1. What is a Civil War and Why Study It

When domestic political conflict takes the form of military confrontation or armed combat we speak of civil war. This is a destructive development: the mean number of deaths in the 146 civil wars that took place between 1945 and 1999 is 143,883 (Sambanis 2004b). Besides direct fatalities, civil war causes many more indirect ones through mass dislocation, epidemics, famines, and the degradation of the state apparatus. Economic costs are also massive, both directly and indirectly. Economic development is stalled or, even, reversed.

Civil war is also a phenomenon prone to serious semantic confusion, even contestation. The description of a conflict as a civil war carries symbolic and political weight since the term can confer or deny legitimacy to a warring party. Indeed the very use (or not) of the term is part of the conflict itself. This is why euphemisms are so common. Civil war is often described through such terms as Troubles, Emergency, or Situation, while rebels are typically described as bandits or, more recently, terrorists (and some civil wars are presented as being instances of the “war on terror”). This sort of semantic and political contestation accounts, in great part, for the fact that the systematic study of civil wars is a rather recent development.

Besides the effects of political contestation, the study of civil war suffered from conceptual competition by cognate phenomena such as revolution and ethnic conflict. Until recently, the study of revolutions was privileged by researchers. This may have reflected a normative preference, or at least sympathy, for social revolutions, whereas civil war is hardly held as a desirable outcome. In the early 1990s, the

*I am grateful for helpful comments and suggestions by Laia Balcells, Nicholas Sambanis, Elisabeth Wood, and the volume’s editors.
violence that broke out following the break-up of the Soviet Union and Yugoslavia caused renewed focus on “ethnic conflict.” The realization that both social revolution and ethnic conflict were part of a broader concept, formerly known as rebellion and now defined as civil war, ushered in a new era of research in the mid-1990s. It is commonplace to state that the study of civil war by political scientists has since boomed. It is much less commonplace, however, to ask why this boom has occurred in the first place.

The boom in the study of civil war has three sources. First, development economists specializing in the study of African economies and funded primarily by the World Bank sought to make the case that civil war was a major impediment of economic development. Second, the quasi-disappearance of interstate wars led scholars of international relations and international security specializing in the study of war to shift their focus to the one instance of war practiced today, namely civil war. Third, the resurgence of ethnic conflict during the early post-Cold War years led students of ethnicity, including sociologists and comparativist political scientists, to focus on all types of intrastate conflict, rather than ethnic conflicts alone.

These three “sources” of interest in civil war correspond to three distinct styles of research: an economic, an international relations (IR), and a comparativist style. While recognizing the complexity of the issue and acknowledging that the causes of civil wars are multiple, economists have primarily stressed the impact of natural resources, IR scholars have pointed to ethnic antagonism, and comparativists have focused on the state.

Several definitions of civil war exist, but they converge around the same key dimensions of the phenomenon. Civil war can be defined as armed combat taking place within the boundaries of a recognized sovereign entity between parties subject to a common authority at the outset of the hostilities (Kalyvas 2006). This definition stresses two key features: the militarization of conflict, requiring at least two competing sides (including a relatively large rebel organization with military equipment and full-time recruits) and differentiating civil war from communal riots, terrorism, crime, and genocide; and a domestic challenge directed against the authority of the current holder of sovereign authority, which distinguishes it from interstate war. In fact internal war (Eckstein 1965) would be a more precise term, but civil war is the dominant term because of its common usage.

Definitional consensus conceals considerable disagreement about operationalization and, hence, divergence in coding practices. A major issue is the definition of internal conflicts crossing the threshold of war. Studies have relied on various fatality thresholds, primarily inspired by the coding rules used in the first major dataset, the Correlates of War (COW), which is also the basis on which most subsequent datasets were built. Disagreements include whether fatality counts are absolute or relative.

---

1 The size of the rebel organization may vary but cannot be reduced to a few dozens of clandestine combatants.
("per capita"), whether they are cumulative over time or yearly, whether fatalities include battle-related deaths only or civilian deaths only (or both), and the distribution of fatalities between competing sides. Coding decisions are crucial in determining the onset and termination of civil wars and distinguishing between a single ongoing but intermittent war or a succession of several distinct ones. The problem is exacerbated by inconsistencies in coding within the same datasets, problematic categories (e.g. “extra-systemic wars”), and the well-known unreliability of fatality data from civil wars: most civil wars take place in impoverished countries where record-keeping bureaucracies are lacking. There is also debate about the usefulness of relying on a dichotomous characterization (war versus peace) versus a continuous conceptualization that would better capture intensity levels, and whether civil war is an independent phenomenon (conceptually speaking) as opposed to being a single “value” of the larger phenomenon of political violence.

2. Macro Findings and Debates

The main method used to identify the determinants of civil war is the statistical analysis of data on all country-years since 1945. The pool of independent variables that have been identified as potential determinants of civil war onset includes the level of economic development, political instability, ethnic heterogeneity, the presence of plentiful natural resources, a history of conflicts, war-prone and undemocratic neighbors, high infant mortality, small military establishments, political regimes that are neither dictatorships nor democracies (“anocracy”), mountainous terrain, large population, diasporas, and oil production. Moreover, there seem to be geographic and time effects (the Middle East and North Africa and the 1960s). Several econometric models have identified a number of factors as potential determinants of civil war onset. Collier et al. (2003, 53–4) summarize a great part of the findings in the following way: “Countries with low, stagnant, and unequally distributed per capita incomes that have remained dependent on primary commodities for their exports face dangerously high risks of prolonged conflict. In the absence of economic development neither good political institutions, nor ethnic and religious homogeneity, nor high military spending provide significant defenses against large-scale violence. Once a country has stumbled into conflict powerful forces—the conflict trap—tend to lock it into a syndrome of further conflict.”

For all its succinctness, this statement conceals several disagreements of emphasis and interpretation. The effect and interpretation of natural resources, geography, ethnic heterogeneity, regime type, inequality, or diasporas has been widely debated (Cederman 2004; Sambanis 2001). Divergence between econometric specifications, estimation methods, measurement procedures, and datasets makes it very hard to evaluate in a definitive way the effect of each variable and arrive at a definitive theory
of civil war onset. Endogeneity remains a major concern. Unfortunately, the identification of instruments remains either extremely difficult or limited, both chronologically and geographically (Miguel, Satyanath, and Sergenti 2004). Perhaps more importantly, several competing and observationally equivalent underlying causal mechanisms account for the observed effect of perhaps the most important variable, poverty. To mention just two, poverty is consistent with both low opportunity costs for joining a rebellion and high grievances and desire for social redress—i.e. opportunity and intention to rebel.

Nevertheless, it is possible to bundle different sets of variables into three types of theoretical arguments, loosely linked to the three schools of analysis outlined above. The story told by international relations scholars stresses ethnic heterogeneity, development economists emphasize the role of natural resources; and comparativist scholars point to the role of state capacity. According to the first version, civil war is primarily an expression of nationalist aspirations and ethnic disputes; the second version stresses civil war as a phenomenon taking place in countries that suffer the curse of plentiful natural resources, and the third one points to weak states as the main prediction of civil war.

Disputes between ethnic groups and nationalism figure prominently in the descriptive and theoretical literature but occupy a minor place in the recent econometric one. Various scholars (Wimmer and Min 2006; Toft 2003; Sambanis 2001; Posen 1993) have pointed at the role of ethnic divisions in leading to civil war. These arguments focus on (ethnic) group dynamics and group demographies. There are at least three stylized stories (and likely several more). According to the first one, state collapse creates a “security dilemma:” in the absence of a state conflict uncertainty about the intentions of the other group inevitably leads to conflict. Since neither group knows the other’s intentions, each has an incentive to build up defensive capabilities to protect itself from an attack by the other group. However, since most defensive capabilities can be used offensively, defensive build-up can appear as signaling aggressive intentions (Posen 1993). According to the second one, civil war is caused by a “commitment problem” that arises when two groups find themselves without a third party that can credibly guarantee agreements between them (Fearon 1998). The third story argues that civil war is caused by ethnic secessionists wishing to carve their own separate state via their capacity to articulate a military challenge based on existing ethnic networks reinforced by patterns of (ethnic) population concentration in specific territory (Toft 2003). Using a global dataset including fixed geographical territories from 1816 to 2001 independent of the political entity in control of the territory in a specific year, Wimmer and Min (2006) expand this last argument over time and find that both interstate and civil war are intimately related to the twin transformations of the rise of empires and the rise of nation-states. They also find that the rise of nation-states is related to civil war onset through political

---

2 In his comparison of several influential datasets, Sambanis 2004b finds that most findings are not robust.
discrimination along ethnic lines and the subsequent demand of new ethnically homogeneous states.

These arguments have been found wanting by the econometric literature because the main indicator used to capture ethnic antagonism, the ethnolinguistic fractionalization index (ELF), tends to be statistically insignificant. However, the jury remains out for a number of reasons. First, ELF has been credibly criticized as inherently problematic (Posner 2004) and unable to capture (or even misrepresenting) the subtle channels through which ethnicity results in civil wars at best. Second, Sambanis (2001) claims that ethnicity should predict the causes of ethnic civil wars, not all civil wars—and argues that once this is taken into account, ethnicity is a predictor of the ethnic civil war onset. Third, it is argued that the relevant indicator is ethnic polarization rather than ethnic fragmentation. Montalvo and Reynal-Querol (2005) show that while ethnic fragmentation is an insignificant factor in civil war, ethnic polarization is highly significant. In a similar vein, Cederman and Girardin (2006) argue, and present some evidence to the effect, that the effect of ethnicity should operate through the channel of ethnic minority rule. Third, existing ethnicity indicators have been criticized as over-aggregate. Buhaug, Cederman, and Rød (2006) geo-coding ethnic demographies to produce a better index of ethnicity on the ground. Sambanis and Milanovic (2006) have been, likewise, at work to produce data using region-years as opposed to country-years.

A more general and theoretical criticism of the effect of ethnicity is that the concept of ethnic conflict is in itself flawed and that ethnicity is constructed and quite mutable (the “constructivist” claim). This is not to deny that ethnicity plays a part in conflict but that it is insufficient to look at the presence of groups (ethnic, religious, sectarian, etc.) to explain the onset of conflict, because the salience of ethnic identities can be itself the result of the conflict. In other words, the salience of ethnicity and the animosity between ethnic groups may be an outcome of the conflict rather than its cause.

Turning to the second argument, Paul Collier et al. (2003; also Collier and Hoeffler 2004) stress the effect of abundant natural resources. While poverty reduces opportunity costs for participation in rebellion, natural resources allow the financing of rebellion which may start with political aims but eventually becomes criminal organizations. The stylized story underlying these findings should be familiar to students of African politics: impoverished countries with large reserves of natural resources (particularly diamonds and oil) generate incentives for rebellion; once rebellion is on, it sustains itself precisely from these resources. Unemployed individuals participate in insurgency primarily driven by the prospect of loot.

The problems with this argument include the proxy used to capture the effect of natural resources (i.e. primary commodity exports), the opaqueness of underlying mechanisms, the lack of empirical validation for the microfoundations posited, and the theoretical assumption informing this research program (Cramer 2002).

3 But see Fearon, Kasara, and Laitin 2006 for counter-evidence.
Humphreys (2005) points to a host of competing mechanisms underlying the correlation between primary commodities and civil war onset: greedy rebels; greedy outsiders (natural resources as an incentive for foreign corporations and states to engage in or even foster the conflict); grievances (dependence on natural resources may cause inequality, vulnerability to trade shocks, processes of social dislocation such as forced migration, and unfair distribution of natural resources); insurgency financing; weakening of states through the “resource curse;” sparse networks. Likewise, Ross (2006) catalogs a host of additional problems of measurement error, spuriousness, endogeneity, and lack of robustness. To cite just one, it may be that poor property rights or weak rule of law cause both dependence on natural resources (through dissuasion of investment in other sectors) and civil war. Collier and Hoeffler (2004, 567) recognize the problems of observational equivalence when they point out that “primary commodities are associated with other characteristics that may cause civil war, such as poor public service provision, corruption and economic mismanagement … Potentially, any increase in conflict risk may be due to rebel responses to such poor governance rather than to financial opportunities.” Qualitative studies have also pointed to the many pitfalls of assuming rather than researching causal mechanisms. For instance, Gutierrez (2004) shows that given the lack of material selective incentives, the sanctions against plundering, and the harsh demands that the Colombian FARC imposes on its members, this group does not fit the “criminal rebels” thesis despite its strong dependence on an illicit natural resource, namely coca. A related problem is the tendency that characterizes this line of research, of extrapolating from contemporary African civil wars to the full set of civil wars, past and present.

In response to these criticisms there have been several attempts to disaggregate natural resources. For example, Lujala, Gleditsch, and Gillmore (2005) propose indicators of natural resources endowment that distinguish lootable from non-lootable natural resources; they specifically distinguish between two types of diamonds (primary or non-lootable and secondary or lootable) and test whether they are diversely associated with civil war onset and incidence. Ross (2006) follows a similar disaggregation course. He constructs more accurate and exogenous measures of oil, diamond, and other mineral wealth, distinguishing non-fuel rents per capita, fuel onshore rents per capita, fuel offshore rents per capita, primary diamonds production per capita, and secondary diamonds production per capita. Results remain inconclusive.

Last, a third story points to the role of (repressive) state capacity, as proxied by GDP per capita, along with conditions favoring rural insurgency, as proxied by an indicator of mountainous terrain (Fearon and Laitin 2003). Both grievances and greed may motivate leaders and followers, but unless they are able to exploit the weakness of the state, they are unable to translate their preferences into civil war. In other words, potential insurgent leaders are more likely to launch rebellions when

---

4 Including the two mechanisms proposed by Collier et al. (greedy rebels and insurgency financing).
they have a better chance of success. One such condition is a state’s inability to fund sufficient police and administrative presence in its hinterlands. The stylized story is one of weak states that police mountainous peripheries poorly and, when rebellions erupt, badly. If a limited rebellion erupts and the state responds with indiscriminate violence, it may feed it rather than stop it. More generally, Fearon and Laitin (2003, 88) have argued that the civil wars of the post-1945 period have structural roots, in the combination of the military technology of guerrilla warfare and decolonization, which created an international system numerically dominated by fragile states with limited administrative control of their peripheries. Hironaka (2005) provides a similar account, stressing the long-term effects of decolonization. Like the previous arguments, this one faces the problem of assuming rather than demonstrating its causal mechanism. As pointed out above, poverty can be an indicator of greed, grievance, and state capacity.

Furthermore, GDP per capita is highly correlated with variables, such as population density or urbanization, which may account for both the weakness of counterinsurgent effort and low state capacity (Kocher 2004). The interpretation of the “rough terrain” variable as signaling low state capacity is also open to questioning since it remains unclear whether countries with rough terrain experience civil wars in precisely those areas (Sambanis 2004a).

Obviously, these three arguments do not exhaust the theoretical accounts of civil war onset. Though unfashionable, grievances also keep attracting theoretical and empirical attention (Regan and Norton 2005). Various accounts of grievance have been proposed. Boix (2004) finds that civil wars are caused by a combination of inequality and capital mobility, while Gurr (1970) argues that relative deprivation through the mechanism of rising expectations, rather than inequality and poverty, causes rebellion and revolutions. Obviously, the problems with grievances as a determinant of civil war onset is that they seem to be much more prevalent than civil war and that they are very hard to measure directly. Ultimately, it is difficult to escape the conjecture that one has to look for combinations of demand for, and supply of, rebellion (or intention and opportunity).

To illustrate some of the complexities in figuring out and sorting out competing causal mechanisms, I discuss below the issue of the relation between a country’s rural dimension and civil war onset.

3. The Rural Dimension

Civil war is associated with a social dimension that is poorly understood and inadequately studied: the rural dimension. Poor societies tend to be rural and insurgencies tend to begin and are fought primarily in the rural countryside (Tong 1991; Brustein and Levi 1987). In contrast, the types of political violence experienced
by developed Western democracies tend to take the form of terrorism in primarily urban settings. Mass ethnic riots seem to be an outlier in this respect as they affect primarily urban areas of poor countries (Varshney 2003; Wilkinson 2004).

However, the rural dimension is simultaneously consistent with several causal mechanisms of civil war onset, including grievances resulting from unequal land distribution, worsening land distribution, or crop failure; the ability of insurgents to hide among rural populations without being denounced because of local norms of solidarity and honor; higher levels of tolerance among rural people to threats of violence; a tradition of peripheral rebellion reinforced by norms of reciprocity which leads to mass participation in anti-state activities ranging from contraband smuggling and banditry to full-fledged rebellion; the fact that an economy based on subsistence farming tends to favor armed resistance more than one based on wage labor, and a pattern of human ecology whereby the dispersion of population settlements in rural environments impedes policing: it is easier to enforce a curfew in a town than in a large rural area because taxing and monitoring hundreds, or even thousands, of hamlets exposes small army detachments to ambush (Kitson 1960, 12; Escott and Crow 1986, 376; Gambetta 1993, 109; Tone 1994, 162–6; Nordstrom 1997, 99; Horton 1998, 126).

A large literature developed mostly in the 1970s, but currently marginalized, has addressed the politics of mass rural rebellion. It is primarily structural, in that it seeks to link forms of ownership and land distribution to the emergence of large-scale peasant rebellion, though some authors have focused on local and individual micro-mechanisms (Stinchcombe 1961; Moore 1966; Wolf 1973; Paige 1978; Popkin 1979; Anderson 1993). According to Scott (1976), peasants rebel motivated by grievances originated in a combination of economic and political structural conditions; peasant behavior is thought to be determined by a set of values related to the right of subsistence and the right and duty of reciprocity that are rooted in the “existential situation” that peasants face. One prediction of this literature is that countries with large numbers of landless peasants, be they agricultural workers or sharecroppers, are likely to spawn violent political unrest. A contradictory prediction is that small landowners may be the source of violent agitation. There are many variations pointing to particular groups such as squatters or migrants (Anderson 1994). Finally, some arguments point to crop types as the relevant dimension. An intriguing implication is that globalization can exacerbate rural conflict by setting wealthy tariff-protected farmers from the West against poor peasants of the developing world (Kirschenmann 2003).

Theorization of the relation between inequality (in general) and violence has a long tradition, but only recently has there been some, admittedly broad, empirical testing of this conjecture (Boix 2003). The evidence is mixed. On the one hand, land property patterns are very hard to measure and are, therefore, usually excluded from econometric studies; on the other hand, the case study evidence is inconclusive (Wood 2003; Wickham-Crowley 1992).

State capacity arguments assume that the rural dimension proxies for processes primarily related to irregular war rather than pre-war grievances. Rural areas tend to
be difficult to police in the first place, which is why insurgencies are likely to be concentrated there. Though grievances may still matter, they cannot result in violent mobilization alone because they are likely to be repressed. Put otherwise, state repression is off the equilibrium path.

This insight allows the reinterpretation of some findings that take ideology or ethnicity as the main causal variables of violence. For example, Gulden (2002) finds that in Guatemala over half of the army killings took place in municipalities in which the Mayas made up between 80 and 90 percent of the population; based in part on this finding, he claims that this instance of mass violence constitutes genocide. However, these municipalities are mostly rural and located far from centers of government control. They could have just as easily been targeted because they were located in areas of guerrilla presence as because they were Mayan. This raises the issue of endogeneity of grievances: did the guerrillas pick their location based on the presence of Maya grievances or did they educate the Mayas who just happened to live in terrain that favored insurgent activity about their plight? Empirical evidence supplied by Stoll (1993, 87) allows a partial separation of the two: the army’s repression did not focus on areas where indigenous organizations (and presumably grievances) were strong but guerrillas had little presence, but rather in areas where the guerrillas were trying to organize despite weak indigenous organizations. In fact, the four areas of greatest government violence follow the insurgents’ swath as it moved south to cut the Pan-American Highway. Trejo (2004) provides additional fine-grained evidence from Mexico, linking the action of the Catholic Church among indigenous communities and the Zapatista insurgency.

An important insight from case studies that goes in a similar direction is that geography may trump pre-war allegiances—an insight that would question the grievances causal mechanism. For example, Confederate guerrillas were strong in the Appalachians, the Cumberlands, and the Ozarks, in the very areas within the Confederacy which most Union sympathizers inhabited (Beckett 2001, 11); following their defeat in the cities, the Chinese communists staged a comeback from backward and isolated “border areas” where their pre-war support was minimal if not nonexistent (Schran 1976); the urban populations in the German-occupied Soviet territories were more likely than rural ones to dislike the occupying authorities, partly because of their closer earlier identification with the Soviet regime and partly because of the more miserable conditions of life and work in the towns; yet, “paradoxically, the partisan movement was largely a rural phenomenon” (Dallin, Mavrogordato, and Moll 1964, 335). The French communist FTP guerrillas were very successful in the rural areas that exhibited very limited pre-war communist support (Kedward 1993, 131). Likewise, the Renamo insurgency against the Frelimo government in Mozambique developed in the same areas where the Frelimo anticolonial insurgency had been strong; in contrast, areas that supported the Portuguese incumbents during the anticolonial war tended to side with the Frelimo incumbents during the Renamo insurgency (Nordstrom 1997, 98–9; Geffray 1990, 41). A high-ranking American officer serving in the Dominican Republic in 1921 argued that the construction of roads would stifle the insurgency: “A highway would bring the people more in
contact with the Capital, thus giving the Central Government an opportunity to control political conditions” (in Calder 1984, 164).

Further confirming the importance of military resources in generating control and hence collaboration is the oft-noted propensity of villages located near central roads to collaborate with incumbents (Sansom 1970, 60–1; Kriger 1992, 208). Whereas “modernizing” villages near main roads in Vietnam had been among the first to respond to revolutionary appeals, they were also more likely to be controlled by the government and “as the risks of political action escalated during the middle and late 1960s, the gap between political attitudes and behavior widened, and many revolutionary sympathizers became inactive when the dangers became too great or, in some cases, adopted a clandestine role so deeply hidden that it often amounted to a temporary cessation of revolutionary activities” (Elliott 2003, 589). The availability of external support for insurgents turns the combination of terrain and proximity to borders into a strong predictor of insurgent control, especially when it comes to ethnic insurgencies (Toft 2003).

The Nicaraguan case allows a type of natural experiment, insofar as it is possible to compare the behavior of the Sandinistas in their successive roles as insurgents and, later, as incumbents. This comparison suggests that popular allegiances were often endogenous to the exercise of territorial control. During the “contra” phase of the war, the (incumbent) Sandinistas firmly controlled the towns but were absent from the mountains: “The only Sandinista presence in the mountains would be a military one” (Horton 1998, 137). As a result, people in those areas supported the contras. In contrast, many mountainous zones, which now collaborated with the contras, had supported the Sandinista guerrillas in the 1970s, even when they were urban activists who had fled the cities for the countryside (1998, 21–2). The opposite is true of the towns, which were controlled by the (Somozista) incumbents in the first phase of the war and the Sandinista (incumbents) in the second one. In Horton’s (1998, 21) words: “Hundreds of Sandinista Army soldiers were stationed in the town of Quilalí and as a result the town itself always remained firmly under FSLN control” (emphasis mine). In other words, whereas the Sandinistas qua insurgents based themselves in inaccessible rural terrain, they faced themselves limited to cities when, qua incumbents, they faced the contra insurgency.

The issues illustrated by the discussion above point to the importance of studying closely the interaction of military, social, and political dynamics of civil wars. At the same time, it is important to recognize that even if individual civil wars can be bundled under the same conceptual category, there may be significant differences between subcategories. For example, Sambanis (2001) has suggested that ethnic civil wars may have different causes than non-ethnic civil wars. The same insight can be derived with regard to state capacity. The fact that there is a civil war in a given country signals (in a tautological sense) that the incumbent state is somehow weak or lacks state capacity. But what does state weakness or insufficient state capacity mean exactly? Consider the Russian state during the late 1990s, facing the Chechen insurgency, the Turkish state of the 1980s, facing a Kurdish insurgency in its eastern periphery, the Liberian state which was facing several peripheral insurgencies during
the early 1990s, and the fledgling Bosnian state, which was facing the Serb secessionist challenge throughout its territory in the early 1990s. All these states were fighting civil wars, and they were all “weak” by the mere fact that they had been unable to prevent these civil wars from erupting. Yet, no one would seriously contend that Russia, Turkey, Liberia, and Bosnia were equally weak.5

4. Type of Onset and Warfare

A way to address this problem is, perhaps, to introduce a measure of induction and distinguish between types of civil war. The problem, of course, is to do so in a way that is informed by either theory or solid empirics, rather than current events.6 For example, Fearon (2004) sets out to distinguish between five types of civil wars in order to explain different patterns of duration. Fearon identifies three types of brief civil wars (civil wars arising out of military coups and popular uprisings, anticolonial wars, and wars arising out of the collapse of the Soviet Union and Yugoslavia) and one type of long civil war (peripheral insurgencies relying on guerrilla warfare) with two particular subtypes (“sons of the soil wars,” i.e. wars between peripheral ethnic minorities and state-supported migrants of a dominant ethnic group, and conflicts where the rebel group has access to natural resources). This distinction has undoubtedly an important heuristic value, but mixes analytical criteria (the war’s origins) with more contextual criteria (the collapse of the Soviet Union and Yugoslavia).

Wars can be classified in many ways: some stress the primary actors involved (e.g. international or domestic), their goals (e.g. offensive or defensive), their world views and societal projects (“greed and grievance”), and so on. A popular way to classify civil wars is by stressing the war’s “master cleavage.” This is the basis, for instance, of the distinction between ethnic and non-ethnic wars. The problem is that uncovering one master cleavage turns out to be much more difficult than seems at first sight because civil wars are highly complex process, where one cleavage potentially hides another (Kalyvas 2003). Of course, complexity is part and parcel of most political and social phenomena and the goal of social scientific explanation is to reduce complexity. However, the problem is magnified in civil wars. Consider the following description of the civil war in Sudan by the journalist Deborah Scroggins (2004: 79–80):

Furthermore, states such as Burkina Faso or Equatorial Guinea would certainly be classified as “weak” by country experts, yet they somehow manage to avoid civil wars. A problem faced by state capacity arguments is to explain how weak states manage to remain stable much more frequently than they face military challenges.

As is the case with the popular distinction between “new” and “old” civil wars (Kaldor 1999) which was inspired by journalistic accounts of ongoing civil wars and a superficial reading of the historiography of past civil wars (Kalyvas 2001).
I have often thought that you need a kind of layered map to understand Sudan’s civil war. A surface map of political conflict, for example—the northern government versus the southern rebels; and under that a layer of religious conflict—Muslim versus Christian and pagan; and under that a map of all the sectarian divisions within those categories; and under that a layer of ethnic divisions—Arab and Arabized versus Nilotic and Equatorian—all of them containing a multitude of clan and tribal subdivisions; and under that a layer of linguistic conflicts; and under that a layer of economic divisions—the more developed north with fewer natural resources versus the poorer south with its rich mineral and fossil fuel deposits; and under that a layer of colonial divisions; and under that a layer of racial divisions related to slavery. And so on and so on until it would become clear that the war, like the country, was not one but many: a violent ecosystem capable of generating endless new things to fight about without ever shedding any of the old ones.

In other words, to analyze the civil war in the Sudan as just (or primarily) an ethnic or a religious war is problematic. An alternative is to rely on analytical criteria based on dimensions of the conflict that combine origins and dynamics. A relevant distinction in this respect is between civil wars associated with processes of state implosion at the center and those associated with peripheral challenges. This distinction has the advantage of combining the criterion of origin with that of warfare, which is essential in understanding the way in which the war is organized and sustained.

A common empirical observation in the literature on civil wars is that most of them are fought by means of irregular (“guerrilla”) warfare rather than conventional warfare. A few civil wars mix irregular and conventional warfare (e.g. Russia, China, Vietnam), while a very small number are fought fully or predominantly as conventional wars (e.g. Spain). All in all, conventional civil wars are “rare instances appearing only under specific and rather exceptional circumstances” (Derriennic 2001). In contrast, almost all interstate wars are fought conventionally. In short, there is a high degree of overlap between civil and non-conventional war on the one hand, and interstate and conventional war, on the other. It follows that the study of civil war must incorporate the dimension of warfare. This is, to a degree, the insight that provides the main microfoundation for the state capacity story (Fearon and Laitin 2003).

The distinction between irregular and conventional war is common and widely accepted, though the terminology varies. Like all distinctions, it is an ideal-typical one with the two types’ edges blending into each other. Conventional warfare entails face-to-face confrontations between regular armies across clear frontlines. This type of warfare requires a commonly shared perception of a balance of power between the two sides in the sense that they are both willing to face each other conventionally, across clearly defined frontlines. In the absence of some kind of mutual consent (which entails some reasonable belief in future victory), no conventional battle can take place (Beaufre 1972). On the other hand, irregular war is a type of warfare that requires a choice by the strategically weaker side “to assume the tactical offensive in

---

Footnote: The very few irregular interstate wars consist mostly of low-intensity border skirmishes, such as the Libya–Chad war and the war between Belize and Guatemala (Harkavy and Neuman 2001).
selected forms, times, and places” (Simons 1999, 84)—in other words to refuse to match the stronger side’s expectations in terms of the conventionally accepted basic rules of warfare. A stylized description of irregular war goes as follows: the state (or *incumbents*) fields regular troops and is able to control urban and accessible terrain, while seeking to militarily engage its opponents in peripheral and rugged terrain; challengers (rebels or *insurgents*) hide and rely on harassment and surprise. Such wars often turn into wars of attrition, with insurgents seeking to win by not losing while imposing unbearable costs on their opponent. There are many variations to this stylized scenario, involving outside intervention or assistance that may lead the insurgents to gradually switch from irregular war to conventional war (e.g. China); conversely, the progressing deterioration of the state may force incumbents to opt for irregular war as well (e.g. Sierra Leone).

In short, irregular warfare is a manifestation of military *asymmetry* between actors—both in terms of their respective power and their ensuing willingness to fight on the same plane: the weaker actor refuses to directly face the stronger one. The main empirical indicator of irregular war is the dearth of large-scale direct military confrontations or “set battles” and the absence of frontlines. Irregular war is not wedded to a specific cause (revolutionary, communist, or nationalist) but can be deployed to serve a very diverse range of goals. Of course, asymmetry is not an exclusive feature of irregular war; it is also compatible with other forms of violence, including the “terrorist” use of indiscriminate violence.

While asymmetry is predominantly expressed in irregular war, the converse is not the case, as often implied: symmetry (or parity) is not synonymous with conventional war. Rather, it is possible to point to a type of warfare that often gets confused with irregular war, which can be dubbed “symmetric non-conventional warfare” (Kalyvas 2005). This type of warfare is often described as “primitive” or “criminal” war and entails irregular armies on both sides in a pattern resembling pre-modern war. Hence the following conjecture: conventional civil war emerges either out of failed military coups or secession attempts in federal or quasi-federal states);

irregular war results from peripheral or rural insurgencies (which may or may not be secessionist in intent); and “symmetric non-conventional warfare” takes place in civil wars that accompany processes of state implosion. State implosion can be sudden or gradual; a way to identify this process is by examining the state of the government army and whether it has become indistinguishable from rival militias in terms of loose organization and fractured chain of command.

In a different formulation, it could be hypothesized that conventional and “symmetric non-conventional” wars tend to result from processes of state implosion, whereas peripheral or rural insurgencies are the likely products of processes whereby the authority of the central state is challenged.

More specifically, conventional civil wars take place when an existing army splits either because of a failed coup (e.g. Spanish Civil War) or because a unit of a federal

---

8 By “quasi-federal” states, I mean states that have devolved a substantial degree of their military authority, particularly through the creation of extensive local and regional militias.
or quasi-federal state, which can claim control over a substantial part of the state’s
armed forces, attempts to secede (e.g. the American Civil War, the Biafran War).
High levels of external support or external intervention in favor of the rebel side may
turn an irregular war into a conventional one: this was the case during the late phases
of the Chinese Civil War and the Vietnam War.

Second, irregular civil wars emerge incrementally and often slowly from the
periphery. They entail a slow and patient process of state building by the rebels.
Geography plays a key role in their onset and conduct. Examples include civil wars in
Malaya, Mozambique during the Portuguese colonization, Kashmir, Aceh (Indonesia), and elsewhere.

Last, “symmetric non-conventional” wars are much less studied and understood;
in fact, they are often bundled with rural guerrilla wars. These wars are fought on
both sides by irregular armies following a process of state implosion. This entails the
disintegration of the state army and its replacement by rival militias which typically
equip themselves by plundering the arsenal of the disbanded army. This type of
warfare differs from conventional civil war because it lacks regular armies and set
battles. At the same time, it differs from irregular war because it displays clear
frontlines. The presence of frontlines, which takes various forms (including road-
blocks and checkpoints), has been stressed in many descriptions of symmetric non-
conventional wars. War often consists of roving militias raiding “enemy” territory
along with killing and plundering. Of course, indiscriminate violence by regular
armies in other types of war often takes the same form (think of the counter-
surgency operations of the German and Japanese armies in occupied countries
during the Second World War). In symmetric non-conventional wars, however, this
becomes the main form of warfare by all sides. Examples include the Lebanese Civil
War, the wars in Congo-Brazzaville, Somalia, Liberia, Sierra Leone, and many civil
wars that erupted in the wake of the Soviet Union’s collapse (Derluguiian 2005). In
some cases, these wars are concluded when a faction manages the transition between
“roving” to “stationary bandit,” thus becoming a state builder. Exactly when and how
this takes place remains to be studied. A fruitful direction of future research along
these lines would consist of relating these types of civil war with various outcomes
(onset, duration, types of settlement, violence, etc.). For example, it could be
hypothesized that conventional civil wars are likely to be longer than coups but
shorter than irregular wars; or that third-party intervention is much easier in
asymmetric, non-conventional wars compared to irregular ones. Likewise, it would
seem that “greed” arguments would correlate with asymmetric non-conventional
wars and state capacity ones with irregular wars. Ethnic animosity arguments could
Correspond to all three types of war: conventional war if minority ethnic groups are
well represented in the state’s army, irregular war if minority ethnic groups
are concentrated in the country’s periphery, or asymmetric non-conventional war
if the war follows a process of state collapse.

In sum, there is a possible relation between the process of civil war onset and the
form of warfare characterizing the war. If this characterization holds it could be
consequential from both a theoretical and a policy perspective. Clearly, we need a
better understanding of warfare, not just onset, especially given the host of dynamics spawned endogenously by civil war.

5. Future Research Agendas

Three highly stylized types of arguments attempt to account for the onset of civil war. All three point to important causal factors and all have a measure of empirical backing: ethnic antagonism, the presence of natural resources, and weak states may all increase the risk of a civil war, especially in poor states. At the same time, each argument faces considerable challenges from alternative methods of statistical estimation, different or improved measures, new data, and novel theoretical and conceptual insights. Sambanis’s (2002, 217) assessment is true: in spite of a recent boom in research, civil war still “represents the most poorly understood system failure in the domestic political process.”

Even when the findings of large cross-national studies are statistically significant and econometrically sound, the likelihood that a country identified as being at risk will experience a civil war in a given year remains very small, which limits the direct policy relevance of this research. More importantly, the actual causal pathways through which the long-term risk of civil war turns into its realization remain unspecified, unknown, and/or untested. Likewise, the stylized facts about the many facets of civil war that motivate econometric stories are usually untested and sometimes false (Cramer 2002). For example, Kalyvas and Kocher (2006) have provided systematic data strongly suggesting that rebels do not always or necessarily face a collective action problem, as is axiomatically assumed in the literature. Because incumbents frequently respond to the flare-up of rebellion with massive indiscriminate violence individual peasants living in targeted areas may find it rational to join the insurgents.

Civil wars are deeply “endogenous” processes (Kalyvas 2006). Collective and individual preferences, strategies, values, and identities are continuously shaped and reshaped in the course of a war, while the war itself aggregates all kinds of cleavages from the most ideological to the most local. Popular loyalty, disloyalty, and support cannot be assumed as exogenous and fixed. Hence, theories which assume actors and preferences to be frozen in their pre-war manifestations and rely on this assumption to explain various aspects of civil wars, such as their onset, duration, or termination, will be likely biased. This bias is reinforced by the tendency to deduce pre-war actors, preferences, and identities from “master narratives” of civil war. To be sure, such narratives simplify the complexity of civil wars. However, the fact that civil wars are also state-building processes means that their “master narratives” are likely to be contaminated by the war’s outcome: they will be distorted and their ambiguities and contradictions will be erased. Often the hegemony of such narratives
is so powerful that even researchers that collect detailed accounts tend to disregard or downplay their findings because they do not fit into existing frames.

Last, it may well be that not all civil wars are caused by the same set of factors. Significant divergences in the ways in which civil wars are conducted can be linked to processes of state implosion as opposed to processes of peripheral challenge. If this is the case, then different types of civil war may emerge from different combinations of causal factors.

These problems help structure the future research agenda. First, research on civil wars will increasingly move toward the specification and testing of disaggregated causal pathways and mechanisms. Second, micro–macro relationships will be studied less through cross-national statistical analyses and more through integrated research designs that make intensive use of fine-grained subnational data—quantitative and qualitative. Third, both these trends call for opening up the black box of civil war and exploring the complex ways in which a military challenge is articulated, emerges successfully, and is countered—i.e. the microfoundations of civil war. Questions such as rebel recruitment, peasant collective action, rebel rule, peripheral state and rebel organization, peripheral state and rebel financing, dynamics of violence will be studied in increasingly sophisticated ways combining ethnographic, archival, and econometric methods (Kalyvas 2006; Arjona and Kalyvas 2006). At the macro level, we are likely to see more studies that embed civil wars into macro-historical processes, but in ways that incorporate insights and findings from both the emerging micro-level literature and older, overlooked literatures (Wimmer and Min 2006; Hironaka 2005; Derluguian 2005; Boix 2004). These trends all point toward studies that are multi-method, take history seriously, and are characterized by more theoretical and empirical depth. In short, the indeterminacy of current findings signals less a declining research program and more the emergence of an exciting research agenda.

References


