *not* present" [emphasis in original].

A.22   Venezuela

EB Venezuela does not mention any precolonial states: "Arawak and Carib Indians were prominent among the groups that arrived later. Nomadic hunting and fishing groups roamed the Lake Maracaibo basin, the Llanos, and the coast. The most technologically advanced Venezuelan Indians lived in farming communities in the Andes." Mahoney (2010, 91) quotes a source that says: "In no part of Venezuela were there large concentrations of urbanized Indians living under centralized political control."

B.1   ADDITIONAL DATA INFORMATION

B.1   ETHNIC CIVIL WARS AND AIMS

The main civil war data used for this paper draw from Fearon and Laitin's (2003) updated civil war data through 2009, and further updated by the author through 2013 by adding new conflicts from the Correlates of War database (COW; Dixon and Sarkees 2015), which also uses a 1,000 death threshold, and the UCDP/PRIO Armed Conflict Database (ACD; Gleditsch et al. 2002). I included every "intra-state war" that COW—which uses a 1,000 battle death threshold for wars—codes as beginning between 2010 and 2013 that ACD also codes as reaching 1,000 battle deaths, one of Fearon and Laitin's key coding rules. I also consulted COW and ACD for conflict termination years for any civil wars that Fearon and Laitin coded as ongoing in 2009. Additionally, Ross and Mahdavi (2015) use a lower population threshold than Fearon and Laitin (2003), which necessitates coding civil wars for smaller countries. ACD does not use a population threshold for deciding which countries to include (see pg. 14 of Version 7.1 of their codebook), therefore providing the needed information—although no civil wars were added through this procedure (Comoros had two conflicts but neither reached 1,000 battle deaths). Finally, I excluded any colonial wars.

Finally, Fearon and Laitin code whether the civil war was center-seeking or separatist. I verified their coding of civil war aims with both COW and ACD, and additional secondary sources when necessary. This enabled assigning aims to the wars that Fearon and Laitin code as mixed or ambiguous. Most cases that they code as mixed are aggregated rebellions that contain distinct rebel groups fighting center-seeking and separatist civil wars (see, for example, the Angola example in the first paragraph of the paper), whereas I further distinguish each case by war aims. By contrast, COW or ACD code each war as *either* center-seeking or separatist, but never both. My coding scheme allows for the possibility of coding a rebellion as exhibiting both aims. However, after disaggregating Fearon and Laitin's civil war entries that contain multiple distinct rebel groups, I only coded two cases as exhibiting both aims (Ethiopia and Sudan). By contrast, in countries such as Burma (coded as mixed war aims by Fearon and Laitin), largely distinct center-seeking and separatist rebellions broke out in 1948, and several other countries such as Angola and India have featured center-seeking civil wars and separatist civil wars at the same time despite not beginning in the same year.

The major advantage of using data based off Fearon and Laitin's (2003) coding procedure rather than ACD is that ACD does not provide a coherent scheme for coding distinct civil wars, and hence civil war *onsets*. Scholars use a lapse rule, typically two years, for translating ACD's incidence data into distinct conflict onsets. If the 25 or 1,000 death threshold (ACD codes both thresholds) is not met for at least two years after being met in the past, then using a two-year lapse rule counts any future year that meets the death threshold as a new civil war. Problematically, this procedure often either undercounts or (more likely) overcounts civil

war onsets, especially when applied to the 25 battle death threshold standard in EPR studies. Fearon and Laitin (2013, 25) summarize:

> "They apply a criterion of one year (or two, or ten, for different codings) with no conflict above their 25 death threshold. This has the advantage of being relatively definite, but the disadvantage of making many long-running, low level conflicts that flit above and below the 25 dead threshold look like many distinct civil wars. In our view they often are more naturally seen as a single, long-running but low level civil conflict, that happens often by chance to get above or below the threshold in some years" (25). (Also see Sambanis 2004, 818-9.)

For an example of overcounting, using the standard two-year lapse coding in ACD2EPR, the Bakongo in Angola fought four different civil wars in the 1990s and 2000s even though the same rebel group was operative during the entire period. Solely using a lapse rule to distinguish conflicts can also undercount civil war onsets. For example, the UCDP Conflict Encyclopedia describes civil wars in the Democratic Republic of the Congo in the 1990s: "In 1996-1997 an armed rebellion led by AFDL and supported by Rwanda and Uganda managed to topple President Mobutu in May 1997. However the new regime was soon at war again [in 1998], this time against RCD and MLC." Although two different sets of governments and rebel groups fought what by any reasonable conceptualization are two distinct wars, the two-year lapse rule does not count a new onset in 1998 for the Tutsi-Banyamulenge because they participated in conflict in the previous year.

Although scholars can also employ lapse rules of other length, coding civil war episodes solely by using lapse rules does not address these problems of undercounting and overcounting. Two of Fearon and Laitin's (2003) coding rules help to guard against these issues. First, "War ends are coded by observation of a victory, wholesale demobilization, truce, or peace agreement followed by at least two years of peace" (Fearon and Laitin 2003, 76, fn. 4; which also states their full set of rules). This directly addresses the concern about overcounting onsets for periodic conflicts, such as Bakongo in Angola, because clear signals of intent to end the current episode of fighting characterize the end of a war. Importantly, this rule still enables coding repeated civil wars with the same rebel group. Second, "If a main party to the conflict drops out, we code a new war start if the fighting continues (e.g., Somalia gets a new civil war after Siad Barre is defeated in 1991)." This addresses the problem of undercounting onsets in cases such as the Democratic Republic of the Congo in the 1990s.

B.2 COVARIATES

- GDP per capita: Annual logged country-level data from Maddison (2008).

- Population: Annual logged country-level data from Maddison (2008).

- Democracy: Annual country-level data from Polity IV's *polity2* variable (Marshall and Gurr 2014).

- Giant oil field: Ethnic group coded as 1 if it has at least one giant oil field within its EPR polygon, or within 250 kilometers offshore and within the group's country's maritime borders. Giant oil field data from Horn (2003).

- Group percentage of population: Ethnic group's share of its country's total population, coded by EPR (Vogt et al. 2015).

- Geographically concentrated: Indicator variable for whether the territorial settlement of an EPR ethnic group is concentrated, coded by EPR (Vogt et al. 2015).

- Noncontiguous: Indicator variable coded by author for whether an EPR ethnic group's territory is noncontiguous from the landmass that contains the country's capital city.

- Distance from capital: Log of the minimum distance between an EPR group's location polygon and the capital city, coded by EPR (Vogt et al. 2015).

- Neolithic transition: Putterman (2008) shows in a global sample that territories experiencing earlier transitions to agricultural production tended to experience higher levels of statehood in the second millennium, a variable measured at the country level. The measure is thousands of years elapsed since an ethnic group's (modern-day) country experienced a transition to agricultural production.

- Precolonial wars: Percentage of years between 1400 and either the last year before colonization or 1900 (whichever occurred earlier) in which a country experienced an internal war. As described in the historical background section of the paper, I used Brecke's (1999) list of conflicts between 1400 and the present to code precolonial internal wars at the level of modern-day countries. I counted only wars fought between groups within the same modern-day country boundaries—hence "internal" wars—as opposed to any war fought across modern-day borders. For example, the entry "Funj-Musabaat Arabs (Sudan), 1747" constitutes an internal war in Sudan in 1747, but "Ethiopia-Funj (Sudan), 1744" does not count as an internal war for either Sudan or Ethiopia.

- Settler colonialism: Indicator variable coded by author that equals 1 if the country's European population ever exceeded 5% of the colonial population, and 0 otherwise (this variable is set to 0 for all non-colonized countries). Data from Easterly and Levine (2016) and Paine (2019a).

- Squiggly borders: Measure from Alesina, Easterly and Matuszeski (2011). As they describe: "The basic idea is to compare the borders of a country to a geometric figure. If a country looks like a perfect square with borders drawn with straight lines, the chances are that these borders were drawn artificially. On the contrary, borders that are squiggly lines (perhaps meant to capture geographic features and/or ethnicities) are less likely to be artificial. Squiggly geographic lines (such as mountains) are likely to separate ethnic groups, for reasons of patterns of communication and migration." The actual measure is based on fractal dimension, as they describe.

- Partitioned group: Indicator variable that equals 1 if an EPR ethnic group is transnational, that is, ethnic kin in a neighboring country. Coded by author using the TEK dataset (Vogt et al. 2015).

C.1 ADDITIONAL TABLES AND FIGURES

Table C.1: Colonial Police Composition

	DV: Police ethnic imbalance		DV: Heavily underrepresented		DV: Heavily overrepresented	
	(1)	(2)	(3)	(4)	(5)	(6)
PCS group	-0.318	-0.125	-1.244	-1.771	-0.357	-0.671
	(1.001)	(0.984)	(1.072)	(1.380)	(0.815)	(0.962)
Ethnic groups	143	94	143	70	143	85
R-squared	0.001	0.046				
Country FE	NO	YES	NO	YES	NO	YES

Notes: Ray (2016) provides the different dependent variables on colonial police composition, and the PCS group variable is that used in the present paper. The sample is British colonies with available police data from Ray (2016). In columns 1 and 2, the dependent variable is the natural log of a group's percent share of the top ranks of the colonial police force in its country of residence on the eve of independence, and the models are OLS. In columns 3 and 4, the dependent variable is an indicator that equals 1 when a group's police imbalance score is less than or equal to the 25th percentile value of the ratio in the sample and 0 otherwise, and the models are logit. In columns 5 and 6, the dependent variable is an indicator that equals 1 when a group's police imbalance score is greater than or equal to the 75th percentile value of the ratio in the sample and 0 otherwise, and the models are logit. The odd-numbered columns include the full sample, and the even-numbered columns model country fixed effects and only include PCS countries. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table C.2: Ruling Group in First Year of Sample

	DV: Ruling group						
	Panel A. Pooled models						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PCS group	1.751***	1.844***	1.980***	1.411***	1.862***	1.808***	1.759***
	(0.377)	(0.407)	(0.463)	(0.472)	(0.409)	(0.421)	(0.422)
SLPCS group	-1.323***	-1.220***	-1.157***	-0.0731	-1.071***	-1.294***	-1.298***
	(0.287)	(0.336)	(0.394)	(0.432)	(0.321)	(0.340)	(0.311)
Ethnic groups	433	433	377	344	391	433	351
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Country FE	NO	NO	NO	NO	NO	NO	NO
	Panel B. Country fixed effects						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PCS group	3.007***		3.235***	1.575**			
	(0.497)		(0.548)	(0.624)			
Ethnic groups	228		193	183			
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Country FE	YES	YES	YES	YES	YES	YES	YES

Notes: Table C.2 summarizes a series of logistic regressions by presenting coefficient estimates for the variables, and ethnic group-clustered standard errors in parentheses. The dependent variable equals 1 if EPR codes the ethnic group's ethnopolitical access as monopoly, dominant, or senior partner. The covariates in each column correspond to those in Table 1. In Panel B, I do not present estimates for specifications in which the only covariates are measured at the country level and are time invariant because they are perfectly collinear with the country fixed effects. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table C.3: Summary Statistics for Main Sample

Variable	Mean	Std. Dev.	Group-years
Separatist CW onset (ongoing set to missing)	0.002	0.047	25646
Center CW onset (ongoing set to missing)	0.002	0.047	26157
SLPCS group	0.492	0.5	26509
PCS group	0.086	0.281	26509
SSA	0.363	0.481	26509
MENA	0.154	0.361	26509
Latin America	0.146	0.353	26509
ln(GDP/pop)	7.833	1.071	25639
ln(Pop.)	16.551	2.116	25857
Democracy	-1.044	6.743	25736
Giant oil field	0.138	0.345	26509
Group % pop.	0.184	0.245	26509
Geo. concentrated	0.847	0.36	26509
Noncontiguous	0.025	0.156	26509
Dist. from capital	3.786	2.65	22954
Neolithic transition	5.239	2.799	25718
Precolonial wars	0.077	0.179	25456
British colony	0.361	0.48	26509
Other colony	0.448	0.497	26509
Settler colonialism	0.106	0.308	26509
Squiggly borders	0.034	0.015	21804
Partitioned group	0.552	0.497	26509

Table C.4: Pre-Colonial Statehood and Center-Seeking Civil War Onset

	DV: Major center-seeking civil war onset						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PCS group	0.263	0.174	0.329	0.0478	0.652	0.236	0.459
	(0.556)	(0.553)	(0.559)	(0.660)	(0.526)	(0.598)	(0.609)
SLPCS group	0.0220	0.0945	0.378	0.299	0.593*	-0.00162	-0.0407
	(0.360)	(0.317)	(0.377)	(0.376)	(0.339)	(0.462)	(0.406)
SSA		2.354***					
		(0.568)					
MENA		1.902***					
		(0.624)					
L. Am.		-0.0496					
		(1.149)					
ln(GDP/pop)			-0.634***				
			(0.212)				
ln(Pop.)			-0.567***				
			(0.108)				
Democracy			-0.117***				
			(0.0260)				
Giant oil field			0.772				
			(0.551)				
Group % pop.				0.663			
				(0.881)			
Geo. concentrated				0.0478			
				(0.674)			
Noncontiguous				-			
Dist. from capital				-0.0417			
				(0.0640)			
Neolithic transition					0.0618		
					(0.0570)		
Precolonial wars					-12.36***		
					(4.305)		
British colony						-0.188	
						(0.406)	
Other colony						0.186	
						(0.438)	
Settler colonialism						-1.921*	
						(1.039)	
Squiggly borders							-9.651
							(15.85)
Partitioned group							-0.182
							(0.312)
Group-years	25,113	25,113	23,958	22,046	24,965	25,113	20,751
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES

Notes: Table C.4 summarizes a series of logistic regressions by presenting coefficient estimates for the variables, and ethnic group-clustered standard errors in parentheses. Column 4 drops every ethnic group with noncontiguous territory because this indicator perfectly predicts the absence of the outcome. ***$p < 0.01$,** $p < 0.05$,* $p < 0.1$.

Table C.5: Assessing Bias from Unobservables using Selection on Observables

(2)	(3)	(4)	(5)	(6)	(7)
-33.1	-9	7.9	-6.1	-44.2	70.1

Notes: Altonji, Elder and Taber (2005) present a commonly used metric that estimates how large the bias from unobserved covariates would need to be for the true coefficient to be 0 in a statistical model, given information from how much adding observable covariates changes the estimates. To compute this measure, Table C.5 compares the coefficient estimate for SLPCS in Columns 2 through 7 of Table 1 with that in Column 1, although because the metric is computed for linear models I re-estimated Table 1 with OLS (findings are qualitatively identical; available upon request). Negative numbers in Table C.5 imply that the coefficient estimate in the specification with covariates exceeds in magnitude the coefficient estimate in the restricted specification. This indicates an estimate highly robust to omitted covariates because the magnitude of the bias of unobserved covariates would need to go in the opposite direction as the bias from omitting observables to drive the coefficient estimate to 0. This occurs for four of the six coefficient estimates shown in Table C.5. In the other two specifications, adding covariates diminishes the magnitude of the theoretically relevant coefficient estimate, but the large positive numbers in Table C.5 show that adding covariates only minimally affects the coefficient estimates: the bias from unobservables would need to be between 7.9 and 70.1 times larger than the bias from omitting the covariates contained in these specifications to overturn the positive coefficient estimate. Altonji, Elder and Taber (2005) calculate a corresponding figure of 3.55 for their own analysis, which they interpret as large in magnitude. Overall, the insensitivity of the coefficient estimates to adding covariates implies that—although it is impossible to control for every possible confounder—if the control variables included the tables are substantively relevant, then there is less reason to believe that covariates not included in any of the specifications would overturn the results.

Table C.6: Cross-Section with Binary Civil War Onset

	Panel A. DV: Any major separatist civil war onsets						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SLPCS group	2.708***	2.427***	2.252***	2.173***	3.019***	2.597***	2.847***
	(0.611)	(0.629)	(0.712)	(0.631)	(0.733)	(0.654)	(0.769)
PCS group	1.179	0.915	0.158	1.552*	1.476	1.101	0.891
	(0.930)	(0.943)	(1.211)	(0.938)	(1.014)	(0.960)	(1.269)
Ethnic groups	442	375	386	364	400	442	360
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES
	Panel B. DV: Any major center-seeking civil war onsets						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SLPCS group	0.476	0.536	1.143**	0.545	0.762*	0.423	0.340
	(0.387)	(0.387)	(0.477)	(0.426)	(0.400)	(0.477)	(0.430)
PCS group	0.739	0.671	1.347**	0.515	0.898	0.661	0.707
	(0.562)	(0.581)	(0.625)	(0.645)	(0.579)	(0.605)	(0.601)
Ethnic groups	442	442	358	383	400	442	360
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES

Notes: Table C.6 summarizes a series of logistic regressions by presenting coefficient estimates for the variables, and robust standard errors in parentheses. The sample consists of every politically relevant ethnic group in the first year in the sample for each country (which for most is the first year of independence) and there is a single observation for each ethnic group. The dependent variable equals 1 if the group experienced a civil war at any point before 2013, and 0 otherwise. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table C.7: Cross-Section with Count of Civil War Onsets

	Panel A. DV: Count of major separatist civil war onsets						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SLPCS group	2.455***	2.148***	1.943***	1.966***	2.653***	2.275***	2.429***
	(0.631)	(0.608)	(0.670)	(0.652)	(0.718)	(0.683)	(0.785)
PCS group	0.859	0.566	-0.181	1.085	1.044	0.732	0.444
	(0.921)	(0.919)	(1.202)	(0.906)	(0.986)	(0.954)	(1.275)
Ethnic groups	442	442	386	397	400	442	360
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES
	Panel B. DV: Count of major center-seeking civil war onsets						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SLPCS group	0.441	0.572*	1.081**	0.606	0.969***	0.381	0.431
	(0.383)	(0.326)	(0.483)	(0.369)	(0.375)	(0.410)	(0.381)
PCS group	0.725	0.831	1.170**	0.592	1.177**	0.709	0.738
	(0.539)	(0.587)	(0.566)	(0.543)	(0.567)	(0.555)	(0.561)
Ethnic groups	442	442	386	397	400	442	360
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES

Notes: Table C.7 summarizes a series of negative binomial regressions by presenting coefficient estimates for the variables, and robust standard errors in parentheses. The sample consists of every politically relevant ethnic group in the first year in the sample for each country (which for most is the first year of independence) and there is a single observation for each ethnic group. The dependent variable equals the number of distinct civil war onsets for the group before 2013. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table C.8: State Antiquity Measure

	Panel A. DV: Major separatist civil war onset						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
State antiquity in 1500	2.183***	2.725***	2.267***	1.913***	3.122***	3.731***	2.156***
	(0.429)	(0.521)	(0.569)	(0.551)	(0.513)	(0.721)	(0.380)
Group-years	22,859	19,388	22,190	20,516	22,711	19,034	20,371
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES
	Panel B. DV: Major center-seeking civil war onset						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
State antiquity in 1500	0.00130	0.838*	0.534	0.258	0.426	-0.275	-0.247
	(0.452)	(0.499)	(0.479)	(0.426)	(0.637)	(0.666)	(0.513)
Group-years	23,303	23,303	22,658	20,302	23,155	21,448	20,751
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES

Notes: Table C.8 summarizes a series of logistic regressions by presenting coefficient estimates for the variables, and ethnic group-clustered standard errors in parentheses. Column 2 of Panel A drops every country in Latin America, and Column 4 of Panel B drops every ethnic group with noncontiguous territory because these indicators perfectly predict the absence of the outcome. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table C.9: ACD2EPR Civil War Measure

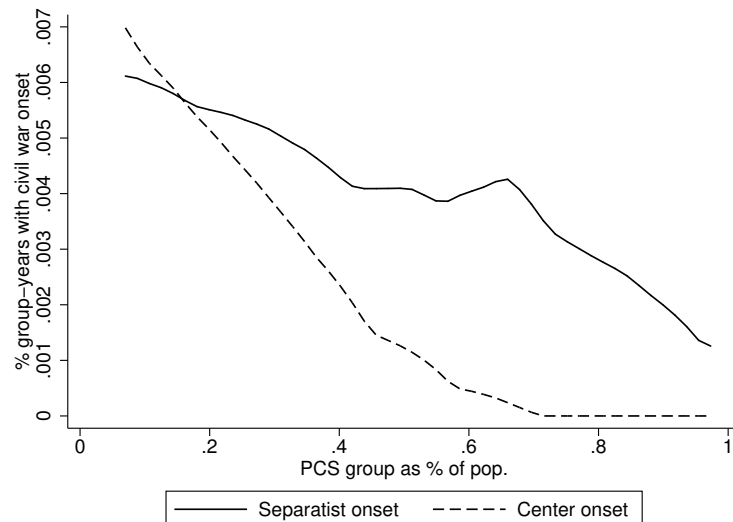|	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	DV: ACD2EPR separatist civil war onset						
SLPCS group	1.380***	1.158***	1.230***	1.053**	1.627***	1.312***	1.487***
	(0.370)	(0.378)	(0.449)	(0.427)	(0.432)	(0.421)	(0.474)
PCS group	-0.611	-0.843	-1.407	-0.220	-0.300	-0.646	-0.961
	(0.787)	(0.780)	(1.092)	(0.817)	(0.813)	(0.808)	(1.115)
Group-years	25,571	21,694	23,728	20,315	23,393	25,571	20,817
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES

Notes: Table C.9 summarizes a series of logistic regressions by presenting coefficient estimates for the variables, and ethnic group-clustered standard errors in parentheses. Column 2 drops every country in Latin America because this indicator perfectly predicts no separatist civil war. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Figure C.1: SLPCS Groups: Civil War Aims by PCS Size

Figure C.1: SLPCS Groups: Civil War Aims by PCS Size

Table C.10: Size of Country's PCS Group and Civil War Onset

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A. DV: Major separatist civil war onset						
SLPCS group (country with small PCS)	2.173***	1.982***	1.980***	1.878***	2.245***	2.057***	2.187***
	(0.541)	(0.507)	(0.593)	(0.562)	(0.533)	(0.606)	(0.589)
SLPCS group (country with medium PCS)	1.812***	1.528***	1.680***	1.391**	1.821***	1.669***	1.581**
	(0.549)	(0.552)	(0.566)	(0.585)	(0.547)	(0.609)	(0.627)
SLPCS group (country with large PCS)	1.014	0.677	0.371	0.586	1.469*	0.866	1.063
	(0.648)	(0.689)	(0.959)	(0.704)	(0.791)	(0.739)	(0.684)
PCS group (country with small PCS)	0.665	0.504	-	0.735	0.766	0.618	-
	(1.080)	(1.068)		(1.013)	(1.074)	(1.132)	
PCS group (country with medium PCS)	-	-	-	-	-	-	-
PCS group (country with large PCS)	-	-	-	-	-	-	-
Group-years	23,076	19,673	21,127	20,707	22,928	23,076	18,356
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES
	Panel B. DV: Major center-seeking civil war onset						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SLPCS group (country with small PCS)	0.701*	0.574*	1.134***	0.872**	1.487***	0.858*	0.518
	(0.387)	(0.348)	(0.420)	(0.378)	(0.396)	(0.521)	(0.454)
SLPCS group (country with medium PCS)	-1.440**	-1.253**	-1.132*	-1.094*	-0.991*	-1.663***	-1.064*
	(0.613)	(0.587)	(0.648)	(0.625)	(0.595)	(0.635)	(0.613)
SLPCS group (country with large PCS)	-	-	-	-	-	-	-
PCS group (country with small PCS)	0.868	0.549	1.110*	0.816	1.429**	0.918	0.947
	(0.612)	(0.603)	(0.639)	(0.603)	(0.571)	(0.691)	(0.662)
PCS group (country with medium PCS)	-0.295	-0.196	-0.903	-0.596	0.107	-0.565	0.00662
	(1.009)	(1.021)	(1.039)	(1.048)	(1.018)	(0.961)	(1.019)
PCS group (country with large PCS)	-	-	-	-	-	-	-
Group-years	21,717	21,717	20,607	19,053	21,569	21,717	17,534
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES

Notes: Table C.10 summarizes a series of logistic regressions by presenting coefficient estimates for the variables, and ethnic group-clustered standard errors in parentheses. Many of the indicators shown perfectly predict the absence of the outcome, and therefore all observations with positive values of those indicators are dropped. Additionally, Column 2 of Panel A drops every country in Latin America, and Column 4 of Panel B drops every ethnic group with noncontiguous territory because these indicators perfectly predict the absence of the outcome. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table C.11: Center-Seeking Civil Wars in SSA

	(1)	(2)	(3)	(4)	(5)	(6)
	DV: Major center-seeking civil war onset					
PCS group	0.547	0.836	0.157	0.859	0.560	1.103
	(0.798)	(0.782)	(0.736)	(0.847)	(0.793)	(0.844)
SLPCS group	0.948*	1.137**	0.713	1.201**	0.977**	1.310**
	(0.487)	(0.476)	(0.456)	(0.497)	(0.475)	(0.641)
Group-years	9,024	8,491	8,476	9,024	9,024	7,422
Covariates	None	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES

Notes: Table C.11 summarizes a series of logistic regressions by presenting coefficient estimates for the variables, and ethnic group-clustered standard errors in parentheses. Column 3 drops every ethnic group with noncontiguous territory because this indicator perfectly predicts the absence of the outcome. The p-value for SLPCS in this column is 0.121. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table C.12: PCS Group Size and Potential Separatist Challengers

	(1)	(2)
	DV: # potential sep. groups	
PCS group as % of pop.	-4.307**	-3.691***
	(1.677)	(1.360)
Countries	49	49
R-squared	0.123	0.135
Modifications	None	DV capped at 10

Notes: Table C.12 summarizes a series of linear regressions by presenting coefficient estimates for the variables, and OLS standard errors in parentheses. In Column 2, the dependent variable is capped at 10 potential separatist groups. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table C.13: Interactive Effect of Territorial Overlap with Dominant Group

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	DV: Major separatist civil war onset						
SLPCS group	2.356***	2.311***	2.298***	2.260***	2.820***	2.171**	2.608**
	(0.801)	(0.780)	(0.823)	(0.810)	(1.061)	(0.846)	(1.095)
PCS group	2.451**	2.377**	1.893	2.338**	2.980**	2.350**	2.334
	(1.007)	(1.008)	(1.218)	(1.023)	(1.216)	(1.031)	(1.494)
% terr. overlap	0.450	1.477	0.664	0.658	0.904	0.478	0.892
	(1.063)	(1.169)	(1.056)	(1.042)	(1.293)	(1.026)	(1.337)
SLPCS group*% terr. overlap	-2.764**	-3.761***	-3.272**	-2.804**	-3.080**	-2.922**	-3.403**
	(1.265)	(1.326)	(1.274)	(1.236)	(1.502)	(1.263)	(1.612)
PCS group*% terr. overlap	-	-	-	-	-	-	-
Group-years	16,683	14,551	15,720	16,389	16,128	16,683	14,106
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES

Notes: Table C.13 summarizes a series of logistic regressions by presenting coefficient estimates for the variables, and ethnic group-clustered standard errors in parentheses. The logit regression drops all cases in which PCS groups exhibit non-zero overlap with the dominant group because these cases did not experience any separatist civil wars. Figure 8 presents the marginal effects plot for the SLPCS indicator in Column 1. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$.

Table C.14: Conditioning on Ethnopolitical Exclusion

			Sample: Excluded group-years				
			Panel A. DV: Major separatist civil war onset				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SLPCS group	1.730***	1.458**	1.880***	1.870***	2.165***	1.809***	1.862**
	(0.639)	(0.632)	(0.666)	(0.668)	(0.781)	(0.688)	(0.768)
PCS group	1.961**	1.738**	1.387	1.699**	2.250**	2.061**	1.731
	(0.865)	(0.807)	(1.183)	(0.840)	(0.920)	(0.854)	(1.261)
Group-years	13,868	11,414	12,931	12,204	12,976	13,868	11,521
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES
			Panel B. DV: Major center-seeking civil war onset				
	(1)	(2)	(3)	(4)	(5)	(6)	
PCS group	1.782**	1.167	2.471***	1.455*	1.660**	1.982***	2.466***
	(0.749)	(0.787)	(0.573)	(0.760)	(0.745)	(0.753)	(0.764)
SLPCS group	0.196	0.118	0.429	0.403	0.818*	0.369	0.378
	(0.409)	(0.415)	(0.517)	(0.402)	(0.430)	(0.520)	(0.475)
Group-years	14,368	14,368	13,495	12,317	13,383	14,368	11,877
Covariates	None	Region FE	Standard	Geography	Precolonial	Colonial	Artificial borders
Lagged conflict incidence?	YES	YES	YES	YES	YES	YES	YES
Event history controls?	YES	YES	YES	YES	YES	YES	YES

Notes: Table C.14 summarizes a series of logistic regressions by presenting coefficient estimates for the variables, and ethnic group-clustered standard errors in parentheses. The p-value for the PCS indicator in Column 2 of Panel B is 0.138. The weaker correlation, coupled with the significance of the SSA dummy, is consistent with the earlier discussion about how small PCS groups are mostly located in Africa. Among excluded groups, PCS groups can explain differences in center-seeking conflict between Africa and other regions, but the correlation is weaker when comparing within regions. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$.

D.1   Formalizing the Logic of Rebellion Aims

A simple formal model clarifies the ideas from Section 2.1 and shows how the lines in Figure 1 can arise as equilibrium outcomes.

D.1   Setup

Suppose that a government G makes a take-it-or-leave-it offer x from a budget normalized to size 1 to a challenger C. G can choose $x \in [0, \beta]$, for $\beta \in (0, 1]$ with higher β reflecting a higher maximum offer that G can make, i.e., higher commitment ability. C can either accept, fight a center-seeking civil war, or fight a separatist civil war. If C accepts, then it consumes x and G consumes $1 - x$.

C's probability of winning depends on the demographic size of G relative to C, $\theta \in (0, 1)$, and on C's war aims. Although other real-world factors can affect the probability of winning a civil war, the model seeks to isolate the effect of group size, and as the text discusses, this factor is also correlated with other contributors to war such as separatist geography. Both players pay a cost $\phi > 0$ if C initiates either type of civil war.

If C initiates a center-seeking civil war, then it wins with probability $1 - \theta$ and loses with complementary probability. For either player, winning yields consumption of the entire budget of 1 and losing yields 0. If C initiates a separatist civil war, then it wins with probability $1 - \alpha \cdot \theta$ and loses with complementary probability, for $\alpha \in (0, 1)$. This parameter captures various factors that can blunt G's strength against C when combating a separatist relative to center-seeking civil war, with lower α corresponding to a government that less effectively translates demographic size into a higher probability of winning. One factor that would lower α is that C's coercive capacity is enhanced when it fights G in the periphery rather than in the capital by expanding the range of techniques that C can use to combat G. One factor that would raise α is the absence of a distinct territory from which it could form an autonomous state (in the extreme case, we could consider a more general probability-of-winning function that sets the probability of winning a separatist civil war to 0 if C lacks a territorial base for secession). For C, the prize of winning a separatist civil war is $\pi \in (0, 1)$, and G consumes the remainder; whereas losing yields 0 for C and 1 for G. The key idea underlying π is that winning a separatist civil war is less valuable than winning a center-seeking civil war because C only gains spoils from its region of the country.

D.2   Equilibrium and Comparative Statics

Solving backwards for subgame perfect Nash equilibria, C will accept any

$$x \geq \max \left\{ 1 - \theta - \phi, \ \left(1 - \alpha \cdot \theta\right) \cdot \pi - \phi \right\},$$

If this inequality is violated, then the type of civil war that C fights depends on θ. If $\alpha < \frac{1}{\pi}$, then there exists a unique threshold value such that C's optimal civil war is center-seeking if θ is small and separatist if θ is large. Regardless of C's binding civil war threat, if possible, G will buy off C because G makes the bargaining offers and pockets the surplus saved by preventing a war. However, only if β is large can G offer enough to induce C to accept. There are a continuum of payoff-equivalent equilibria.

Proposition D.1 (Equilibria strategy profile). *Assume $\alpha < \frac{1}{\pi}$.*

- *If $\beta > \max \left\{ 1 - \theta - \phi, \ \left(1 - \alpha \cdot \theta\right) \cdot \pi - \phi \right\}$, then G offers $x = \max \left\{ 1 - \theta - \phi, \right.$ $\left. \left(1 - \alpha \cdot \theta\right) \cdot \pi - \phi, 0 \right\}$. If this inequality is violated, then G is indifferent among all $x \in [0, \beta]$.*

72

- C accepts any $x \geq \max\left\{1 - \theta - \phi, \ (1 - \alpha \cdot \theta) \cdot \pi - \phi\right\}$ and fights otherwise. If C fights, it chooses center-seeking if $\theta < \frac{1-\pi}{1-\alpha\cdot\pi} \in (0,1)$ and separatist otherwise.

> **Proof.** The only part of the claim that does not follow immediately from assumptions is that the unique threshold claim for θ requires that the difference between C's expected utility to a initiating center-seeking civil war and C's utility to initiating a separatist civil war strictly decreases in θ, which is true given the assumption $\alpha < \frac{1}{\pi}$. $\blacksquare$

Proposition D.2 (Equilibrium outcomes)**.** *Assume* $\alpha < \frac{1}{\pi}$ *and* $\beta < \frac{(1-\alpha)\cdot\pi}{1-\alpha\cdot\pi} - \phi$.

- *If* $\beta > 1 - \phi$*, then war does not occur.*

- *If* $\beta \in \left(\frac{(1-\alpha)\cdot\pi}{1-\alpha\cdot\pi} - \phi, 1 - \phi\right)$*, then a center-seeking war occurs if* $\theta < 1 - \phi - \beta$ *and no war occurs otherwise.*

- *If* $\beta < \frac{(1-\alpha)\cdot\pi}{1-\alpha\cdot\pi} - \phi$*, then a center-seeking war occurs if* $\theta < 1 - \phi - \beta$*, a separatist war occurs if* $\theta \in \left(1 - \phi - \beta, \frac{1}{\alpha} \cdot \left(1 - \frac{\beta+\phi}{\pi}\right)\right)$*, and no war occurs otherwise.*

> **Proof.** If G can buy off C at $\theta = 0$, then war does not occur for any value of θ because C's expected utility to both types of war weakly decreases in θ. Proposition D.1 shows that C prefers center-seeking at $\theta = 0$. Therefore, if $\beta > 1 - \phi$, then war does not occur in equilibrium.
>
> If $\beta < 1 - \phi$, then whether or not war occurs and its aims depends on θ. C chooses a center-seeking civil war if $1 - \theta - \phi > \max\left\{\beta, (1 - \alpha \cdot \theta) \cdot \pi - \phi\right\}$, which simplifies to $\theta < \min\left\{1 - \phi - \beta, \frac{1-\pi}{1-\alpha\cdot\pi}\right\}$. C chooses a separatist civil war if $(1 - \alpha \cdot \theta) \cdot \pi - \phi > \max\left\{\beta, 1 - \theta - \phi\right\}$, which simplifies to $\theta \in \left\{\frac{1-\pi}{1-\alpha\cdot\pi}, \frac{1}{\alpha} \cdot \left(1 - \frac{\beta+\phi}{\pi}\right)\right\}$. C accepts G's offer if $\beta > \max\left\{1 - \theta - \phi, (1 - \alpha \cdot \theta) \cdot \pi - \phi\right\}$, which simplifies to $\theta > \max\left\{1 - \phi - \beta, \frac{1}{\alpha} \cdot \left(1 - \frac{\beta+\phi}{\pi}\right)\right\}$.
>
> The separatist range is non-degenerate if and only if $\frac{1-\pi}{1-\alpha\cdot\pi} < \frac{1}{\alpha} \cdot \left(1 - \frac{\beta+\phi}{\pi}\right)$. Algebraic rearranging shows this is true if and only if $\beta < \frac{(1-\alpha)\cdot\pi}{1-\alpha\cdot\pi} - \phi$. $\blacksquare$

Figure 1 depicts the thresholds from Proposition D.2. It sets $\phi = 0.4$, $\alpha = 0.6$, and $\pi = 0.8$. It sets $\beta = 0.6$ for the top solid gray line (high government commitment ability), and sets $\beta = 0.1$ for the lower solid gray line (low government commitment ability).

REFERENCES

AbuKhalil, As'ad. 1998. *Historical Dictionary of Lebanon*. Scarecrow Press.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8(2):155–194.

Alesina, Alberto, William Easterly and Janina Matuszeski. 2011. "Artificial States." *Journal of the European Economic Association* 9(2):246–277.

Altonji, Joseph G, Todd E. Elder and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113(1):151–184.

Anthony, Michael. 1997. *Historical Dictionary of Trinidad and Tobago*. Scarecrow Press.

Baxter, Craig and Syedur Rahman. 1996. *Historical Dictionary of Bangladesh, Section Edition*. Scarecrow Press.

Black, Jan Kippers and Edmundo Flores. 1989. Historical Setting. In *Panama: A Country Study*, ed. Sandra W. Meditz and Dennis M. Hanratty. Washington, D.C.: Federal Research Division, Library of Congress pp. 1–66.

Blood, Peter R. 1990. Historical Setting. In *Sri Lanka: A Country Study*, ed. Russell R. Ross and Andrea Matles Savada. Washington, D.C.: Federal Research Division, Library of Congress pp. 1–56.

Brecke, Peter. 1999. "Violent Conflicts 1400 AD to the Present in Different Regions of the World." 1999 Meeting of the Peace Science Society.

Burki, Shahid Javed. 1991. *Historical Dictionary of Pakistan*. Scarecrow Press.

Burrowes, Robert D. 2010. *Historical Dictionary of Yemen*. Rowman & Littlefield.

Commins, David. 1996. *Historical Dictionary of Syria*. Scarecrow Press.

Copper, John F. 2000. *Historical Dictionary of Taiwan (Republic of China)*. Rowman & Littlefield.

Corfield, Justin. 2010. "Historical Dictionary of Singapore, New Edition.".

Corfield, Justin J. and Laura Summers. 2003. *Historical Dictionary of Cambodia*. Scarecrow Press.

Dixon, Jeffrey S. and Meredith Reid Sarkees. 2015. *A Guide to Intra-State Wars: An Examination of Civil, Regional, and Intercommunal Wars, 1816-2014*. Washington, DC: CQ Press.

Easterly, William and Ross Levine. 2016. "The European Origins of Economic Development." *Journal of Economic Growth* 21(3):225–257.

Encyclopædia Britannica. 2019. "Encyclopædia Britannica Online.".

Englebert, Pierre. 1996. *Burkina Faso: Unsteady Statehood in West Africa*. Westview Press.

Ertan, Arhan, Martin Fiszbein and Louis Putterman. 2016. "Who was Colonized and When? A Cross-Country Analysis of Determinants." *European Economic Review* 83:165–184.

Fearon, James D. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8(2):195–222.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–90.

Fearon, James D. and David D. Laitin. 2013. "How Persistent is Armed Conflict?". Working paper, Department of Political Science, Stanford University. Available at https://web.stanford.edu/group/fearon-research/cgi-bin/wordpress/wp-content/uploads/2013/10/persistenceofarmedconflict.pdf. Accessed 11/16/17.

Ghareeb, Edmund A. and Beth Dougherty. 2004. *Historical Dictionary of Iraq*. Scarecrow Press.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg and Håvard Strand. 2002. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39(5):615–637.

Gubser, Peter. 1991. *Historical Dictionary of the Hashemite Kingdom of Jordan*. Scarecrow Press.

Guillermo, Artemio R. 2011. *Historical Dictionary of the Philippines*. Scarecrow Press.

Gunn, Geoffrey C. 2010. *Historical Dictionary of East Timor*. Scarecrow Press.

Hawley, Donald. 1970. *The Trucial States*. George Allen & Unwin.

Horn, Myron K. 2003. Giant Fields 1868-2003 (CD-ROM). In *Giant Oil and Gas Fields of the Decade, 1990-1999*, ed. M.K. Halbouty. Houston: AAPG Memoir 78.

Iliffe, John. 1979. *A Modern History of Tanganyika*. Cambridge University Press.

Kaur, Amarjit. 2001. *Historical Dictionary of Malaysia*. Scarecrow Press.

Lal, Brij V. 2015. *Historical Dictionary of Fiji*. Rowman & Littlefield.

Leonard, Thomas M. 2014. *Historical Dictionary of Panama*. Rowman & Littlefield.

Levine, Robert M. 1979. *Historical Dictionary of Brazil*. Vol. 19 Scarecrow Press.

Maddison, Angus. 2008. "Statistics on World Population, GDP and Per Capita GDP, 1-2008 AD.". Available at http://www.ggdc.net/maddison/oriindex.htm. Accessed 4/10/16.

Mahoney, James. 2010. *Colonialism and Postcolonial Development: Spanish America in Comparative Perspective*. Cambridge University Press.

Majul, Cesar Adib. 1999. "Muslims in the Philippines: A Historical Perspective.".

Mansingh, Surjit. 1996. *Historical Dictionary of India*. Scarecrow Press.

Marshall, Monty G. and Ted Robert Gurr. 2014. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2013." http://www.systemicpeace.org/polity/polity4.htm.

Marshall, Peter James. 1987. *Bengal: The British Bridgehead: Eastern India 1740–1828*. Vol. 2 Cambridge University Press.

McCann, Frank D. 1998. Historical Setting. In *Brazil: A Country Study*, ed. Rex A. Hudson. Washington, D.C.: Federal Research Division, Library of Congress pp. 1–86.

McFarland, Daniel Miles and Lawrence Rupley. 1998. "Historical Dictionary of Burkina Faso.".

Mukhtar, Mohamed Haji. 2003. *Historical Dictionary of Somalia*. Scarecrow Press.

Naylor, Phillip C. 2015. *Historical Dictionary of Algeria*. Rowman & Littlefield.

Nazzal, Nafez and Laila A. Nazzal. 1997. *Historical Dictionary of Palestine*. Scarecrow Press.

Ofcansky, Thomas P. and Rodger Yeager. 1997. *Historical Dictionary of Tanzania*. Scarecrow Press.

Oliver, Roland and Anthony Atmore. 2005. *Africa Since 1800*. Cambridge, UK: Cambridge University Press.

Olsson, Ola. 2009. "On the Democratic Legacy of Colonialism." *Journal of Comparative Economics* 37(4):534–551.

Paine, Jack. 2019*a*. "Democratic Contradictions in European Settler Colonies." *World Politics* 71(3):542–585.

Paine, Jack. 2019*b*. "Ethnic Violence in Africa: Destructive Legacies of Pre-Colonial States." *International Organization* 73(3):645–683.

Parke, Aubrey. 2014. Fijian Society: The Islands of Fiji (General). In *Degei's Descendants: Spirits, Place and People in Pre-Cession Fiji*, ed. Matthew Spriggs and Deryck Scarr. Australian National University Press pp. 1–17.

Peebles, Patrick. 2015. *Historical Dictionary of Sri Lanka*. Rowman & Littlefield.

Perkins, Kenneth J. 1997. *Historical Dictionary of Tunisia, Second Edition*. Scarecrow Press.

Putterman, Louis. 2008. "Agriculture, Diffusion and Development." *Economica* 75(300):729–748.

Ray, Subhasish. 2016. "Sooner or Later: The Timing of Ethnic Conflict Onsets After Independence." *Journal of Peace Research* 53(6):800–814.

Ross, Michael L. and Paasha Mahdavi. 2015. "Oil and Gas Data, 1932–2014. Harvard Dataverse.". Harvard Dataverse. Accessed 7/17/17.

Sambanis, Nicholas. 2004. "What is Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48(6):814–858.

Scarritt, James R and Shaheen Mozaffar. 1999. "The Specification of Ethnic Cleavages and Ethnopolitical Groups for the Analysis of Democratic Competition in Contemporary Africa." *Nationalism and Ethnic Politics* 5(1):82–117.

Stuart-Fox, Martin. 2001. *Historical Dictionary of Laos*. Scarecrow Press.

Suchlicki, Jaime. 2002. Historical Setting. In *Cuba: A Country Study*, ed. Rex A. Hudson. Washington, D.C.: Federal Research Division, Library of Congress pp. 1–88.

Turner, Ann. 1994. *Historical Dictionary of Papua New Guinea*. Scarecrow Press.

Vandewalle, Dirk. 2012. *A History of Modern Libya*. Cambridge University Press.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." *Journal of Conflict Resolution* 59(7):1327–1342.

Warren, James Francis. 2007. *The Sulu Zone, 1768-1898: The Dynamics of External Trade, Slavery, and Ethnicity in the Transformation of a Southeast Asian Maritime State*. National University of Singapore Press.

Worden, Robert L. 1991. Bhutan. In *Nepal and Bhutan Country Studies*, ed. Andrea Matles Savada. Washington, D.C.: Federal Research Division, Library of Congress pp. 251–344.