Modeling Endogeneity in Civil War:
An Agent-Based Approach

Matthew Reichert
Harvard University
mreichert@g.harvard.edu

Miguel Garces
Lustick Consulting
miguelag@lustickconsulting.com

Ian S. Lustick
University of Pennsylvania
ilustick@sas.upenn.edu

A significant theoretical innovation in the civil war literature has been the recognition and explicit theorizing of endogenous processes that interact authority, identity, and violence. However, this has also posed a hefty methodological challenge. Developing creative ways to implement or approximate quasi–experimental designs offers a promising solution – but at the cost of sacrificing attention to external validity. In this paper, we propose agent-based modeling as mediator between these two groups of research. Agent-based simulations explicitly model endogenous processes, while providing a ‘sandbox’ in which researchers interested in causal inference can test counterfactual claims. We demonstrate this role by modeling the emergence, disintegration, and co-evolution of insurgent orders in a dynamic agent-based space. We test hypotheses drawn from Kalyvas and Kocher’s "endogenous cleavages thesis," and use the results to generate scope conditions for potential quasi-experimental designs.

Support for the research reported in this paper was received from the Office of Naval Research (Contract #s N00014-12-C-0042 and N00014-14-C-0387).

Researchers interested in replicating findings should contact the authors for access to the necessary platform, templates, and protocols. Open-source software for constructing agent-based models, as well as published documentation and theoretical footnoting for the models appearing in this paper, is available at www.lustickconsulting.com.
Introduction:
In the civil war literature, the recognition and explicit theorizing of endogenous processes has been a red thread running through approximately the past ten years of research. Understanding endogeneity in civil war has represented both a theoretical innovation and a methodological challenge, particularly for making strong causal inferences. Over the past five years, a body of work on civil war has cut through the problem of endogeneity by developing creative ways to implement or approximate experimental designs. The problem is that findings from contemporary experimental work often feel theoretically unsatisfying, frequently receiving the critique that they lack external validity.

In this paper, we suggest that agent-based modeling can fill a useful and needed role as mediator between these two groups of research – theories that emphasize endogeneity, and quasi-experimental approaches that generate highly specific findings. Agent-based simulations explicitly model endogenous processes, while providing a ‘sandbox’ in which researchers interested in causal inference can test counterfactual claims. We demonstrate this role with an agent-based model of a generic civil war environment, designed to operationalize and interact theories of identity, authority, and violence prominent in the civil war literature.\(^1\) We test hypotheses drawn from Kalyvas and Kocher (2007)’s "endogenous cleavages thesis," and use the results to generate propositions about the scope conditions under which findings from quasi-experimental designs should and should not have external validity.\(^2\)

The Problem of Endogeneity in Civil War
Not all studies of civil war argue that wartime process are endogenous to one another – in fact, until the mid-2000s, in most of the civil war literature the word "endogenous" rarely appeared at all.\(^3\) During the 1990s and early 2000s, researchers usually imported well-established theoretical mechanics from international relations, or conducted large-N "ground-clearing exercises" to establish broad macro-level relationships.\(^4\) Only afterward, between 2000 and 2010, did new

\(^{1}\) Specifically, we try to capture theoretical insights from Kalyvas and Kocher (2007), Staniland (2012), Peterson (2002), Chandra (2012) and Cederman et. al. (2010).

\(^{2}\) Kalyvas and Kocher, "Ethnic Cleavages and Irregular War," 204-10.

\(^{3}\) Except when a researcher had an endogeneity problem.

\(^{4}\) For examples, see Posen, "The Security Dilemma and Ethnic Conflict;" Kaufmann, "Possible and Impossible Solutions to Ethnic Civil Wars;" Walter, "The Critical Barrier to Civil War Settlement;" Collier and Hoeffler, "Greed and Grievance in Civil War;" Fearon and Laitin, "Ethnicity, Insurgency, and Civil War;" Sambanis, "What is a Civil War?"
theoretical innovations tear down the conceptual categories implicit in that work and replace them with a new research agenda based on a recognition that identity, authority, and violence are endogenous to one another in civil war contexts. This section will briefly overview this research agenda, which we will refer to together as “meso-theories of civil war,” emphasizing particularly explicit treatments of authority and identity as endogenous variables.

Kalyvas’ 2001 article overturning the distinction between "new" and "old" civil wars marked the beginning of this kind of thinking. Kalyvas argued that post-Cold War "new" civil wars (which are loot-seeking, predatory, and ethnicized) and pre-Cold War "old" civil wars (which are justice-seeking, political, and ideological) present very differently at the macro-level of analysis than at the micro-level, where they often resemble each other. Implicit here was an argument for endogeneity: to explain why we ended up with these poor categories, Kalyvas pointed to how "locally segmented cleavages often aggregate in misleading ways." Aggregation was responsible for the poor categories of the 1990s because "individual loyalties in old civil wars, as in new civil wars, are often informed . . . by fluid, shifting, and often locally based cleavages." The task for researchers, then, was to go from micromotives to macrobehavior, to identify the particular ways that "multiple, highly fragmented conflicts" aggregate under particular conditions versus others. The following sections briefly review how subsequent scholarship has gone about fulfilling this task by theorizing the endogenous interactions of authority, identity, and violence. Ultimately, this scholarship has generated strong theory that is parsimonious, logically coherent, and possesses face validity; however, on its own this scholarship has lacked strong empirics.

**Endogeneity of Authority**

In his 2006 book *Logic of Violence in Civil War*, Kalyvas takes a large step in the direction of this research task by providing a foundational building-block for a new research program organized around the question of endogeneity. To this end, he states a puzzle that, in its basic ontological premises, is a sharp break from earlier questions that asked about onset, cessation, or duration of civil wars – his puzzle is “a disjunction between the macro-level causes of war and the micro-level

---

5 Kalyvas, "'New' and 'Old' Civil Wars."
6 Ibid., 112.
7 Ibid., 110.
patterns of violence.” In more detail, “almost every macrohistorical account points to the importance of preexisting popular allegiances for the war’s outcome, yet almost every microhistorical account points to a host of endogenous mechanisms, whereby allegiances and identities tend to result form the war or are radically transformed by it.” So, Kalyvas (2006) is explicitly beginning the task posited by Kalyvas (2001) to go from micromotives to macrobehavior.

Kalyvas’ approach to accomplishing this is to wield the assumptions of rationality and imperfect information to build a strategic logic of wartime violence. Given a strategic environment of intimacy and a technology of irregular warfare, Kalyvas posits as his jumping-off point the distribution of territorial control between state incumbents and rebel insurgents. Where information on the identity of combatants is highly imperfect and intimacy privileges local knowledge, civilian collaboration and defection become crucial mechanisms of military effectiveness. However, civilians in Kalyvas’s story are not passive recipients of violence or authority – they are self-interested agents that strategically respond to incentives and coercion. So in this formulation, the relative credibility of coercive threats made by competing authorities is a function of territorial control and generates analytically distinct patterns of civilian collaboration and defection, which generate patterns of (selective) violence, which in turn affect territorial control and the ability to enact authority with credible coercion.

Staniland, 2012 builds on this foundation to show how state capacity is endogenous to political relationships, so that particular arrangements of insurgent and state authority generate particular modes of competition, which in turn generate new arrangements of authority. This is what is meant by the "co-evolution" of states and orders, and a serious application of "the oft-repeated claim that civil wars represent competitive state building," in which violence and state-building are mutually productive. The conceptual vocabulary that Staniland provides is not only an attempt to match a new theoretical premise with its own dependent variable; it is also a way of grappling with the

---

9 Kalyvas, The Logic of Violence in Civil War, 3.
10 Ibid., 4.
11 Distribution of territorial control, as part of an endogenous process, is both an independent and a dependent variable. However, in order to make any analytic progress, Kalyvas must bracket the dependent character of this variable and treat it exclusively as an independent variable.
12 Ibid., 173-209; Subsequent work has adopted Kalyvas’ zone-based schema. For example: James, Frontiers and Ghettos: State Violence in Serbia and Israel, or Bhavnani, Miodownik, and Choi, "Three two tango: Territorial control and selective violence in Israel, the West Bank, and Gaza."
problems of simple description posed by endogeneity in civil war. As Staniland explains, "this conceptualization allows us to measure the dynamics of conflict at any given point in time...this fine-grained dependent variable can measure shifts in trajectories of conflict, rather than coding a whole campaign a single un-variegated observation." The problem with coding a whole campaign as a single observation is that it overlooks endogenous processes; Staniland’s typology is meant to break the observation down into smaller pieces in an analytically useful and theoretically meaningful way. The next step is to work out the (causal) pathways by which we get one of Staniland’s flavors of wartime order over another. This means identifying the conditions under which particular micro-level interactions tend to aggregate towards particular meso- and macro-level patterns versus others.

Endogeneity of Identity

The conventional wisdom on patterns of identity in civil war paints a picture of "simple binary conflicts between organizations crystallizing popular support and popular grievances along well-defined cleavages," such that patterns of identity and patterns of authority are almost deterministically congruent. Kalyvas and Kocher problematize this conventional wisdom, based on the premise that "ethnic cleavages are further activated and deepened by the war, rather than war merely reflecting already deep ethnic cleavages," so that "ethnic affiliation [is] partially endogenous to the course of the war itself." So not only is authority endogenous to violence, but identity is as well.

Overturning conventional conceptualizations of binary and immutable ethnicized conflict is a deconstructive theoretical act. Kalyvas and Kocher complement it with a constructive one: theorizing pathways by which endogenous conflict processes generate particular distributions of

14 Ibid., 253.
15 Ibid., 244.
16 Kalyvas, "‘New’ and ‘Old’ Civil Wars," 113.
17 Kalyvas and Kocher, "Ethnic Cleavages and Irregular War," 186, 189; To substantiate this claim, Kalyvas and Kocher hypothesize that within one conflict we should expect to see change in the relationship between particular identity groups and government control over time. They test this hypothesis on 12 ethnolinguistic and six religious groups that participated in the Vietnam War, and find a sign change in the probabilistic relationship between identity and control. The Vietnam War provides Kalyvas and Kocher with a conflict of long duration and, therefore, significant variation over time in patterns of violence (as well as geo-coded survey data at the hamlet level of analysis) (Ibid., 190-1). They further argue that not only is identity endogenous to the conflict, but it is also tied up in local forms of order. For example, Vietnamese Communist party members drew on ethnic backgrounds to "preserve independent power bases and aggressively pursue their political interests" (Ibid., 197). It is important to keep in mind here that what is being tested as endogenous is not the existence of identity groups, but the relevance of identity as a predictor for authority by one side or another.
identity versus others. They hypothesize that "civil wars between a strong state and a weaker but well-organized challenger will take the form of irregular war; in contrast, weak, fragmented states lower the threshold of military contestation and invite badly organized challengers; both may be forced to rely on the ‘cheapest’ mobilizable networks, including ethnic ones." This is a first step past simply positing endogeneity - breaking that endogeneity down into falsifiable and scoped predictions.

Kalyvas, 2008 extends this thinking by providing actual mechanisms of identity construction and destruction that flow from conflict dynamics to generate meso-level aggregate patterns of violence: identity shift and ethnic defection. For both mechanisms, an important nuance is that identity changes occur not necessarily via the complete acquisition and discard of individual "hard" identities, but rather though "the manipulation of cleavage salience." Manipulation requires a manipulator; thus, Kalyvas points to "the organizational demand for ethnic defection," "the expansion of territorial control," and "the logic of revenge." Thus, decisions about identity become endogenous to territorial control, and Kalyvas, 2008 leads us right back to Kalyvas, 2006 (Logic of Violence).

Considered together as a body of research, the chief contribution of these meso-theories of civil war has been a set of concepts – distribution of territorial control, types of authoritative order, relative organization of state versus challenger, and identity shift and ethnic defection – that we can use to delineate distinctive sets of micro- and meso-level conditions. At the bottom of the balance sheet for this research agenda, we have: (1) a substantiated claim that authority and identity are endogenous to one another and to violence, (2) a few reasonable and specified pathways and mechanisms, but also (3) a difficult methodological challenge in endogeneity.

---

18 Ibid., 214.
19 Identity shift is "the acquisition by individuals of a new ethnic [or national] identity that replaces the old one" and ethnic defection is "a process whereby individuals join organizations explicitly opposed to the national aspirations of the ethnic group with which they identify" (In addition to building on Kalyvas and Kocher, 2007, because identity shift and ethnic defection are functions of territorial control, these mechanisms also tie back to arguments about the endogeneity of authority made in Kalyvas, 2006 [Kalyvas, "Ethnic Defection in Civil Wars," 1045]).
20 Ibid., 1050.
21 Ibid., 1051.
Why Endogeneity is a Problem

Because one of the major contributions of these meso-theories of civil war is to establish that a number of important variables are endogenous in civil war, and because contemporary political science as a field is often interested in understanding causal relationships, it is useful to note briefly why endogeneity poses a methodological challenge for causal inference. There are three closely related ways that endogeneity inhibits our ability as researchers to make definitive causal claims. The first and most intuitive way is ambiguity of causal direction – if X and Y are endogenous so that X causes Y but Y also causes X, then if X and Y are observed to covary, it is unclear whether that correlation is due to the effect of X on Y or the effect of Y on X.

The second way that endogeneity poses a challenge for causal inference is by generating omitted variable bias. Given two different observations on Y, denoted Y₁ and Y₂, imagine that Y₁ causes X which in turn causes Y₂. If we study the relationship between X and Y₂ and infer that X causes Y₂, then Y₁ will be acting as an omitted explanatory variable on the relationship between X and Y.

Finally, endogeneity can obscure the causal effects of its component relationships. Imagine again that two variables, X and Y, are endogenous to one another. However, here imagine that the effect of X on Y is positive so that an increase in X generates an increase in Y, but the effect of Y on X is negative, so that an increase in Y generates a decrease in X. A researcher studying the relationship between X and Y may observe a decrease in X correlated with an increase in Y, and infer that the causal effect of X on Y is negative. However, in fact the researcher was measuring the negative effect of Y on X, and overlooking the actual positive relationship between X and Y.

22 King, Keohane, and Verba (KKV) use the example of the endogenous relationship between party fragmentation and proportional representation to illustrate this issue (in the context of explaining the rise of the Nazi party in Weimar Germany). If an inference is made that proportional fragmentation causes party fragmentation when in fact the two variables are endogenous to one another, then prior party fragmentation will be performing the role of an omitted explanatory variable. KKV explain: "Thus, an endogeneity problem can become omitted variable bias. That is, prior social fragmentation is an omitted variable that causes proportional representation, is causally prior into it, and led in part to the fall of Weimar" (King et. al., Designing Social Inquiry, 189-91).

23 KKV illustrate this point with an example of the effect of constituency service on electoral advantage. They show that although researchers seeking to estimate the causal impact of service on vote share often see a negative effect, it would be unreasonable to conclude that increased constituency service causes decreased vote share. If constituency service in fact increases vote share, but expected strong vote share also decreases constituency service (the endogenous relationship), then the researcher might infer a negative relationship between service and vote share when in fact, in the component part of the relationship that the researcher is interested in, the relationship is positive (Ibid., 186).
Each of these methodological challenges posed by endogeneity is closely related to the other, and all three are derivative of the same basic problem, which is that the causal arrow does not always point clearly in only one direction.

**Parsing Causality from Endogeneity**

*Achievements and Deficiencies in Current (Quasi-)Experimental Work*

If the body of theoretical literature discussed here is thought of as a research agenda, then based on the methodological challenges posed by endogeneity, it is an especially challenging research agenda. The picture of civil war painted is one in which everything causes everything else. Stathis Kalyvas makes two prescriptions for gaining empirical traction on this kind of intense complexity. First, he recommends focusing on the micro-level of analysis of individual behavior and intra-community dynamics. Second, he prescribes fine-grained historical and quasi-anthropological work. Both serve the same purpose: to isolate individual relationships in their complex and shifting contexts, while maintaining acute awareness of those contexts.

Research on civil war over the past (approximately) five years from scholars expressing a renewed and intense interest in robust and creative methods for making strong causal inferences has followed the spirit, if not the letter, of Kalyvas’ recommendation. We will roughly categorize these approaches under the shared methodological heading “quasi-experimental approaches to civil war.” The chief payoff from quasi-experimental approaches has been strong identification of narrowly defined causal relationships, independent of pre-existing conceptual baggage, and often located at Kalyvas’ micro- and meso- levels of analysis. This group of literature can remain puzzling when considered as a whole, especially when comparisons of closely related studies reveal contradictory findings – a sign of a chronic inability to control enough of the variables that are at play. However, each individual study nevertheless represents a sharp cut through the tangle of endogeneity revealed by the previous wave of scholarship.

Quasi-experimental approaches have devoted less attention to the second piece of Kalyvas’ recommendation - context. While contributing stronger identification of causal relationships, this work has tended to do so at the expense of isolating those relationships from their theoretical

25 Ibid.
context. This is why a common - almost becoming cliché - criticism of quasi-experimental studies of civil war has been that they lack external validity. In other words, their generalizability to new cases is suspect. Jason Lyall’s 2009 article "Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya" provides a useful example. With insufficient attention devoted to external validity, when a methodologically robust quasi-experiment like Lyall’s overturns some conventional wisdom in the literature like the effect of indiscriminate violence on retaliatory attacks from Kalyvas (2006), there is little in the study to prevent Kalyvas from claiming that the findings do not generalize to the contexts he was theorizing.

Situating (Quasi-)Experimental Work in its Theoretical Context

If we can situate the specific empirical findings generated from quasi-experimental work in the context of theories that address endogeneity, there is significant promise for amplifying the utility of these findings. However, as in the example bringing Lyall’s 2009 article into conversation with Kalyas, 2006, external validity concerns consistently get in the way. Here, we will briefly illustrate this difficulty by locating one quasi-experimental study in the relevant body of theoretical work.

Lyall cuts through the endogenous relationship between authority and violence in civil war. He does so by identifying a specific causal relationship between indiscriminate state violence – associated in Kalyvas (2006) with poor authority - and retaliatory insurgent violence. By exploiting near-randomness in targeting by Russian artillery strikes in the second Chechen civil war, Lyall is able to infer what he argues is a counter-intuitive relationship: that near-random artillery shelling reduces retaliatory violence from Chechen rebels.

The broader implication of Lyall speaks to the relationship between state authority and insurgent violence: considered within the theoretical framework of Kalyvas (2006), indiscriminate violence occurs when selection criteria are rough and precise information is unavailable, which in turn describes areas where states lack territorial control. This is consistent with Lyall's Chechen case, where the Russian forces using artillery do not physically occupy the targeted villages. Furthermore, Lyall's findings shed light on indeterminate and endogenous dynamics from Kalyvas' framework: if indiscriminate violence imposes costs on civilians who would otherwise "free ride" on rebel violence, it may motivate civilians to seek protection from either insurgents or incumbents. This is

the theoretical indeterminacy that Lyall seeks to clear up – with authority endogenous to violence, the direction of the causal arrow from violence to authority is unclear.\textsuperscript{27} This will in turn affect the coercive capabilities of either side and therefore the distribution of territorial control, which in turn affects the decision to use indiscriminate violence. So it is informative in Lyall’s case that we see, under particular conditions, indiscriminate violence causing civilian collaboration with incumbent forces.

However, although we can draw a direct connection between Lyall's findings and Kalyvas' theorizing on the endogeneity of violence to authority, we cannot be fully confident in the exercise due to issues of external validity. For example, Lyall's Russian artillery strikes only produce a handful of victims - at least 265 deaths over 5 year period, or about 4 per month. Insurgent violence is also low - treated (shelled) villages saw an average of 2.11 attacks pretreatment and 1.5 attacks post-treatment within a 90-day treatment window.\textsuperscript{28} Is this violence of sufficient magnitude that, in the stories told by Kalyvas, it should drive civilians to participation, thereby feeding authority of one side or the other? Does this violence make “free riding” sufficiently costly? Without a clear understanding of this scope condition from Lyall, it is difficult to make a reasonable decision on whether these support or refute Kalyvas (2006) or Kalyvas and Kocher (2007)’s theoretical claims about the endogenous relationship between authority and violence.\textsuperscript{29}

**Why External Validity is a Problem**

It is important to note that there is nothing inherent in the quasi-experimental approach that places undue or disproportionate restrictions on external validity. It is true that a quasi-experiment will

\textsuperscript{27} Kalyvas and Kocher, "How Free is 'Free Riding' in Civil Wars?" 186-7; Lyall, "Does Indiscriminate Violence Incite Insurgent Attacks?" 334-7.

\textsuperscript{28} Lyall, "Does Indiscriminate Violence Incite Insurgent Attacks?" 349-51.

\textsuperscript{29} We can perform a similar exercise by attempting to generalize Lyall (2010) and Lyall et. al., (2014). These quasi-experimental studies examine the causal role of co-ethnic bias in wartime informing, though in very different contexts. Lyall (2010) exploits variation in the ethnicity of military units conducting "sweep operations" for Russian forces in the second Chechyan war, and finds that sweeps from military units comprised primarily of recruited Chechens were associated with a reduction in subsequent insurgent attacks compared to sweeps from military units comprised primarily of recruited Russians (Lyall, "Are Co-ethnic More Effective Counter-Insurgents?" 3-4). Lyall et. al. (2014) intervene in an International Security Assistance Force in Afghanistan (ISAF) program called "Guardians of Peace" (GP), which promotes civilian collaboration with a “See Something, Say Something”-type campaign. Lyall et. al. vary the ethnicity of public faces endorsing the GP program, and use surveys to measure support for the program as a proxy variable for likelihood that a respondent would provide information to state forces (Lyall et. al., "Coethnic Bias and Wartime Informing." 9-11). Lyall, 2010 examines ethnic defection and the identification problem in zone 2 of Kalyvas, 2006’s typology, where incumbent control is incomplete; Lyall et. al., 2014 examines ethnic defection and the identification problem in zone 1, where incumbent control is very close to complete. However, for each it is unclear how far beyond those zones, strictly defined, each finding travels.
necessarily take place in a single idiosyncratic setting. Although experiments or quasi-experiments will often take place in particularly idiosyncratic settings - especially when studying civil war, where field work is often limited to where U.S. military or development work is present - this is only an artifact of the practical measures that the method entails, not an inherent characteristic of the method.

One difference is that the comparative case approach has developed, at least in its most persuasive applications, very conscious attention to the careful specification of scope conditions. As Daniel Ziblatt explains, "With the aim of avoiding the pitfalls of ‘overreaching’ [or under-reaching], middle-range theorists identify the parameters within which a theory can be expected to be valid, thereby establishing a reasonable standard against which the work can be judged." Unfortunately, methodological discussions of quasi-experimental approaches have not had the benefit of receiving as much deliberate attention to the issue of generalizability and scope conditions. For example, the chapter on “design-based inference” in Brady and Collier’s *Rethinking Social Inquiry*, a methodological text that appears on many introductory graduate syllabi in political science, devotes only a few sentences to the issue of external validity. However, there is very little that precludes comparable attention to scope conditions for experimental and quasi-experimental work. To illustrate, Dunning’s characterization of external validity as more broadly asking "whether the intervention…in fact yields causal inferences about the real causal hypothesis of concern, and for the units we would really like to study" is not very different from Ziblatt’s version of the question, "what are the outer limits or ‘reach’ of a theory’s applicability?"

Since the application of quasi-experimental designs does not necessitate undue or disproportionate restrictions on external validity, how do we explain why these studies have devoted comparably less attention to the issue of generalizability and scope relative to, for example, the comparative case study approach? One explanation may be that for many quasi-experimental studies, undertheorizing scope conditions may simply be the opportunity cost associated with limited space and a responsibility to very thoroughly describe treatments, interventions, or identification strategies. If

---

30 Brady and Collier, *Rethinking Social Inquiry*, 300; However, this is also true of any case study. In fact, in a traditional comparative case design that adheres to Mill's method, cases are chosen firstly according to their appropriateness for method of agreement or difference designs, and only secondly according to their generalizability (Sekhon, "Quality Meets Quantity," 282-4).
31 Ziblatt, "Of Course Generalize, but How?" 9.
quasi-experimental designs are to keep their comparative advantage in making strong causal inferences, then it would not be useful to reduce the amount of attention devoted to research design in published studies.

What is needed then is some alternative, complementary way of providing theoretical context that sheds light on external validity without sacrificing gains in internal validity. We propose agent-based modeling as a means to that end.

Agent-based Modeling

*What it is*

Agent-based modeling is a type of formal modeling related to, but distinct from, game theoretic modeling. As in game theoretic modeling, agent-based modeling formalizes a theory by operationalizing simple rules of interaction between internally coherent, boundedly rational actors (here, agents). However, instead of maintaining the parsimony of a two- or three-actor model, the agent-based approach models a *landscape* of hundreds or thousands of simultaneously interacting agents. The consequence of this difference is that, although rules of interaction at the micro-level are themselves parsimonious and deliberately specified, patterns that emerge at the meso- and macro-levels - that is, across whole parts or the entirety of the model landscape - are the result of complex chains of interaction and, though logically consistent, are not predictable at the outset by pure deduction alone. This is why agent-based modeling is a natural approach for studying endogeneity, and why it is often termed a "bottom-up" approach that captures "emergent" phenomena in descriptions like this one in Michael Laver and Ernest Sergenti’s book on political party competition:

"Agent-based models (ABMs) are "bottom-up" models that typically assume settings with a fairly large number of autonomous decision-making agents. Each agent uses some well-specified decision rule to choose actions, and there may well be considerable diversity in the decision rules used by different agents. Given the analytical intractability of the decision-making environment, the decision rules that are specified and investigated in ABMs are typically based on adaptive learning, rather than forward-looking strategic analysis, and agents are assumed to have bounded rather than perfect rationality (Gigerenzer and Selten 2001; Rubinstein 1998; Simon 1957). ABM is a modeling technology that is ideally suited to investigate outcomes that may emerge when large numbers of boundedly rational agents, using adaptive decision rules selected from a diverse portfolio of possibilities, interact with each other continuously in an evolving dynamic setting (MacGregor et al. 2006)." [emphasis mine]33

Most researchers in political science who use ABM adhere to Robert Axelrod’s "Keep It Simple Stupid" (KISS) principle. This approach to the model design minimizes the number of parameters

operationalized in order to explore the macro-level effects of small or simple variations in decision-making rules at the micro-level in an abstract, non-representative setting. Here, we use an ensemble-type model, which is more complex than Axelrod's type and is intended to capture a theoretically relevant class of cases, but without going so far as to represent a specific empirical context with the use of real data. With this modeling decision, what we lose in parsimony we gain in verisimilitude; this suits our purpose of testing the implications of theories of civil war already in the literature.

How it can be Useful

There are a number of methods we could turn to for testing civil war theories, contextualizing quasi-experiments, and mediating between the two. There are even other approaches to modeling that we could employ. Why not use those? We contend that agent-based modeling possesses an ontological congruence with theories of civil war that emphasize endogenous processes, which more well-established and widely-used methodologies lack. In other words, the basic conceptual pieces that feature prominently in theories of endogeneity in civil war, are also the basic conceptual pieces that anchor the agent-based modeling approach.

Scholarship that theorizes endogenous dynamics in civil war tends to operate at the intersection between two levels of analysis: the micro-level of individual interaction, and the meso-level of group dynamics. This is particularly true of recent scholarship: Staniland (2012) describes "the pivotal ‘meso-level’ analytical space between the local and the national where states are forged or shattered, populations displaced, and new forms of governance built," Christia (2012) "takes this meso-level approach and looks at the role of local elites," and Kalyvas (2006) contends that "the relation between political actors and underlying populations must be problematized rather than assumed away - which defines the content of the second (or meso) level." In contrast, the empirical findings offered by quasi-experimental work have tended to speak to the micro-level of individual motives. The emphasis on the micro-level of analysis is not an inherent feature of the quasi-experimental approach; rather, it is an artifact of the particular methods that are typically available to researchers.

34 Lustick and Miodownik, "Abstractions, Ensembles, and Virtualizations."
35 Staniland, "States, Insurgents, and Wartime Political Orders," 246; Christia, Alliance Formation, 42; Kalyvas, The Logic of Violence in Civil War, 11.
36 Blair et. al., "Poverty and Support for Militant Politics;" Blattman and Annan "Can Employment Reduce Lawlessness and Rebellion?" Lyall et. al., "Explaining Support for Combatants in Wartime." For example – Blattman and Annan, 2014 address individual opportunity costs for potential high-risk rebel participants; Lyall et. al., 2013 explain individual attitudes toward combatants in Afghanistan; Blair et. al., 2013 explain similar individual attitudes amongst impoverished subjects towards combatants in Pakistan.
that aspire to strong causal inference in conflict settings. For example survey data, which comprise a large portion of the data collected from quasi-experimental designs, necessarily provide information on micro-level individual motives and behaviors. Any method that seeks to bring this type of work into conversation with the theories of endogenous dynamics needs to be able to go from micro-motives to meso-behavior.

Formal game theoretic models target the micro-level, but go from micro-motives to micro-behavior. In doing so, they reproduce the kind of information that we gain from quasi-experimental work. While game theoretic models may be useful in providing the logics behind particular experimental findings, without modeling endogeneity they are less helpful for providing the kind of scope conditions that we are looking for. Large-n work that uses regression analysis has featured prominently in the civil war literature, and was especially useful for broad ground-clearing exercises in the late 1990s and early 2000s. However, this kind of work goes from macro-motives to macro-behavior, so it would be inappropriate as a mediator between meso-level theory and micro-level experimental findings. In fact, generalizing macro-level findings to lower levels analysis can be counter-productive, and has been criticized for being "a premise akin to that of ecological fallacy: in the absence of individual-level data about particular acts of violence, we tend to extrapolate from the aggregate down to the individual level." For this reason, agent-based modeling offers a comparative advantage relative to existing approaches in its capacity to model micro-to-meso level dynamics.

In addition and related to this congruence of levels of analysis, there are two characteristics of agent-based models that make them a particularly appropriate candidate for intervening between theories of endogeneity and experimental and quasi-experimental findings. First, for any causal claim, the model allows the researcher to see the counterfactual.

37 Collier and Hoeffler, "Greed and Grievance in Civil War;" Fearon and Laitin, "Ethnicity, Insurgency, and Civil War;" Sambanis, "What is Civil War?"
39 For a very up-to-date understanding the counterfactual model of inference as it is often used in quasi-experimental approaches (and of current contributions to political science from econometrics more generally), see Angrist and Pischke, Mostly Harmless Econometrics, 2009 and Angrist and Pischke, Mastering Metrics, 2014. For a less technical overview of the role of counterfactuals in making causal inferences, see Fearon, "Counterfactuals and Hypothesis Testing," 1991. For an exploration of counterfactual thinking and the role that agent-based modeling as an instrument for counterfactual experiments, see Lustick, “Tetlock and Counterfactuals,” 2010.
the inductive approach, the modeler runs the model some large number of times, identifies those runs with and without the independent variable of interest (or above and below some theoretically specified threshold if the independent variable of interest is continuous), and then observes differences in the dependent variable between the two newly created subsets. A second approach is deductive, in which the modeler changes the independent variable of interest before running the model, so that a true "treatment" and "control" group is created. The proof-of-concept provided here uses the first, inductive approach.

The deductive approach, in which the model is used to simulate a kind of counterfactual experiment, is not one exclusive to agent-based modeling. However, the ability to conduct counterfactual analysis in this way is especially useful when combined with the second advantage posed by agent-based modeling – that endogeneity can be modeled explicitly. This is the payoff from modeling a landscape of agents instead of the small number that would appear in a formal game theoretic model – although parameter specifications are parsimonious and tractable, how those specifications interact, combine, and overlap to generate criss-crossing chains of causation and non-intuitive second- and third-order effects is not predictable by pure deduction alone. While a game theoretic model may formalize an intuitive logic, an agent-based model shows the wider array of outcomes possible from that logic, without sacrificing logical coherence (or else the model would not run). This allows the researcher to conduct counterfactual experiments that play with the interactions between multiple theoretical propositions and frameworks.

It is important to keep in mind, however, that because an agent-based model is a model, it is inherently limited by the boundaries of the theoretical parameters used to specify it. Outcomes that the model generates will always be outcomes given particular decisions about model design and parameter settings. For this reason, the model is best used as a complementary tool rather than in isolation, to test existing theoretical hypotheses and to generate results that inform further empirical work. This is exactly what we propose here. That agent-based models explicitly operationalize endogenous processes speaks directly to those theorists of civil war that emphasize the endogeneity of identity, authority, and violence. That the models can be used to test counterfactual claims speaks to those

---

40 It should be noted that although using this method allows causal claims (initial conditions led to these differences), potential intervening variables could still introduce a difficulty of understanding causes more proximate than the initial treatment.

41 For example, game-theoretic models are often used with this approach.
researchers interested in performing quasi-experiments. As a kind of "sandbox" for testing theoretical hypothesis, locating the conditions under which those hypothesis hold or change, and deriving scope conditions, agent-based modeling can serve as a formal method for adjudicating between the two bodies of work.

Model Design
To demonstrate the utility of an agent-based model in this methodological role, we provide a proof-of-concept with a model of an abstract civil war environment. The burden of proof for this model as a demonstration of the method comes in three parts: first, to prove that an agent-based model can generate results that have reasonable face validity and usefulness; second, that the development and operation of an agent-based model can serve as an analytically useful exercise for disciplining the operationalization of complex theories of civil war; third, that an agent-based model can provide guidance as a mediator between meso-theories and potential quasi-experimental work.

To these ends, we present a model design that targets those areas of theorizing that most strongly emphasize endogenous dynamics in civil war. Our design interacts closely-knit groups of theory on identity, authority, and political violence in civil war. Since we intend to use our results as guidelines for more narrowly focused and experimentally oriented research, interacting groups of theory is more informative than trying to maintain absolute fidelity to a single theory. This approach will allow us to propose quasi-experimental work that can, in a more informed way, weigh the relative explanatory power of some theoretical premises over others under differing conditions. Finally and most importantly, the model will provide empirical predictions with clear scope conditions along multiple dimensions. Clear scope conditions along multiple dimensions is key, because it will be the basis on which we would assess the comparative usefulness or relevance of potential or completed quasi-experimental findings.

Our dependent variable is patterns of violence, observed at the meso-level of analysis. Our model results should answer the question: on what political basis is violence performed? For example, model results may show the organization of agents into ethnic blocks, with violence perpetrated across ethnic identities but not within them. Or, model results may show the organization of agents into local factions that cross-cut ethnicity, with violence having an indeterminate relationship with identity. This dependent variable, which we look for at the group-level, is common across multiple
theories of endogeneity. In the world of Stathis Kalyvas, patterns of violence are likely to emerge along frontiers between insurgent and incumbent; in the world of Paul Staniland, Ana Arjona, and Zachariah Mampilly, patterns of violence are likely to follow boundaries between competing wartime authoritative orders. For some theorists of ethnic conflict, group-level violence occurs between pre-defined identity groups: for Barry Posen and Chaim Kauffman, horizontally between opponents in a security dilemma; for Roger Peterson and Lars-Erik Cederman, vertically across political hierarchies. This dependent variable directly captures Kalyvas and Kocher (2007)’s endogenous cleavages thesis that "civil war may cause violence in a way that is relatively autonomous from the cleavages that led to the war in the first place."42

The chief independent variables in the model are the distribution of authority and identity. We describe each here in turn.

Identity is operationalized as the basic logic by which agents at the micro-level interact with one another, and is modeled to capture a basic 'thin' version of constructivism reflected in the body of scholarship represented by David Laitin and Kanchan Chandra. This strand of constructivist identity theory sees identity affiliation as an attribute that is fungible, multiple, variable in salience or 'stickiness,' but also instrumentalizable. To operationalize this large body of theory in a simple way, we endow each agent with a repertoire of identities, reflecting the idea of an "operative" repertoire of identity categories from Chandra, 2012 and Christia, 2012.43 Each agent actively publicizes one identity at a time while simultaneously including multiple identities in its repertoire.44 Borrowing Chandra’s vocabulary, an agent may "switch" those identities included in its repertoire, or "reclassify" to activate on a new identity category already included in its repertoire.45

42 Kalyvas and Kocher, "Ethnic Cleavages and Irregular War," 204.
43 Christia, Alliance Formation, 47; Chandra, Constructivist Theories of Ethnic Politics, 1-50.
44 This operationalization is identical to that described in Lustick, “Agent-based modeling of collective identity”
45 Chandra, Constructivist Theories of Ethnic Politics, 20-1.
Reclassify and switching identities is the mode by which agents interact with one another; at each timestep, an agent surveys its neighborhood, assesses the relative political desirability of the identities in its repertoire as determined by hierarchal position in one or another authoritative order (described in the following section), and switches or reclassifies those identities in order to maximize political desirability. Because identity determines relative position in authoritative orders, it is the interaction of identity and authority that determines political behavior. This operationalization, in which identity affiliations can change in their salience for determining political behavior and can be congruent with or cross-cut patterns of authority, is how this model design allows identity and authority to be endogenous to one another. This formulation can generate meso-level model outcomes that look like the world of Posen or Kaufman, where authority and identity are congruent and identities are ethnicized and polarized, as well as outcomes that look like the world of Stathis Kalyvas, where (ethnic) identity is minimally relevant and ethnic defection is frequent. For example, "switching" makes possible what Kalyvas, 2008 called identity shift ("the acquisition by individuals of a new ethnic [or national] identity that replaces the old one") and "reclassification" makes possible ethnic defection ("a process whereby individuals join organizations explicitly opposed to the national aspirations of the ethnic group with which they identify").

To capture the image of wartime order theorized by scholars like Staniland, Arjona, and Mampilly, we model orders as locally defined, fluid and competing. Agent behavior is governed by mutually exclusive orders that vary in breadth over time and space. This modeling decision seeks to operationalize Staniland (2012)'s insight that "states and insurgents are not simple-minded maximizers of monopoly but instead are optimizers of authority in complex, often counterintuitive, interaction with other armed actors." Orders have differential capacities to elicit compliance from those agents that belong to them, reflecting the distinction that Staniland, 2012 makes between segmented versus fragmented within-order distributions of control, as well as the variation in insurgent rank-and-file discipline theorized by Humphreys and Weinstein (2006). Orders also have

\[46\] Kalyvas, "Ethnic Defection in Civil Wars," 1045.
\[47\] Staniland, "States, Insurgents, and Wartime Political Orders," 244.
\[48\] Ibid., 247; Humphreys and Weinstein, “Handling and Manhandling Civilians in Civil War.”
differential capacities to produce violence, reflecting the distinction that Staniland, 2012 makes between violence and control, and will direct violence both inwardly and outwardly against those identity groups at opposing ends of the order’s own internal hierarchy.\textsuperscript{49} Although orders elicit compliance on a range of political behaviors, to maintain parsimony in our model orders only affect (1) who agents listen to and (2) against whom agents perform violence.\textsuperscript{50} Finally, orders have differential capacities to persist as coherent entities. Agents are allowed to join and leave one order for another, affecting at the micro-level with whom that agent will interact and "listen to," and at the meso level the overall breadth of one or another order.

Reflecting on work by Cederman and Peterson on the role of identity in civil conflict, we operationalize identity groups as hierarchically arranged within authoritative orders. A particular agent’s position in an order is determined the identities that agent holds – specifically by the degree to which an agent shares identities with the group that dominates that particular order. Still drawing on Peterson and Cederman, it is politically desirable to be closer to the top of a hierarchy, so that agents at the bottom of a hierarchical order will move upward if they can, either by allying with their own dominant group (‘voice,’ crudely conceived) or by joining a different order (‘exit,’ crudely conceived).\textsuperscript{51} Thus, the logic of identity groups seeking dominance in hierarchies also provides the logic by which authoritative orders grow, shrink, cohere, and collapse. More specifically in the model, agents can choose to “exit” (or alternatively “enter”) a new zone if they are isolated from the political center in their own zone.\textsuperscript{52} This allows the size of the zone to change over time depending on how inclusive either zone (state or insurgent) tends to be. In this way, authority and political violence are rendered endogenous to identity, and vice versa.

\textit{Counterfactual Experiment Design}

As a loose proof-of-concept that an agent-based model can produce results with reasonable face validity, can be an analytically useful exercise for disciplining the operationalization of complex theories of civil war, and can provide some guidance in adjudicating between theories of endogeneity.

\textsuperscript{49} Staniland, "States, Insurgents, and Wartime Political Orders," 245.
\textsuperscript{50} Arjona, "Wartime Institutions: A Research Agenda;" Staniland, "States, Insurgents, and Wartime Political Orders."
\textsuperscript{51} Hirschman, \textit{Exit, Voice, and Loyalty}.
\textsuperscript{52} Since our agents are only boundedly rational, they may find that their new zone is also isolating and they may also switch back. This operationalization of authority is a modified version of the Dynamic Political Hierarchy described in Lustick, “From Theory to Simulation.”
and potential experimental work, we ran the following counterfactual experiment. The experiment uses a model designed according to the specifications described earlier.

An agent-based model – particularly the ‘ensemble’-type model that interacts multiple theories – will generate data using multiple operationalized variables (as well their simple statistical transformations). Each variable exists in at least two dimensions: ‘vertically’ through time in a single model run, and ‘horizontally’ across multiple model runs at a single moment in time (a ‘time-step’). For this reason, the options for investigating one "batch" of model output are numerous. We choose to show a small handful of simple visualizations as a rough ‘plausibility probe’ that tests the usefulness of the model for testing a hypothesis from the theoretical literature and for providing some kind of useful information on the external validity of potential quasi-experimental work.

We explore a hypothesis taken directly from Kalyvas and Kocher (2007). Kalyvas and Kocher hypothesize that "unified states are likely to subsume ethnic cleavages into irregular wars, whereas the fragmentation of the state structure is more likely to allow the transformation of ethnic cleavages into ethnic war."\(^5\) It should be noted that this hypothesis is not specified as holding absolutely across all conditions – they write: "civil wars between a strong state and a weaker but well-organized challenger will take the form of irregular war; in contrast, weak, fragmented states lower the threshold of military contestation and invite badly organized challengers; both may be forced to rely on the ‘cheapest’ mobilizable networks, including ethnic ones."\(^5\) The model results presented below explore this hypothesis. It is important to note that what the model is doing is not validating or invalidating Kalyvas and Kocher’s hypothesis.\(^5\) Rather, it is illustrating the conditions under which the hypothesis holds, given our particular theoretically grounded choices in model design.

We ran multiple versions of the model, varying the "distribution of authority" independent variable, but here we report only on an experiment initialized with a medium-sized insurgency. The model was run 1,000 times, with each run comprising 100 time steps. We chose to focus on the following proposition from Kalyvas and Kocher’s hypothesis: that, “weak, fragmented states lower the

\(^5\) Ibid., 214
\(^5\) Validating or invalidating Kalyvas and Kocher’s hypothesis would be most convincing with empirical work, as in experimental and quasi-experimental approaches. However, it should be noted that if we were to observe very few or zero model runs in which the hypothesis holds, and if we were reasonably confident in the model design, then we would have some reason to be skeptical about the validity of the hypothesis.
threshold of military contestation and invite badly organized challengers.” We operationalized relative organization of state versus challenger by measuring state control within its own zone and insurgent control within its own zone. We find that there is a positive relationship between these two variables, but also that his trend does not hold absolutely across all conditions.

**Results**

First we will provide some descriptive observations. In the model design, we allowed agents to join and leave one order for another, to operationalize differential capacities for orders to persist as coherent entities. Figure 3 illustrates the aggregate consequence of this dynamic in our model runs. The figure shows the size of the insurgent order over time, measured as a percentage of agents governed by that order over total agents within that order.

![Size of Insurgent Zone over Time](image)

*Figure 3*

Figure 3 is a modified box-and-whisker plot, where the x-axis is time, the y-axis is the size of the insurgent order, and instead of outlier points, we provided outlier lines graphed over time for each individual run. Each box-and-whisker summarizes insurgent order size at that moment in time across all 1,000 runs. All of the runs start with an insurgent order that occupies about 6.5% of the landscape. In general, the median size of the insurgent zone decreases over time; however, we also see a significant number of runs where the insurgent zone can expand to almost 25% of the
landscape. This initial descriptive look tells us that fragmented authority does not deterministically encourage or quell insurgent success, and that there is significant variation in insurgent zone size across the experiment.

Our model design stipulated a second dimension of insurgent order coherence - the differential capacity to elicit compliance from those agents that belong to the order. This concept of “insurgent control” is operationalized as the percentage of the insurgent-controlled zone that affiliates with that zone’s most dominant identity. Figure 4 compares size of insurgent zone with this operationalization of insurgent control. The overall decrease moving across the bar chart from left to right suggests a trade-off between size and control in the insurgent zone. In order for insurgent zone control to be above 50%, zone size must suffer.

![Size of Zone by Insurgent Control](image)

*Figure 4*

State control is operationalized in the same way as insurgent control: as the percentage of the state-controlled zone that publically affiliates with that zone’s most dominant identity.

Violence in the model can be categorized simply according to whether it originates from agents in the state order or from agents in the insurgent order. Figure 5 graphs these two categories of violence across an x-axis of insurgent control.
From this graph we can see how attacks from members of each order trend when compared against insurgent control. We can account for the upside-down “U” shapes in Figure 5 by considering the location of potential targets. Violence may be targeted either at agents belonging to the opposing zone, or at hierarchically lower agents within the same zone (potential “defectors” that, according to the model design, may switch loyalties). On the left side of the graph we would expect to see more internally-directed violence against potential defectors (where insurgent control is low) and the right side we would likely see more externally-oriented violence against the state (high insurgent control.) In the middle, however, we see the accumulation of both resulting in more overall attacks. From the state’s point of view, on the left side of the graph, the more poorly controlled insurgent zone presents a greater number of targets (because control tends to be negatively associated with zone size); on the right side of the graph, increased insurgent control implies a smaller, more coherent insurgent zone, but alongside a larger, less controlled state zone with more potential targets at the opposing end of the state’s internal hierarchy. Again, in the middle we see the accumulation of both resulting in more overall attacks from the state order.

With a general sense of the model output in hand, we can investigate Kalyvas and Kocher’s hypothesis about the effect of state fragmentation on organization of challengers. Figure 6 compares the level of state control against insurgent control.
In general, the model results show a positive relationship between state control and insurgent control. This finding substantiates generally the claim that a badly organized state invites badly organized insurgencies. The model dynamic driving this general trend is the relationship between the control by an authoritative order and that order’s capacity to deploy violence. On the left side of the graph when state control is low, the state order is unable to direct sufficient violence to quell the insurgency. As we move from left to right, stronger state control only allows those insurgencies that are most well-organized to survive, so that insurgencies with low control are selected out. Put differently: disorganized insurgencies are allowed to exist when state control is low but are quickly defeated when state control is high, leaving only the most well-organized insurgencies.

However we can also see that this finding does not hold true consistently across all levels of state control. Although the trend appears on the left and right sides of the graph, the middle section implies that state control has no effect on insurgent control. This plateau in the middle of the graph represents a scope condition on the Kalyvas and Kocher hypothesis; where the state is “semi-organized,” with these model dynamics, the hypothesis does not hold. In our model results, variation in state control within the category of “semi-organized states” has no impact on the types
of challengers that arise, but highly disorganized states invite badly organized challengers, and well-organized states invite well-organized challengers.

So according to our model, in which quasi-experimental settings would we expect Kalyvas and Kocher’s hypothesis to hold? Where a quasi-experiment is conducted in the context of total state failure, we should expect to see badly-organized challengers and, by extension according to the theoretical logic, ethnicized conflict. However, where a quasi-experiment is conducted in the context of a weakly organized but still viable state, we should plan to place stricter limitations on the external validity of our findings. Furthermore, we may be able to specify these takeaways more precisely by continuing to explore these modeling results in creative ways, and by testing alternative operationalizations and parameter settings.

Conclusion

In this paper, we used an agent-based model to interact theories of identity, authority, and violence in civil war. A hypothesis about the relationship between authority and ethnic violence from Kalyvas and Kocher, 2007 was operationalized, and results were used to generate some theoretically informed expectations for scope conditions on potential quasi-experimental work.

This proof-of-concept illustrated a wider methodological argument: that agent-based modeling can serve a useful role as mediator between theories that emphasize the endogeneity of identity, authority, and violence in civil war on one hand, and quasi-experimental findings that make strong causal inferences on the other. By providing to theorists of civil war a way to operationalize endogeneity in a disciplined way, agent-based modeling can help move forward a research agenda that has posed a difficult methodological challenge. By providing a "sandbox" for simulating counterfactual experiments to researchers, this type of modeling can help mitigate questions of external validity associated with quasi-experimental designs. This paper makes a first step in the direction of locating a useful methodological place for agent-based modeling in the research agenda on civil war.
Works Cited


