Articulating the Geo-Cultural Logic of Nationalist Insurgency

Lars-Erik Cederman
Swiss Federal Institute of Technology (ETH)
Center for Comparative and International Studies (CIS)
Seilergraben 49
8092 Zurich Switzerland
+41-1-6326759
lcederman@ethz.ch
http://www.icr.ethz.ch

Version 2.0
Comments welcome!

April 9, 2004

Abstract:
Logistical factors have assumed an increasingly important role in studies of civil wars. In particular, political economists have recently proposed models of insurgency that depend on states’ capacity to project power in rugged terrain rather than on ethnic allegiances. Yet, a large gap remains between observed macro-level patterns and postulated rationalistic micro-level explanations. Aiming to reproduce similar, logistical macro-level patterns, this paper offers an alternative account that interprets insurgency as a process of nationalist mobilization leading to violence when the power balance favors the peripheral communities over the centers. To articulate such identity-based mechanisms, I rely on an agent-based model, whose output is analyzed with cross-national regressions. The computational findings suggest that the search for explanations of civil wars needs to go beyond purely rationalist conjectures.

*) Paper prepared for presentation at the Joint Sessions of Workshops, ECPR, Uppsala, April 13-18, 2004, and at the conference on “Order, Conflict, and Violence,” Department of Political Science, Yale University, April 30-May 1, 2004. An earlier version of this paper was prepared for presentation at the workshop on “Origins and Patterns of Political Violence I: Violence in Civil Wars,” Santa Fe Institute, January 15-19, 2004, and at the Centre for the Study of Civil War, PRIO, Oslo, February 12, 2004. I would like to thank the participants of those workshops and Nicholas Sambanis for excellent comments, which have only been responded to partly in this version. Without the computational advice and support of Luc Girardin this paper would not have been possible.
Introduction
As interstate wars become less frequent, academic attention has shifted increasingly to internal conflict. In recent years, an exciting literature on the determinants of civil wars has emerged. Political economists, relying on cross-national statistics and rational-choice modeling, have played a prominent role in this debate. In contrast to most many past attempts to account for domestic unrest, they tend to explain outbreaks of civil wars in materialist and logistical, rather than cultural, terms.

There can be no doubt that this kind of analysis has advanced the research frontier considerably. Instead of sweeping generalizations based on diffuse and scattered case-study evidence, political economists have brought the phenomenon in sharper focus thus allowing for a more precise evaluation of their and others’ hypotheses. Yet, it is somewhat premature to draw definitive theoretical conclusions from these studies, because there is a considerable gap between their macro-level findings and the rationalistic micro-level mechanisms that they posit as explanations. As the latter remain unsupported by direct evidence, one cannot reject alternative causal mechanisms that deviate quite radically from these scholars’ materialist conjectures.

Questioning the political economists’ heavy reliance on materialist factors, this paper articulates alternative causal mechanisms that bring both politics and culture back to the fore. Such a task calls for deeper and more systemic explanations than those associated with standard methods and assumptions of microeconomics. Therefore, I rely on computational modeling, which is better suited to trace historical path-dependency and to capture intangible entities such as national identities.

Contrary to standard rationalistic assumptions, I show that it is possible to create plausible scenarios in which logistical variables such as state strength and terrain matter, although the effect is primarily mediated through geo-cultural mechanisms. The main goal is to create an artificial world within which peripheral insurgents manage to form coalitions, the effectiveness of which depends on the inclusiveness of their shared national identity. Before creating such interactions in greater detail, however, it is necessary to “grow” a comprehensive, geo-cultural system, which includes geography, political organizations, cultural “raw material,” and national identities.

Based on a simple geo-cultural logic, the model generates results that are very similar to those found in Fearon and Laitin’s (2003) theory of insurgency, both in terms of state strength and rugged terrain. Unlike the expectations of these authors, however, the model’s insurrectional mechanisms hinge directly on national identities, and thus indirectly on the ethnic preconditions of the state in question. To confirm the power of nationalist identity-formation while controlling for materialist factors, I run multivariate regressions on the model’s output. This analysis suggests that it is often hard, if not even impossible, to separate political-economy interpretations from constructivist theorizing merely based on quantitative macro-level analysis. Clearly, further empirical investigations will be necessary to resolve this debate.

In the following, I first review the relevant literature before introducing the four main phases of the computational model. After having specified the model, the attention turns to the main computational findings, followed by a section subjecting them to sensitivity analysis. A concluding section elaborates the theoretical and empirical consequences of the present study.
The contemporary literature on civil wars
There can be no doubt about the policy relevance of civil wars. Whereas interstate warfare has declined in recent years, intrastate conflict shows few signs of abating (Sollenberg and Wallensteen 2001; Gleditsch et al 2002). It is therefore not surprising that scholars, both in comparative politics and international relations, have joined forces in a series of systematic attempts to account for why civil wars break out and why they persist (Lake 2003). Much of this research draws on the availability of large-N data from sources such as Gurr’s Minorities at Risk project (Gurr 1993) and improved, more fine-tuned measurements of internal conflict (Gleditsch et al. 2002).

Marrying such quantitative tools with rational-choice modeling, political economists have been leading the efforts to advance candidate explanations along rationalistic lines. These studies tend to ascribe more explanatory power to materialist incentives than to ethnic divisions of the state in question (for a trenchant review, see Sambanis 2003). While several studies have appeared that propose variations on this theme (see e.g. Collier and Hoeffler 2002; Skaperdas 2003), a recent article by Fearon and Laitin (2003) stands out as particularly well-researched and provocative.

Contending that state strength and the logistical conditions of the rebels play the main role in determining the probability of civil war, Fearon and Laitin (2003, p. 75) define insurgency as “a technology of military conflict characterized by small, lightly armed bands practicing guerilla warfare from rural base areas.” Based on this definition, they “hypothesize that financially, organizationally, and politically weak central governments render insurgency more feasible and attractive due to weak local policing or inept and corrupt insurgency practices” (pp. 75-76). Furthermore, such a logic implies that insurgency should become easier to carry out in rough and mountainous terrain (see also Collier and Hoeffler 2002). However, more controversially, the authors find that the ethnic composition of the state in question has no impact on the likelihood of conflict.

Less than a year after its publication, Fearon and Laitin’s article has already attracted considerable attention. Representing the state of the art in terms of methodology, it serves as a useful reference point for critiques and re-analyses. Relying on Bayesian analysis to correct for measurement error, Quinn, Hechter, and Wibbels (2003) confirm Fearon and Laitin’s contention that weak states are more violence prone, but cast doubt on ethnic diversity’s alleged irrelevance. Sambanis (2003) probes the causal reasoning of the authors based on an impressive series of case studies rather than on statistics, but comes to similar critical conclusions. Also preoccupied with measurement error caused by Fearon and Laitin’s reliance on imperfect proxies, he shows that their postulated mechanisms deviate from the empirical record in many of the cases. Most importantly, the political economists’ tendency to aggregate all instances of civil war into one category obscures the diversity of causal logics. There is no a-priori reason to believe that the same mechanisms explain both coups within stable borders and secessionist attempts. Indeed, disaggregation demonstrates that in wars over self-determination, ethnic cohesion among the peripheral rebels supports collective mobilization against the central government (Sambanis and Zinn 2003). It is equally plausible that several mechanisms, rational or not, co-exist in the same civil war (Kalyvas 2003). 1

What are we to make of this intriguing, if somewhat contradictory, picture? True to their rationalistic inclinations, the political economists have relied on rational-choice models

---

1 In his dissertation research, Kocher (2003) also uses Fearon and Laitin’s study as a point of departure. His aim is to build a more refined theory of insurrectional civil war that relies more explicitly on ecological analysis, especially the rural-urban balance. For another study that focuses on the configuration of settlement patterns, see Toft (2003).
as a way to bridge the micro-macro gap. Although there is obviously no complete consensus among these scholars as regards the postulated mechanisms, their theoretical reconstructions are internally consistent and often elegant. Nevertheless, due to the looseness of the macro-level indicators and proxies used, and in the absence of extensive micro-level validation, it would be unwise to place too much confidence in such conjectures. In his review of the literature, Sambanis (2003, p. 1) sums up the situation cogently:

The gap between micro-level behavior and macro-level explanation is large. It is magnified when micro-macro relationships are studied solely through cross-national statistical analyses. What is often lost in such studies in information about causal pathways that link outcomes with causes.

After all, the observed macro-level patterns could have been produced by radically different causal mechanisms, possibly even including the cultural ones that Fearon and Laitin downplay. Therefore, it would seem reasonable to heed Sambanis’ (2003, p. 12) call urging researchers to “consider a wider array of micro-level theories, including how emotions, ideology, revenge, or private material interest interact to produce collective action that culminates in a civil war.” Similarly, Quinn, Hechter, and Wibbels (2003, p. 13) stress that their “demonstration that cultural variables are important determinants of civil strife does not elucidate the mechanisms that are responsible for their efficacy. Causal mechanisms of outcomes like civil war require evidence about units at a lower level of aggregation than the state.”

Nevertheless, it remains unclear where to turn for alternative theoretical theories to fill the conceptual void left open by the political economists. For sure, none of the critiques considered here offers very clear guidance as to where to look for coherent theories in which to embed non-materialist mechanisms. Moreover, constructivist theory has not managed to live up to the same high standards of internal consistency and systematic validation as have their rationalistic counterparts (Fearon and Laitin 2000). Being preoccupied with mental events rather than with macro-level processes, constructivists have had just as great difficulties at advancing complex relational theories as do their rationalist colleagues (McAdam, Tarrow, and Tilly 2001; Tilly 2003; Cederman 2002b; Cederman and Daase 2003).

My contention is that historical sociology goes a long way toward providing the missing conceptual pieces in the explanatory jigsaw. There are striking parallels between the cases at hand and Rokkan’s (1999) classical theory of European state-formation and nation-building. At the heart of his model is a center-periphery relationship not unlike the one proposed by Fearon and Laitin. Rokkan insists that “any analysis of variations among political systems must start from notions of territory” and proceeds by specifying his three main building blocks as centers, peripheries, and transactions, the latter linking the centers to the peripheries. Because transactions have to respect logistical constraints, it is necessary to study the physical, technological, military, economic, and cultural conditions of communication. Even based on this brief enumeration it becomes evident that Rokkan’s framework steers clear of the political economists’ materialist reductionism. For “peripherality” in Rokkan’s conception means much more than a physical struggle for power. In his more nuanced picture, peripheries are characterized by distance, difference, and dependence:

A periphery is located at some distance from the dominant centre or centres, and its transactions with the latter are fraught with costs. A periphery is also different from the central areas on one or more scores: while the degree of distinctiveness will vary, being to some extent a function of distances and dependence, there will
invariably be some minimum level and sense of separate identity. Finally, a periphery is dependent upon one or more centres … in political decision-making, in cultural standardization, and in economic life (p. 115).

By stressing identity-formation, Rokkan’s account belongs to a long and prominent series of studies that draws the link between cultural standardization and political mobilization. In a classic study of rural France in the 19th century, E. Weber (1976) adopts a multi-dimensional approach to political center’s cultural penetration of the periphery with the aid of road building, educational reforms, and other institutional mechanisms. Other seminal accounts along similar lines include Hintze (1975), Deutsch (1953), and Gellner (1983). In addition, there are excellent, empirical studies that stress identity-formation in a regional context. Focusing on Southern Asia, Weiner’s (1978) identifies a pattern that he refers to as the “sons of the soil,” which spawns conflicts between settlers from the capital clashing with ethnically distinct peripheral populations. Likewise, there is a cultural aspect to Herbst’s (2000) stimulating study of failed nation-building in the geographical periphery of African states.

To say that nation-building is important does not mean that one has proposed a full-fledged theory. The strength of Rokkan’s system building lies at the macro level rather than at the level of micro-level mechanisms. Moreover, his research did not focus on violent conflict as much as on cleavage structures and party politics. In recent years, however, considerable conceptual progress has been made in the study of “contentious politics,” both as regards an ontological shift toward relational theorizing and an increased epistemological stress on causal mechanisms as the key to explanation (McAdam, Tarrow, and Tilly 2001; Tilly 2003). It is in the spirit of that promising research program that I embark on the task of articulating the geo-political logic of nationalist insurgency.

A computational model of nationalist insurgency
Agent-based modeling has emerged as an increasingly popular complement to conventional, rational-choice methods. By representing distinct actors in their social and often spatial context, this computational technique enables the analyst to represent complex, macro-level processes without losing the relevant micro-level mechanisms out of sight. Moreover, intangible concepts, such as culture and social identities, which do not easily fit into a rationalistic framework, can be readily formalized and integrated into agent-based models (Cederman 1997, Chap. 3). As Ian Lustick (2000) nicely puts it:

Difficulties of amassing and manipulating collective identity data into theoretically potent comparisons are among the reasons that agent-based modeling can play an important role in the elaboration, refinement, and testing of the kind of specific and logically connected theoretical claims that constructivists have been faulted for not producing. Because the models are run on computers there is no room for ambiguity in the specification of the model’s underlying rules.

This statement describes precisely what this paper aims at, although it is important not to interpret the reference to testing too literally. While it is easy to generate unlimited amounts of “data” with a simulation model, such information should not be confused with real-world evidence. Thus, computational modeling cannot serve as a full-fledged substitute for empirical research. Rather, agent-based models constitute artificial, and

---

2 Interestingly, Fearon (2002) offers a partly inductive and disaggregated analysis of the duration of civil wars that confirms the power of Weiner’s explanation.
indeed simplified, worlds in which the plausibility and consistency of specific causal mechanisms can be evaluated in a context that is more complex than that of standard, rationalistic modeling tools, but still much simpler than the real world. Serving as a stepping stone between micro and macro analysis, such models can help untangle interacting mechanisms, that together generate the phenomenon to be explained. This perspective employs generative research strategy that starts from such patterns and moves backwards in time to the candidate mechanisms that could have generated them (Cederman forthcoming).

In our case, the main configuration to be reproduced is constituted by the statistical finding that rebellions in weak and mountainous states are more frequent than elsewhere. What would it take to construct a set of plausible, a predominantly identity-based mechanism that generates such findings without resorting to purely materialist explanations? As indicated in the previous section, any reconstruction of this kind must specify a process of identity-formation that creates center-periphery relationships within which civil wars occur.

Such configurations feature the following building blocks:

- Political organizations. The main protagonists are the sovereign states, whose centers dominate their peripheries.
- Terrain. Interaction has to proceed in a realistic space with obstacles, such as mountains and other types of rough terrain.
- Cultural map. Identity formation presupposes a multi-dimensional substrate of cultural traits.
- National identities. Nationalist behavior requires identities, which are constituted as combinations of cultural traits that define coalitions.

 Needless to say, a model that comprises all these components would have to be very complex. In fact, no existing model combines all these properties in the same framework. Most computational models of cultural politics make no distinction between culture and identities, thus violating a fundamental tenet of constructivist theory. Typically, culture is modeled as fixed and immutable properties influencing behavior (Epstein 2002) or as endogenous vectors whose traits all matter (Axelrod 1997, Chap. 7). Identities, however, are more selective, because only politically relevant culture enters the power calculus. Allowing for such variation, Lustick’s (2000; 2002) agent-based models ABIR and PS-I feature endogenous repertoires composed by sets of identities, but the component identities stand in no specific relationship to each other and thus do not describe a coherent cultural space. Moreover, this, and almost all other models of this type, fail to provide an explicit representation for formal political organizations or terrain.

Inspired by Bremer and Mihalka’s (1977) pioneering computational application to International Relations, I have developed a modeling architecture, which I have come to call Geosim, that is better suited to solve the research problem studied in this paper (Cederman 1997). Embedded in a dynamic state system constituted by formal political organizations, it contains a constructivist logic that captures culture as a tunable attribute space, with identity mappings as schemata based on this substrate (Cederman 1997: Chap. 8; 2001; 2002a). Once having generated national identities of the centers and peripheries, these can be brought to bear on behavior, thus avoiding cultural determinism. What is missing here, however, is a representation of rugged terrain. The current model

---

3 Whereas my first models were based on Pascal, in more recent years, I have relied exclusively on Repast, the leading Java-based package for agent-based modeling, see http://repast.sourceforge.net/. This is the first of my models that has been implemented in a distributed environment developed by Luc Girardin.
thus needs to incorporate a topological map beneath the cultural and political representations.

In macro-historical theorizing, it is impossible to circumvent the question of where to start. Strictly speaking, the contingency of world history forces us to depart from the dawn of human history. Yet, such a research design would obviously be impractical. Instead, I propose a telescoped, staged process that attempts to combine dynamics at different time scales. Complex systems are known to contain processes operating according to vastly different time tables (Simon 1981; Braudel 1980). While sometimes lasting for a long time, civil wars typically outpace nation-building. National identities, in turn, usually emerge much faster than the cultural systems within which they are situated.

In order to reduce endogeneity to manageable proportions, the staged process creates a system starting with the most slow-moving processes before shifting the focus to faster dynamics. As the faster-moving phases are added, the configurations resulting from the previous phases are held constant. This facilitates the design and calibration of the framework as well as the tracing of causal pathways unfolding during each stage.

More precisely, the current model features four phases, each one corresponding to a distinct set of historical processes:

**Phase I: Initialization**

The first stage initializes the model by generating three basic ontological building blocks, namely a state system, terrain, and a cultural landscape, all embedded within a square grid. These basic entities are parametric and result from algorithms that make no claims to achieving realism in terms of the dynamics of the system. Yet, the parametric nature allows the experimenter to retain full control over the configuration of each component.

First the initialization routine generates a geographic map that allocates a level of ruggedness to each cell in the grid up from zero to \( \text{maxRugged} = 3 \). Initially, a tunable number of mountain peaks are distributed randomly across the grid. Then a diffusion process connects these peaks with their surrounding sites, thus creating relatively smooth mountainous terrain gently sloping down to the plains. Later the ruggedness levels will be used as weights slowing down interactions between the center and the periphery (see the Appendix for details).

Second, a state system, equipped with a tunable number of states, here 30, is created. Again, the algorithm starts by randomly selecting the same number of sites, from where a recursive process absorbs the surrounding provinces. This accelerated process of geopolitical consolidation is slowed down in mountainous areas, which means that the state borders will respect geographical obstacles at least to some extent. Note that this process is a simplified and drastically compressed version of the more realistic series of conquest in the second phase of the model.

**Figure 1**

Figure 1 shows the standard 30 x 30 grid. Each cell possesses a specific level of ruggedness, which constitutes mountain chains in gray shades, with darker shades corresponding to the roughest areas. In this case, a large part of the grid is mountainous. On top of this geographic map, thirty sovereign states reside, with the black lines corresponding to their borders and the circles to their capitals.

\[ \text{maxRugged} = 3 \]

4 The parameter values used in this section are the default values as indicated in the Appendix.
Finally, the initialization phase produces a cultural map that is a multi-dimensional array allocating a symbol string to each site of the grid. The algorithm orchestrates a simulated “age of migrations” as in early Medieval Europe. As in the previous types of maps, this one starts with a selection of random locations, this time corresponding to the homeland of each ethnic community. Each of these ethnic communities is allocated an entirely stochastic symbol string, such as \(2 \ 3 \ 1 \ 4 \ 4 \ 2 \ 2 \ 2\) with eight traits and four possible values. Two parameters drive this process: the number of ethnic groups \(\text{numEthnies} = 5\) that participated in the cultural settling of the landscape and the parameter \(\text{culturalDrift} = 0.1\) defining the cultural differences within each tribe’s territory. Cultural drift describes a process of spontaneous change, as exemplified by linguistic shifts and religious fragmentation (Axelrod 1997, p. 170). This entire algorithm, which is also affected by geographic obstacles, guarantees that there will be both dialectal nuances as one moves from province to province as well as more abrupt ethnic cleavages representing language, or even civilizational, borders (Rokkan 1999, p. 172).

Instead of depicting geographical obstacles, the shading in Figure 2 portrays the cultural differences among the local sites. While the darker shades correspond to areas characterized by great cultural differences, the brighter ones refer to cultural “plains” in the cultural landscape. Overlaid on top of this cultural map is a simplified version of the geographic contours, coloring all areas of roughness above one unit in brown, semi-translucent shades.

**Phase II: State formation and assimilation**

The next phase operates as a geo-cultural “pressure cooker” at much slower speed than the first phase (see the Appendix for details). First, because the initial stage does not guarantee that the states are even remotely geopolitically viable, it is important to expose them to some internal and external competition. Once unleashed, the sovereign states will start interacting locally with their immediate territorial neighbors and with their own provinces. This means that both interstate wars and secession take place, although the latter is less likely because the would-be rebels cannot resort to nationalist coordination as in Phase IV.

Interaction can be of two types: either the actors peacefully exist side by side or they engage in combat. As explained in the Appendix, conflict is initiated according to a simple decision rule, which depends on the local power balance between the two parties (see also Cederman 2003a). Roughly speaking, the actors play a “grim trigger” strategy with another. This means that they normally reciprocate whatever their neighbors did to them on their respective fronts. But whenever the power balance in their favor exceeds a preset, stochastic threshold that governs the offense dominance of the system, they launch unprovoked attacks.

Because interactions among states unfold as conventional battles with clearly marked defense lines, the midpoint of the attacker’s stochastic threshold is located at 2.5, requiring this actor to be more than twice as powerful as its victim before an unprovoked attack can be launched. This gives the defender a considerable advantage. In internal conflict, however, the threshold is set to one, because in insurrectional fighting, it is less likely that there are fixed defense lines requiring the aggressor to be decisively superior. In these cases, only the provinces engage in unprovoked attacks, and then always with the aim of breaking out of the state. In view of the dominated peripheries’ revisionist
interests as opposed to the centers’ status-quo orientation, this is a reasonable first approximation. Yet, the capitals always fight back should there be an insurrection.5

Identical stochastic criteria govern the outcome of these battles. In interstate wars, a victorious attacker is allowed to absorb the province fought over. If the defending state wins, however, the status quo results. Civil wars always end in secession if the breakaway province prevails, otherwise there is no change to the geopolitical map.

The system is animated by a primitive resource “metabolism” directly based on territorial size with some stochastic noise added to the extraction process. Crucially, the centers’ resource extraction declines as a function of effective distance discounting for terrain obstacles. Even in the plains, transportation incurs a cost per distance unit. In the mountains, however, moving through each site adds a logistical penalty proportional to the roughness of the area.6

It is convenient to model the logistical constraints with a loss-of-strength gradient (Boulding 1963; Gilpin 1981). Figure 3 displays such a logistical curve, which slopes gently from 100 % extraction at zero distance from the capital down to 30% at large distances. This function also governs power projection of all actors. This is especially important, because it should be difficult for the governments to fight far away from their capitals, especially in rugged terrain. The provinces, however, never have to worry about terrain discounting, because all of their power extraction and projection take place “at home.”

In addition to wars and secessions, this phase also features assimilation by the capitals to their own cultural makeup. Although both the political and cultural maps produced by Phase I are both influenced by geographical conditions, they are not directly connected. Yet, throughout the ages, states have had varying degrees of influence over the cultural makeup of their populations. The assimilation process, introduced in the phase, proceeds by allowing each capital to assimilate one cultural trait of a province with a constant probability per “visit” (see the Appendix). In our test run, we will let the assimilation rate assimil be 0.01 only, which means that 1000 time periods do not suffice to iron out all differences.

The result of this geo-cultural process can be inspected in Figure 4. At the end of this period, i.e. at time step 1000, three of the smallest states in Figure 2 have disappeared leaving behind them some two-unit fragments. Otherwise, the state system looks very much like it did at the beginning of the phase, thus reflecting the evolutionary stability of the surviving states. In terms of culture, there are more important changes. Although the rate of assimilation is quite low in this case, the sharp “civilizational” differences have been partly smeared out. This can be seen clearly in the state in the lower right corner of the grid. The capital has managed to iron out some of the cultural “ruggedness.”

---

5 This may be a more problematic assumption, because in the real world, many centers choose to make concessions to the peripheral areas, including granting them autonomy arrangements and sometimes even full independence.

6 If the ruggedness is three, for example, the effective distance increases by three units from one to four. A ruggedness of zero yields a factor one, and thus no penalty.
Phase III: Nation-building

Once having created a plausible geo-cultural environment, we are ready to introduce national mobilization. Rather than treating nationalism as a constant “law”, I model the phenomenon as a macro-historical process that emerges exogenously in the entire system at a specific point in history. This point occurs at time step 1000 and the process of nation-building continues for another 200 time steps.

In keeping with the staged research design, I freeze all previous social configurations, including the state system and the culture map, assuming that they represent slow-moving changes that can be treated as exogenously. This means that no interstate warfare or secession can take place during this stage. Furthermore, assimilation ceases to operate as well. This allows us to focus entirely on national identity-formation in a stable geo-cultural framework.

In designing the nationalism part of the model, I have striven to stay as close as possible to modern, constructivist theories of nationalism (Cederman 2002b). Most importantly, this modeling extension introduces a fundamental distinction between the underlying cultural landscape and the national identities as political coalitions. Although the attribute space has an impact on identity formation, there is no one-to-one correspondence between culture and national identities, because only politically relevant traits count in national identity-formation (Gellner 1983).

Fortunately, there is a computational solution to this conceptual problem. Letting sets of symbol strings be used in “schemata,” John Holland (1995) introduces wildcards (#) for those traits that could be of any value (Cederman 1997, Chap. 8; 2001; 2002a). For example, the string \{4 # # 2 3 # # #\} represents an identity template (see the Appendix for details). It is convenient to represent national identities in a similar way. Let us assume that national communities form in the public domain and that states could become members of them as long as their culture strings match them. These constraints generate a hypothetical feasible set of identities to which a given actor could belong.

National identities are represented as distinct groups to which any state or province can belong if the nation’s identity matches the actor’s own culture. While both capitals and provinces are eligible members of nations, the probability of launching a nationalist movement depends crucially on the geopolitical status of the actor in question. Thanks to their cultural and material resources, capitals have much higher chances of founding their own nations, but provinces may sometimes create distinct nationalist platforms in opposition to their respective capitals. The likelihood of joining an existing nation, however, is quite high, reflecting the lower difficulty of such an identity selection. In addition, all mobilized actors, whether provinces or capitals, attempt to recruit new members to their identities by visiting neighboring sites, discounted by the terrain roughness. Once mobilized, it is assumed that an actor never loses its national identity.

[Figure 5 about here]

---

7 This simplification is strictly motivated by the research design rather than by the framework’s inherent possibilities. In reality, these geopolitical processes can be expected to be heavily influenced by the onset of nationalism (Cederman 2001; 2002a). Fortunately, the model allows for simultaneous, co-evolving processes. This is an obvious extension to which I will return in future research.

8 Rokkan (1999, p. 173) states that “there is an enormous number of distinguishable languages in the world, but only a few of these rate a realistic chance to be transformed into standards of written communication. Such a standard cannot be developed without a centre or network, or a network of centres… Territorial consolidation increased the opportunities for elite encounters over long distances.”

---
Showing the grid at time 1200, at the end of the nation-building phase, Figure 5 displays the national identities as colored disks, complemented with a number beneath each disk corresponding to the respective nation’s index number. It is clear that most capitals have already become nationally mobilized. In several cases, national communities spill over state borders. For example, three capitals shares the same nation on the upper part of the grid (see nation #7).

Nearly perfect nation-states appear in the plains. In the lower right corner, there are two nation-states (see nations #3 and #21). The mountainous regions, in contrast, offer less hospitable conditions to nation-builders. Consider the state in the upper left corner whose capital belongs to nation #9. It is clearly bi-national, because apart from four members of that nation including the capital, there are also nine members of nation #19. As we will see in the next phase, such differences influence the likelihood of civil wars.

Phase IV: Civil war

Focusing on internal conflict, the final phase of the model pits provinces against capitals while freezing everything else, including the national identity-formation process of the previous phase. This means that national identity configurations and the states’ outer borders have to remain intact. Culture no longer has any direct impact on the behavior of the model. Having ruled out interstate wars and secession, only rebellions are allowed to happen. Whether the capital or the province prevails, the status quo results, and the rebels are allowed to regroup.

It may seem that this exogenization unduly restricts the analysis, but it makes some sense for the period following 1945. With the very important exceptions of imperial collapse associated with the decolonization process and the disintegration following the end of the Cold War, secession has been a relatively rare event in world politics (see e.g. Hechter 1992), especially in post-colonial Africa (Herbst 1989). Likewise, as already noted, there have been few interstate wars as well. However, the exclusion of these geopolitical transformations should be seen as a simplifying, analytical assumption rather than a statement about how the world is constituted. Moreover, as a first approximation, the model attempts to explain violence leading to secession without necessarily tracing the entire process. Of course, endogenous analysis will be necessary to reach firmer conclusions about secessionist violence in its entirety.

The current phase retains the mechanism that governs secession attempts in Phase II, though without the geopolitical consequences. While obviously fought interactively between the centers and the peripheries, such conflicts can only be initiated by the provinces in this version of the model. Still, it is assumed that the capitals fight back if attacked. The previous domestic decision-making threshold at one (rather than 2.5) operates in this phase as well. The main difference pertains to the resource calculation of the mobilized provinces. Once belonging to a nation, a province will attempt to recruit nationalist support from other co-national provinces within the same state. The potential rebel organizes such recruitment campaigns before deciding whether or not to attack the capital. Controlling how many of the nationalists that feel obliged to help their kin, the parameter $natOblig$ (with default value 0.5) determines the stochastic proportion of success in such cases. After adding up the resources of the potential coalition, the province in question is more likely to trigger a civil war, because thanks to nationalist

---

9 In the current configuration, these mobilization attempts are based on a fixed identity. In reality, however, identities often adapt to the center-periphery power balance, with thinner identities becoming selected or chosen should the rebels encounter stiff resistance from the center (see Hannan 1979; Cederman 1997, Chap. 8).
solidarity, the balance can shift more decisively in favor of the periphery. With this logic in place, it is harder for the capital to “divide and rule” within its territory.

If the leading province decides on action, all co-national units that pledged support will make good of their promise.\textsuperscript{10} Fighting on several fronts will force the capital to redistribute its resources, thus in many cases opening opportunities for other provinces to join the fray even if they do not belong to the same nation. In some cases, unmobilized actors, who could be bandits or other opportunistic rebels, may decide that the power balance has shifted so much in their favor that they also want to fight the capital.

Combat then proceeds according to the same rules as in Phase II. Depending on the local power balance for each center-periphery dyad either the capital or the province prevails. Who wins, and for how long the fighting lasts, is immaterial for the purposes of this paper, because as already stated, all geopolitical consequences have been inhibited for methodological reasons. What matters here is whether there is fighting at all or not.

Switching back to a pure geographic ruggedness map, Figure 6 illustrates how all this works. At time 1278, civil wars rage in two states. Each local battle is marked by a red cross. The state in the upper left corner of the grid is under attack by six nation #19 insurgents and one unmobilized province. In the second rightmost state at the bottom of the grid, three nation-#1 actors have taken up arms against the capital, which belongs to nation #4. In the plains, however, peace reigns.

[Figure 6 about here]

This particular snapshot suggests that fighting happens predominantly in rugged terrain. Is this a recurrent pattern? To find out, we have to study the behavior of the model more systematically. Fortunately, agent-based modeling allows us to rerun history counterfactually as often as we want, and it is to this task that we turn in the next section.

**Replication results**

The telescoped histories constitute artificial worlds that can be subjected to controlled experiments. This section applies a combination of experimental and quasi-experimental methods to the system. Relying on the former, the sensitivity analysis in the next section varies several key parameters experimentally. In the current section, which presents results that mimic conventional, quasi-experimental methodology, cross-sectional data need to be collected. Then the simulated conflict outcomes are regressed on the states’ determinants to find out how well alternative explanations fare. This approach allows us to assess what quantitative analysis would have told us had the model coincided with the real world. Note that the task of measurement is considerably easier in an artificial world, because there is no room for measurement errors. Moreover, thanks to the telescoped design, endogeneity becomes more manageable as well.

The research design is simple. For each parameter configuration, 30 runs featuring entire state systems are executed. The replications are based on ten randomly generated geographical maps with their associated state systems as in Figure 1. For each such configuration, the model is run with three independently generated cultural maps.\textsuperscript{11} Measurement proceeds during Phase IV, which lasts from time step 1200 through step

\textsuperscript{10} This is obviously a questionable assumption that needs to be endogenized in future versions of the model (see Gates 2002).

\textsuperscript{11} This is achieved through a separation of the stochastic processes, so that the cultural mechanisms rely on random numbers drawn from a separate generator that is seeded apart from the main seed.
The number of observations amounts to almost 900, because there are up to 30 states in each of the 30 runs (excluding micro states that are smaller than five units). Regressions are run on all of these observations in a pooled sample. Data for the following variables are collected for each (sufficiently large) state in each run:

- The dichotomous dependent variable *civilwar* is set to one if at least one center-periphery battle occurred within the territory of the state in question throughout the entire Phase IV, otherwise the variable remains zero. While a crude measure of conflict, the absence of time series means that temporal interdependence can be ignored in the regression analysis.

- The first independent variable *size* indicates the territorial size of the state measured as the number of provinces plus the capital.

- The second independent variable *strength* holds the resource level of each state.

- The third independent variable *ruggedness* measures the average ruggedness level based on the state’s entire territory including both the capital and all provinces (see the *res* variable in the Appendix).

- The fourth independent variable *natfrac* refers to a nationalist fractionalization index, computed according to the standard definition $1 - \Sigma p_i^2$ where $p_i$ stands for the proportion of each nationalist group (cf. Fearon 2003 for a discussion of various fractionalization measures).

- The fifth independent variable *natdist* reveals the average terrain-adjusted distance from the capital of all the oppositional nationalist provinces if there are any, otherwise the variable is set to zero.

Whereas the first three independent variables tap materialist properties of the states, the last two ones capture cultural determinants relating to national identities. It is therefore possible to separate materialist from identity-based hypotheses.

Let us start by considering what one would expect from a rationalistic perspective. The hypothesis would be that the first three materialist variables should be significant as opposed to the nationalist indicators. Indeed, Fearon and Laitin (2003) find that population size (which can be seen as a proxy for territorial state size), state weakness (operationalized in terms of GDP per capita), as well as rough terrain, to be strong predictors of internal conflict.

In contrast to this view, a relationist interpretation inspired by Rokkan would lead us to anticipate the last two variables to have a strong impact independently of the power of the materialist ones. First of all, the more nationally fractionalized the state, the more opportunities to rebel. This hypothesis goes beyond ethnic fractionalization, which is what Fearon and Laitin (2003) measure. Indeed, nationalism presupposes a desire to form a state rather than mere cultural belonging (Weber 1978; Gellner 1983). It is therefore questionable whether Fearon and Laitin really “test” theories of nationalism in their

---

12 This size-related truncation of the sample was introduced to prevent very small states from distorting the effect of the size variable. Otherwise, the findings are reasonably robust to changes in this arbitrary limit.

13 Note, however, that a very high degree of fractionalization may make the task of rebelling harder (see Cederman 1997, Chap. 8). I rely on a fractionalization index because it is a standard indicator in the literature, not because I think it is a suitable way to measure the opportunities of collective action in insurgencies. For example, it is clearly a deficiency that the standard indices are symmetric in that the ignore the distinction between the state-controlling groups and the peripheral ones.
article. Second, as suggested by Weiner’s (1978) “sons of the soil” mechanism, it is not merely the average ruggedness that matters, but more specifically whether the peripheral nationalists can take advantage of rough terrain in their struggle against the center. Thus, the location of the groups is central in this mechanism.

To investigate the influence of these factors, I ran multivariate logit regressions with and without the nationalist variables. The column labeled Model 1.1 in Table 1 lists the results from a regression that includes only the logistical variables size, strength and ruggedness. The table reveals that all three variables are strongly significant, with signs in the expected directions. First, larger states are more conflict prone than smaller ones. Second, state strength reduces significantly the likelihood of insurgency. Finally, as expected, the ruggedness variable works in the opposite direction: the rougher the terrain, the more likely civil unrest becomes. Purely based on this data, political economists would be forgiven for believing that their theories suffice to account for the behavior of this model.

However, this picture changes drastically when nationalist mobilization is controlled for. Model 1.2 makes clear that the constructivist variable natfrac accounts for much of the variation in conflict behavior. As would be expected, fractionalized states with high natfrac scores are especially conflict prone. On average, peripheral groups that benefit from rough terrain find more opportunities to rebel.

This effect does not change when the second nationalist variable natdist is added to the regression analysis (see Model 1.3). This model shows that the nationalists’ terrain-adjusted distance from the capital plays a major role, regardless of the state’s ruggedness. It seems that the logistical obstacles confronting the center in its exchanges with the periphery have a direct impact on the probability of civil war.

Interestingly, the materialist variables remain pertinent in most cases, though their effect becomes somewhat reduced. The main exception is the size variable, which loses much of its power once natdist is introduced into Model 1.3. Apparently, in the current geopolitical setup, large states are not just associated with a larger number of rebellions in general, but also with more serious logistical challenges confronting the capitals.14

In view of these findings, it is clear that statistical accounts of the system that incorporate appropriate identity-related variables perform better than materialist ones. But is this general finding robust? To find out, the next section presents a preliminary sensitivity analysis.

### Sensitivity analysis

Although a more thorough analysis would have called for a combinatorial sweep of parameters, this section varies retains the default settings of the previous section while varying one parameter at a time.15 Given the telescoped nature of the model, it is important to alter variables belonging to all phases. This procedure will enhance our confidence in the robustness of the findings across radically different time scales. Let us now consider each of the other phases in turn, starting with the most fundamental one.

---

14 Additional tests, including modifications of the states’ resource updating (not reported here), suggest that size may become significant in alternative configurations. Yet, the significance is usually reduced by the introduction of natdist.

15 Obviously, a more thorough sensitivity test would have to include combinations of parameter values.
Variations in Phase I
Empirical scholars can only dream of manipulating environmental background factors, such as terrain. In an artificial world, by contrast, it is just as easy to alter the altitude and shape of mountain chains as it is to change one individual’s behavior. Thus, the first sensitivity test entails “lowering” and “elevating” the mountains by running simulations where the maximum ruggedness is one and five respectively, rather than three as in the default system (see Models 2.1 and 2.3). To facilitate comparison, Model 2.2 displays the regression pertaining to the default system in Model 1.3 (see Table 1).

Table 2 presents the findings of this exercise. It is obvious that the two cultural variables remain robust, whereas strength is the only materialist variable that holds consistently across the test values. As would be expected, reducing the geographic obstacles makes the ruggedness variable less significant. At the same time, these variations influence the frequency of civil wars quite drastically. Whereas only 11% of the states experience internal conflict in the less rugged systems of Model 2.2, this frequency climbs to 48% in Model 2.3. This finding confirms the real-world observation that regions with rough terrain, like the Balkans or the Caucasus exhibit more than average domestic unrest.

Having studied the effect of changes in the terrain, it is now time to introduce variation in the structure of the initial cultural map. This is done along two dimensions. First, the number of ethnic communities populating the initial grid is varied (5 or 15 rather than the default value 10). Second, the analysis continues with changes of the rate of cultural drift, which determines the gradual, village-to-village cleavages (here the alternative levels are 0.05 and 0.2 as opposed to the default system’s 0.1).

As with the terrain parameter, we find that the more “rugged” the initial cultural landscape, the less internal conflict there is. This effect applies to both the cultural polarity and drift. Thus, there is a strong parallel between cultural and geographic obstacles. Despite these changes, Tables 3 and 4 suggest that the parameter manipulation makes a rather limited difference in terms of the regression results, although the impact of the cultural variables declines somewhat with increasing cultural fragmentation (i.e. more initial ethnic groups and more cultural drift among them). In contrast, the materialist indicators exhibit less consistency across the parameter values. Territorial size becomes more significant with more ethnies (see Model 4.1) and with both less and more cultural drift compared to the default values (see Models 4.1 and 4.3). The level of ruggedness also loses some of its impact as the number of ethnic groups increases. The strength variable, however, remains robust in all settings.

Variations in Phase II
Having found the main results quite robust to limited alterations of Phase I, we now proceed to Phase II, where assimil, the rate of state-led cultural assimilation, plays the main role. What happens if we tune this parameter down to zero from the default value of 0.01? Will there be any change should this value be set to 0.1? Table 5 has the answers to these questions.

---

16 A shift from 5 to 15 ethnies decreases the civil-war frequency from 40% to 32% and a change of the rate of cultural drift from 0.05 to 0.2 cuts the rate from 42% down to 29%.
This time the behavioral output changes more as a result of experimental manipulation. Given that the default value is quite low, removing assimilation matters less than accelerating the process to its maximum speed. It should be recalled that the latter configuration produces a cultural “level playing field” within the states, with virtually no deviations from the capitals’ culture. Such flat cultural landscapes create almost perfect nation states, which feature less violent, domestic politics. As a consequence, civil wars become extremely rare, which helps explain the lower significance of the estimates in this case (3% of the cases in Model 3.3 as opposed to 47% and 36% in Models 3.1 and 3.2 respectively). Despite this fact, the cultural variables remain the most robust, but even they lose some explanatory power as the assimilation reaches its maximum value. Again, the ruggedness variable becomes insignificant at the expense of natdist. The size variable, on the contrary, regains significance for the alternative assimilation rates.

**Variations in Phase III**

The next step of our sensitivity analysis concerns the nation-building phase. I have chosen to control the minNumTraits parameter, which determines the national identities’ minimal “thickness.” This variable stipulates how many active symbols (rather than wildcards) a national identity needs to hold. The lower the number, the more inclusive the identity, which should facilitate peripheral nationalist mobilization. This expectation is confirmed by the negative correlation between identity thickness and civil war frequencies (52% for minNumTraits = 3 down to 23% for minNumTraits = 5, to be compared with 36% in the default system).

In terms of parameter estimates, these limited deviations from the default value minNumTraits = 4 seem to matter little, as shown by Table 6. There is a slight decline of the cultural variables’ impact with increasing thickness, but the significance remains very solid despite the variation. A similar reduction of significance can be observed for the materialist variables as well, especially for the size variable.

**Variations in Phase IV**

We have now reached the last stage of our sensitivity experiments, which corresponds to the civil war phase of the research design. At this point, it is appropriate to study the impact of different levels of nationalist solidarity by altering the values of the parameter natOblig both downward to 10% and upward to 100% (see Table 7). As before, the middle column represents the default setting of 50% (see Model 7.2).

The pattern of the previous stages recurs in this case too. We find that whereas the nationalist variables hold up well across the parameter values, their logistical counterparts

---

17 Note, however, that assimilation is assumed to happen before nationalist mobilization. Once a unit has been mobilized in the model, it cannot be “converted” to a different nation. Deutsch (1953) and many other theorists have hypothesized that centers that attempt to accelerate the assimilation of already mobilized peripheral populations may provoke more, rather than less, civil unrest (see also Cederman 1997, Chap. 7).

18 Of course, more drastic manipulations lead to important changes of behavior. With minNumTraits = 1, the frequency of civil wars grows to 71% but all the explanatory variables remain significant, except ruggedness, for which p = 0.066. In the other extreme case, with all identities being maximally thick (minNum Traits = 8) as few as 9% of the states are involved in internal conflict. Even in this case, the variables retain most of their impact, but here size changes sign.
exhibit a less consistent picture, with size becoming significant only for low levels of nationalist activity (see Model 7.1) though at expense of the ruggedness variable.

In sum, we have found that the main computational findings reported in Table 1 survive a series of robustness tests. In the absence of a more thorough, combinatorial sensitivity analysis, I thus conclude that the results are far from “knife-edge.” Both the key materialist and identity-related variables retain their impact across the board, with the exception of size and in some cases of ruggedness.

Conclusion
What are the theoretical lessons to be drawn from this paper? Let us start by stressing what it does not purport to say. First, it does not claim to have refuted rationalist theories of civil wars. It deserves to be reiterated that the statistical analysis shown above relates to an artificial world, and certainly not to real-world evidence. In fact, I offer no more micro-level validation than do most political economists, who also usually advance untested (but testable) theoretical interpretations at that level. Second, nor does this essay argue that the current model offers the best formalization of relational and constructivist theories of national insurgencies. After all, this is the first model of its kind. To my knowledge, geography and identity formation have never been investigated together in an integrated, dynamic model. Thus, I ask for the reader’s patience with the sometimes drastic simplifications introduced.

It is easy to think of important extensions of the model that could bring it closer to the empirical record. One of the most urgent tasks would be to let political and cultural boundaries co-evolve (Cederman 2002b). Fortunately, this can be readily done within the current computational framework (Cederman 2001; 2002a). Both qualitative and quantitative studies have shown convincingly that ethnic conflict has a tendency to spill over state borders due to irredentism and other diffusion effects (Weiner 1971; Sambanis 2001; 2003). Yet, sometimes one step backward means that we can take two steps forward in the future. Despite this obvious loss of realism, the telescoped, phased design of the model introduces a powerful way of handling drastically different time scales without getting swamped by endogeneity.

Another part of the model that calls for elaboration is the behavioral interaction between the center and the periphery. The current specification lets the provinces take the lead, but this obscures an often delicate strategic “dance” going back and forth between the main protagonists, which features prominently counter-insurgency initiatives on the part of the capitals. Moreover, the current design stresses opportunity structures rather than actual motivations. Rational-choice theory offers many cues about the microfoundations of internal conflicts that could be integrated into the current computational framework (see e.g. Gates 2002). Nor is any claim to realism made as regards the clustering of fighting into civil wars, the frequency and duration of which can be validated empirically. More realistic strategies, together with cluster-finding algorithms, could facilitate such measurements, as illustrated by Cederman (2003a) in the case of interstate wars.

Having stated what this study does not say, I owe the reader a few final words about what I believe to be the actual consequences for future research. Most importantly, the paper discourages students of civil wars from drawing too far-reaching inferences from political-economy studies such as Fearon and Laitin (2003). While micro-level interpretations of macro-level data are indispensable to most explanations, the current paper illustrates that such mechanisms do not have to be narrowly rationalistic and materialist in the political economists’ sense. Moreover, it should be clear that a model
with a powerful geo-cultural logic as the one proposed here produces results that, at the surface, are very hard to separate from the expectations of rationalist scholars.

More specifically, two of the Fearon and Laitin’s (2003) key variables, namely state strength and rough terrain, seemingly vindicate their logistical theory of insurrections. If interpreted differently, however, these explanatory factors may turn out to be correlated with mechanisms of nationalist identity-formation. Crucially, the outbreak of violence appears to be linked to the settlement patterns of peripheral nationalist groups. In particular, these groups are predominantly located in rough terrain. Where the center is unable to mobilize more than a limited nationalist following due to logistical constraints, its resistance to insurgencies suffers accordingly. Far from being a matter of ancient, primordialist conditions, nationalist identity formation has evolved in the last couple of centuries, and in many cases, much more recently in conjunction with the decolonization process.

Rather than imposing a false choice between identity-related and materialist explanations, the computational model offers guidance as to how these mechanisms could be disentangled conceptually. Yet, though this is easily done in a computational model, it is much harder to carry out empirically. The political economists’ statistical controls for “nationalism theory” often fail to grasp the logic of these alternative perspectives, partly because the theories to be refuted are rendered too sweepingly, and partly because the macro-level indicators used are inappropriate. There is therefore an acute need for better articulated theories of nation-building.

The main goal of this paper has been start closing this gap by articulating candidate mechanisms that can be more readily exposed to empirical tests. For sure, we still have some way to go both empirically and theoretically before a close match can be achieved, but I would submit that the current paper makes some progress in this direction. Fortunately, thanks to the considerable interest that civil wars are attracting, new data are constantly made available, especially at the micro level. I have already mentioned Sambanis’ (2001) systematic series of case studies, Sambanis and Zinn’s (2003) new data base focusing on center-periphery interactions, and Buhaug and Gates’s (2002) efforts to use geographic data to pinpoint the exact location of conflict. It would be very helpful if the conflicting parties could be pinpointed geographically as well, especially in relation to existing communication infrastructures.

Adding a new tool to our toolbox in the study of civil wars, agent-based modeling will help relationist theorists sharpen their hypotheses and test empirical indicators in methodologically underdeveloped areas that have so far escaped the grasp of conventional approaches. Thanks to their “colonization” of the literature on civil wars, the political economists have raised the explanatory bar considerably. Now the main task is to prevent their methodological advances from leading to premature theoretical closure. After all, much is at stake, and not merely scientifically, but also in terms of policy making. While factoring in macro-processes may seem defeatist from a policy perspective, it is more likely to lead to more realistic expectations about what is politically desirable and feasible in the short run, and what is not. Failure to see this difference may have serious consequences, as illustrated by the difficulties posed by “nation-building” in post-war Iraq. In an era characterized by ambitious democratization projects with uncertain consequences for political mobilization, historical insensitivity and cultural “color-blindness” are likely to worsen the problems of domestic violence and terrorism already haunting the world’s most troubled neighborhoods.
Appendix: Technical notes on the model

A variable-size grid constitutes the Geosim’s core. Constituting the basic unit of the framework, a primitive actor reside permanently in each square of the grid. They can never be destroyed, but they do expand territorially and can also lose their sovereignty as other actors come to dominate them hierarchically.

All actors, whether primitive or compound, keep track of their geopolitical context with the help of a portfolio holding all of their current relationships. These can be of three types:

- **Territorial relations** point to the four territorial neighbors of each primitive actor (north, south, west, and east).
- **Interstate relations** refer to all sovereign neighbors with which an actor interacts.
- **Hierarchical relations** specify the two-way link between provinces and capitals.

The behavior of the geopolitical base model depends crucially on these three types of relationship. Whereas all strategic interaction is located at the political level, territorial relations become important as soon as structural change happens. Combat takes place locally and results in hierarchical changes that will be described below.

The assumption that provinces retain relations only with their capitals represents a strong idealization of sovereignty in that the subordinate units cannot have political relations with other provinces or foreign actors. This simplification rules out the possibility of civil wars among the provinces not involving the capital. Secessionist civil wars between a capital and a province can occur, however.

As in my previous models, the logic of execution is quasi-parallel with double buffering (Cederman 1997). As opposed to picking one actors at a time and giving it all the “limelight”, the process follows stages in which each actor executes code (cf. Cusack and Stoll 1990). To achieve quasi-parallel execution the list of actors is scrambled each time structural change occurs. The actors keep a memory of one step and thus make up an advanced Markov process (though with varying transition probabilities).

In the following, technical notes will be provided on each of the four phases, supplementing the specification given in the main text:

**Technical notes on Phase I**

The text above describes most of the details pertaining to the initialization routines of the geographical, political, and cultural configurations. However, a few additional notes on the first of these are in order. First, the terrain module allocates heights to every cell in the grid up to $\text{maxRugged} = 3$. The algorithm starts by creating a random selection of mountain summits constituting a fraction $\text{propSummits} = 0.05$ of all unitary cells. Then a recursive algorithm is run repeatedly 20 times smoothening the height of the surrounding cells. This “brush” continues to a cell in the von Neumann neighborhood with a probability of 0.8. For each “visit”, the algorithm sets the neighboring call to a weighted average 0.7 of the initial cell and its previous value. From an intuitive standpoint, this algorithm is similar to water flowing from the mountain peaks dragging with it soil that leads to a smoother landscape around the summits.

**Technical notes on Phase II**

The main simulation loop of Phase II contains six steps that will be presented in the following. In the first phase, assimilation takes place. Second, the actors update their resources. Third, both capitals and provinces allocate these resources to their fronts. Fourth, a decision procedure follows, during which they decide on whether to cooperate or defect in their neighbor relations. Fifth, the interaction phase determines the winner of
each battle, if any. Sixth, the structural change procedure carries out conquest and other border-changing transformations.

**Step 1. Assimilation**
During each time step, with a probability of \( \text{assimil} = 0.01 \) each compound state gets the chance to assimilate each one of its provinces. A trait of the province’s culture string is selected randomly and set to the same value as the capital’s corresponding cultural trait. If the traits are already the same, no change is effected.

**Step 2. Resource updating**
As the second step in the actual simulation loop, the resource levels are updated. The simple “metabolism” of the system depends directly on the size of the territory controlled by each capital. It is assumed that all sites in the grid are worth one resource unit. A sovereign actor \( i \) begins the simulation loop by extracting resources from all of its provinces. It accumulates a share of these resources determined by a distance-dependent logistical function \( \text{dist} \) (see Figure 3 above):

\[
\text{dist}(d) = \text{distOffset} + \frac{1-\text{distOffset}}{1+(d/\text{distThresh})^{(-\text{distSlope})}}
\]

where \( \text{distOffset} = 0.3 \) sets the flat extraction rate for long distances, and \( \text{distSlope} = 3 \) the slope of the curve (higher numbers imply a steeper slope) and \( \text{distThresh} = 5 \) the threshold value of the curve. Based on this logistic curve, the tribute gained from each province is discounted in proportion to the province’s geography-corrected distance \( d \) from the capital. The resources of the actor \( \text{res}(i,t) \) can now be computed by adding the non-discounted resources of the capital together with the sum of all tribute from the provinces. Then some normally distributed noise is added, with a standard-deviation \( \text{resSD} = 0.02 \) times the resource level. Finally, the resources \( \text{res}(i,t) \) of actor \( i \) in time period \( t \) can be computed by factoring in the new resources multiplied by a fraction \( \text{resChange} = 0.5 \):

```plaintext
res = \text{resUnit} \\
\textbf{for all provinces } j \text{ of state } i \textbf{ do} \\
res = res + f(dist(i,j)) \\
\text{res}(j,t) = (1-f(dist(i,j))) \times \text{resUnit} \\
res = \text{Normal}(\text{res}, \text{resSD} \times \text{res}) \\
\text{res}(i,t) = (1-\text{resChange}) \times \text{res}(i,t-1) + \text{resChange} \times \text{res}
```

**Step 3. Resource allocation**
Before the states can make any behavioral decisions, resources must be allocated to each front. For unitary states, there are up to four fronts, each one corresponding to a territorial relation (those located in the corners of the grid have a few as two neighbors). Resource allocation proceeds according to a hybrid logic. A preset share of each actor’s resources are considered to be fixed and have to be evenly spread to all external fronts. Yet, this scheme lacks realism because it underestimates the strength of large actors, at least to the extent that they are capable of shifting their resources around to wherever they are needed. The remaining part of the resources, here \( \text{propMobile} = 0.5 \), are therefore mobilized in proportion to the opponent’s strength and the previous activity on respective front. The current model features a simplified version of Proportional Resource Allocation. Mobile resources are proportionally allocated to fronts that are active (i.e. where combat occurs will occur), but also for deterrent purposes in anticipation of a new
attack. Resources are always allocated under the assumption that no more than one new attack might happen.

For example, a state with 50 mobile units could use them in the following way assuming that its two provinces control 10 and 15 units, and three neighboring states 20, 25, and 30 respectively. If the previous period featured civil war with the second province and interstate war with second of the neighbors, these two fronts would be allocated $\frac{15}{(15+25)} \times 50 = 18.75$ and $\frac{25}{(15+25)} \times 50 = 31.25$. Under the assumption that one more wars could start, the other province, and the remaining two states would be allocated respectively: $\frac{10}{(15+25+10)} \times 50 = 10$, $\frac{20}{(15+25+10)} \times 50 = 20$, and $\frac{30}{(15+25+10)} \times 50 = 30$.

Formally, resource allocation for state $i$ starts with the computation of the fixed resources for each relationship $j$. A preset proportion of the total resources $\text{res}$ are evenly spread out across the $n$ fronts:

$$\text{fixedRes}(i,j) = (1-\text{propMobile}) \times \text{res} / n$$

The remaining part $\text{mobileRes} = \text{propMobile} \times \text{res}$ is allocated in proportion to the activity and the strength of the opponents. To do this, it is necessary to calculate all resources that were targeted at actor $i$:

$$\text{enemyRes}(i) = \sum_j \{\text{res}(j,i)\}$$

The algorithm of actor $i$’s allocation can thus be summarized:

```plaintext
for all relations $j$ do
  in case $\text{enemyRes}(i) = 0$ then [actor not under attack]
  $\text{res}(i,j) = \text{fixedRes}(i,j) + \text{mobileRes}$
  in case $i$ and $j$ were fighting in the last period then
  $\text{res}(i,j) = \text{fixedRes}(i,j) + \frac{\text{r}(j,i)}{\text{enemyRes}(i)} \times \text{mobileRes}$
  in case $i$ and $j$ were not fighting the last period then
  $\text{res}(i,j) = \text{fixedRes}(i,j) + \frac{\text{r}(j,i)}{(\text{enemyRes}(i)+\text{r}(j,i))} \times \text{mobileRes}$
```

The provinces allocate all their resources to the front against the capital.

**Step 4. Decisions**

Once each actor has allocated resources to its external and possibly internal fronts, it is ready to make decisions about future actions. This is done by recording the front-dependent decisions in the corresponding relational portfolio. Since two-level action is allowed, both sovereign states and provinces execute their algorithms in each period. As with resource allocation, this happens in parallel (or to be more precise in randomized order) without any reference to other actors’ decisions during the current time step.

All actors start by playing unforgiving “tit-for-tat” with all their neighbors, a strategy that resembles the game-theoretic strategy “grim trigger” in conventional iterated games. The power-seeking aspect of the strategy relates to the second step featuring an unprovoked attack provided that no combat is already under way. First a potential victim $j'$ is selected.

Provinces only have one option: to attack the capital. Capitals, however, never launch an unprovoked attack against a province, but may well launch combat action against
external enemies. The decision to attack depends on a probabilistic criterion to which we will return below. If an attack is approved, the aggressor chooses a battle path consisting of an agent and a target province. The target province is randomly chosen as any primitive actor of \( j^* \) (including the capital) that borders on \( i \). Also stochastically selected, the agent province is a province of \( i \) (including the capital) that borders on the target.

In summary, the decision algorithm for sovereign states subscribes to the following logic:

Decision-rule of sovereign state \( i \):

for all internal and external fronts \( j \) do
  if \( i \) or \( j \) played D in the previous period then
    act\((t,j)\) = D
  else
    act\((t,j)\) = C  [Grim Trigger]

if there is no action on any front then
  randomly select a neighbor state \( j' \)
  with \( p(i,j') \) do
    change to act\((t,j')\) = D [launch attack against \( j' \)]
    randomly select target\((i,j')\) and agent\((i,j')\)

Since provinces have only one relationship to deal with, namely that with the capital, they use a simpler decision-rule:

Decision-rule of province \( i \) with capital \( j \):

if \( i \) or \( j \) played D in the previous period then
  act\((t,j)\) = D
else
  act\((t,j)\) = C  [Grim Trigger]

if there is no action with parent \( j \) then
  if \( p(i,j) \) then
    change to act\((t,j)\) = D [attack parent]
    set target\((i,j)\)=\( j \) and agent\((i,j)\)=\( i \)

The precise criterion for attacks \( p(i,j') \) remains to be specified. The model relies on a stochastic function of logistic type in which the local power balance plays the main role:

\[
bal(i,j') = \frac{\text{dist}(d)\times\text{res}(i,j')}{\text{dist}(d')\times\text{res}(j',i)}
\]

where \( \text{dist}(d) \) is the distance function described above and \( d \) and \( d' \) the respective distance from the capitals of \( i \) and \( j \) to the battle site. This discounting introduces distance-dependence with respect to power projection. Hence, the probability can be computed as:

\[
p(i,j') = \frac{1}{1+(bal(i,j')/\text{supTresh})^{(-\text{supSlope})}}
\]

where \( \text{supTresh} \) is a system parameter specifying the threshold that has to be transgressed for the probability of an attack to reach 0.5, and \( \text{supSlope} \) a tunable parameter that determines the slope of the logistic curve, which is set to 15 for the runs
reported in this paper. Note that supTresh = 2.5 for interstate interactions, and supTresh = 1 for domestic conflict.

Step 5: Interaction
After the decision phase, the system executes all actions and determines the consequences in terms of the local power balance. The outcome of combat is determined probabilistically. If the updated local resource balance \( \text{bal}(i,j') \) tips far enough in favor of either side, that side wins the battle. In the initial phase, the logistical probability function \( q(i,j') \) has the same shape as the decision criterion with the same threshold set at \( \text{victTresh} = 2.5 \) and with an identical slope: \( \text{victSlope} = 15 \):

\[
q(i,j') = \frac{1}{1+(\text{bal}(i,j')/\text{victTresh})^{(-\text{victSlope})}}
\]

This formula applies to attacks among capitals. In accordance with the strategic rule-of-thumb that stipulates, in interstate wars, that an attacker needs to be about three times more powerful than a defender to prevail, the threshold of the latter is set to \( 1/\text{victTresh} \). No such inversion is needed in civil wars because in those cases \( \text{victTresh} \) is set to one.

Each time-step of a battle can generate one of three outcomes: it may remain undecided, one or both sides could claim victory. In the first case, combat continues in the next round due to the grim-trigger strategy in the decision phase. If the defending state prevails, all action is discontinued. If the aggressor wins it can advance a territorial claim, which is processed in the structural change phase.

Step 6: Structural change
Structural change is defined as any change of the actors’ boundaries. The current solution closely follows the mechanism introduced in Cederman (1997, Chap. 4). Thus structural change only affects one primitive unit at a time, and can thus be seen as a matter of territorial combat rather than a diplomatic game at the political level.

To render the system manageable, the underlying assumption governing structural change enforces states’ territorial contiguity in all situations. As soon as the supply lines are cut between a capital and a province, the latter becomes independent.

The units affected by any specific structural claim are defined by the target\((i,j)\) province set in the decision phase:
- if target\((i,j)\) is a unitary actor, then the entire actor is absorbed into the conquering state,
- if target\((i,j)\) is the capital province of a compound state, then the invaded state collapses and all its provinces become sovereign,
- if target\((i,j)\) is a province of a compound state, then the province is absorbed. If, as a consequence of this change, any of the invaded states’ other provinces become unreachable from the capital, these provinces regain sovereignty.

Secession works in a similar way. For each sovereign state \( i \), the algorithm also goes through all of the provinces to see if any of them has raised a claim. If this is the case, the province becomes sovereign while retaining all its previous resources. Sometimes the independence of a province makes other provinces unreachable from the old capital, in which case these provinces are made sovereign as well. By contrast, if the capital advances a claim in the interaction phase, no structural change occurs, but the actions and memories are reset to \( C \) thus ending that particular secessionist attempt.
**Technical notes on Phase III**

Reflecting the distinction between culture and nations, nations are implemented as independent computational objects with their own lists of members. A nation is born as soon as a state decides to create one. It is also possible to join or leave an already existing nation. Nations without members are deleted. A state can only be a member of one nation at a time.

When a nation is created, its creator determines its identity based on the actor’s own culture string. This is done stochastically by setting each of the nation’s identity trait to the value of the creator’s corresponding cultural trait with probability $\text{probTrait} = 0.2$. If not set, the trait remains irrelevant and is marked by a “wildcard” #. A total number of $\text{minNumTraits} = 4$ have to be specified. This procedure is repeated until at least one of the traits becomes relevant. For example, if the creator has the culture string \{4 1 2 3 3 2 4 1\}, valid nations include \{4 # 2 3 3 # # #\}, \{# 1 2 3 3 2 4 1\} and the whole culture string itself. However, the identities \{4 # # # # # # #\} cannot be created, the first one because it is too short, and the second one because it does not fit.

An actor can join a nation if there is no incompatibility between the national identity’s relevant traits and the actor’s corresponding cultural trait. Thus, a state with culture \{4 1 2 3 3 2 4 1\} can join nations identities \{# # 2 3 # # 4 1\} but not \{# 2 3 3 4 # # 2\}. Note that a potential member does not have to be adjacent to any co-national before joining. Nations, as opposed to territorial states, are predominantly ideational entities that can span over vast spaces without there necessarily being contiguity.

Nationalist mobilization hinges on the status of the primitive unit $i$, which may be a capital or a province. If it is not yet mobilized, parameterized probabilistic criteria govern the creation or joining of nations. Thanks to their resources, capitals are much more likely than other units to create new nations $\text{prCapitalCreateNation} = 0.01$ than provinces, in which case the probability is as low as $\text{prCapitalCreateNation} = 0.00001$. If these low-probability events do not apply, the likelihood of a units joining an existing nation (to the extent that one exists with the right fit) is very high: $\text{prJoinNation} = 0.3$. Should the unit in question already belong to a nation, it will attempt to recruit the immediate territorial neighbors. With probability $\text{prRecruitToNation} = 0.3$, the unit in question will move to an adjacent site in its VonNeumann neighborhood and invite this primitive unit to join to the extent an identity fit exists. Here is a summary of the algorithm in pseudo code:

```python
if i is unmobilized then
    if i is capital with prob. $\text{prCapitalCreateNation}$ then
        i creates a new nation
    else if i is province with prob. $\text{prProvinceCreateNation}$ then
        i creates a new nation
    otherwise with prob. $\text{prJoinNation}$
        i attempts to join a nation
else if i is mobilized with prob. $\text{prRecruitToNation}$ then
    i attempts to recruit a territorial neighbor to its nation
```

**Technical notes on Phase IV**

This phase features modifications in the behavior of the provinces, which are now able to form nationalist coalitions with nationalist kin provinces within the same state, as long as the capital does not belong to the same nation as the provinces in question. Nationalist

---

19 For the time being, mergers and splits of nations are not supported but could be easily implemented.
insurrections operate through a modified balancing mechanism, which facilitates the insurrectionists’ search for a solution to the collective choice dilemma. If contemplating a rebellion, a province will attempts to recruit support. The probability of any other kin-province’s accepting such an invitation is $p_{NatOblig} = 0.5$.

The modified power-balance criterion for a possible attack by province $i$ against capital $j'$ is:

$$
\text{bal}'(i,j') = \begin{cases} 
\frac{\text{res}(i,j')}{\text{res}(j',i)} & \text{if capital } j' \text{ is not in } i's \text{ nation} \\
\sum k \frac{\text{res}(k,i)}{\text{res}(j',i)} & \text{if capital } j' \text{ is in } i's \text{ nation}
\end{cases}
$$

where $\sum k \text{ res}(k,i)$ are the pooled resources of the nationalist coalition. All provinces that accepted the call make good of the pledge of fighting against the capital.
References


Tables

Table 1. Logit regression of simulated data in the default system

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1.1</th>
<th></th>
<th>Model 1.2</th>
<th></th>
<th>Model 1.3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coeff.</td>
<td>std.err.</td>
<td>prob.</td>
<td>coeff.</td>
<td>std.err.</td>
<td>prob.</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.700</td>
<td>0.442</td>
<td>0.113</td>
<td>-2.947</td>
<td>0.589</td>
<td>0.000</td>
</tr>
<tr>
<td>State Size</td>
<td>0.109</td>
<td>0.016</td>
<td>0.000</td>
<td>0.071</td>
<td>0.019</td>
<td>0.000</td>
</tr>
<tr>
<td>State Strength</td>
<td>-0.028</td>
<td>0.004</td>
<td>0.000</td>
<td>-0.027</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>0.804</td>
<td>0.209</td>
<td>0.000</td>
<td>1.148</td>
<td>0.262</td>
<td>0.000</td>
</tr>
<tr>
<td>Natfrac</td>
<td></td>
<td></td>
<td></td>
<td>7.049</td>
<td>0.586</td>
<td>0.000</td>
</tr>
<tr>
<td>Natdist</td>
<td></td>
<td></td>
<td></td>
<td>0.113</td>
<td>0.022</td>
<td>0.000</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td></td>
<td></td>
<td></td>
<td>-418.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>828</td>
<td></td>
<td>828</td>
<td>828</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Logit regression of simulated data for different levels of rugged terrain

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 2.1 (maxRugged= 1)</th>
<th></th>
<th>Model 2.2* (maxRugged= 3)</th>
<th></th>
<th>Model 2.3 (maxRugged= 5)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coeff.</td>
<td>std.err.</td>
<td>prob.</td>
<td>coeff.</td>
<td>std.err.</td>
<td>prob.</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.410</td>
<td>0.857</td>
<td>0.005</td>
<td>-3.442</td>
<td>0.624</td>
<td>0.000</td>
</tr>
<tr>
<td>State Size</td>
<td>0.127</td>
<td>0.041</td>
<td>0.000</td>
<td>0.016</td>
<td>0.021</td>
<td>0.439</td>
</tr>
<tr>
<td>State Strength</td>
<td>-0.040</td>
<td>0.008</td>
<td>0.000</td>
<td>-0.018</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>1.487</td>
<td>0.790</td>
<td>0.060</td>
<td>1.064</td>
<td>0.271</td>
<td>0.000</td>
</tr>
<tr>
<td>Natfrac</td>
<td>5.968</td>
<td>0.828</td>
<td>0.000</td>
<td>6.540</td>
<td>0.621</td>
<td>0.000</td>
</tr>
<tr>
<td>Natdist</td>
<td>0.195</td>
<td>0.050</td>
<td>0.000</td>
<td>0.113</td>
<td>0.022</td>
<td>0.000</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td></td>
<td></td>
<td></td>
<td>-204.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>876</td>
<td></td>
<td>828</td>
<td>789</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*) Settings of the default system (see Model 1.3 in Table 1).

Table 3. Logit regression of simulated data for different numbers of ethnies

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 3.1 (numEthnies= 1)</th>
<th></th>
<th>Model 3.2* (numEthnies= 3)</th>
<th></th>
<th>Model 3.3 (numEthnies= 5)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coeff.</td>
<td>std.err.</td>
<td>prob.</td>
<td>coeff.</td>
<td>std.err.</td>
<td>prob.</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.075</td>
<td>0.628</td>
<td>0.000</td>
<td>-3.442</td>
<td>0.624</td>
<td>0.000</td>
</tr>
<tr>
<td>State Size</td>
<td>0.024</td>
<td>0.022</td>
<td>0.284</td>
<td>0.016</td>
<td>0.021</td>
<td>0.439</td>
</tr>
<tr>
<td>State Strength</td>
<td>-0.019</td>
<td>0.004</td>
<td>0.000</td>
<td>-0.018</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>1.348</td>
<td>0.274</td>
<td>0.060</td>
<td>1.064</td>
<td>0.271</td>
<td>0.000</td>
</tr>
<tr>
<td>Natfrac</td>
<td>6.601</td>
<td>0.610</td>
<td>0.000</td>
<td>6.540</td>
<td>0.621</td>
<td>0.000</td>
</tr>
<tr>
<td>Natdist</td>
<td>0.116</td>
<td>0.023</td>
<td>0.000</td>
<td>0.113</td>
<td>0.022</td>
<td>0.000</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td></td>
<td></td>
<td></td>
<td>-293.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>828</td>
<td></td>
<td>828</td>
<td>828</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*) Settings of the default system (see Model 1.3 in Table 1).
Table 4. Logit regression of simulated data for different levels of cultural drift

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 4.1 (culturalDrift=0.05)</th>
<th>Model 4.2* (culturalDrift=0.1)</th>
<th>Model 4.3 (culturalDrift=0.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coeff.</td>
<td>std.err.</td>
<td>prob.</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.841</td>
<td>0.547</td>
<td>0.000</td>
</tr>
<tr>
<td>State Size</td>
<td>0.050</td>
<td>0.020</td>
<td>0.012</td>
</tr>
<tr>
<td>State Strength</td>
<td>-0.022</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>0.848</td>
<td>0.247</td>
<td>0.001</td>
</tr>
<tr>
<td>Natfrac</td>
<td>6.794</td>
<td>0.578</td>
<td>0.000</td>
</tr>
<tr>
<td>Natdist</td>
<td>0.078</td>
<td>0.018</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Log-likelihood: -314.00, -290.49, -272.91
Sample size: 828, 828, 828

*) Settings of the default system (see Model 1.3 in Table 1).

Table 5. Logit regression of simulated data for different levels of assimilation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 5.1 (assimil=0.0)</th>
<th>Model 5.2* (assimil=0.01)</th>
<th>Model 5.3 (assimil=0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coeff.</td>
<td>std.err.</td>
<td>prob.</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.05</td>
<td>0.487</td>
<td>0.000</td>
</tr>
<tr>
<td>State Size</td>
<td>0.042</td>
<td>0.019</td>
<td>0.026</td>
</tr>
<tr>
<td>State Strength</td>
<td>-0.020</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>0.536</td>
<td>0.222</td>
<td>0.016</td>
</tr>
<tr>
<td>Natfrac</td>
<td>3.655</td>
<td>0.495</td>
<td>0.000</td>
</tr>
<tr>
<td>Natdist</td>
<td>0.147</td>
<td>0.022</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Log-likelihood: -367.00, -290.49, -18.73
Sample size: 828, 828, 828

*) Settings of the default system (see Model 1.3 in Table 1).

Table 6. Logit regression of simulated data for different levels of identity exclusiveness

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 6.1 (minNumTraits=3)</th>
<th>Model 6.2* (minNumTraits=4)</th>
<th>Model 6.3 (minNumTraits=5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coeff.</td>
<td>std.err.</td>
<td>prob.</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.925</td>
<td>0.581</td>
<td>0.000</td>
</tr>
<tr>
<td>State Size</td>
<td>0.047</td>
<td>0.022</td>
<td>0.037</td>
</tr>
<tr>
<td>State Strength</td>
<td>-0.025</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>0.797</td>
<td>0.255</td>
<td>0.002</td>
</tr>
<tr>
<td>Natfrac</td>
<td>7.274</td>
<td>0.627</td>
<td>0.000</td>
</tr>
<tr>
<td>Natdist</td>
<td>0.114</td>
<td>0.025</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Log-likelihood: -300.71, -290.49, -245.69
Sample size: 828, 828, 828

*) Settings of the default system (see Model 1.3 in Table 1).
Table 7. Logit regression of simulated data for different levels of nationalis solidarity

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 7.1 (natOblig= 0.1)</th>
<th>Model 7.2* (natOblig= 0.5)</th>
<th>Model 7.3 (natOblig=1.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coeff.</td>
<td>std.err.</td>
<td>prob.</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.341</td>
<td>0.841</td>
<td>0.686</td>
</tr>
<tr>
<td>State Size</td>
<td>0.092</td>
<td>0.0367</td>
<td>0.012</td>
</tr>
<tr>
<td>State Strength</td>
<td>-0.066</td>
<td>0.010</td>
<td>0.000</td>
</tr>
<tr>
<td>Ruggedness</td>
<td>0.237</td>
<td>0.330</td>
<td>0.471</td>
</tr>
<tr>
<td>Natfrac</td>
<td>5.440</td>
<td>0.729</td>
<td>0.471</td>
</tr>
<tr>
<td>Natdist</td>
<td>0.169</td>
<td>0.029</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Log-likelihood: -195.16  -290.49  -307.60
Sample size: 828  828  828

*) Settings of the default system (see Model 1.3 in Table 1).
Figure 1. Varied terrain inhabited by 30 states at time at the end of Phase I at $t=0$. 
Figure 2. A cultural map shown together with mountain contours at the end of Phase I at $t=0$. 
Figure 3. Terrain-adjusted loss-of-strength gradient affecting power extraction and projection.
Figure 4. The system at the end of Phase II at $t=999$, after some geopolitical consolidation and cultural assimilation.
Figure 5. Mobilized populations after the end of Phase III at $t=1199$. 
Figure 6. Civil wars raging in two states during Phase IV at $t=1278$ (see red crosses).