When Do Civil Wars Begin?
Evidence from Markov Chains

Iris Malone*

May 2019

Abstract

When do armed groups escalate their campaigns to civil war? Some armed groups operate for years before launching insurgencies, while other armed groups rapidly transition to war. I explain this variation in terms of when an armed group begins guerrilla operations. These violent activities have two competing effects: it builds a group's capabilities to fight, but also reveals information about those capabilities to the state. Consequently, an armed group is most likely to escalate its campaign to civil war early on because the state is most uncertain over whether to act. I measure variation in armed group activities by leveraging overlaps in the population of armed groups across different conflict datasets. I apply Markov Chain modeling techniques to estimate the probability of civil war at any given time given information about the current state of an armed group's campaign. I also examine how the timing of civil war differs across rebel groups. The results show the probability of civil war spikes during the first two to three years of a campaign then dissipates. These findings advance scholarly understanding about stochastic processes and conflict escalation.

*Department of Political Science, 100 Encina Hall West, Stanford University, Stanford, CA 94305.

1 Introduction

When do armed groups escalate their campaigns to civil war? Armed groups often operate for years before civil wars erupt, but when and why they escalate remains relatively unclear. For example, the Tamil Tigers first launched its campaign in 1972, but its efforts did not reach conventional levels of civil war violence until 1984. Similarly, the Somali Democratic Force formed in 1978, but did not appear in conventional civil war datasets until its campaign intensified in the early 1980s. The timing of these wars represents an underdeveloped aspect of conflict escalation research.

Conventional wisdom suggests that civil wars begin following an uptick in political violence.[1] Escalation occurs when an armed group accumulates enough strength to surpass the violence threshold demarcating civil war. In order to amass the capabilities to meet this requirement, armed groups use guerrilla and terrorist tactics to acquire supplies, attract support, and intimidate opponents (Crenshaw 1978, 1981; Merari 1993; Carter 2016). As an armed group conducts more attacks, it signals that its capabilities are growing to the point where the group is posed to enter civil war territory (Kydd and Walter 2006). Yet, this view of escalation leaves a troubling paradox. If a record of violent activity reveals a group's growing capabilities, then the state can infer when an armed group will reach the civil war threshold. If this is so, then the state should be able to use this information to act early enough to suppress the armed group. Counterintuitively, this predicts that a history of political violence is inversely related to the timing of civil war.

I offer an alternative explanation: campaigns are most likely to escalate to civil war in their first few years of operation. Guerrilla and terrorist tactics are a double-edged sword. They can help an armed group grow, but also make an armed group increasingly susceptible to detection and destruction the longer it operates. For example, hit-and-run guerrilla attacks against military posts can help an armed group secure weapons, but these attacks

[1] See, for example, Collier and Sambanis (2002), Walter (2004), Findley and Young (2012), and Daly (2012) whom all find previous conflict predict future conflict. For an exception, see Sambanis (2008) and Sanchez-Cuenca and La Calle (2011).

can also reveal information about the group, making it vulnerable to counterterrorism and counterinsurgency operations. These two competing effects of guerrilla violence – building a group's capabilities and revealing those capabilities to the state - affect the timing of civil war. The more an armed group can grow before it begins to use violence, the sooner its campaign will escalate.

The risk of escalation peaks shortly after an armed group first emerges when it is relatively unknown. Hampered by uncertainty over an armed group's potential capabilities, the state risks missing its window of opportunity to suppress the group in one of two ways. First, the state may wait too long to react, allowing the problem to fester and boil over into civil war. Second, the state may act too early, cracking down prematurely and provoking a sudden backlash. I identify different catalysts that cause these mistakes to escalate into civil war. In some cases, the precipitating event is simply the low-intensity accumulation of guerrilla violence or an unexpected shock in the political environment. In other cases, it is a sudden high-intensity attack by the rebel group or state.

I test this logic through a Markov chain analysis of civil war escalation. I identify four possible states of an armed group's campaign, which correspond to different levels of militant activity. I create a series of Markov Chains and estimate the probability of civil war in the next year given information about the current state of an armed group's campaign. I use multinomial logistic regression to determine which types of armed groups are more likely to escalate through different catalysts. Finally, I estimate the window of time when different rebel campaigns are most likely to turn into civil wars.

I show three results. First, the probability of civil war spikes during the first two to three years of a campaign before dissipating. The cumulative probability distribution predicts that approximately 50% of rebel groups escalate their campaigns in the first three years. Second, the probability of civil war begins to decline once an armed group reveals information about its potential capabilities by initiating the use of violence. Third, the expected time to civil war is shortest when rebel groups have external bases or are already organizing at the time of a shock.

This research makes three general contributions to conflict escalation research. It offers one of the first analyses on the precise timing of conflict. Civil conflicts are most likely to emerge when the state has the least amount of information about an armed group's capabilities. This generally falls towards the beginning of an armed group's campaign. It also provides a new way to think about repeated interactions between an armed group and the state. Markov processes model the different sequences of interactions between parties and the long-run probabilities of conflict between these dyads. Finally, it introduces the utility of Markov chain methods to track the evolution and timing of armed conflict.

The rest of this paper proceeds in four parts. In the next section, I review existing theories of conflict escalation from the interstate war literature and explore where they can and cannot apply to civil war. I present a new explanation for conflict escalation which predicts that the risk of conflict is highest shortly after an armed campaign starts. The theory makes observable predictions both about when an armed group's campaign should escalate to civil war as well as how and when different types of rebel groups escalate. The third section conducts a Markov chain analysis in order to trace the predicted probability of civil war during the course of a campaign. I complement this with evidence on how different types of rebel groups escalate and why the timing of civil war varies among them. The paper concludes with avenues for future research on conflict escalation.

2 Existing Explanations

Despite a robust literature on conflict escalation dynamics, scholars still know relatively little about when these events occur. A central challenge to answering this question has been the selection bias between which dyads are at risk for conflict and which actually escalate. Comparing escalation patterns across dyad with different propensities for conflict risks underestimating when and how these conflicts can occur.

Early work focused on two-stage selection models to navigate this challenge. It sought to identify what conditions made certain dyads disposed for 'rivalry' and then what conditions made these rival dyads more likely to experience militarized disputes or interstate war (Diehl 1992; Reed 2000; Lemke and Reed 2001; Kinsella and Russett 2002). A key finding was that conditions which made rivalry more likely seemed to paradoxically make escalation less likely. Reed (2000) conjectured asymmetric information about state resolve would initially increase the likelihood of a rivalry by raising tensions. Repeated interactions between states would reveal information during the course of a rivalry, delaying escalation from occurring. However, this finding was largely disproven on methodological grounds because of assumptions made about the correlation between the error term in the two equations. When this assumption is changed, the same factors that predict rivalry also predict conflict escalation (Sartori 2003).

For years, there was no commensurate set of explanations in the civil war literature because of a similar selection bias. This has begun to change. Bartusevicius and Gleditsch (2019) perform a similar two-stage analysis and argue that ethnic grievances predict incompatabilities, but not conflict onset. Condra (2010) and Cederman, Wimmer, and Min (2010) show certain ethnic groups are more prone to civil conflict than others. However, grievances and ethnic groups may still be too imprecise to solve to selection problem. Armed groups have incentives to hide and avoid detection, making it hard to identify a population of armed groups at-risk for civil war (Lewis 2017).

If there is a prevailing view in the civil war literature about when civil wars begin, it is that that they follow from low-level protests, political instability, or violence. Terrorism and insurgency scholars have characterized the use of low-level political violence as strategic because it can help an armed group signal its capabilities and grow (Lake 2001; Pape 2003; Kydd and Walter 2006). These attacks do not happen in a vacuum: an armed group strategically condition these attacks based on the state's anticipated response (Carter 2016). As a result, current violence tend to positively predict future conflict dynamics. Empirical evidence bears this out. Findley and Young (2012) show higher levels of terrorism precede civil war onset. Walter (2004) and Daly (2012) both find areas with a history of conflict are more likely to see future violence.

This research conjectures whether a potential conflict will escalate or not. However, it does not explain *when* this conflict will escalate. The closet set of literature on the timing of conflict is the literature on when states choose to enter negotiations. Zartman (2001) argues that negotiations occur when there is a relative stalemate between parties. Others argue negotiations occur after fighting reveals information about the relative capabilities of both parties and facilitates opportunities for a negotiated settlement (Powell 2004; Reiter 2009; Min Working). This paper builds on this work by examining how information and uncertainty explain changes in the probability of conflict.

The theory developed here is closest to Axelrod (1979). Axelrod develops a model to determine when an actor chooses to use a resources that can bolster their expected payoff. He finds that the actor's optimal time to use the resource is when the expected stakes meet or exceed some threshold. Our explanations differ on the expected costs of using this resource. Axelrod assumes the actor pays a cost for not using the resource; I assume the actor also pays a cost for launching a violent attack. Axelrod treats the threshold in terms of the stakes of a conflict; I treat the threshold in terms of the capabilities an armed group must accrue to fight an insurgency.

In my theory, I explore when an armed group chooses to start using one resource – political violence – given both its existing capabilities and potential gains from using violence. The armed group wants to maximize the probability that using violence helps them grow enough to surpass the threshold for civil war. If violence does not benefit the group enough to get over the threshold, then it pays an informational cost which makes it vulnerable to repression, mitigating the potential gains from violence.[2] An armed group's best response to this decision problem is to wait to use violence until it is as close to the threshold as it possibly can get.

[2]The sequence of events is thus memoryless.

3 Theory

I argue conflict escalation occurs when the state misses its window of opportunity to act. Emerging insurgent threats are often shadowy organizations, which creates uncertainty about their potential capabilities. This makes it difficult to determine precisely when to move against an armed group, increasing the risk of miscalculation.

Assembling a response against an armed group is costly so the state prefers to preempt a group that will become a larger threat and ignore a group that will stop using violence on its own. In order to sift through the noise, the state collects intelligence to assess when an an armed group will become a threat. As part of this, the state monitors an armed group's violent activities to learn about the group's current capabilities.[3] The number or pacing of attacks can sometimes reveal how close the armed group is to surpassing the threshold for civil war. It can help alert the state whether it needs to act immediately or not.

If the state can learn about the group based on current guerrilla activities, then it can condition the timing of its efforts accordingly. Low-level violence is increasingly useful for the state if it clarifies the armed group's potential capabilities.[4] Low-level violence is increasingly inefficient for an armed group if it reveals more information about its current capabilities than the additional strength it gains from an attack. It predicts that – once an armed group initiates violence – the probability of civil war escalation begins to decline.

HYPOTHESIS 1: Once violence begins, the probability of civil war escalation is decreasing over time.

Instead, this suggests civil wars should erupt shortly after an armed group initiates violence. When an armed group first forms, it can choose to commence violent operations immediately or lay low and attempt to organize discretely. The latter may seem like an ostensibly tricky way to amass strength, but waiting is advantageous. When a group refrains

[3] In the remainder of the paper, I refer to any low-level political violence as guerrilla violence.

[4] Relatedly, Thomas (2014) finds that rebel groups in Africa which engage in more violence are able to achieve better outcomes because violence clarifies the bargaining range between parties. On the perspective that political violence does not pay for an armed group, see Abrahms (2006) and Fortna (2015).

from violence, it can grow without revealing significant information about its strength to the state. This element of surprise gives it an initial advantage in escalating against the state. The probability of civil war increases as the group quietly grows and the state delays its response. Once the group transitions to using violence, the state acquires information about the group's capabilities and counteracts. At this point, the probability of civil war should begin to decline.

HYPOTHESIS 2: Before violence begins, the probability of civil war escalation is increasing over time.

If violence is inefficient, than an armed group has incentives to refrain from using violence as long as possible. However, an armed group may not always be able to pursue this strategy as long as it prefers. When an armed group forms, it is often weak and vulnerable; it prefers to organize behind-the-scenes in order to avoid destruction by state forces. Initial efforts may prove fruitful in acquiring popular support for a campaign. Yet, an armed group can only grow so large without attracting the attention of the state or using violence.

The need to amass the capabilities to credibly challenge the state creates competing incentives to start guerrilla operations. That is, even if an armed group would prefer to wait to use violence, attacks against police stations, warehouses, or military posts become necessary to mount a militarized challenge against the state. These incidents can force an armed group to rationally and prematurely disclose its presence before it is strong enough to fight an insurgency.

Consequently, the timing of civil war varies as a function of how long and how effectively the group grows during this latent stage of existence. When an armed group forms and refrain from using violence, the probability of civil war escalation is increasing over time. Once armed groups begin to use violence, the probability of civil war escalation decreases. While this predicts most civil wars will begin soon after formation, the nature of these escalations differ.

Repeated interactions between an armed group and the state can lead to de-escalation,

escalation, or no change in the intensity of conflict (Table 1). Escalation can arise rapidly if an unknown rebel group launches a sudden, high-intensity attack or the state cracks down suddenly against a group, provoking backlash. Conversely, civil war can arise slowly if the state ignores the slow, low-intensity accumulation of guerrilla violence. The likelihood of different types of catalysts comes from when the state and rebel group choose to act.

Table 1: **Types of Conflict Shifts.** Changes in the conflict intensity arise from when and how the state and armed group interact with each other over time. In the first period, the state and armed group can choose to use violence (act) or not (wait). In the second period, the armed group can choose to act or not. This creates eight potential outcomes corresponding to changes in the intensity of conflict.

	State	
Armed Group	Wait	Act
Wait, Wait	Neither side takes action; peace prevails. *No escalation.*	*Proactive.* State acts against latent armed group; efforts maintain the peace. *No escalation.*
Wait, Act	*Surprise.* Latent armed group quietly prepares then launches surprise attack, surpassing threshold. *High-intensity escalation.*	*Provocation.* State acts against latent armed group, but efforts backfire. This triggers militant violence that surpasses threshold. *High-intensity Escalation.*
Act, Wait	*Unclear.* Armed group stops using violence in next period despite no state action. *De-Escalation.*	*Destructive.* State acts against violent armed group; efforts deter militant violence in next period. *De-Escalation.*
Act, Act	*Procrastination.* Continuous militant violence slowly builds across periods. With no state action, it eventually surpasses threshold. *Low-Intensity Escalation.*	*Reactive.* State acts against violent armed group, but efforts do not deter militant violence in next period. *No escalation.*

First, civil wars can escalate suddenly when the state waits too long to act against an armed group patiently organizing behind the scenes. If an armed group refrains from the use of violence, then it can launch a sudden attack and overrun the state before the state even knows it has an opportunity to act.

For example, the Syrian Muslim Brotherhood went underground after a failed rebellion in 1964 and went underground, allegedly operating out of Iraq or Jordan for the next decade.

When the group re-emerged in 1976, it began launching high-scale attacks against military training schools and other security targets. It quickly escalated past the 25-battle death threshold. In another case, the Revolutionary United Front (RUF) was able to take advantage of the violence in Liberia to mask its own operational activities in the years before it escalated. After two years of guerrilla training in Liberia, the RUF launched a cross-border invasion into Sierra Leone, surprising and overrunning local security forces. The large-scale attack was enough to surpass the violence threshold and prompt a civil war. This implies that access to external bases or outside sponsors provide a unique opportunity for armed groups to effectively grow strong enough to challenge the state without using guerrilla violence. This resource grants them the privilege to challenge the state on their own timing.

HYPOTHESIS 3: Transnational groups are more likely to escalate through large-intensity militant violence.

Second, civil wars can escalate suddenly when the state miscalculates the group's capabilities and cracks down too hard against a latent group. The state may prefer to preempt an armed group when it first emerges in order to deter future challengers from launching violent campaigns (Walter 2009). Acting early and often against newly-formed armed groups can be beneficial even if an armed group has not yet presented any demonstrable threat or used violence. In some cases, these efforts can completely destroy the proto-campaign. In other cases, this can backfire and lead to enough violence to surpass the threshold. If the state acts too early, then it risks drawing attention to an armed group and inadvertently empowering it to grow. For example, Azerbaijan leader Ayan Mutalibov – with assistance from Soviet forces – launched "Operation Ring" against a small sect of Armenian militants in Nagorno-Karabkh in spring 1991. The joint military operation backfired causing more people to join the Armenian militias and leading to increased intense violence during the summer and fall of 1991.

State-initiated attacks are most likely to lead to civil war when the state acts against armed groups already endowed with certain capabilities. For example, the state might have reputational incentives to act against separatist groups, but this can recoil if a separatist group already has a territorial base, population, or local resources to effectively counteract the state's response. Similarly, armed groups that form with the support of foreign fighters, former rebels, or ex-military also have an organizational advantage in responding to the state's efforts. Their combat experience allows them to survive state repression efforts and counter-mobilize a response.

HYPOTHESIS 4: Separatist and veteran groups are more likely to escalate through large-intensity state violence.

Even though most civil war escalations are characterized by high-intensity catalysts, some conflicts emerge slowly. Escalation can occur if the state waits too long to act against a potential threat despite relatively good information about the armed group's capabilities. Even though the group's ex ante capabilities are not very large, its ability to systematically grow – year after year – in an unrestrained environment allows it to eventually transform into a full-blown insurgency.

There is no clear catalyst for the conflict and the escalation is slow. Escalation slowly occurs through the gradual accumulation of capabilities and intensification of militant violence. An armed group's behavior in the past year should be a strong predictor of escalation in the present state. These 'unclear' cases should also be most likely in states where the state is unable to act, preoccupied by other threats, or believes low levels of guerrilla violence are endurable.[5]

The state might wait too long because it lacks the bandwidth to deal with the group.

[5] Institutional constraints may also limit the state's ability to respond against an armed group when it wants. For example, bureaucratic procedures may induce delay in intelligence and operational planning against armed groups. This happened with parts of the Iraqi insurgency in 2003. As early as August 2003, the United States detected Moqtada Al-Sadr organizing in parts of Baghdad for an insurgency; the military issued a warrant for his arrest. An internal White House debate throughout the fall debated whether the United States should take advantage of this warrant in order to get ahead of the problem. However, delays arose as officials debated the arrest would hurt the nascent Mahdi Army or drive more support to the movement. Days passed and the administration waited. Some felt Sadr "had crossed the line" and needed to be arrested (Rumsfeld 2011, 537). Others believed "the political risk of doing so always seemed to outweigh the benefits" (Rice 2011, 273). Eventually, they waited sufficiently long enough for the Mahdi Army to launch a series of high-profile attacks in April 2004.

For example, when the Afghan Taliban formed as a low-level militia in 1994, it was able to take advantage of the larger Afghan Civil War. It operated relatively openly and was able to quickly and effectively acquire enough support from locals to escalate to civil war the following year. The state simply did not have the bandwidth nor the capacity to learn about and respond to the armed group. In other cases, the state has the capacity to act, but prefers to gamble on the risk of civil war because guerrilla violence is beneficial. For example, in the Philippines, Ferdinand Marcos leveraged – and even hyped up – guerrilla violence by the Communist Party of the Philippines in order to consolidate power around his rule. When the CPP threw a bomb onto the stage of a political rally in Manila, Marcos imposed martial law and other draconian measures despite common knowledge that the regime was exaggerating the threat of the Communist movement "investing it with a revolutionary aura that only attracted more supporters" (Kessler 1991, 40). This violence empowered Marcos with an opportunity to seize more power, consolidating his rule and suspending political opponents in the name of national security.

In the above cases, civil war emerged from shifts in the rebel's capabilities. The rebel group was able to grow strong enough through violent activities and other tactics in order to surpass the requirements for civil war. An alternative catalyst, which sometimes drives either the state or rebel's decision to attack, is an exogenous shock to the armed group's capabilities. These shocks can suddenly shift a group's capabilities and enable groups that have been operating for different amounts of time to get over threshold. A shock has the effect of enriching a group with the capabilities to fight an insurgency without imposing any of the informational costs associated with guerrilla attacks. Domestic political shocks often involve leadership turnovers, elections, or attempted coup d'etats. International political shocks involve neighboring instability, leadership turnover, or other types of conflicts. Shocks might be hard to anticipate so it is unclear when and whether group would prefer to refrain from guerrilla violence and gamble on the possibility of waiting for a favorable shock.

Among rebel groups, these catalysts suggest variation in the timing of escalation. Armed groups that more effectively mobilize during the latent stage have an advantage when they transition to using violence. External bases, access to diasporas, or outside sponsors make it easier for latent armed groups to organize more effectively. Similarly, shocks can create an impromptu window for rebellion, shortening the lead time to civil war.

HYPOTHESIS 5: The expected time to civil war is shortest for latent armed groups with external bases and/or latent armed groups at the time of shocks.

Collectively, this leads to a series of predictions both about *how* armed groups will escalate their campaigns and *when*. In the next section, I test these intuitions using Markov chains.

4 Data

I test these hypotheses using a series of existing armed conflict datasets. There is a surprisingly amount of overlap in the population of armed groups that appear in different conflict datasets. These commonalities provide an opportunity to link scattered incident reports to paint a broader picture of campaign violence from start to finish. In addition, I pair dyadic data on the characteristics of 964 armed group campaigns with new data on the precipitating event of 241 campaign escalations between 1970 and 2012. I research the onset of every case in order to determine why the conflict escalated and create a typology of catalysts.

The population of armed groups comes from the Armed Group Dataset, which records information about the group's formation, violent activity, goal, ideology, and other important group characteristics from 1970-2012 (Malone working). This dataset builds on the population first introduced in the Terrorist Organization Crosswalk project (Asal et al. 2015). Each group in the data satisfies three criteria: (1) a unique name identifier, (2) politicized opposition to the state, and (3) record of sustained violent activity. While the crosswalk initially includes over 2600 names, these three criteria lead to a population of 964 armed groups in 72 countries from 1970 to 2012 of which 241 escalate to civil war.

Incident-level data on armed group violence comes from the Global Terrorism Database

(GTD) and the Uppsala Armed Conflict Database (ACD). While media reports often discount political violence as the work of criminals or bandits, these armed groups often evolve into credible insurgent threats. As a result, I do not differentiate between armed groups that engage in lower levels of violence, sometimes apocryphally referred to as terrorists, and armed groups that engage in civil war levels of violence. Rather, this paper builds on growing work that suggests insurgent and terrorist organizations should not be studied in isolation from one another, but rather treated as synonymous actors whom all use violence in the pursuit of political ends (Fortna 2015; Asal et al. 2015). I treat the violence perpetrated by groups in one dataset as comparable to the violence recorded in the other dataset. This leads to a collection of violent attacks across 12,264 dyad-years which differ in their scale and intensity. It includes information from 38,743 GTD attacks, 241 episodes of low-level escalation, and 75 episodes of high-level escalation.

By itself, this incident data is sufficient to test when civil conflicts begin. In fact, a huge feature of Markov chain analysis its ability to estimate the probability of transitioning to another stage of conflict without inputting any information about the armed group itself. However, in order to see whether transition probabilities vary across different types of armed groups, I also code how these transitions to civil war occur.

The typology of conflict catalysts is developed by first identifying which armed groups escalate their campaigns to civil war and second when they do so. The ACD measures the year of escalation as the year the group first engages in enough violence with the state to meet or surpass the 25-battle death threshold. For each armed group campaign that eventually surpasses this threshold, I research how and why violence escalated in a given year. When there are multiple precipitating events such as a domestic shock which triggered a government crackdown, I code the catalyst as the most proximate event to the date of violence.

I use the narratives available from the Armed Group Dataset to identify the group's outcome and the state's response to the group when it initially formed. In many cases, these narratives include information about a particular rebel or government-initiated incident which pushed these campaigns beyond the threshold. For example, the Liberian Civil War started in 1989 when Charles Taylor launched a large cross-border invasion with the NPFL from his external base in Guinea-Bissau. The Lal Masjid siege in 2007 by Pakistani security forces backfired abruptly as the violence catalyzed the armed group Tehrik-e-Taliban Pakistan to launch retaliatory attacks against military outposts in Waziristan.

In 30% of cases, the precipitating event is a domestic or external shock such as a leadership turnover, neighboring conflict, or series of large-scale protests. In other cases, the precipitating event revolves around a change to an ongoing conflict. For example, the Contras in Nicaragua was an umbrella movement, which formally organized in 1982 after low-level, scattered violence among different armed groups against the Sandinista government. Other cases involves groups which join an ongoing conflict (e.g. Moro Islamic Liberation Front) or groups which choose to keep fighting even after several factions of their group have laid down arms (e.g. Real IRA). A summary of these catalysts is in Table 2.

Table 2: Typology of Civil War Catalysts.

Intensity	**Conflict**	**Precipitating Event**	**N**
Low-Intensity	New	Unclear (Procastination)	30
	Ongoing	Rebel Reorganization	13
	Both	Other	3
High-Intensity	New	Rebel-Initiated Attack (Surprise)	23
	New	State-Initiated Attack (Provocation)	43
	New	Domestic Shocks	48
	New	External Shocks	23
	Ongoing	Recurring - Join Ongoing Conflict	41
	Ongoing	Recurring - Reject Ceasefire or Peace Agreement	17
TOTAL			241

As a scope condition, I restrict my analysis to only the start of new conflicts. The causes of conflict recurrence may revolve around a different set of issues or processes than the ones theorized above. This leads to a set of 93 armed campaigns that eventually surpass the low-level civil war threshold of 25-battle deaths; 32 eclipse the higher-level 1000-battle death threshold. While most armed groups only operate in one country, some are transnational and express politicized opposition against multiple governments. Thus, the unit of analysis is the dyad-year.

While these datasets are imperfect and often underestimate the amount of political violence perpetrated by a particular group, they also provide a means to gain large cross-national leverage on variation in armed group campaigns. However, as an extra measure of precaution, I also randomly sample and reserve 35% of the observations for a test set. I compare the predictive accuracy of these results with the test set findings in order to quantify the external validity of the results.

5 Markov Chain Analysis of Conflict Escalation

A Markov chain describes a set of states that armed groups move between with different transition probabilities p_{ij}. The sequence of events for each group describe how the armed group has acted in the past. The transition probabilities describe the probability of moving between various states. In this section, I test the first two hypotheses using different Markov chains.

5.1 Why Markov?

A Markov chain analysis of armed groups is a natural extension of existing research on conflict escalation. The Markov model is a stochastic process – named after mathematician Andrey Markov – which describes how system of states change over time (Amemiya 1985; Ross 1996). The model relies on the property of memoryless, or the assumption that the probability of future events only depends on the current state – not the sequence of prior events. As a result, it is a useful tool to make probabilistic statements about future conflict levels given information about current conflict.

The model is also theoretically useful for testing theories of conflict escalation. First, it captures sequential decision-making between actors and has been employed, which makes it ideal for describing changes in dyadic relationships over time. Conflicts are not homogenous events, but change in scale and intensity over time. Markov chains can reveal when and how long distinct stages of conflict last (Eberwein 1981). Second, it can model random walks, which makes it ideal to test informational theories of conflict. In the economics literature, for example, Markov models have been used to test the efficient market hypothesis by modeling the time to reach certain stock prices based on both a gradual drift – or growth rate – and a volatile shock component (Black and Scholes 1974; Almgren and Chris 2001).

Quantitative conflict studies have sporadically employed Markov chains in the past. There is precedent for the method, whether it involves describing diplomatic tensions between China and India (Duncan and Siverson 1975) or modeling alliance durability (Midlarsky 1983). In each of these cases, scholars explored the sequence of events diplomatic cables or patterns of violence predict the probability of an increase in tensions in the current period. Some simply compare the transition matrices of different dyad subsets to note whether the transition probabilities are significantly different from one another. Others employ a generalized additive model with a logit link specification where the dependent variable is the transition probability of moving from one stage to another (Jackman 2000).

Other work uses Markov chains to explicitly model changes in the intensity of conflict (Kiefer 1988; Geller 1993; Toft and Zhukov 2012). Kiefer (1988) compares the correlates of moving from a state of peace to a state of war with the reverse case, finding that arms races, defense spending, and other conflicts positively affect the probability of transition. Geller (1993) analyzes trends in militarized interstate disputes between rival dyads in order to estimate the long-run probabilities of how often a dispute will occur. He finds significant transition probability differences between dyads with relative power parity and asymmetric capabilities. Toft and Zhukov (2012) explore how different counterinsurgency strategies in the previous year affect the risk of insurgent violence. They find that lethal punishment strategies are associated with a transition to a more violent stage while denial strategies are associated with a transition to a more peaceful stage.

These papers demonstrate the large potential for modeling conflict as part of a Markov process. However, they do not estimate the timing of conflict, in part, because they do not include absorbing states in their analysis. States can cycle between nonviolent and violent stages of behavior indefinitely. Within states, this assumption does not always hold. Armed groups can 'die' or disappear – and do so with great frequency. Understanding when civil wars first occur also introduces a second absorbing or transient state into the chain, which these analysis do not include.

A key assumption under-riding this approach is the Markov property, otherwise known as the memoryless property. That is, the armed group's future state of conflict only depends on its current state. This can be summarized by the equation:

$$P\{S_t = j \mid S_{t-1} = i\} = p_{ij}$$

The theory laid out here supports this assumption. If the state conditions its efforts in each current period, then the armed group's future capabilities should only depend on what its current capabilities are. The state is updating and trying to act on whatever information it currently has. The first two hypotheses can be conceptualized in terms of this conditional probability relationship. The first hypothesis examines the probability of civil war in the next period conditional on guerrilla violence today. The second hypotheses examines the probability of civil war in the next period conditional on nonviolent activity today.

5.2 States of Insurgency

In the simplest scenario, assume four possible states (S_i) of existence for an armed group: latent, guerrilla, civil war, and death: $S_i = \{S_{\text{Latent}}, S_{\text{Guerrilla}}, S_{\text{War}}, S_{\text{Death}}\}$.[6] These stages of conflict map onto the phases of civil conflict identified by guerrilla theorists decades ago (e.g. Mao 1937, McCuen 1966).

I operationalize whether an armed group is operating in each state using available incident

[6] I collapse low-intensity and high-intensity civil conflicts into one category for ease of interpretability.

data. If a group has no recorded militant activity, it is recorded as nonviolent or in the latent stage. If a group perpetrates at least one act of "violence by a non-state actor to attain a political, economic, religious, or social goal" and less than 25-battle deaths, then its campaign is described as operating in the guerrilla stage. If a campaign leads to at least 25-battle deaths in a given year, it is recorded as being in a state of low-intensity conflict with the state. Finally, if a campaign leads to at least 1000-battle deaths in a given year, it meets the state of high-intensity conflict. After a group is inactive for at least three years, I treat it as 'dead' or unlikely to ever mount an attack again.[7]

States are recurring if an armed group can leave and return to them with any positive probability. States are absorbing if – once entered – cannot be left; they are, in a sense, "game-ending." In this scenario, latent and guerrilla stages are recurring because an armed group can cycle between them from year to year depending on whether it launches any violent attack. Civil war and death are absorbing because an armed group cannot be 'resurrected' or return to engaging in lower levels of guerrilla violence.

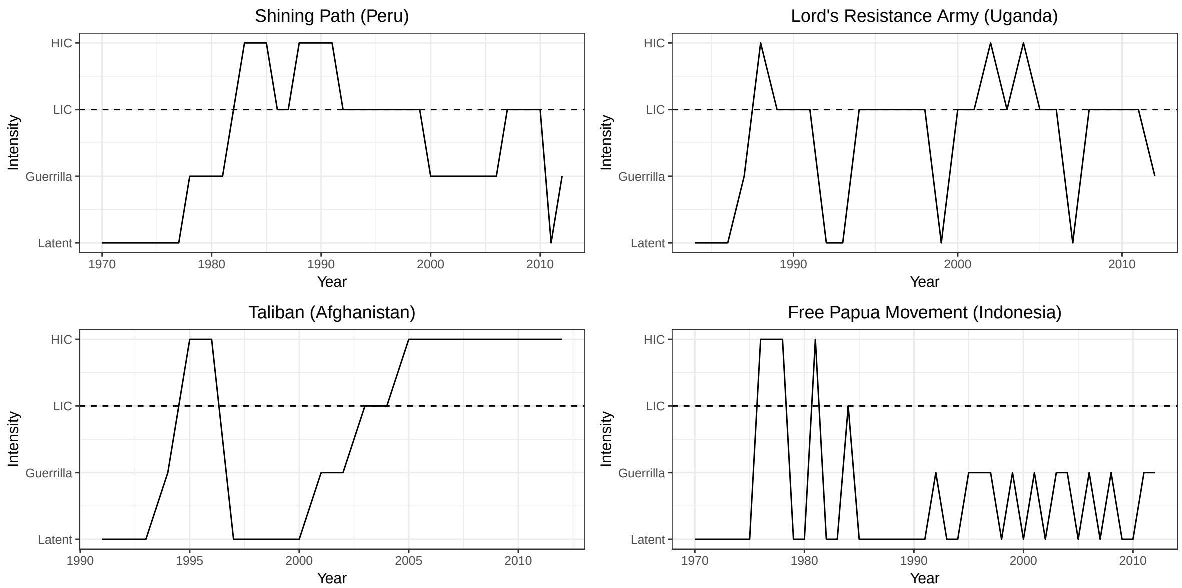
In order to simplify the analysis, I assume that there are no hidden states – or other states beyond those laid out here – which an armed group can transition to. I also treat the process as a discrete-time Markov Chain meaning that an armed group evolves through distinct time steps (years) and that changes cannot happen at any time t. This is consistent with the conflict literature using the dyad-year as the principal unit of analysis. These assumptions also allow for a slightly simpler analysis of the problem although future work could consider more complicated variations.

There is substantial variation in when different rebel campaigns intensify in violence (Figure 1). For example, the Taliban and Shining Path both gradually increased their activities until they reached the state of high-intensity conflict. The Lord's Resistance Army in Uganda, by contrast, cycled between states of little to no conflict and large-scale conflict

[7] In the larger population of armed groups, there is only one recorded armed group that returns to using violence after a six year window of inactivity. The vast majority are only inactive for two to three years at a time.

– seemingly spending very little time in the guerrilla stage of operations. The separatist Free Papua Movement in Papua New Guinea escalated shortly after formation in 1965 and then again several times in the 1970s as it clashed with Indonesian security forces. These cases demonstrate that armed groups frequently move between different states of violence during their campaigns.

Figure 1: **Conflict Sequence of Four Rebel Campaigns.** The evolution and escalation of four different campaigns is shown here. During the years shown, rebel campaigns cycle between several states of conflict. The dashed line is the conventional cutoff for civil war.



We can describe these general sequences of events as a Markov chain. In the simplest case, we might assume that after armed groups form, they can cycle between nonviolent and lower-level violent activity until one of two events happen. First, an armed group could simply stop using political violence and 'die' or cease to exist. Second, an armed group could escalate its campaign to civil war. In this scenario, civil war and death are both absorbing states. Once an armed group enters this stage, it cannot return to earlier forms of violence. In contrast, latent and guerrilla stages are recurrent. Armed groups can return to these stages over and over before they eventually absorbed.

Groups move around these states with different transition probabilities p_{ij}:

State_{t-1} / State_t	Latent	Guerrilla	Civil War	Death
Latent	3898/5138 (0.76)	862/5138 (0.17)	70/5138 (0.01)	304/5138 (0.06)
Guerrilla	769/1148 (0.67)	336/1148 (0.29)	43/1148 (0.04)	0/1148 (0.00)
Civil War	0/0 (0.00)	0/0 (0.00)	113/113 (1.00)	0/0 (0.00)
Death	0/0 (0.00)	0/0 (0.00)	0/0 (0.00)	304/304 (0.00)

This can be equivalently written as the transition probability matrix:

$$P = \begin{pmatrix} 0.76 & 0.17 & 0.01 & 0.06 \\ 0.67 & 0.29 & 0.04 & 0.00 \\ 0.00 & 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

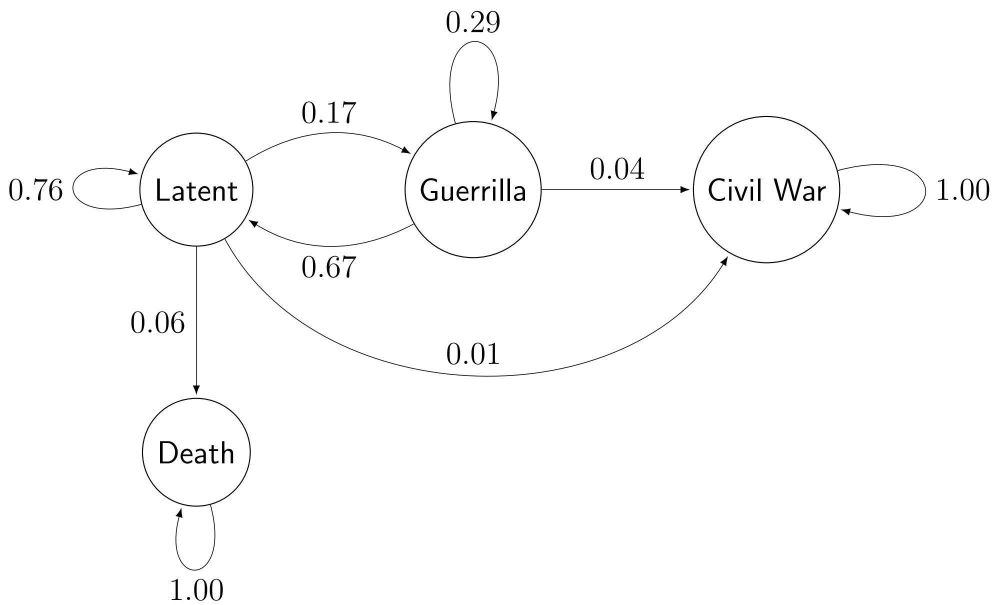
The probabilities capture the probability of moving from a state in period $t-1$ to another state in time t. For example, the transition probability of an armed group moving from the latent to guerrilla stage is 0.17. The transition probability of moving from the guerrilla stage to civil war is 0.04. Concurrently, the transition probability of moving directly from the latent to civil war stage is only 0.01. The last column reveals that approximately 6% of dyads transition to the death stage each year.

A transition diagram helps visualize the probabilities that an armed group travels between these different states (Figure 2). The results demonstrate that approximately one third of armed groups in the sample data – 70 armed groups – escalate from the latent to the civil war stage in one step. In all other cases, 113 armed groups cycle through the guerrilla stage at least once before escalating.

5.3 Hypothesis 1 and 2

The first two hypotheses predict that the probability of conflict is a function of when an armed group initially forms and begins to use violence. If the theory is correct, then armed

Figure 2: **Two Absorbing State Transition Diagram of Armed Group Campaigns.** The diagram illustrates the transition probabilities of moving from one state to another state in one step. An armed group initially starts in either the latent or guerrilla stage. An armed group eventually ends up in either the death or civil war stage. Armed groups only die after three years of constant inactivity.



groups which refrain from the use of violence should initially have a higher probability of escalation to civil war because the state knows little about their capabilities. Once an armed group initiates the use of guerrilla violence, its organizational advantage dissipates as the state learns more about the group and the If an alternate explanation prevails – that guerrilla activities allow an armed group to grow – then the probability of civil war escalation conditional on current militant activity should be increasing over time.

I test the first two hypotheses by examining the predicted probability of civil conflict for an armed group. The Chapman-Kolmogorov equation calculates the n-step transition probabilities given information about an armed group's initial state and transition matrix.

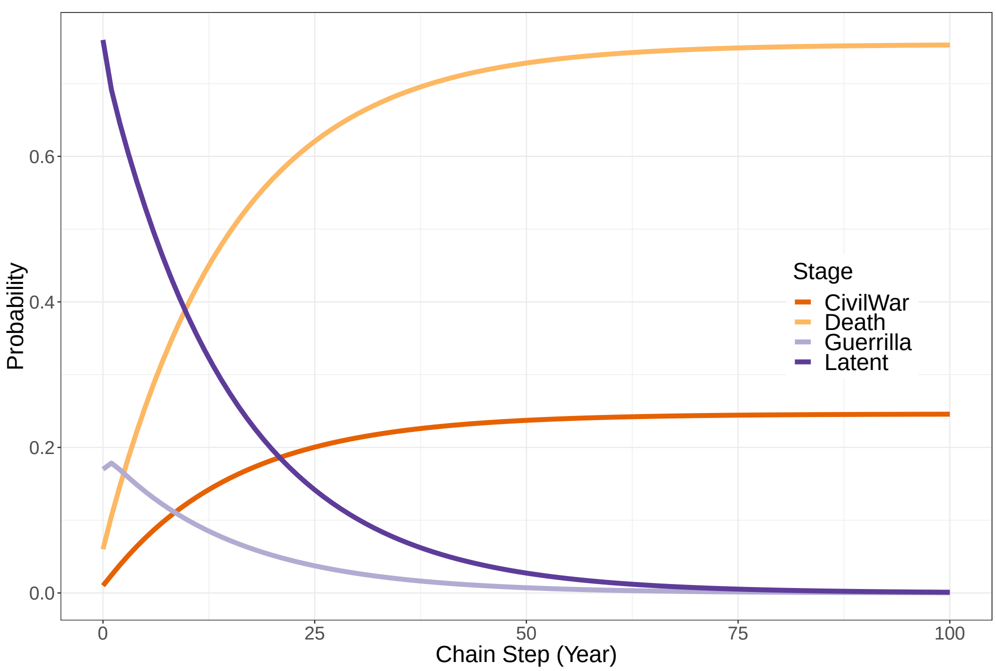
$$p_{ij}^{n} = \sum_{k \in S} p_{ik}^{n} p_{kj}^{m}$$

The equation states that the transition probability of moving from state i to state j in period $n+m$ is the product of the transition probabilities over time as an armed group moves from its initial state to other states. The probability of transitioning to state j in period $n+m$ is the product of all the different ways an armed group can travel to state j over m periods when it starts in state i in period n.

On the most basic level, this equation provides insight into what proportion of the sample will eventually escalate to civil war. If we assume that that an armed group can take m periods to travel from state i to state j, then this becomes equivalent to taking the limit of the transition matrix as m grows arbitrarily large. The result is that in the long run the probability of reaching any future state is the same regardless of the armed group's initial state. This simplifies to a 'steady state' vector of probabilities.

The steady state relates the long-run proportion of dyads that will end up in each of the two absorbing states. In other words, it explains what proportion of armed groups eventually escalate to civil war(Figure 3).

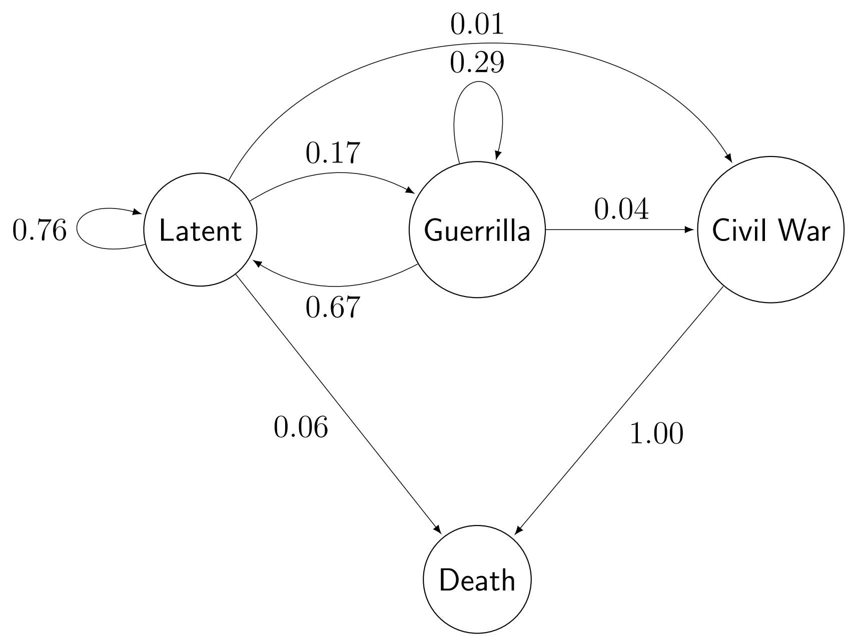
Figure 3: **Steady State Probability of Armed Group Transitions.** This figure plots the approximate vector of steady state probabilities after 100 periods. The results show that the long-run probabilities of transitioning to either 'death' or 'civil war' regardless of the starting state.



Since there are two absorbing states, the model here suggests that in a long enough timeline, 28.8% of dyads will escalate to civil war and 71% will not. For comparison, the out-of-sample test set of observations predicts that 25.5% of dyads will escalate to civil war and 74.5% will not. This reveals substantiation variation across dyads in which armed group campaigns escalation. Understanding which dyads are most likely to transition to civil war, but others do not is examined elsewhere (Malone Working).

In order to estimate how the probability of escalation to civil war changes over time, I make a slight tweak to the Markov chain. In the period after an armed group escalates to civil war, I now assume it transitions to the 'death' stage with probability 1.00. That is, once an armed group escalates to civil war, it cannot do so again.[8] The new transition diagram for this model is presented in Figure 4.

Figure 4: **One-Absorbing State Transition Diagram of Armed Group Campaigns.** The diagram illustrates the transition probabilities of moving from one state to another state in one step. An armed group initially starts in either the latent or guerrilla stage. Armed groups enter the 'death' stage after three years of inactivity or after they escalate to civil war for the first time.



[8]In the appendix, I examine two other Markov Chains. The first is a simple chain that restricts the ability to travel between states. I find statistically significant differences in the probability of war in period t conditional on latent activities in period $t-1$. The second is a dynamic chain that treats civil war as a recurrent, rather than absorbing state. The results show that the probability of escalating to civil war peaks slightly later because many armed groups which escalate to civil war once tend to do it again later. This is consistent with existing scholarship on conflicts traps and patterns of civil war recurrence.

I apply the Chapman-Kolmogorov equation to this Markov Chain. The theory predicts that whether an armed group begins to use violence immediately or not affects the transition probability of civil war. In order to draw comparisons, I analyze the Markov Chain two different ways depending on whether the group's initial state was latent or guerrilla.

Figure 5: **Predicted Transition Probabilities of Armed Group Campaigns.** This figure illustrates the predicted probability of transitioning to any stage of an armed group's campaign after a certain number of steps. The probabilities are differentiated by whether an armed group's initial state is latent (nonviolent) or guerrilla (violent).

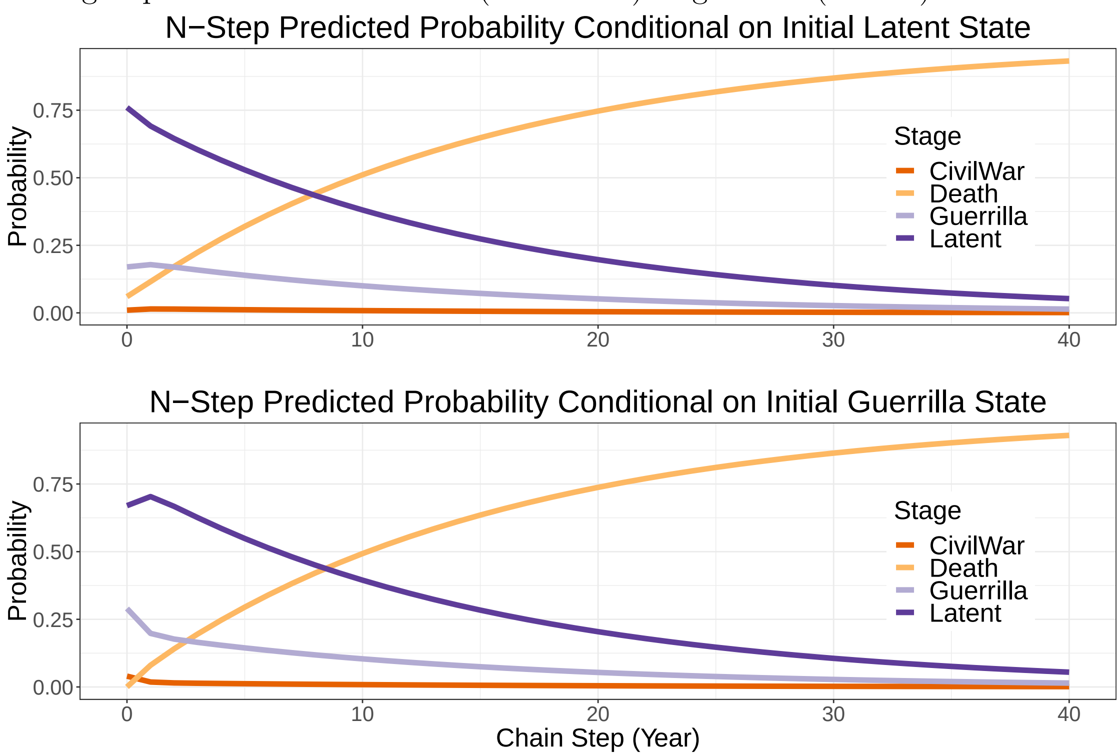
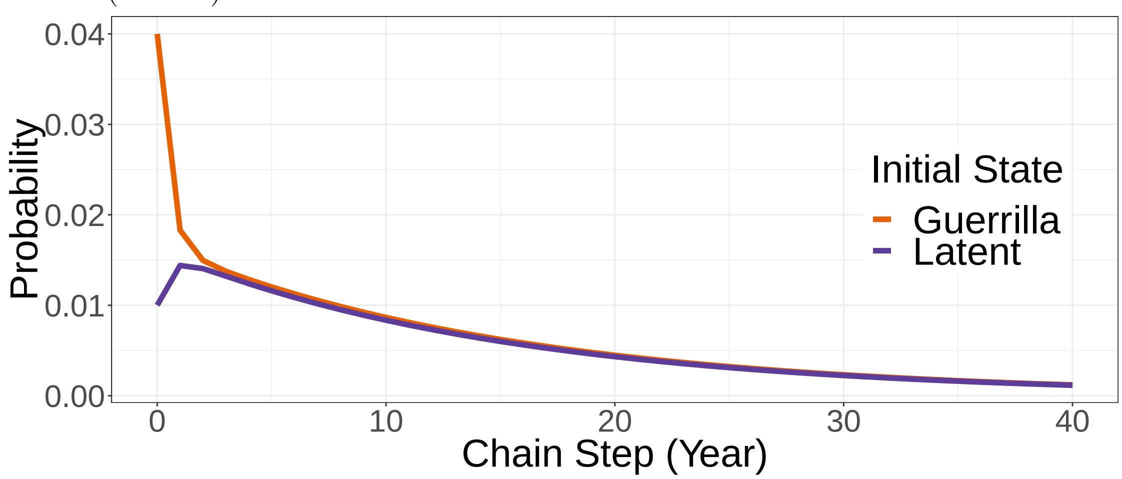


Figure 5 shows the estimated probability of transitioning to any of the four states. The probabilities provide the estimated likelihood of being in each state after a certain number of years depending on the armed group's initial state. In either case, the probability of transitioning to civil war is small, reflecting its relative rare occurrence.

In the first case, I treat the armed group's initial state as nonviolent. This captures the case when an armed group forms and is initially nonviolent. After just two years, the probability of being "dead" supersedes both the probability of civil war and guerrilla violence. After ten years, over 50% of dyads are dead. In the second case, an armed group forms and immediately commences guerrilla operations. The predicted probability of transitioning to civil war in the first year exceeds the probability of transitioning to death. The armed group is likely to avoid immediate destruction by the state.

Figure 6 shows the estimated probability of reaching the state of civil war over approximately 40 years. The two lines are differentiated by whether an armed group's initial state is violent ("guerrilla") or nonviolent ("latent"). The results provide evidence consistent with the first two hypotheses.

Figure 6: **Armed Group Transition Probabilities for Civil War.** This figure illustrates the predicted probability of transitioning to civil war after a certain number of steps. The probabilities are differentiated by whether an armed group's initial state is latent (nonviolent) or guerrilla (violent).



The transition probability of civil war varies by an armed group's initial state. Guerrilla violence initially gives the armed group an initial advantage in growing faster than the state can degrade it. The highest probability of civil war escalation is in the first year an armed group uses violence. However, over time, the probability of civil war declines. Once an armed group reveals its presence to the state, the state is able to weaken the group faster than it can acquire the strength to fight a civil war.

Conversely, if the group's initial state is nonviolent, the probability of escalating is initially low. The group is weak and avoids engaging with state forces which could destroy it. In the next year, the probability of war increases. If the group uses its latency period effectively, then it can grow strong enough to challenge the state while avoiding detection. The probability of civil war peaks as armed group campaigns transition from latent to guerrilla operations. The highest probability of escalation is in the second and third year of operations, respectively. After 3 years, the probability of escalating to civil war begins to decline again.

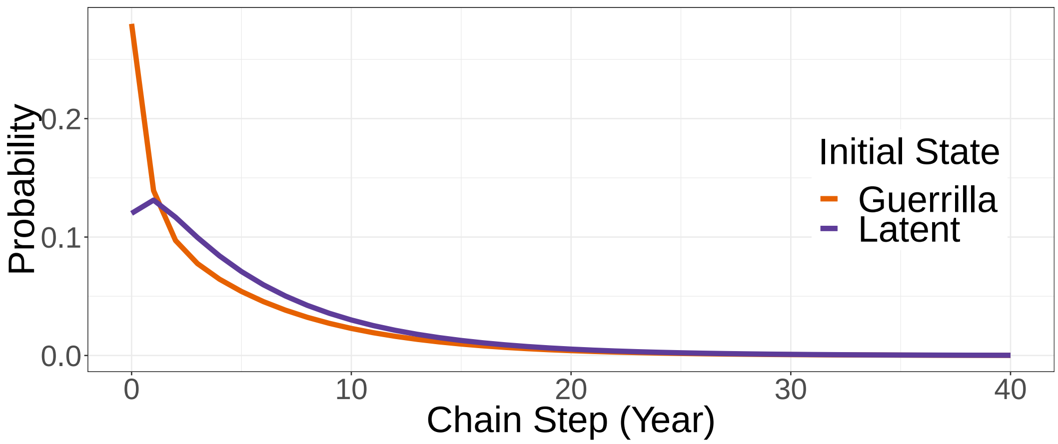
The transition probabilities of civil war begin to converge after four years. That is, the transition probability of civil war quickly becomes independent of an armed group's initial state.

There are two potential sources of measurement error here. The first involves the specified measure of guerrilla violence. The guerrilla state is based on GTD incidents, but GTD has a series of known methodological issues (LaFree and Dugan 2007; LaFree et al. 2014). Principally, the dataset relies on a series of different coding rules across time, which lead to measurement error both in the quality and quantity of information it records about each attack in the dataset. There is a serious concern that the data is not systematically coded across countries and under-counts the amount of political violence in the world. However, this would bias against finding a spike in the transition probability of civil war. If the dataset recorded no violence, but the state had information about the group, then the state could act early and suppress the group. The expected result would be a constant decreasing probability of civil war instead of the nonmonotonic result we observe.

A second concern is that these results mask key differences between armed groups that escalate and those that do not. For example, if proto-rebel groups are ex ante stronger than other types of armed groups, then their transition probabilities might differ, affecting the timing of civil war. Proto-rebel groups could benefit more from guerrilla violence or be able to offset the informational costs of acting, negating the likelihood they escalate early.

As a robustness check, I restrict the sample to dyads that only include eventual rebel groups (Figure 7). The result looks similar. The probability of civil war initially spikes and then declines the longer an armed group operates. There are two key differences. First, the probability of civil war declines at a slower rate for armed groups that initially start off nonviolent. Second, the size of the transition probability is larger: a larger proportion of rebel groups escalate early.

Figure 7: **Rebel Group Transition Probabilities for Civil War.** This figure shows the predicted probability of rebel groups transitioning to civil war after a certain number of steps. The probabilities are differentiated by whether proto-rebel initial states are latent (nonviolent) or guerrilla (violent).

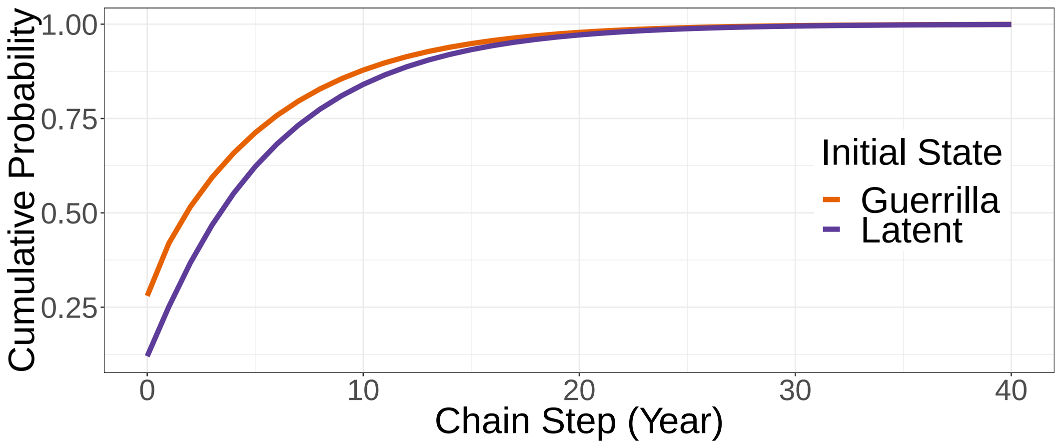


Collectively, this narrows down the window of time when armed campaigns are most likely to escalate to civil war. A cumulative distribution plot highlights this further (Figure 8). If rebel groups form and immediately begin using violence, then approximately 51.6% will escalate in the first three years. Concurrently, if rebel groups form, but waits to use violence, then approximately 55.2% will escalate in the first four years.

I assess the accuracy of this prediction by comparing it to the test set. On average, proto-rebel groups operate for approximately 3 years before escalating their campaign to civil war. In the test set, approximately 25% of rebel groups escalate less than a year after formation while another 25% take at least 4 years from the time they initiate violence. Thus, these results are close, but just slightly overestimate when civil wars occur.

These results provide collective evidence that armed groups tend to escalate their campaigns to civil war early. However, it is not just enough to know when armed groups are likely to escalate their campaign. It is also important to know what catalysts prompt these armed groups to surpass the threshold.

Figure 8: **Cumulative Distribution Plot of Civil War Timing.** This cumulative distribution plot shows the probability of rebel groups that are expected to escalate to civil war after operating for a certain number of years.



6 Multinomial Analysis

In order to test the second set of hypotheses, I look at which armed groups are more likely to escalate according to different catalysts. The analysis is simple. The primary dependent variable is the type of catalyst that prompts an armed group to surpass the 25-battle death threshold for the first time. I differentiate between state-initiated attacks, rebel-initiated attacks, shocks, and unknown or unclear catalysts. It is coded based on the year the campaign escalates to civil war and zero otherwise. The omitted category is any dyad-year where no civil conflict erupts.

For the independent variable, I include country-level and group-level characteristics that explain which catalyst leads to civil war. The 'unclear' mechanism predicts that previous guerrilla violence should be predictive of an unclear catalyst for escalation. The 'preparation' mechanism implies that less visible groups should be more likely to escalate by launching high-intensity attacks when the state waits too long to act. I thus include measures of whether an armed group has an external base and whether the government is authoritarian and has incentives to delay acting against a group. Finally, the 'provocation' mechanism predicts that groups with separatist or veteran bases are susceptible to high-intensity government crackdowns due to their ex ante capabilities. I thus include group-level information on whether they had separatist aims or initial members with veteran or combat experience.

The model is a multinomial logit with country and year fixed effects. The reference category are dyad-years in which there is no escalation to civil war. I cluster the standard errors at the country level in order to control for heterogeneity across countries. Country fixed effects adjust for country-invariant factors that increase the susceptibility of certain countries to different civil war catalysts. Year fixed effects adjust for common shocks to the international system.

Table 3: **Multinomial Analysis of Civil War Catalysts.** The results of a multinomial logistic analysis. The dependent variable is how an armed group campaign escalates to civil war in a given year. The reference category is dyad-years with no civil war escalation.

Catalyst	(1) Unclear	(2) Rebel Attack	(3) State Attack	(4) Shock
GUERRILLA$_{t-1}$	2.05^{***}	6.65^{***}	1.31^{***}	1.12^{***}
	(0.53)	(2.40)	(0.46)	(0.41)
SEPARATIST	0.26	-5.53^{**}	1.27^{**}	1.07^{**}
	(0.78)	(2.53)	(0.57)	(0.46)
VETERANS	0.36	-5.20^{**}	0.85^{*}	0.06
	(0.58)	(2.22)	(0.45)	(0.44)
EXTERNALBASE	0.02	4.15^{**}	-0.75	-0.72
	(0.59)	(1.82)	(0.55)	(0.50)
AUTOCRACY$_{t-1}$	-1.31	2.67	-1.36	-0.58
	(0.99)	(2.57)	(0.88)	(0.50)
Country FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
AIC	1650.77	1650.77	1650.77	1650.77
BIC	4795.11	4795.11	4795.11	4795.11
Log Likelihood	-357.39	-357.39	-357.39	-357.39
Deviance	714.77	714.77	714.77	714.77
Num. obs.	6116	6116	6116	6116

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ SE clustered by country.

The results in Table 3 demonstrate differences in how armed groups escalate to civil war. The coefficients reveal the log-odds of escalating to civil war through a particular mechanism. Exponentiating these coefficients provides the relative risk of escalation. Model 1 examines the correlates of escalation through unknown or unclear reasons. As expected, there are very few systematic trends in the data. The only correlate which explains whether an armed group will escalate is whether it engaged in guerrilla violence in the previous year. However, this factor is rather uninformative as it is also a statistically significant and substantive factor for all other escalation processes. This is consistent with the prediction that whether an armed group escalates its campaign through the gradual accumulation of violence does so for relatively unclear reasons.

Model 2 details the correlates of rebel-initiated attacks. The coefficient on separatist is negative and statistically significant. Center-seeking armed groups are more likely to escalate their campaigns through a high-intensity concentrated attack than separatist groups. The coefficient on external base is also positive and statistically significant. Armed groups with access to an external base are able to organize an insurgency without provoking the attention of the state. This gives them an advantage in escalating their campaigns immediately from

In contrast, Model 3 finds a set of different factors explain state-initiated attacks. Separatist groups typically escalate their campaigns to civil war because the state crackdowns on them. In some cases, a crackdown on a concentrated ethnic group allows a large number of groups to escalate as the 1990 crackdown by Jagmohan Malhotra did in the Kashmir. In other cases, a crackdown is isolated, only enabling one armed group to escalate such as the Kosovo Liberation Army in 1998. Similarly, armed groups with veteran bases are more likely to escalate due to a government crackdown. This may occur because the government employs a large amount of force, but the group has the necessary capabilities available to survive these efforts and reorganize. In other cases, a veteran group might be able to strategically avoid the brunt of the government's response at the expense of noncombatants in their area of operations experiencing violence.

Model 4 attempts to identify which armed groups are likely to escalate according to shocks. Separatist groups are slightly more likely to escalate in response to shocks than other armed groups. If separatist groups are ex ante strong, then a shock to their relative capabilities could be sufficient to push them over the threshold. The same might be true for groups with veteran bases; the coefficient is positive, but not significant.

In the Appendix, I explore the robustness of these results further. I remove year fixed effects since the effect of any local shock could be mediated by common global shocks. I also add additional covariates measuring GDP/capita and the number of violent armed groups in a country. Omitting these variables could bias results if weak countries or countries fighting a larger number of challenges discounted a threat more often. The results do not substantially change across various specifications.

7 Expected Time to Civil War

In the final section, I examine the expected time to civil war for different types of rebel groups. The results above demonstrate that armed groups with external bases are more likely to escalate to civil war through a rebel-initiated attack. Separatist and veteran groups are more likely to escalate to civil war through a state-initiated attack.

If the theory is correct, rebel groups that can effectively amass capabilities and avoid detection are likely to escalate more quickly once they start using violence. Rebel groups that rely on the use of sustained violence to achieve their aims and grow stronger will be much slower to escalate.

The theory also implies differences in how well these estimated windows of escalation generalize. If shocks emerge in unpredictable intervals, then the age of a rebel group at the time of its escalation should be a poor predictor of escalation. In contrast, armed groups that escalate due to endogenous shifts in their capabilities over times should be more easy to predict.

I disaggregate the transition diagrams for different rebel groups depending on how they escalated to civil war. In the training set, 21 rebel groups have unclear escalations, 15 escalate via high-intensity rebel attacks, 31 escalate via high-intensity state attacks, and 48 escalate due to shocks. The transition matrices for each of these four categories are:

Training Set Transition Matrices

$$P_{\text{Unclear}} = \begin{pmatrix} 0.79 & 0.16 & 0.05 & 0.00 \\ 0.42 & 0.40 & 0.18 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix} \qquad P_{\text{StateAttack}} = \begin{pmatrix} 0.73 & 0.17 & 0.10 & 0.00 \\ 0.44 & 0.19 & 0.36 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

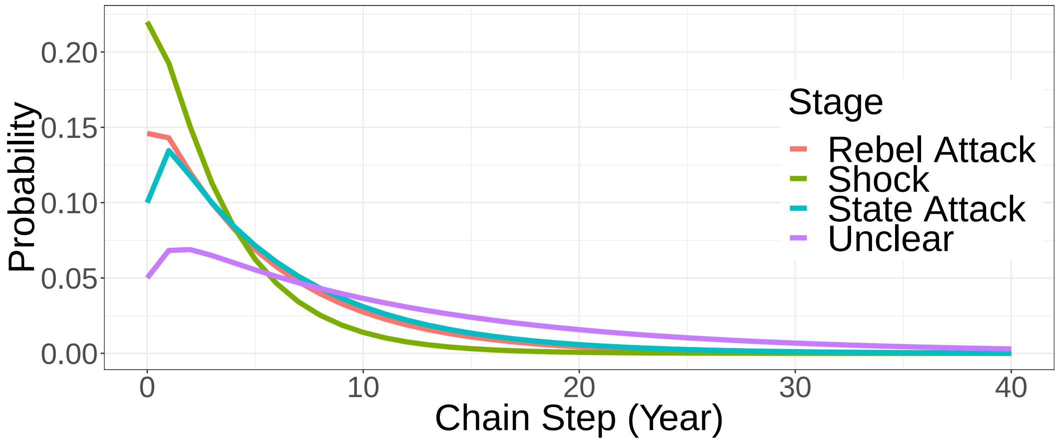
$$P_{\text{RebelAttack}} = \begin{pmatrix} 0.76 & 0.09 & 0.15 & 0.00 \\ 0.56 & 0.11 & 0.33 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix} \qquad P_{\text{Shock}} = \begin{pmatrix} 0.59 & 0.19 & 0.22 & 0.00 \\ 0.27 & 0.40 & 0.33 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

The transition matrices provide suggestive evidence that the probability of civil war in period t conditional on current nonviolent activities is much higher for rebel groups that escalate via shocks or rebel-initiated attacks. However, a multinomial confidence interval result at the 95% level shows slight overlap in the probability across different rebel groups.

I estimate the n-step transition probability for each matrix and plot the results in Figure 9. The results shows some differences in the expected timing of civil war across rebel groups. Namely, the expected probability of rapid escalation is highest after a shock. Rebel-initiated and state-initiated attacks both peak shortly after formation before declining at similar rates. The long-run probability of transitioning to civil war for an unclear reason is larger than the probability of transitioning to civil war through some other mechanism.

This leads to suggestive evidence that shocks cause rebel groups to escalate the quickest while unclear catalysts take longer. In order to estimate the expected time to civil war, I look at the expected steps of each stage before reaching the absorbing state of death.

Figure 9: Predicted Transition Probability of Different Rebel Group Types.



Another way to write the transition matrix P is

$$P = \begin{pmatrix} Q & R \\ \mathbf{0} & I \end{pmatrix}$$

Q is the transition matrix for transitioning between different transient states. R is the transition matrix for transitioning to the absorbing state (civil war). We can calculate the expected passage time by solving:

$$N = (I - Q)^{-1}$$

The matrix N is the fundamental matrix for P. Each entry n_{ij} gives the expected number of times the process is in the transient states.

$$N = \begin{pmatrix} n_{ii} & n_{ij} \\ n_{ji} & n_{jj} \end{pmatrix}$$

The variance of each entry is:

$$Var(N) = (2N - I)\mathbf{t} - \mathbf{t}_{\text{sq}}$$

From these two equations, I calculate the expected number of years that a rebel group is in either the latent or guerrilla stages of operations conditional on its initial starting state. The sum of these two expectations is the rebel group's expected time to civil war or the most likely 'window of opportunity' for the state to act (Table 4 and 5).

Table 4: **Expected Time to Civil War - Initial Nonviolence.** This table describes the the expected time it takes a rebel group to escalate to civil war after it forms. The initial state is latent. 95% confidence intervals for each stage are in parentheses below.

Catalyst	**E(Latency)**	**E(Guerrilla)**	**E(Time to Civil War)**	**Classification Accuracy**
Unclear	10.20 (9.02, 11.39)	2.72 (1.65, 3.79)	12.93 (6.77, 19.08)	0.33
Rebel-Attack	5.47 (4.53, 6.42)	0.60 (0.01, 1.20)	6.08 (2.64, 9.54)	0.38
State-Attack	5.67 (4.99, 6.35)	1.20 (0.58, 1.81)	6.87 (4.04, 9.69)	0.42
Shock	3.08 (2.87, 3.30)	0.98 (0.48, 1.47)	4.06 (2.47, 5.64)	0.08
All	5.35 (5.17, 5.53)	1.28 (1.10, 1.46)	6.63 (5.07, 8.19)	0.15

The results show that the expected time to civil war is shortest for rebel groups which escalate through shocks. The expected time to civil war is longest for rebel groups which escalate through unclear mechanisms. I use the 95% confidence intervals for expected time to civil war to see how well these estimates predict an out-of-sample set of observations. The results show that the classification accuracy varies across groups.

The expected time to civil war based on a rebel- or state-initiated attack accurately predicts between 38-42% of civil war timing in the test set. Newer proto-rebel groups are more likely to escalate their campaigns to civil war via a high-intensity attack. Older proto-rebel groups are more likely to escalate their campaigns to civil war via an unclear mechanism. Most of this variation is driven by differences in how long a proto-rebel group conducts guerrilla attacks. Rebel groups most likely to escalate via a high-intensity attack – mainly those with an external base – have the shortest guerrilla period.

The expected time to civil war via a shock only predicts 8% of civil war timing in the test set. This follows that the timing of shocks is unpredictable; the current age of an armed group's campaign cannot accurately predict whether it will escalate due to a shock.

Table 5: **Expected Time to Civil War - Initial Violence.** This table describes the the expected time it takes a rebel group to escalate to civil war after it forms. The initial state is guerrilla. 95% confidence intervals for each stage are in parentheses below.

Catalyst	**E(Latency)**	**E(Guerrilla)**	**E(Time to Civil War)**	**Classification Accuracy**
Unclear	7.14 (5.93, 8.36)	3.57	10.71 (4.70, 16.73)	0.44
Rebel-Attack	3.42 (3.00, 3.85)	1.50	4.93 (1.63, 8.22)	0.13
State-Attack	3.12 (2.65, 3.60)	1.90	5.02 (2.37, 7.68)	0.33
Shock	1.39 (0.91, 1.87)	2.11	3.47 (1.98, 5.00)	0.25
All	3.11 (2.93, 3.30)	2.24	5.35 (3.86, 6.85)	0.12

The second set of results condition on whether a rebel group initiates violence immediately upon formation. Overall, the expected time to civil war is slightly shorter. The shorter sojourn time to civil war arises because the expected amount of time a rebel group is nonviolent is much shorter. In other words, once an rebel group begins to use violence, it typically does not stop using violence until it escalates to civil war. The classification accuracy for these expected windows is larger than the other case. This suggests that a large proportion of rebel groups immediately begin using violence instead of refraining from the use of violence.

As a first-cut, this provides suggestive evidence that it is possible to predict when an armed group will escalate to civil war given knowledge about its organizational characteristics and previous state of behavior. However, the misclassification rate is still large, suggesting there is ample room for refining these estimates.

8 Conclusion

When do civil wars begin? The results presented here suggest a puzzling answer. The probability of escalation to civil war peaks shortly after an armed group's formation. During these first few years, the armed group's ability to grow faster than the state can degrade it carve out a small window of opportunity to escalate. Once the armed group initiates guerrilla violence, the results suggest that this violence actually worsens the odds of escalation. Violence reveals more information about the group than it helps the group grow stronger.

The results show four different catalysts can precipitate civil war escalation. Different types of rebel groups escalate to civil war through these mechanisms. As a result, the expected time to civil war varies across rebel groups. Rebel groups with access to external bases or conducting operations at the time of a broader political shock are posed to escalate their campaigns right away.

For scholars, this research advances both theoretical and empirical debates on the causes of war. Theoretically, it contributes a new explanation about the timing of civil conflict based on the uncertainty and information surrounding an armed group's potential capabilities. While it shows that the broad set of cases involve rapid escalation, there is still substantial variation within these cases. Future extensions involve better theorizing under what conditions different types of armed groups escalate at different rates. If the state is inclined to delay, then an armed group has incentives to operate openly and overexaggerate its potential capabilities in order to negotiate a better outcome. If the state is strong, then an armed group has incentives to hide and underexaggerate its potential capabilities in order to avoid detection. It may sporadically conduct attacks that are not commensurate with its true capabilities.

Empirically, the Markov chain analysis here presents a novel way to examine the timing and escalation of conflicts. While scholars have made substantial progress on this question in the past using hazard models and survival analysis, Markov chains allow for a more nuanced look at conflict sequencing. The process measures both the probability of civil war occurrence and the expected duration of these conflicts. Beyond the insurgency and civil conflict literature, Markov chains provides a new framework to understand interstate conflict escalation. A Markov chain analysis could, for example, could empirically test spiral models of conflict by modeling whether tensions gradually ratchet up or emerge more idiosyncratically. There is a rich set of future research to better understand when conflicts occur.

References

[1] Abrahms, Max. "Why Terrorism Does Not Work." *International Security* 31, no. 2 (2006): 42-78.

[2] Almgren, Robert, and Neil Chriss. "Optimal Execution of Portfolio Transactions." *Journal of Risk* 3 (2001): 5-40.

[3] Amemiya, Takeshi. *Advanced Econometrics.* Harvard University Press, 1985.

[4] Arieff, Alexis, Lauren Ploch Blanchard, and Tomas F Husted. "The Lord's Resistance Army: The U.S. Response." *Congressional Research Services*, 2011, 23.

[5] Asal, Victor, Ken Cousins, and Kristian Skrede Gleditsch. "Making Ends Meet: Combining Organizational Data in Contentious Politics." *Journal of Peace Research* 52, no. 1 (2015): 134-138.

[6] Axelrod, Robert. "The Rational Timing of Surprise." *World Politics* 31, no. 2 (1979): 228-246.

[7] Bartusevicius, Henrikas, and Kristian Skrede Gleditsch. "A Two-Stage Approach to Civil Conflict: Contested Incompatibilities and Armed Violence." *International Organization* 73, no. 1 (2019): 225-248.

[8] Black, Fischer, and Myron Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81, no. 3 (1973): 637-654.

[9] Carter, David B. "Provocation and the Strategy of Terrorist and Guerrilla Attacks." *International Organization* 70, no. 1 (2016): 133-173.

[10] Crenshaw, Martha. *Revolutionary Terrorism: The FLN in Algeria, 1954-1962.* Vol. 196. Hoover Institution Press, 1978.

[11] Crenshaw, Martha. "The Causes of Terrorism." *Comparative Politics* 13, no. 4 (1981): 379-399.

[12] Daly, Sarah Zukerman. "Organizational Legacies of Violence: Conditions Favoring Insurgency Onset in Colombia, 1964-1984." *Journal of Peace Research* 49, no. 3 (2012): 473-491.

[13] De la Calle, Luis, and Ignacio Sanchez-Cuenca. "What We Talk about When We Talk about Terrorism." *Politics and Society* 39, no. 3 (2011): 451-472.

[14] Diehl, Paul F. "What Are They Fighting for? The Importance of Issues in International Conflict Research." *Journal of Peace Research* 29, no. 3 (1992): 333-344.

[15] Duncan, George T., and Randolph M. Siverson. "Markov Chain Models for Conflict Analysis: Results from Sino-Indian Relations, 1959-1964." *International Studies Quarterly* 19, no. 3 (1975): 344-374.

[16] Eberwein, Wolf-Dieter. "The quantitative study of international conflict: Quantity and quality? An assessment of empirical research." Journal of Peace Research 18, no. 1 (1981): 19-38.

[17] Findley, Michael G., and Joseph K. Young. "Terrorism and Civil War: A Spatial and Temporal Approach to a Conceptual Problem." *Perspectives on Politics* 10, no. 2 (2012): 285-305.

[18] Fortna, Virginia Page. "Do Terrorists Win? Rebels' Use of Terrorism and Civil War Outcomes." *International Organization* 69, no. 3 (2015): 519-556.

[19] Geller, Daniel S. "Power Differentials and War in Rival Dyads." *International Studies Quarterly* 37, no. 2 (1993): 173-193.

[20] Jackman, Simon. "In and out of War and Peace: Transitional Models of International Conflict." Department of Political Science, Stanford University, 2000.

[21] Kessler, Richard J. *Rebellion and Repression in the Philippines.* Yale University Press, 1989.

[22] Kiefer, David. "Interstate Wars in the Third World: A Markov Approach." *Conflict Management and Peace Science* 10, no. 1 (1988): 21-36.

[23] Kinsella, David, and Bruce Russett. "Conflict Emergence and Escalation in Interactive International Dyads." *The Journal of Politics* 64, no. 4 (2002): 1045-1068.

[24] Kydd, Andrew H., and Barbara F. Walter. "The Strategies of Terrorism." *International Security* 31, no. 1 (2006): 49-80.

[25] LaFree, Gary, and Laura Dugan. "Introducing the Global Terrorism Database." *Terrorism and Political Violence* 19, no. 2 (2007): 181-204.

[26] LaFree, Gary, Laura Dugan, and Erin Miller. *Putting Terrorism in Context: Lessons from the Global Terrorism Database.* Routledge, 2014.

[27] Lake, David A. "Rational Extremism: Understanding Terrorism in the Twenty-First Century." *Dialogue IO* 1, no. 1 (2002): 15-28.

[28] Larson, Jennifer M., and Janet I. Lewis. "Rumors, Kinship Networks, and Rebel Group Formation." *International Organization* 72, no. 4 (2018): 871-903.

[29] Lemke, Douglas, and William Reed. "War and Rivalry among Great Powers." *American Journal of Political Science*, 2001, 457-469.

[30] Lewis, Janet I. "How Does Ethnic Rebellion Start?" *Comparative Political Studies* 50, no. 10 (2017): 1420-1450.

[31] Malone, Iris. Armed Group Dataset. Working Dataset. Stanford University. 2019.

[32] Malone, Iris. "Uncertainty and Civil War Onset." Working Paper.

[33] McCuen, John J. *The Art of Counter-Revolutionary War: The Strategy of Counter-Insurgency.* Faber and Faber, 1966.

[34] Midlarsky, Manus I. "Absence of Memory in the Nineteenth-Century Alliance System: Perspectives from Queuing Theory and Bivariate Probability Distributions." *American Journal of Political Science* 27, no. 4 (1983): 762-84.

[35] Merari, Ariel. "Terrorism as a Strategy of Insurgency." *Terrorism and Political Violence* 5, no. 4 (1993): 213-251.

[36] National Consortium for the Study of Terrorism and Responses to Terrorism. *Global Terrorism Database [Data File]*. 2016th ed., n.d. `https://www.start.umd.edu/gtd`

[37] Pape, Robert A. "The Strategic Logic of Suicide Terrorism." *American Political Science Review* 97, no. 3 (2003): 343-361.

[38] Reed, William. "A Unified Statistical Model of Conflict Onset and Escalation." *American Journal of Political Science* 44, no. 1 (January 1, 2000): 84-93.

[39] Reiter, Dan. *How Wars End.* Princeton University Press, 2009.

[40] Ross, Sheldon M. *Stochastic Processes.* Vol. 2. Wiley New York, 1996.

[41] Sambanis, Nicholas. "Terrorism and Civil War." *Terrorism, Economic Development, and Political Openness* 2008, 174-206.

[42] —. "What Is Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48, no. 6 (2004): 814-858.

[43] Sartori, Anne E. "An Estimator for Some Binary-Outcome Selection Models without Exclusion Restrictions." *Political Analysis* 11, no. 2 (2003): 111-138.

[44] Slantchev, Branislav L. "The Power to Hurt: Costly Conflict with Completely Informed States." *American Political Science Review* 97, no. 1 (2003): 123-133.

[45] Thomas, Jakana. "Rewarding Bad Behavior: How Governments Respond to Terrorism in Civil War." *American Journal of Political Science* 58, no. 4 (2014): 804-818.

[46] Toft, Monica Duffy, and Yuri M Zhukov. "Denial and Punishment in the North Caucasus: Evaluating the Effectiveness of Coercive Counter-Insurgency." *Journal of Peace Research* 49, no. 6 (November 1, 2012): 785-800.

[47] Tse-Tung, Mao. *On Guerrilla Warfare.* Courier Corporation, 1937.

[48] Walter, Barbara F. "Does Conflict Beget Conflict? Explaining Recurring Civil War." *Journal of Peace Research* 41, no. 3 (2004): 371-388.

[49] Walter, Barbara F. *Reputation and Civil War: Why Separatist Conflicts Are so Violent.* Cambridge University Press, 2009.

[50] Zartman, I. William. "The Timing of Peace Initiatives: Hurting Stalemates and Ripe Moments." *The Global Review of Ethnopolitics* 1, no. 1 (2001): 8-18.

[51] Zhukov, Yuri M. "Roads and the Diffusion of Insurgent Violence: The Logistics of Conflict in Russia's North Caucasus." *Political Geography* 31, no. 3 (March 1, 2012): 144-56.

9 Appendix A. Markov Chains

In this section, I delve into several other Markov chains and analyses. The results are used to show the difference in expectations as assumptions about the states of insurgency change. I also analyze several conditional logit models to see which country-level and group-level characteristics explain transitions between insurgency.

9.1 Simple Markov Chain

An alternative model assume only three states: $S_i = \{S_{\text{Latent}}, S_{\text{Guerrilla}}, S_{\text{War}}\}$. Once an armed group initiates violence for the first time it transitions to the guerrilla stage. This model assumes that an armed group cannot return to the latent stage once it reveals its presence to the state. The state gleans sufficient information from In other words, S_{Latent} is transient, $S_{\text{Guerrilla}}$ is recurrent, and S_{War} is absorbing.

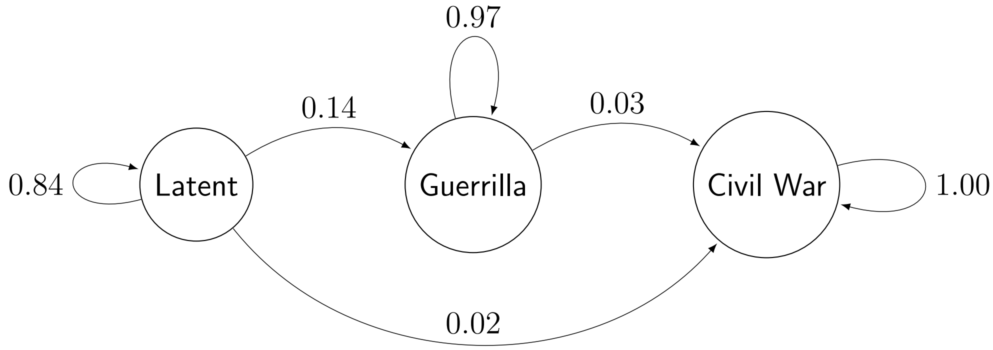
The transition matrix records information about transitioning between these three states. For example, p_{LG} is the probability that an armed group initiates violent activities in the next period. Similarly, p_{GW} is the probability that an armed group transitions from guerrilla violence to civil war in the next period.

$$\begin{array}{r} \\ \text{Latent}_{t-1} \\ \text{Guerrilla}_{t-1} \\ \text{War}_{t-1} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Latent}_t & \text{Guerrilla}_t & \text{War}_t \end{array} \\ \begin{pmatrix} p_{LL} & p_{LG} & p_{LW} \\ p_{GL} & p_{GG} & p_{GW} \\ p_{WL} & p_{WG} & p_{WW} \end{pmatrix} \end{array}$$

The transition matrix and diagram for armed groups in this sample are:

$$P = \begin{pmatrix} 0.84 & 0.14 & 0.02 \\ 0.00 & 0.97 & 0.03 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}$$

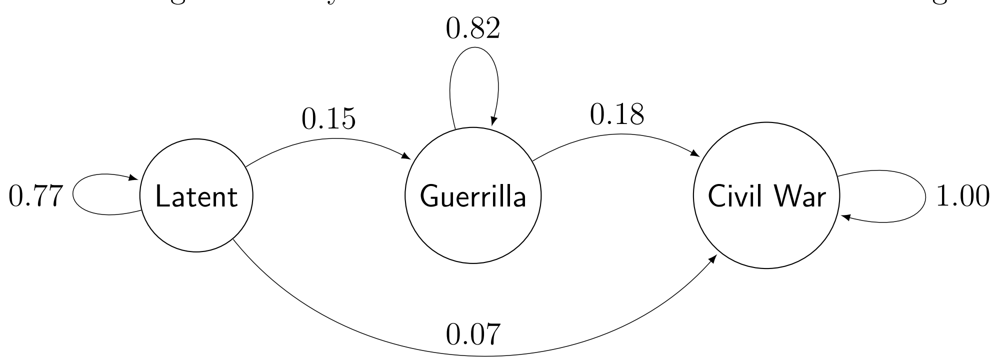
Figure 10: **Simple Transition Diagram of Armed Group Evolution.** This is the transition diagram for an overly simplified model. Armed groups enter each stage after they first reach the level of violence demarcating each stage. Armed groups that never escalate to civil war stay in the absorbing stage until they leave the sample.



The transition matrix and diagram for armed groups that eventually reach the civil war threshold are:

$$P = \begin{pmatrix} 0.77 & 0.15 & 0.07 \\ 0.00 & 0.82 & 0.18 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}$$

Figure 11: **Simple Transition Diagram of Rebel Group Evolution.** This is the transition diagram for only armed groups that eventually reach the civil war threshold. Armed groups enter each stage after they first reach the level of violence demarcating each stage.



I disaggregate the transition diagrams for different rebel groups depending on how they escalated to civil war. In the training set, 21 rebel groups have unclear escalations, 15 escalate via high-intensity rebel attacks, 31 escalate via high-intensity state attacks, and 48 escalate due to shocks. The transition matrices for each of these four categories are:

Training Set Transition Matrices

$$P_{\text{Unclear}} = \begin{pmatrix} 0.89 & 0.11 & 0.00 \\ 0.00 & 0.88 & 0.12 \\ 0.00 & 0.00 & 1.00 \end{pmatrix} \qquad P_{\text{StateAttack}} = \begin{pmatrix} 0.79 & 0.14 & 0.07 \\ 0.00 & 0.77 & 0.23 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}$$

$$P_{\text{RebelAttack}} = \begin{pmatrix} 0.74 & 0.13 & 0.13 \\ 0.00 & 0.81 & 0.19 \\ 0.00 & 0.00 & 1.00 \end{pmatrix} \qquad P_{\text{Shock}} = \begin{pmatrix} 0.58 & 0.27 & 0.16 \\ 0.00 & 0.75 & 0.25 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}$$

I analyze whether there are any differences. First, a multinomial confidence interval helps test whether p_{LW} is statistically significant across different types of mechanisms. While the transition probability is nearly double that of others, the 95% confidence intervals overlap.

Second, I examine differences in the expected duration of being in each state. I calculate the expected amount of time a rebel group is in the latent and guerrilla stage when it it initially nonviolent. The sum of these two expectations is the rebel group's expected time to civil war or the most likely 'window of opportunity' for the state to act (Table 7).

Table 6: **Expected Time to Civil War - Initial Nonviolence.** This table describes the the expected time it takes a rebel group to escalate to civil war after it forms when it is initially nonviolent (latent). 95% confidence intervals for each stage are in parentheses below.

Catalyst	**E(Latency)**	**E(Guerrilla)**	**E(Time to Civil War)**	**Classification Accuracy**
Unclear	9.09 (7.75, 10.43)	8.33 (6.28, 10.39)	17.42 (11.22, 23.63)	0.22
Rebel-Attack	3.85 (3.19, 4.50)	2.63 (1.49, 3.77)	6.48 (3.36, 9.59)	0.38
State-Attack	4.76 (4.20, 5.32)	2.90 (2.08, 3.72)	7.66 (5.49, 9.83)	0.42
Shock	2.37 (2.30, 2.44)	2.53 (1.88, 3.18)	4.90 (3.43, 6.37)	0.42
ALL	4.42 (4.29, 4.56)	3.76 (3.53, 4.00)	8.19 (6.97, 9.41)	0.10

The results show that the expected time to civil war is shortest for rebel groups which escalate through shocks. The expected time to civil war is longest for rebel groups which escalate through unclear mechanisms. I use the 95% confidence intervals for expected time to civil war to see how well these estimates predict an out-of-sample set of observations. The results show that the classification accuracy varies across groups. The expected time to civil war via a high-intensity catalyst accurately predicts between 38-42% of civil war timing in the test set. The expected time to civil war via a low-intensity catalyst only predicts 22% of civil war timing. There is room for improvement.

Table 7: **Expected Time to Civil War - Initial Violence.** This table describes the the expected time it takes a rebel group to escalate to civil war after it forms when it is initially violent (guerrilla). 95% confidence intervals for each stage are in parentheses below.

Catalyst	**E(Latency)**	**E(Guerrilla)**	**E(Time to Civil War)**	**Classification Accuracy**
Unclear	0.0 (0, 0)	8.33 (7.07, 9.59)	8.33 (4.15, 12.51)	0.44
Rebel-Attack	0.0 (0, 0)	5.26 (4.33, 6.19)	5.26 (2.53, 8.00)	0.13
State-Attack	0.0 (0, 0)	4.35 (3.84, 4.86)	4.35 (2.88, 5.82)	0.25
Shock	0.0 (0, 0)	4.00 (3.54, 4.46)	4.00 (2.67, 5.33)	0.25
ALL	0.0 (0, 0)	5.56 (5.39, 5.72)	5.56 (4.57, 6.54)	0.08

9.2 Dynamic Markov Chain of Conflict

I build on the Markov chain introduced in Section 4, but relax the assumption that civil war is an absorbing state. Instead, I assume that the intensity of a campaign can de-escalate after it reaches civil war . Armed groups die either after three years of inactivity of – if they are a rebel – after the last known year of violence.

$$P = \begin{pmatrix} 0.76 & 0.16 & 0.02 & 0.05 \\ 0.60 & 0.33 & 0.05 & 0.03 \\ 0.09 & 0.09 & 0.78 & 0.04 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

I estimate the n-step predicted probabilities using this transition matrix. The results are below.

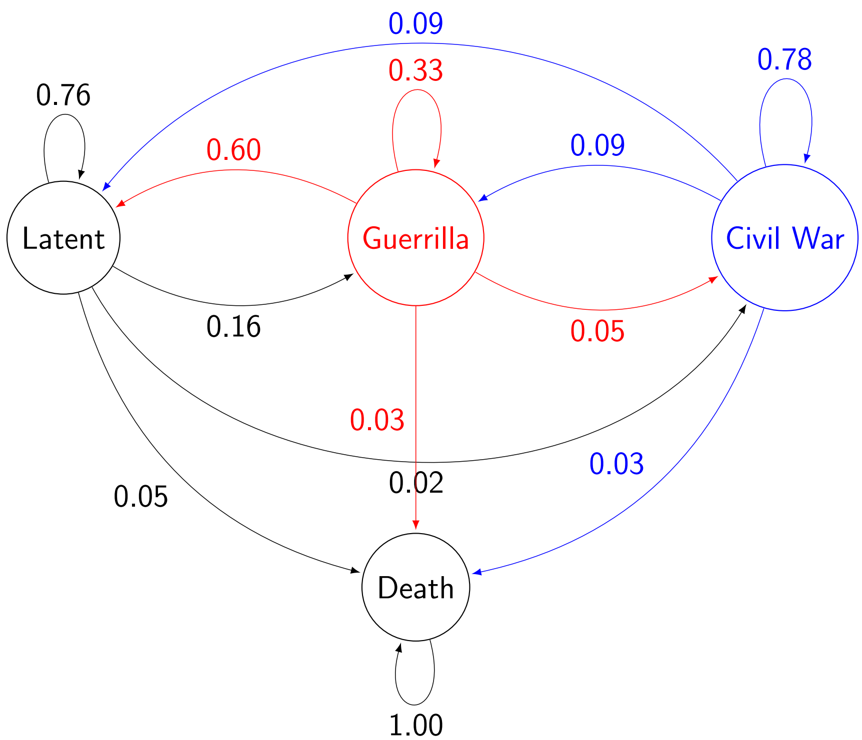
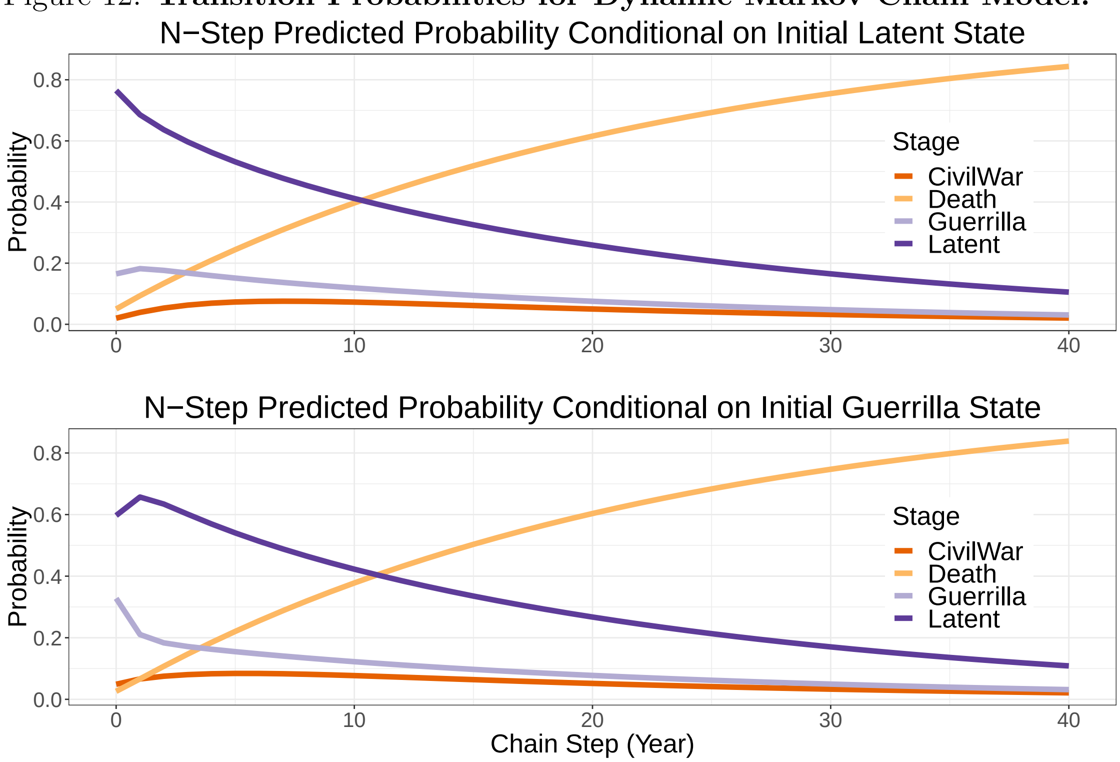
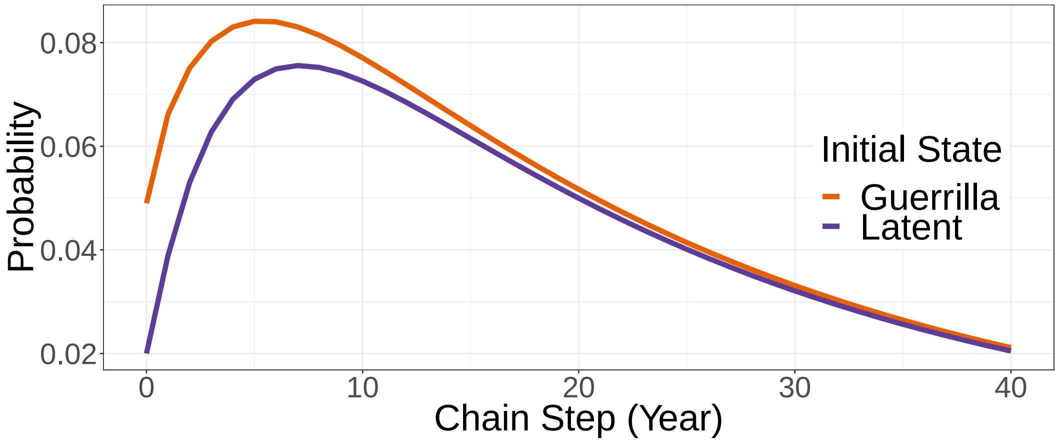


Figure 12: **Transition Probabilities for Dynamic Markov Chain Model.**



When civil war is a recurrent state, the probability of civil war peaks later on during an armed group's campaign. When an armed group starts from the latent stage, the probability of civil war peaks around 7 years into its campaign. When an armed group starts from the guerrilla stage, the probability of civil war peaks around 5 years into its campaign.

Figure 13: **Civil War Transition Probability for Dynamic Markov Chain.**



9.3 Dynamic and Disaggregated Markov Chain of Conflict

The last Markov Chain disaggregates civil war levels of conflict. It relies on the following assumptions:

- There are four states of insurgency: $S_i = \{S_{\text{Latent}}, S_{\text{Guerrilla}}, S_{\text{LIC}}, S_{\text{HIC}}\}$
- The Markov chain is regular and ergodic (irreducible).
- Armed groups are latent whenever there is no recorded violence in that year.
- Armed groups do not die.

This process assumes that armed groups cycle along four different conflicts. I disaggregate between low-intensity civil conflict, which produces between 25-999 battle-deaths in a single year, and high-intensity civil conflict, which produces over 1000 battle deaths in a single year. The transition matrix and diagram for this chain is below.

Table 8: Transition Probabilities of Dynamic, Disaggregated Markov Chain.

State_{t-1} / State_t	Latent	Guerrilla Activity	Low-Intensity (25-999bd)	High-Intensity (1000+bd)	N
Latent	4931 (0.823)	958 (0.160)	90 (0.015)	10 (0.002)	5989
Guerrilla Activity	870 (0.623)	458 (0.328)	63 (0.045)	5 (0.004)	1396
Low-Intensity (25-999bd)	78 (0.144)	61 (0.113)	363 (0.672)	38 (0.071)	540
High-Intensity (1000+bd)	13 (0.080)	4 (0.025)	34 (0.208)	112 (0.687)	163

The transition probabilities from this table can be summarized by the transition matrix P.

$$P = \begin{pmatrix} 0.823 & 0.160 & 0.015 & 0.002 \\ 0.623 & 0.328 & 0.045 & 0.004 \\ 0.144 & 0.113 & 0.672 & 0.070 \\ 0.080 & 0.025 & 0.209 & 0.687 \end{pmatrix}$$

Figure 14: **Transition Diagram of Dynamic, Disaggregated Markov Chain.** This chain assumes four stages of conflict, which all communicate with each other. Every stage is recurrent and armed groups cannot die.

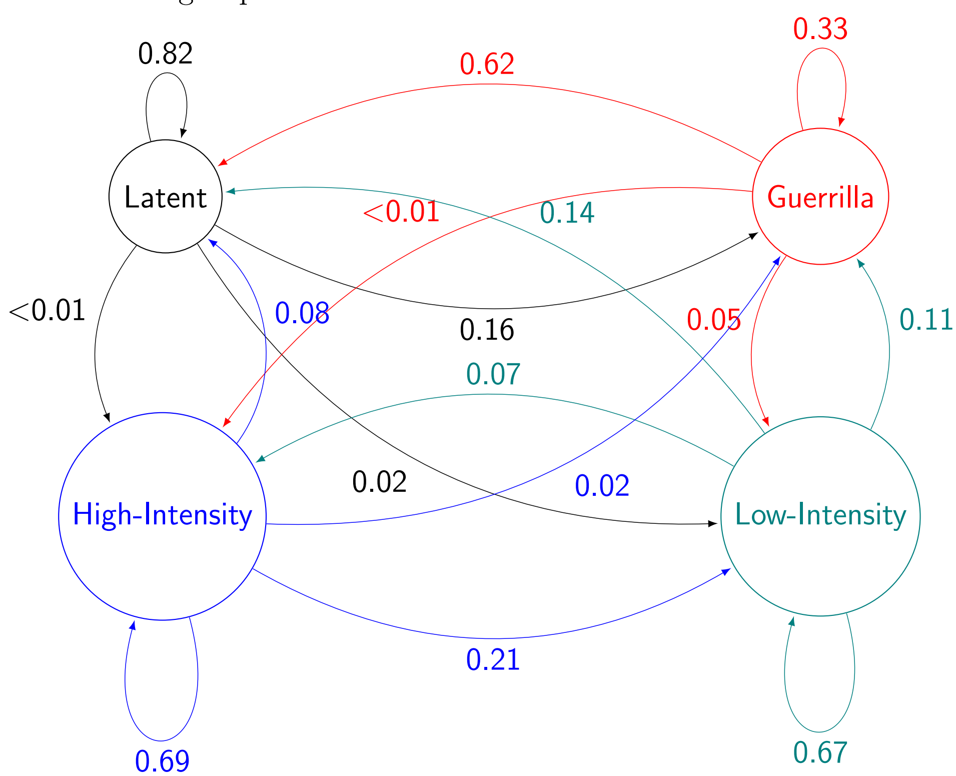
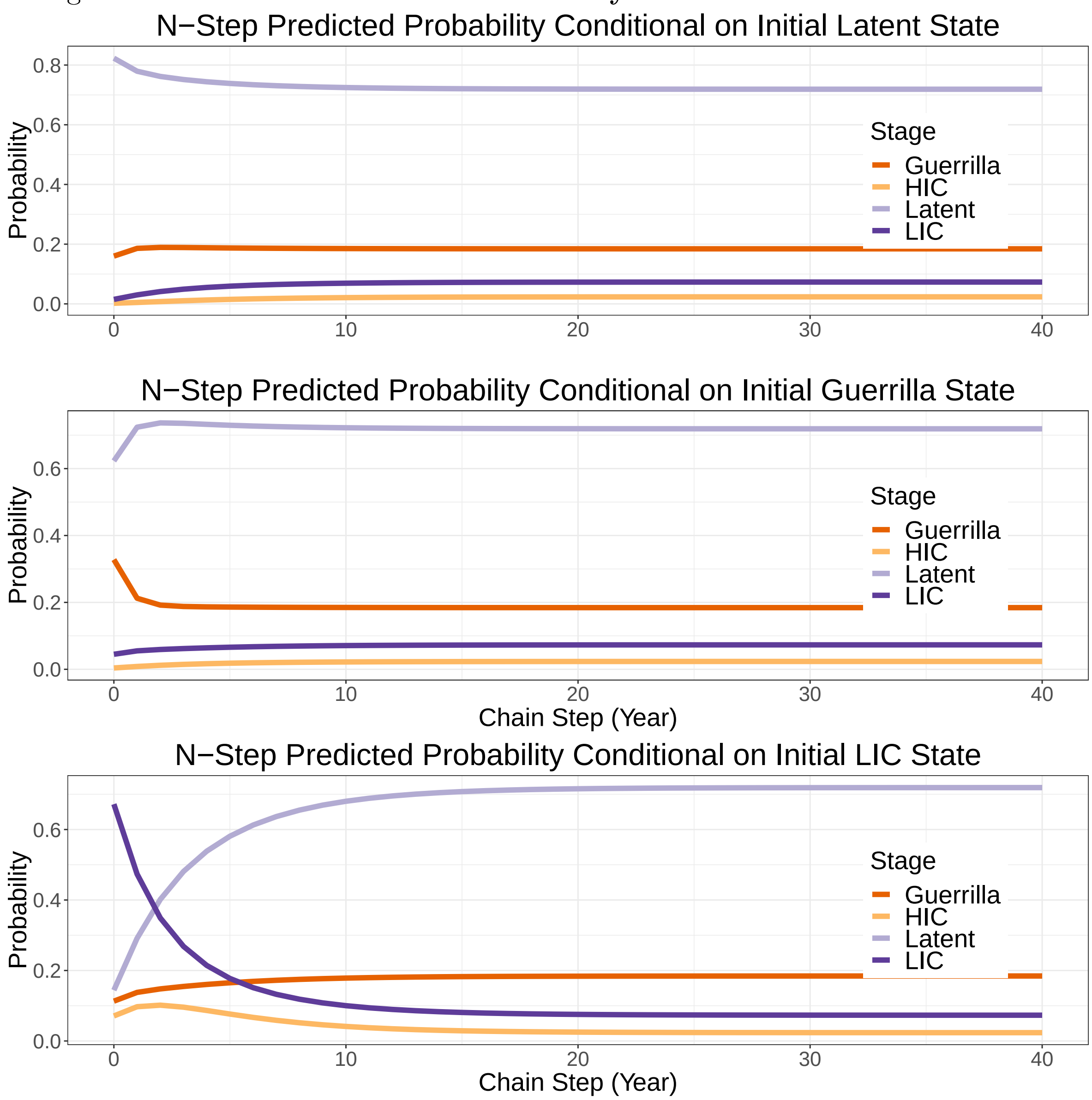
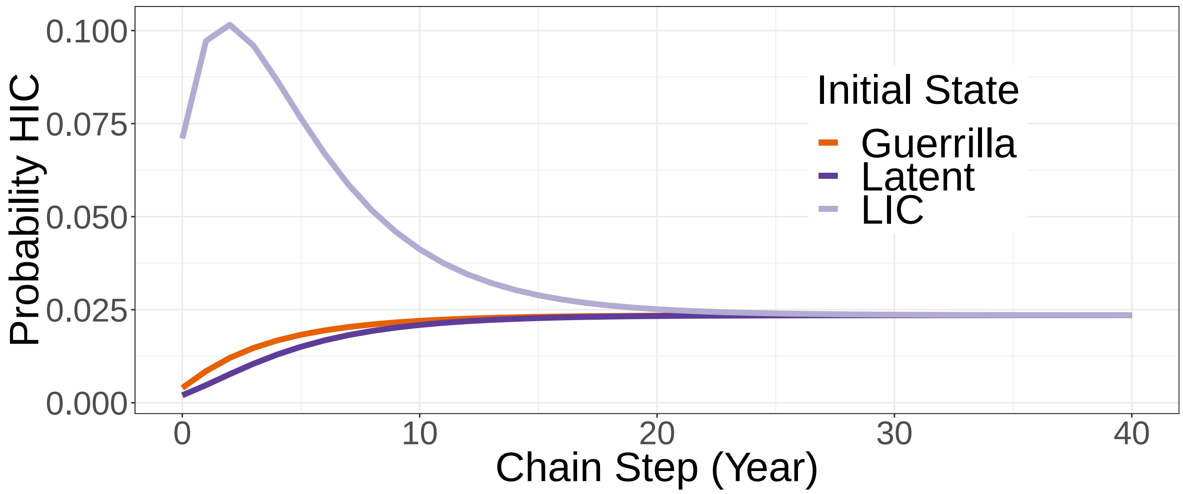


Figure 15: **Transition Probabilities for Dynamic Markov Chain Model.**



I estimate the transition probabilities for high-intensity civil war over n-steps. Unsurprisingly, the results show that the probability of transitioning to high-intensity conflict is highest when an armed group is already conducting a low-intensity insurgent campaign against the state. The long-run transition probability that an armed group escalate its campaign to a high-intensity civil war is approximately 2.5%.

Figure 16: **High-Intensity Civil War Transition Probability for Dynamic Markov Chain.**



9.4 Transitional Models of Conflict

An assumption in the models above is that the probability of transitioning to civil war only depends on the current state of conflict. This claim of relative exogeneity may strain credulity if we believe that certain types of dyads are more likely to transition to certain states. In order to assess how different correlates affect various transition probabilities, I estimate a series of conditional logit models.

Jackman (2000) recommends using the functional form of the estimating equation to estimate conditional changes in the probability of war:

$$\text{logit}[Pr(y_{it} = j \mid y_{i,t-1})] = y_{i,t-1}\alpha + x_{it}\beta_0 + \epsilon$$

These models condition the probability of transitioning to civil war both on the current state as well as additional covariates which could be correlated with both y_{it} and $y_{i,t-1}$. The dependent variable is the conditional transition between states. I run a pooled and fixed effects version of the model. The results are presented below.

The pooled results show that the probability of transitioning directly from latency to civil conflict is higher for separatist groups than for center-seeking groups. There are no comparable results governing the transition from guerrilla to civil war. The coefficients on GDP per capita and ongoing civil conflicts are both negative and statistically significant for p_{LW} suggesting that weaker countries are more likely to see armed groups immediately jump to civil war.

The pooled results are not ideal for two reasons. The group-level variables are not time-varying so they are not informative in estimating a change in future conflict levels. Further, the dependent variable examines all different types of transitions in the full training sample. It pools together different types of transitions, which the results above suggests is inappropriate.

I also conduct a fixed effects analysis. Within states, changes in authoritarianism and ongoing civil conflicts affect the probability of transitioning from latency to civil war.

Table 9: Pooled: Transition Probability Model

	Model 1	Model 2	Model 3	Model 4	Model 5
	L→ L	L→ G	L→ W	G→ G	G→ W
NONVIOLENT$_{t-1}$	8.21***	17.02***	16.09***		
	(0.84)	(0.23)	(0.19)		
GUERRILLA$_{t-1}$				0.00	19.46***
				(0.06)	(0.26)
SEPARATIST	−0.18*	0.23***	0.68*	−0.00	0.41
	(0.10)	(0.08)	(0.36)	(0.27)	(0.45)
EXTERNALBASE	0.42***	−0.25***	−0.02	−0.00	−0.07
	(0.08)	(0.09)	(0.29)	(0.14)	(0.38)
VETERANS	−0.12	0.16*	0.42	−0.00	0.55
	(0.08)	(0.08)	(0.27)	(0.11)	(0.35)
AUTOCRACY$_{t-1}$	0.25	−0.32*	0.22	−0.00	0.31
	(0.17)	(0.17)	(0.29)	(0.32)	(0.42)
LN(GDP/CAP)$_{t-1}$	−0.03	−0.00	−0.30***	−0.00	−0.38***
	(0.05)	(0.04)	(0.06)	(0.14)	(0.12)
OTHERCIVILCONFLICT$_{t-1}$	0.08	−0.02	−1.13***	0.00	−1.06**
	(0.10)	(0.09)	(0.37)	(0.24)	(0.41)
Country FE	N	N	N	N	N
Year FE	N	N	N	N	N
AIC	5425.11	4490.12	669.71	16.00	363.13
BIC	5478.85	4543.87	723.46	69.75	416.87
Log Likelihood	-2704.55	-2237.06	-326.86	-0.00	-173.56
Deviance	5409.11	4474.12	653.71	0.00	347.13
Num. obs.	6116	6116	6116	6116	6116

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. SE clustered by country.

Table 10: Fixed Effects Conditional Probability Transition Model

	Model 1	Model 2	Model 3	Model 4	Model 5
	L→ L	L→ G	L→ W	G→ G	G→ W
NONVIOLENT$_{t-1}$	8.32***	24.14***	17.93***		
	(0.87)	(2.06)	(0.32)		
GUERRILLA$_{t-1}$				-0.00	70.61***
				(0.00)	(3.49)
SEPARATIST	-0.40^{***}	0.42***	0.91	0.00	0.38
	(0.13)	(0.09)	(0.56)		(0.65)
EXTERNALBASE	0.34***	-0.20^{**}	-0.26	0.00	-0.22
	(0.08)	(0.09)	(0.44)		(0.50)
VETERANS	-0.02	0.14	0.26	0.00	1.64***
	(0.10)	(0.11)	(0.35)		(0.56)
AUTOCRACY$_{t-1}$	0.39**	-0.45^{*}	-1.06^{**}	0.00	-1.15
	(0.18)	(0.27)	(0.50)	(0.00)	(0.81)
LN(GDP/CAP)$_{t-1}$	0.01	-0.01	-0.03	0.00	-1.20
	(0.05)	(0.05)	(0.11)	(0.00)	(1.17)
OTHERCIVILCONFLICT$_{t-1}$	0.09	-0.01	-1.12^{***}	-0.00	-2.40^{***}
	(0.11)	(0.13)	(0.31)	(0.00)	(0.52)
AIC	5362.19	4528.40	711.02	240.00	407.94
BIC	6168.41	5334.62	1517.24	1046.22	1214.16
Log Likelihood	-2561.10	-2144.20	-235.51	-0.00	-83.97
Deviance	5122.19	4288.40	471.02	0.00	167.94
Num. obs.	6116	6116	6116	6116	6116
Country FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
AIC	5425.11	4490.12	669.71	16.00	363.13
BIC	5478.85	4543.87	723.46	69.75	416.87
Log Likelihood	-2704.55	-2237.06	-326.86	-0.00	-173.56
Deviance	5409.11	4474.12	653.71	0.00	347.13
Num. obs.	6115	6115	6115	6115	6115

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. SE clustered by country

9.5 Predicting When and How Civil Wars Escalate

As a descriptive exercise, I attempt to see if it is possible to predict the timing of civil conflict across different types of armed groups. I run a multinomial model that omits the previous state of conflict and uses regional fixed effects instead of country fixed effects. The results are comparable. I use these results predict the expected catalyst for conflict, if any, among a test set of armed groups. I then estimate the transition probabilities of this test set and compare it to the transition probabilities of the actual catalysts.

Comparing the different catalysts reveals slight variation in the expected transition probabilities. For example, the probability of war in the next period conditional on current latent activity is similar across all groups. The probability of war conditional on current guerrilla activity is more distinct. Escalation through a sudden high-intensity rebel attack or state attack is associated with a higher probability than escalation through an unclear mechanism.

The predicted and actual transition matrices are:

Predicted Transition Matrices

$$P_{\text{Unclear}} = \begin{pmatrix} 0.78 & 0.16 & 0.01 & 0.04 \\ 0.54 & 0.38 & 0.08 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

$$P_{\text{RebelAttack}} = \begin{pmatrix} 0.88 & 0.13 & 0.00 & 0.00 \\ 0.75 & 0.00 & 0.25 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

$$P_{\text{StateAttack}} = \begin{pmatrix} 0.80 & 0.11 & 0.07 & 0.03 \\ 0.54 & 0.33 & 0.13 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

$$P_{\text{Shock}} = \begin{pmatrix} 0.74 & 0.16 & 0.01 & 0.08 \\ 0.57 & 0.38 & 0.05 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

Actual Transition Matrices

$$P_{\text{Unclear}} = \begin{pmatrix} 0.81 & 0.14 & 0.05 & 0.00 \\ 0.52 & 0.27 & 0.21 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

$$P_{\text{RebelAttack}} = \begin{pmatrix} 0.84 & 0.11 & 0.05 & 0.00 \\ 0.00 & 0.33 & 0.67 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

$$P_{\text{StateAttack}} = \begin{pmatrix} 0.74 & 0.15 & 0.11 & 0.00 \\ 0.38 & 0.33 & 0.29 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

$$P_{\text{Shock}} = \begin{pmatrix} 0.67 & 0.11 & 0.22 & 0.00 \\ 0.25 & 0.55 & 0.20 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}$$

I apply the Chapman-Kolmogorov equation to these predicted transition probabilities and compare them to the actual set. As before, I differentiate between whether an armed group starts off nonviolent or violent. The results are in Figure 17 and Figure 18.

The plots show three important result. First, there is substantial room for improvement in the model's classification accuracy; it systematically underestimated the probability of transitioning to civil war across all four bins. Second, when an armed group initially forms and refrains from violence, the probability of civil war peaks during the first two to three years before declining. The clear exception is when a shock occurs. Under this circumstances, the expected probability of civil war is constantly declining in each successive year. This complements prior work that shocks produce transitory windows of opportunity to act. Once the effect of the shock dissipates, the rebel's advantage in challenging the state disappears.

Finally, the probability of rapid escalation varies by groups. Rapid escalation via a shock or high-intensity attack is much larger than the probability of other types of escalation. The area under the curve for the first five years of a rebel group's campaign is larger for high-intensity attacks than the area under the curve for unclear attacks. This is weakly suggestive evidence that groups with an external base, separatist aims, or veteran bases escalate more quickly than other groups. In other words, groups which can amass capabilities without attracting the state's attention are more likely to escalate soon after formation.

Figure 17: **Predicted Transition Probability of Civil War by Catalyst Risk and Initial Nonviolence.** This plots the predicted transition probability of civil war for armed groups when the initial state is nonviolent (latent). The results are differentiated by how the armed group is most at risk to escalate.

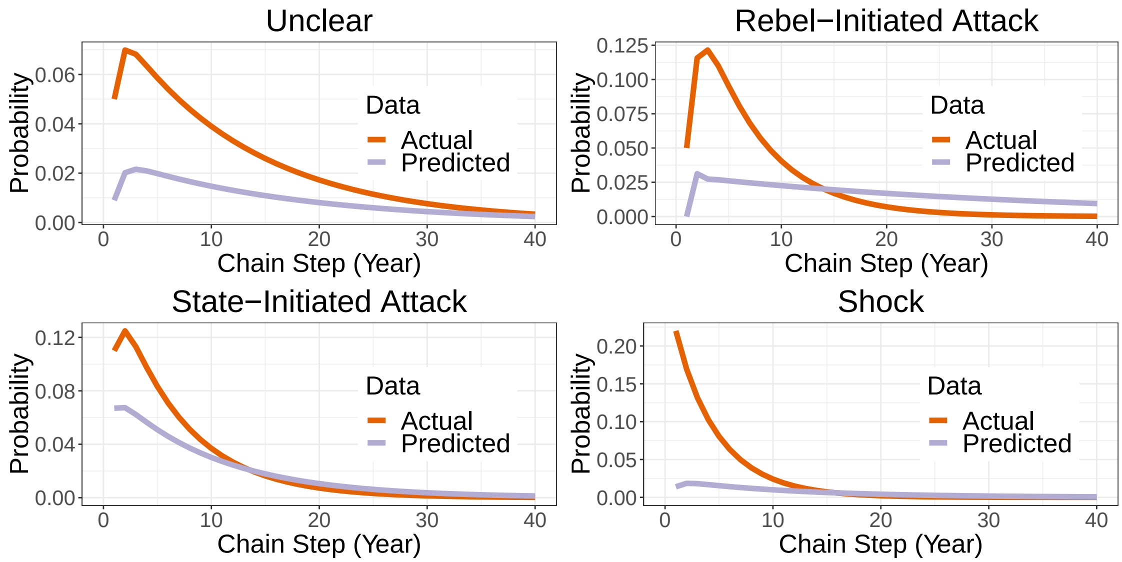


Figure 18: **Predicted Transition Probability of Civil War by Catalyst Risk and Initial Violence.** This plots the predicted transition probability of civil war for armed groups when the initial state is violent (guerrilla). The results are differentiated by how the armed group is most at risk to escalate.

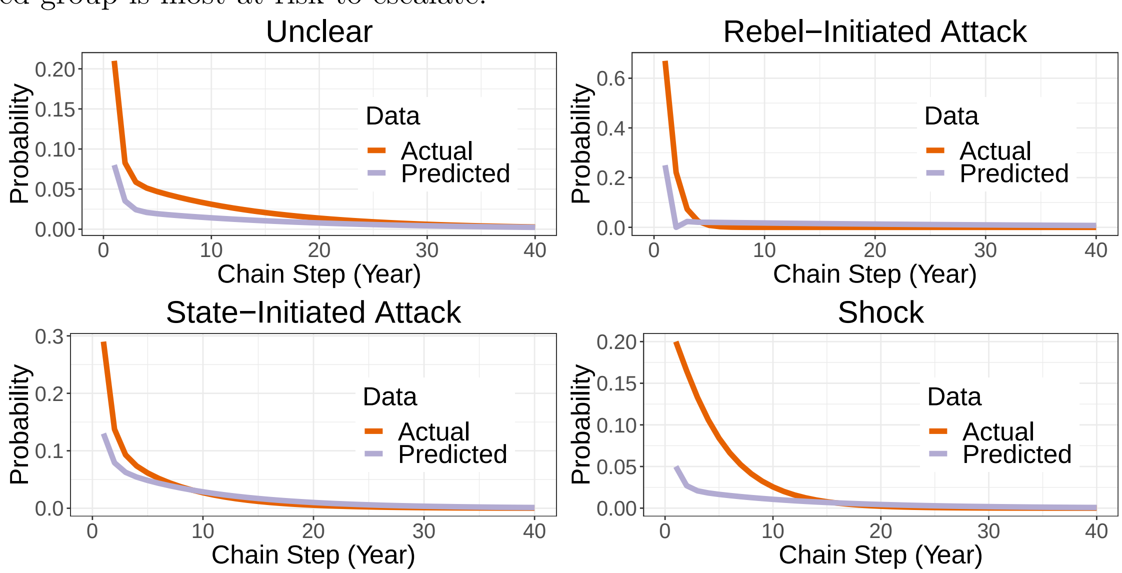
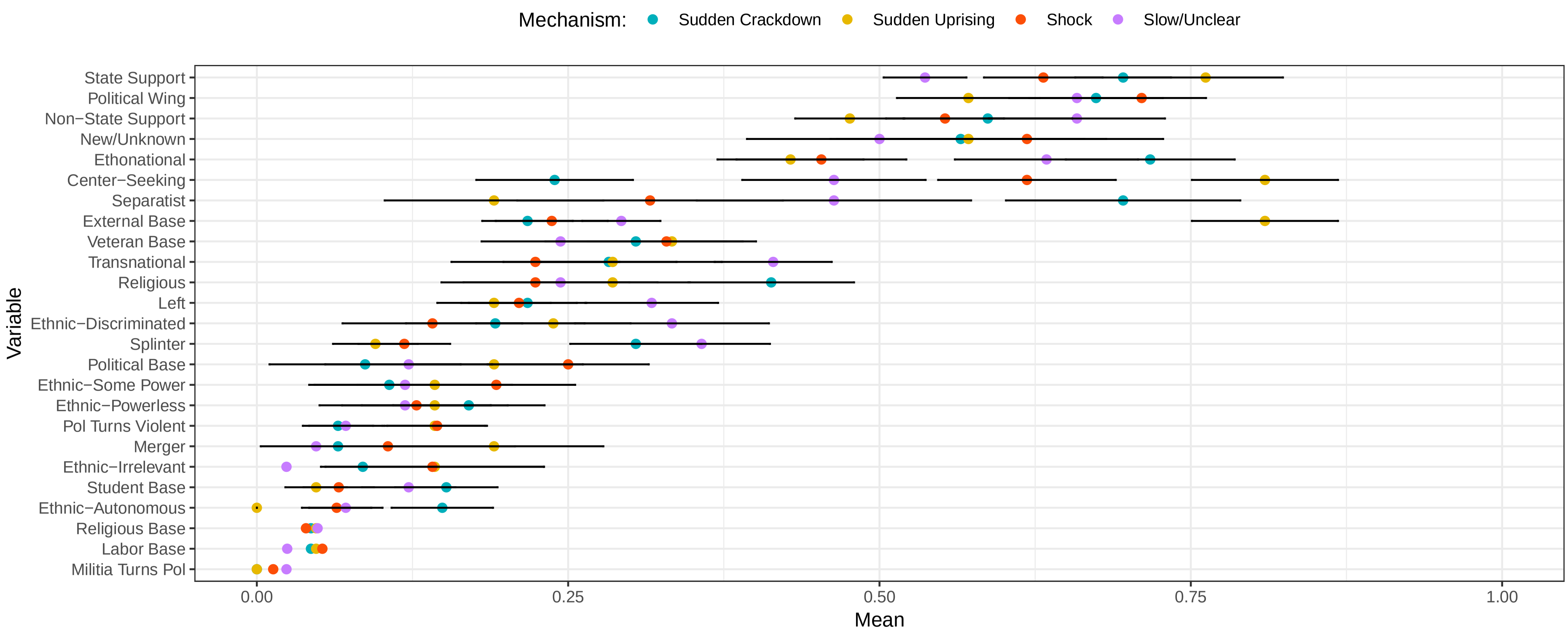


Figure 19: **Means and Confidence Intervals of Catalysts.** This plots the mean and standard error of different organizational characteristics by the four types of catalysts.



10 Appendix C. Robustness Checks

I conduct the following robustness checks on what predicts different types of civil war escalation:

- Remove Year Fixed Effects: I remove year fixed effects since the effect of any local shock could be mediated by common global shocks. This could lead the original model to underestimate the effect of different group-level and country-level characteristics on civil war. The size of the coefficients decreases slightly, but otherwise does not change the main results.

- Add Additional Covariates: I add two variables measuring GDP/capita and a binary variables measuring whether at least one other armed group within the country was fighting above the 25-battle death threshold in the previous year. Omitting these variables could bias results if weak countries or countries fighting a larger number of challengers discounted a threat more often. When I re-run the analysis with these results, the size of the coefficients decrease, suggesting they are correlated with both the primary explanatory and outcome variables. The overall main results still hold. The sign on prior civil conflict is negative and statistically significant for whether an armed group escalates via a shock. This implies that increases in the number of armed groups decreases the effect of a shock on any one armed group's relative capabilities. As a result, it is harder for any one armed group to benefit enough from the shock to escalate its campaign to civil war.

- Alternate Measure of Ongoing Conflict: I rerun the analysis above, but replace the binary variable demarcating a civil conflict in the previous year with the logged number of active violent armed groups within the country instead. This could more precisely measure the number of potential challengers a state faces even if not all of them are able to escalate their campaigns to civil war. Rerunning the analysis with this alternate measure decreases the coefficient size of the primary explanatory variables. The coefficient on external base is also no longer statistically significant. The direction of the variable is negative – the same as any prior conflict. However, in no model does the number of armed groups have a statistically significant effect on the outcome of interest.

Table 11: Robustness Check - No Year Fixed Effects.

Catalyst	(1) Unclear	(2) Rebel Attack	(3) State Attack	(4) Shock
GUERRILLA$_{t-1}$	1.80***	2.16***	1.39***	0.87**
	(0.47)	(0.72)	(0.40)	(0.35)
SEPARATIST	0.43	-2.66^{**}	1.18**	0.83*
	(0.73)	(1.33)	(0.53)	(0.43)
VETERANS	0.15	-2.24^{**}	0.75*	-0.22
	(0.55)	(1.03)	(0.42)	(0.42)
EXTERNALBASE	-0.10	1.40*	-0.85	-0.55
	(0.55)	(0.83)	(0.53)	(0.45)
AUTOCRACY$_{t-1}$	-0.58	2.37**	-1.04	0.18
	(0.87)	(0.94)	(0.80)	(0.42)
Country FE	Y	Y	Y	Y
Year FE	N	N	N	N
AIC	1550.68	1550.68	1550.68	1550.68
BIC	3593.15	3593.15	3593.15	3593.15
Log Likelihood	-471.34	-471.34	-471.34	-471.34
Deviance	942.68	942.68	942.68	942.68
Num. obs.	6116	6116	6116	6116

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. SE clustered by country.

Table 12: Robustness Check - No Country Fixed Effects.

Catalyst	(1) Unclear	(2) Rebel Attack	(3) State Attack	(4) Shock
GUERRILLA$_{t-1}$	1.90***	1.47**	1.15***	0.82**
	(0.46)	(0.64)	(0.40)	(0.34)
SEPARATIST	0.17	-1.07	1.25***	0.18
	(0.47)	(0.83)	(0.42)	(0.33)
VETERANS	0.27	-0.64	0.91**	0.43
	(0.50)	(0.81)	(0.40)	(0.36)
EXTERNALBASE	0.34	2.11***	-0.32	-0.47
	(0.49)	(0.71)	(0.48)	(0.38)
AUTOCRACY$_{t-1}$	0.21	1.03	-0.25	0.67**
	(0.53)	(0.67)	(0.52)	(0.33)
Country FE	N	N	N	N
Year FE	Y	Y	Y	Y
AIC	1408.82	1408.82	1408.82	1408.82
BIC	2671.93	2671.93	2671.93	2671.93
Log Likelihood	-516.41	-516.41	-516.41	-516.41
Deviance	1032.82	1032.82	1032.82	1032.82
Num. obs.	6116	6116	6116	6116

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. SE clustered by country.

Table 13: Robustness Check - Add Additional Covariates.

Catalyst	(1) Unclear	(2) Rebel Attack	(3) State Attack	(4) Shock
GUERRILLA$_{t-1}$	2.02***	6.84***	1.42***	1.33***
	(0.52)	(2.61)	(0.47)	(0.42)
SEPARATIST	0.25	−5.75**	1.37**	1.28***
	(0.79)	(2.62)	(0.57)	(0.47)
VETERANS	0.40	−5.11**	0.92**	0.15
	(0.58)	(2.23)	(0.46)	(0.45)
EXTERNALBASE	0.01	3.93**	−0.73	−0.76
	(0.59)	(1.86)	(0.55)	(0.51)
AUTOCRACY$_{t-1}$	−1.43	3.35	−1.45	−0.76
	(1.02)	(2.97)	(0.90)	(0.54)
OTHERCIVILCONFLICT$_{t-1}$	0.04	−0.26	−0.62	−1.91***
	(0.76)	(2.09)	(0.61)	(0.54)
LN(GDP/CAP)$_{t-1}$	1.31	1.50	−0.44**	−0.20
	(1.27)	(2.19)	(0.22)	(0.26)
Country FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
AIC	1645.79	1645.79	1645.79	1645.79
BIC	4843.87	4843.87	4843.87	4843.87
Log Likelihood	-346.89	-346.89	-346.89	-346.89
Deviance	693.79	693.79	693.79	693.79
Num. obs.	6116	6116	6116	6116

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. SE clustered by country.

Table 14: Robustness Check - Add Additional Covariates.

Catalyst	(1) Unclear	(2) Rebel Attack	(3) State Attack	(4) Shock
GUERRILLA$_{t-1}$	1.78***	2.22***	1.49***	0.88**
	(0.47)	(0.74)	(0.41)	(0.36)
SEPARATIST	0.42	−2.39*	1.32**	0.84*
	(0.73)	(1.36)	(0.53)	(0.44)
VETERANS	0.11	−2.37**	0.82*	−0.21
	(0.55)	(1.08)	(0.43)	(0.42)
EXTERNALBASE	−0.09	1.19	−0.93*	−0.56
	(0.55)	(0.88)	(0.53)	(0.45)
AUTOCRACY$_{t-1}$	−0.55	2.31**	−1.15	0.14
	(0.87)	(0.99)	(0.78)	(0.44)
LN(ARMEDGROUPS)$_{t-1}$	0.22	−0.83	−0.54	−0.18
	(0.87)	(0.81)	(0.53)	(0.48)
LN(GDP/CAP)$_{t-1}$	0.29	0.26	−0.32*	−0.05
	(0.60)	(0.98)	(0.19)	(0.23)
Country FE	Y	Y	Y	Y
AIC	1560.88	1560.88	1560.88	1560.88
BIC	3657.11	3657.11	3657.11	3657.11
Log Likelihood	-468.44	-468.44	-468.44	-468.44
Deviance	936.88	936.88	936.88	936.88
Num. obs.	6116	6116	6116	6116

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. SE clustered by country.

Table 15: **Predicted Transition Model.** These models classify how armed groups escalate to civil war. The dependent variable is whether an armed group ever escalates to civil war according to one of the four catalyst categories.

Catalyst	(1) Unclear	(2) Rebel Attack	(3) State Attack	(4) Shock
SEPARATIST	0.70^{***}	0.04	2.43^{***}	0.25
	(0.15)	(0.27)	(0.23)	(0.20)
EXTERNALBASE	-0.29^{*}	1.16^{***}	-0.08	-0.30
	(0.15)	(0.26)	(0.18)	(0.20)
VETERANS	-0.29	0.06	-0.23	0.09
	(0.19)	(0.27)	(0.22)	(0.21)
AUTOCRACY_{t-1}	0.07	0.98^{***}	0.51^{**}	0.78^{***}
	(0.15)	(0.26)	(0.20)	(0.18)
MENA	12.62	21.56^{***}	0.63	-0.26
	(89.94)	(0.27)	(0.44)	(0.39)
SSA	13.36	23.96^{***}	2.44^{***}	1.58^{***}
	(89.94)	(0.25)	(0.44)	(0.38)
ASIA	12.49	22.82^{***}	1.52^{***}	0.38
	(89.94)	(0.23)	(0.41)	(0.37)
AMERICAS	12.40	-46.13^{***}	0.31	0.81^{**}
	(89.94)	(0.00)	(0.72)	(0.40)
Region FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
AIC	5164.12	5164.12	5164.12	5164.12
BIC	6501.27	6501.27	6501.27	6501.27
Log Likelihood	-2382.06	-2382.06	-2382.06	-2382.06
Deviance	4764.12	4764.12	4764.12	4764.12
Num. obs.	5918	5918	5918	5918

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. SE clustered by country.