

The Prospect of Punishment and the Strategic Escalation of Civil Conflicts*

Casey Crisman-Cox[†]

June 23, 2020

Abstract

Why do some intrastate disputes escalate to conflict or war while others do not? One answer to this question that has not been considered in past work on intrastate conflict is whether would-be rebels face punishments for backing down from demands. I consider this explanation using recently released data on the outcomes of non-violent disputes and newly-developed tools for identifying the costs for backing down in disputes. I find that, on average, prospective rebels face large costs for backing down. Furthermore, I find that changes to these punishments produce novel results for understanding civil conflict. In most observations there is a *U*-shaped relationship between punishments and the likelihood of initiating a dispute. Additionally, in just over half of all observations, raising the punishments for backing down increases the likelihood that a group initiates a dispute against the state, while in the remaining observations such an increase raises the likelihood of civil peace. However, increasing these punishments reduces the likelihood that a government resists a threat in every observation. Overall, this suggests that the existence of these punishments strengthens a group's bargaining position and factors that raise these punishments, which include an increase in the number of other potential challengers to the state, may be detrimental to civil peace.

*Thanks to Tyson Chatagnier, Michael Gibilisco, Brad Smith, and the participants of the 2020 Texas Triangle Conference for their helpful remarks. Additionally, Niels Appeldorn and Swarup Das provided invaluable research assistance for this project. All errors are my own.

[†]Texas A&M University. Email: c.crisman-cox@tamu.edu

1 Introduction

When an intrastate group challenges the state, it is undertaking a risk.¹ After all, should the state stand firm against their demands, the group faces the prospect of either escalating into a costly civil conflict or backing down from their challenge. To the extent that backing down is a costly act, a strategic trade off appears and further questions are raised.² Specifically, does the prospect of punishment deter potential challengers to the state or does it embolden them by endowing their threats with enhanced credibility? Additionally, how do the strategic aspects of these punishments affect decisions to not only challenge the state but also to escalate from non-violent disputes to full-fledged civil conflict? The possibility of these strategic concerns leads to a broader theoretical and empirical question: Do groups expect to be punished for backing down and what effect does the expectation of punishment have on the onset and escalation of intrastate disputes?

The prospect of facing punishment for backing down in disputes has a rich history in the study of international relations where this mechanism is frequently discussed in terms of an audience cost or audience-cost-like penalty. However, such punishments are rarely considered in the study of intrastate conflict escalation. There are both theoretical and practical reasons for this omission. Theoretically, the prospect of an audience cost-like punishment for backing down is (at least historically) thought to be unique to democratic governments (Fearon 1994; Partell and Palmer 1999). Under this assumption, it would appear unlikely that intrastate groups would be constrained by their internal membership in any meaningful way. However, recent work has found increased evidence for audience cost-

¹I use the term “group” to refer to an intrastate actor who is capable of (but may or may not actually initiate) a violent or non-violent challenge the authority of the state. Groups become rebels if they decide to make this challenge violent.

²It remains an open question as to whether backing down during a dispute is costly to intrastate groups. However, there various examples of groups punishing their leadership at other points in the conflict process. One example is the case of Michael Collins who in 1921 signed a treaty with the British that fell short of Irish demands for independent and unified 32 county Irish republic. In doing so, he commented he had “signed [his own] death warrant” (quoted in Walsh 2002, 84). His prediction was borne out by both the subsequent Irish Civil War (1922-3) and his assassination. Whether or not such punishments affect pre-conflict intra-group dynamics is currently unknown.

like penalties across all regime types (Crisman-Cox and Gibilisco 2018; Weeks 2008). The discovery that autocratic leaders also face the prospect of punishment when backing down from disputes suggests that punishment mechanisms apply to a wider set of institutions and settings than just interstate disputes involving a democracy. Indeed, new studies have begun looking for whether the threat of punishment effects the outcomes and duration of civil conflict (Prorok 2016, 2018). One natural question that emerges is whether punishments (or, to be more precise, the prospect of them) affect the initiation of civil disputes and their escalation to civil conflict.

Practically, punishments for backing down have not appeared in work on the onset of civil conflict because the data seldom consider groups who take an “off ramp” and back down rather than fight. Standard civil conflict datasets distinguish between violent internal disputes and peace, but almost never include the cases where the dispute ends before violence begins. Some recent work relies on partially observed methods to sidestep the problem, which is the best an analyst can do when the data do not distinguish cases where rebels back down from their challenges to the state (e.g., Chatagnier and Castelli 2019; Nieman 2015). These data limitations hide the possibility of punishments from the analyst. As a result, the ability to consider what happens when groups back down from a demand has been historically limited.

In this paper, I use newly introduced data from Bartusevičius and Gleditsch (2018) to push deeper into the study of how civil conflicts escalate. These data focus on both violent and non-violent civil disputes and allow us to consider why some internal disputes escalate to conflict. I do this by fitting a formal model of dispute-escalation to these data to address new questions about the existence and effects of punishment in the pre-conflict stage of an intrastate dispute. These questions tie the intrastate conflict literature into the broader conflict literature in a new and more direct way by exploring how the prospect of punishment (long known to be an important driver of interstate dispute escalation) affects the dynamics of intrastate disputes. Specifically, I look at the decision making processes of governments

and groups from the moment of contestation to form a unified theoretical and statistical model of how civil disputes turn into civil conflicts.

Several results emerge from this approach. First, I find evidence that groups face the prospect of punishment when deciding whether to initiate an intrastate dispute. Second, the costs have a *U*-shaped relationship with the probability of dispute onset. Specifically, when the punishments are smaller than average, increasing them deters a group from initiating a threat. This deterrence effect can be thought of as a direct effect that expected punishments have on a group as the increase makes groups less willing to risk initiating a dispute against the state. However, as these punishments get larger, this trend flips and increasing punishments makes dispute onset more likely. The reason for this second trend is that large punishments serve as a commitment device that makes threats more credible. This enhanced credibility in turn has a coercive effect on governments, wherein raising the punishments for backing down increases the likelihood that the government grants a concession to the rebel groups. Overall, these punishments appear to generate an incentive to gamble among groups, wherein because governments become more likely to give in as these punishments increase, groups are more likely to challenge. In nearly all observations this leads to an overall increase in the risk of observed civil conflict, further demonstrating that the coercive benefit that groups gain from an increase in the expected punishment for backing down dominates the deterrence effect.

With these results, I contribute to the study of civil conflict in several ways. First, I provide direct evidence that potential rebel groups expect to be punished if they back down from a challenge against the state. This evidence builds on recent work that has found that groups expect to be punished for backing down within conflict (Prorok 2016), by showing that these punishment concerns appear throughout and affect the entire the civil conflict process. Second, I find a novel and non-monotonic result on civil dispute onset, where potential rebel groups weigh the prospects of punishment against the costs and benefits of dispute onset. Furthermore, while increasing the size of these expected

punishments sometimes deters a group from challenging the state, more often than not an increase makes a dispute more likely. In other words, raising the prospective punishment for backing down tends to actually embolden would-be rebel groups by instilling their demands with an enhanced credibility, which in turn makes governments more likely to accommodate rather than resist their demands. Additionally, these results suggest that selectorate theory can help in understanding civil dispute escalation and the internal dynamics of intrastate groups, supporting analysis by Cunningham and Sawyer (2019) on groups within active conflicts. Overall, variation in the expectation of punishment can explain why some intrastate disputes escalate to conflict while others do not as they directly affect decisions to initiate and escalate disputes.

2 Punishment, selection, and conflict

Past research on civil conflict has focused mainly on the direct transition from peace to armed violence. In recent years, however, there has been a push to better understand the pre-conflict or escalation stage that exists between peace and conflict. For example, Cunningham (2013) considers how various political and economic factors affect a potential rebel group's choice between conventional politics, violent conflict, and a non-violent protest, while Bartusevičius and Gleditsch (2018) consider a selection model for groups who first decide whether to initiate a dispute against the state and then whether that dispute will be violent or non-violent. This growing push to include non-violent disputes as an additional outcome to the peace-conflict dichotomy provides a fuller picture of the conflict process, and it builds on a long tradition of interstate conflict scholarship. Both the theoretical and empirical work on interstate conflict provide guidance on how to consider this intermediate stage between peace and armed conflict.

On the theoretical side, there is a long history of crisis bargaining models starting with Fearon (1994) and Schultz (1998) that provide the workhorse framework for understanding the escalation from peace to crisis to dispute. On the empirical side, Reed (2000) applies

Heckman-style selection models to capture these intermediate stages between peace and conflict (similar to what Bartusevičius and Gleditsch (2018) do for intrastate conflict). Building on this escalation framework, the theoretical and empirical models used by Signorino and Tarar (2006), Lewis and Schultz (2003), Crisman-Cox and Gibilisco (2018), and others focus on the strategic logic of dispute escalation by fitting formal models that capture the steps between peace and conflict. This line of work highlights the multi-stage bargaining process that leads to conflict, wherein actors face choices about either escalating or deescalating. As such, multi-stage empirical tools that match our theoretical understanding of the conflict process can provide important insight into this process.

The idea that conflict results from bargaining failure is not unique to the interstate system; observed civil conflict is also frequently framed as bargaining problem as well (e.g., Arena and Hardt 2014; Crisman-Cox 2020). One of the more common reasons why actors in the interstate system fail to find an agreeable alternative to fighting is that they are constrained by threats of punishment (Fearon 1994; Partell and Palmer 1999; Schultz 1998). Such costs typically result from forces within the state that punish actors who are caught making empty threats that hurt the state’s reputation or credibility. Indeed, the prospect of internal punishment is commonly thought to lead interstate leaders to avoid initiating disputes while leading them into costly wars when they do. However, only recently have scholars considered what these punishments might look like for non-state actors.

Whether or not (actual or potential) rebels face the threat of internal punishments remains an open question. However there are several reasons to suspect that the prospect of internal punishment exists in the intrastate setting. Recent work has increasingly found that groups can be and are punished by their own members or the population they represent. For example, we know that rebel groups can and do change leadership and that these changes in leadership can be in response to conflict or negotiation performance (Prorok 2016). Additionally, Cunningham and Sawyer (2019) note that many active rebel groups have selection mechanisms for choosing leaders that include institutions like elections or a selectorate of

high ranking officials. They further argue that even among tightly controlled groups where leaders choose their own successors, the choice is done with at least the tacit approval of a selectorate. In such cases, it becomes clear that internal punishment opportunities exist within groups and that leaders need to be cognizant of the support they hold among these groups.

The most direct look at the question of internal group punishments comes from Prorok (2016) who considers whether rebel leaders are punished for backing down during a civil conflict and granting concessions to the state. Specifically, she finds that rebel leaders received some form of internal punishment when conflict outcomes were disappointing to their intra-group audiences (2016, 73). Additionally, Prorok (2018) finds the threat of internal punishment leads rebel leaders to engage in the same type of “gambling for resurrection” dynamics that are common to state actors. These findings provide the most direct evidence to date that the prospect of punishment is a factor in determining how and when civil conflicts end. However, the existence of such costs during a violent conflict raises an important question about their role, if any, at the selection step into conflict. If the prospect of these punishments leads rebels to continue fighting in an otherwise costly conflict, we should expect the same punishments to be a strategic factor when groups decide whether to make a demand or challenge against the state in the first place. This leads to our main hypothesis of interest:

H1 (Backing down is costly): On average, groups expect to be punished for backing down in a dispute against the state.

Supposing that groups expect to face punishment when they back down rather than escalating to armed conflict, we want to know how these costs affect their decision to challenge the state. However, prior expectations about the effect of audience cost-like punishment on prospective challengers is very unclear. With little past work on the strategic escalation of civil disputes, we turn to the interstate conflict literature for theoretical guidance on what effect these punishments might have on a group’s decision to initiate a dispute and government’s response. On the one hand, the threat of punishment makes would-be rebels

more credible, leading opponents to give in rather than risk escalation to a full-fledged conflict. This coercion is a somewhat indirect effect that was originally noted by Schultz (1999), while Crisman-Cox and Gibilisco (2018) provide some of the first empirical support for this idea. The logic here is that, conditional on making a demand, groups will be more likely to stand firm when facing larger punishments. As a result, governments recognize that resisting them will likely lead to armed conflict. To the extent that governments want to avoid an armed fight, they will become more likely to make a concession to groups who face large punishments. Thus, punishment mechanisms strengthen a group's bargaining position and as the government becomes more willing to grant a concession, groups become more willing to initiate a challenge. The overall effect then becomes that groups will be more likely to challenge if the costs of backing down increase. This set of expectations can be thought of as a coercive advantage, where the threat of punishment results in less more disputes, which generates two additional hypotheses.

H2 (Emboldening effect): On average, increasing the expected size of punishments that groups face for backing down makes groups more likely to initiate an internal dispute.

H2a (Coercive effect): On average, increasing the expected size of punishments that groups face for backing down makes government concessions more likely.

On the other hand, a more direct effect may dominate a group's strategic calculus. The direct effect suggests that increasing the punishment for backing down will likely deter groups from challenging the state in the first place (as in Kurizaki and Whang 2015; Weeks 2012). Here, the threat of having to pay these costs maintains civil peace. Groups who are considering a dispute weigh the risk that they may find themselves choosing between costly armed conflict or paying a hefty punishment for backing down. As such, groups will be less likely to initiate a dispute and will instead choose to preserve the status quo. This leads to an additional hypothesis:

H3 (Deterrence effect): On average, increasing the expected size of punishments that groups receive for backing down makes internal disputes less likely.

Of course punishments do not exclusively come from within. Making threats and demands undoubtedly raises the ire of the state who may be interested in pursuing their own punishment against groups who challenge their authority. However, compared to those groups that escalate to armed conflict, it is perhaps unclear whether the state will want to pursue strong punishments against groups who back down. However, if states are looking to impose punishments against a group for the act issuing a challenge, then these would appear as a belligerence cost that groups pay after making a demand (as in Kertzer and Brutger 2016). When analyzing punishments below, we are able to hold these belligerence costs constant and look at only the effects of raising the punishments incurred when the group backs down to the state. As such, I frame the main discussion in terms of internal punishments, while in fact being agnostic about the actual source of punishment and also acknowledging that the results capture any external punishment for backing down beyond a belligerence cost (if any).

3 A formal and statistical model of civil dispute escalation

As in the interstate setting, civil conflict emerges as the result of bargaining failure. Such a failure tends to be costly for both sides, making the failure to avoid conflict all the more puzzling (Fearon 1995; Lewis and Schultz 2003). One of the best theoretical tools we have for understanding the process that leads from peace to conflict is the crisis-bargaining model. This framework accounts for the stage between “dispute onset” and “escalation to conflict” that international relations has long acknowledged as an important phase of the conflict process (Fearon 1994; Huth 1996).

Crisis bargaining models, while typically associated with interstate conflict, enjoys a broader range of applications including economic sanctions (e.g., Lacy and Niou 2004; Whang, McLean and Kuberski 2013), terrorism (e.g., Lapan and Sandler 1993), and civil

conflict (e.g., Arena and Hardt 2014). Its usefulness in the civil conflict setting comes from its ability to describe how threats can lead to either violent or non-violent disputes. Throughout, I stick with Lewis and Schultz’s (2003) version of the crisis bargaining game, as it (i) focuses directly on both the dispute-to-conflict escalation process, (ii) includes the prospect of punishment as a main parameter of interest, and (iii) is designed with empirical estimation in mind.

Figure 1 presents an outline of the theoretical model. Here we have a group R bargaining with a government G over some good or policy that G controls. At the start of the interaction, R decides if it wants to make a threat or challenge to the government. In cases where R does not initiate a dispute, the game ends with the status quo intact (outcome SQ). When the status quo is reached, the group and the government receive payoffs of S_R and S_G , respectively.

However, if R challenges the government, then G decides to either give in to the group (outcome CL) or to resist the group’s demand. If the government concedes to the threat, then the group receive a victory payoff V_R and the government receives a concession payoff C_G . If the government does not concede, then the group has to decide if they want escalate to civil conflict (outcome CC) or end the dispute without receiving a concession (BD). In the case of the former, armed conflict begins and the group and the government receive their conflict payoffs of W_R and W_G , respectively. In the case of the latter, the group receives a payoff a for backing down that includes any internal or external punishments incurred, while the government receives a victory payoff for retaining control of the good V_G .

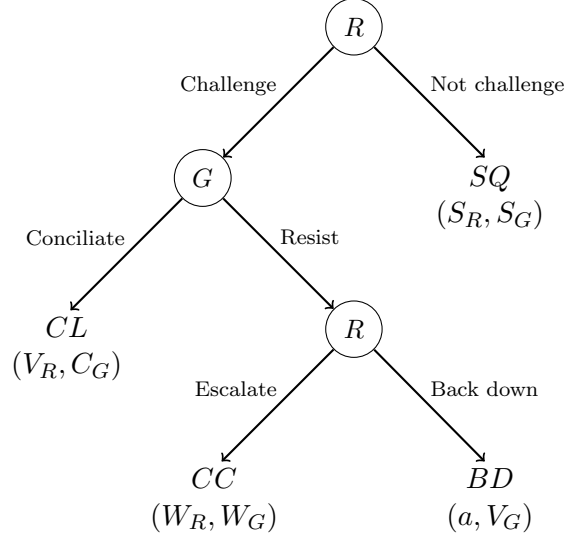
Following Lewis and Schultz (2003), I introduce private information in the form of three additively separable payoff shocks:

$$W_R = \bar{W}_R + \varepsilon_R$$

$$W_G = \bar{W}_G + \varepsilon_G$$

$$a = \bar{a} + \varepsilon_a,$$

Figure 1: The dispute escalation game



where each ε term is drawn from a standard normal distribution and represents private information that actors have about their costs of civil conflict $(\varepsilon_G, \varepsilon_R)$ and the expected punishment for backing down (ε_a) . At the start of the game, Nature reveals ε_R and ε_a to the group and ε_G to the government. These shocks ensure that every outcome is reached with positive probability. The set of structural parameters to be estimated is $\theta = (\bar{a}, C_G, (S_i, V_i, \bar{W}_i)_{i=R,G})$.

A perfect Bayesian equilibria (PBE) to this game can be represented using by a vector of choice probabilities. In particular, let p_r be the probability that G resists a demand, p_c is the probability that R challenges, and p_e is the probability that R escalates given that it made a demand. Jo (2011) shows that a PBE, $p^* = (p_c^*, p_r^*, p_e^*)$, to this game can be characterized as a solution to a system of three equations $p^* = \Psi(p^*; \theta)$. The full expression of Ψ can be found in Appendix E and full derivation details can be found in Lewis and Schultz (2003) and Jo (2011).

Before proceeding, it is worth taking the time to consider what the parameter \bar{a} means in the civil conflict context and how the empirical model can support or falsify the main hypothesis. Depending on the estimated value of \bar{a} , it can be a cost or benefit that the group receives for backing down in an internal dispute relative to its other payoffs. Following

Kurizaki and Whang (2015), I focus on the relationship between the backing down payoff \bar{a} and the status quo payoff S_R , which is to say that punishment is conceptualized as a receiving a worse payoff for challenging and backing down than having just accepted the status quo. Within the context of Hypothesis 1, we can think about three possibilities. First, we might have $\bar{a} < S_R$, that is groups, on average, face a penalty for making a threat against the state and not following through on it. In this case, civil conflict emerges at this final decision node by group leaders who decide that such a punishment is worse than taking their group into conflict.

However, it is not at all obvious that backing down should be all that costly. After all, both the state and internal group members may wish to actually encourage groups to not take their group into conflict. As such, it is entirely possible that groups may actually benefit from making demands/threats and not following through, in this case we would observe $\bar{a} > S_R$ and we would reject the main hypothesis. A group that appreciates the media exposure associated with just airing their demands might fall into this category. For example, Mahoney (2018) argues that groups might benefit from bluffing in situations where they want to intimate a larger audience or air their grievances on the cheap. Here, disputes escalate to civil conflicts when there is also a high value to fighting, perhaps from the existence of oil, or some kind of exogenous shock makes conflict more attractive (as in Fearon 2004).³

Under the remaining hypotheses we are interested in comparative statics on \bar{a} . The model is flexible enough to produce estimates that are associated with a variety of different comparative statics. As such, the theoretical structure does not “bake in” support for any of the remaining hypotheses. This setup allows the data to speak for itself in assessing the effect of \bar{a} on the conflict process.

³A third possibility is that we find $\bar{a} = S_R$, which is to say punishments are statistically indistinguishable from the status quo. Such a finding would not be strong enough evidence to reject the null hypothesis associated with the main hypothesis.

3.1 Empirical strategy and data

Overall, our goal is to consider the strategic selection that goes into the decision of potential rebels to escalate all the way to violence. Past efforts to consider these effects rely on Heckman-style selection models. While such efforts are important for bringing the escalation processes into the empirical study of civil conflict, there are a few advantages to adopting a structural approach based on a standard conflict model.⁴

First, structural approaches can be used to test for the presence of unobserved, but fundamentally interesting, parameters. There is a rich history in political science of using structural models to answer questions about theoretically interesting parameters such as ideal points (Poole and Rosenthal 1985) or audience costs in both interstate dispute (Crisman-Cox and Gibilisco 2018; Kurizaki and Whang 2015) and economic sanction threats (Crisman-Cox and Gibilisco 2019). Continuing this tradition by using a structural model as a tool to search for punishment effects in the intrastate setting is only natural.

Second, the structural approach offers a more direct match between theory and data than standard selection models. Accounting for the strategic nature of dispute escalation provides a theoretically principled way to make inferences from observational data. While no theoretical model will exactly capture the true data generating process, a reasonable model allows the researcher to capture important quantities of interest and conduct counterfactual experiments. Such model-based inference is particularly relevant in situations where (i) design-based inference (e.g., experimental manipulation) is infeasible and (ii) the quantity of interest is inherently theoretical and, as such, unobservable in nature.

Additionally, this particular structural approach provides an advantage over other structural approaches to studying the escalation to civil conflict. Specifically, data limits have led recent scholars to use partially observed strategic models to allow their “peaceful” outcomes to reflect both civil peace and civil disputes that do not escalate to conflict (Chatagnier and

⁴Note that the Heckman selection model is itself a structural model, but not one that was designed with crisis escalation in mind.

Castelli 2019; Nieman 2015). However, such approaches do not allow for the identification of punishments for backing down as they do not allow analysts to distinguish among the non-conflict outcomes. Fortunately, newly available data from Bartusevičius and Gleditsch (2018) allows us to move beyond partially observed structural models as it records a number of non-violent disputes.

Figure 2 reprints the game but in terms of regressors and parameters to be estimated. Notice that S_G is omitted as it does not enter into any either player’s strategic calculations, making it unidentifiable, and V_G is normalized to zero making it a reference category for the government’s other payoffs. It is well known that direct methods of estimating the β parameters using maximum likelihood are problematic due to the possibility that any particular guess at β may be associated with multiple equilibria (Jo 2011).⁵ Simply put, the multinomial log-likelihood associated with this formal model is a correspondence rather than a function and ordinary optimization methods will frequently converge to incorrect estimates.⁶ As such, I use the nested pseudo-likelihood estimator from Crisman-Cox and Gibilisco (2019), which they show to work well with this particular model. Additional details can be found in Appendix E.

3.2 Dependent variable

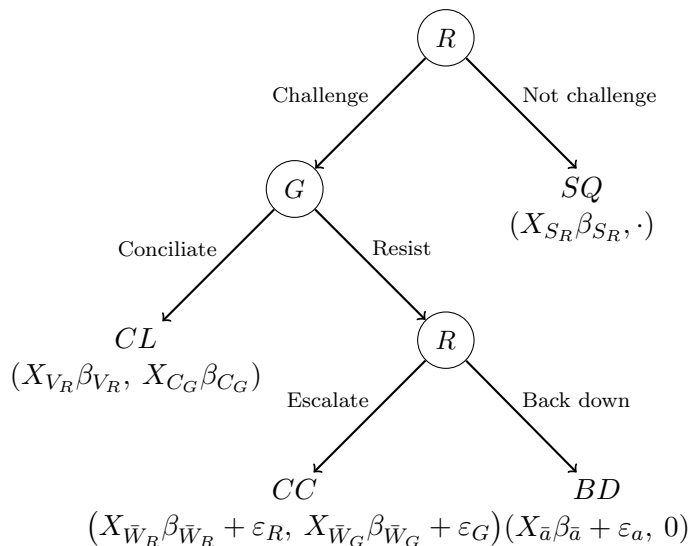
Case selection comes from two separate sources: the CONIAS data introduced by Bartusevičius and Gleditsch (2018) and the Minorities at Risk (MAR) project.⁷ The former are used to record cases of violent and non-violent disputes, while MAR groups that do not appear in CONIAS are used to add cases of status quo (non-dispute) outcomes. These two datasets are combined and built into dyad-decade observations, such that each observation records the interaction between a state and an actual or potential internal challenger. Over-

⁵Which is to say that the equations $p = \Psi(p; \beta)$ may have multiple solutions for any fixed value of β .

⁶Incorrect in the sense that the estimates will not actually maximize the log likelihood even when a successful convergence code is issued.

⁷The CONIAS data are produced by the Heidelberg Institute for International Conflict Research. These predecessor to these data, called the KOSIMO dataset, has been used by previous conflict scholars (e.g., Chiozza 2002).

Figure 2: Statistical version of the dispute escalation game



all, this approach closely follows Whang, McLean and Kuberski (2013) who also use dyad decades to mitigate measurement concerns when using the empirical crisis-escalation model. Specifically, any particular dispute can unfold over the course of several months or years before reaching its terminal node.⁸

Any particular country-group-decade is coded as ending in a status quo if no dispute onset appears in the CONIAS data during that decade. This selection of status quo cases identifies a plausible set of could-be rebel groups that decide not to initiate a dispute over the time period considered (as in Walter 2006). Defining status quo outcomes in this way provides us with a set of “politically relevant” cases where we consider all the cases where internal disputes occur at some point in time alongside a set of cases where a dispute was plausible (Maoz and Russett 1993). Including both dyads that do and do not experience a dispute in forming the status quo mitigates concerns about selection bias that would emerge if we only considered cases where a dispute appeared. This setup deliberately omits any states without a politically active minority presence for both theoretical (these states are

⁸As an alternative to the dyad-decade, I also consider an episodic approach where an observation is either a CONIAS dispute or a MAR group that does not appear in CONIAS in the online appendix. This approach has the advantage of having fewer aggregations and potentially questionable status quo cases, but has the drawback of perhaps understating the number of status quo outcomes. The main results hold with both the dyad-decade and the episodic data.

not playing this game) and technical reasons (estimates on the non-status quo payoffs would trend towards $-\infty$). As a result, inference here is conditional on the existence of a plausibly organized internal opposition (violent and non-violent).

For non-status quo cases CONIAS records what they call “political conflicts,” which they describe as violent and non-violent disputes between a government and an opposition actor that are beyond the scope of ordinary politics (Trinn, Wencker and Schwank 2016). They code these cases on a five-point scale based on their final outcome. The first two levels of this scale are non-violent events, which means that they end with either the government offering concessions to the group (*CL*) or with the would-be rebels backing down from their demands (*BD*). These cases were individually researched and hand-coded to determine which of these two outcomes best describes the situation.⁹ The main distinction used to code these cases is whether or not the dispute ends with major changes to the status quo policy in favor of the challenger. Examples of cases that involve this non-violent settlement in favor of the challenger include situations such as countries challenging and exiting the Soviet Union (without violence) at the end of the Cold War, peaceful democratic transitions, bloodless coups, and other non-violent cases of policy changes. These are all situations where a demand was made and a noticeable change occurred without an escalation to violence.

Alternatively, backing down cases include situations where groups initiate a challenge that, but neither escalate to violence nor get meaningful concessions from the state by the time the dispute ends. Some examples here include such cases as the Khmer people in Vietnam and various non-LTTE Tamil groups in Sri Lanka that threatened violence but never followed through. These cases are typically less well known as they do neither escalate to civil conflict nor result in new policies or government concessions. Again, the main coding depends on searching for pro-group changes to the status quo at the time COINAS records a dispute as over. Cases where the coding was ambiguous or could otherwise could not be determined are excluded here but considered in the online appendix. The value added of the

⁹Vignettes were produced by graduate student research assistants to aid in this coding and are available upon request.

CONIAS data is that it tracks these nonviolent cases, even those that result in little-to-no policy changes.

Note that this understanding of backing down is relatively broad and potentially overcounts the number of conflicts where a group backs down as it includes groups that disbanded or merged without receiving any kind of concession. These cases are clearly neither concessions by the government nor an escalation to conflict and as such they reflect a group that initiated a dispute but exits it without policy change. As such, this coding may not ideal as it could reflect cases where groups challenge but fail due to a lack of resources rather than a deliberate backing down, although it is entirely plausible that internal audiences would still punish this outcome as it is still a failure and likely damages a group's reputation.¹⁰ However, classifying these cases as backing down works against the main hypothesis as overcounting the number of *BD* outcomes will move the estimate of \bar{a} away from negative infinity (i.e., make backing down appear less costly than it might actually be).

A similar concern may be in recording a conflict as ending at *BD* because a government crack down leads the group to back down rather than escalate (as in cases like the 2007 Khmer protests in Vietnam). In these cases we might be worried that \bar{a} would be overly influenced by the belligerence costs imposed by the state for initiating a dispute (i.e., external punishments that incurred before backing down rather than a punishment for backing down). I account for this in the analysis by comparing the expected punishments (\bar{a}) to the expected costs of civil conflict (\bar{W}_R) as both will contain pre-conflict repression costs. If the expected backing down punishments are larger than the costs of escalation to conflict, then we can say that there is something more than just repression being captured by the parameters at this end node. Overall, a broad approach to classifying backing down outcomes as situations where there was neither an escalation to civil conflict nor a policy change provides a conservative approach to categorizing these cases that will at worst bias the estimate away from negative

¹⁰Indeed, the broad understanding matches the continuum of outcomes from Fortna (2015) who argues that “fizzling out” is a clear sign of group failure; this kind of failure would be a prime candidate for internal punishment.

infinity and work against the main hypothesis.

The remaining three CONIAS categories are used to denote a civil dispute that escalates to violence (outcome CC). In these cases, the dispute is treated as ending at civil conflict (CC) with neither side backing down during the escalation phase. Splitting the CONIAS cases in this way matches Bartusevičius and Gleditsch (2018), who demonstrate that this split maps into common understandings of civil conflict and reliably matches the violence conflicts recorded by the Uppsala Conflict Data Program (UCDP)

3.3 Independent variables and specification

The independent variables are standard civil war control variables that are thought to reflect the strategic escalation decisions of governments and groups. Specifically, I focus on controls that are known to both parties prior to the start of conflict and plausibly affect the decisions to both make a demand and follow through on it. As in Lewis and Schultz (2003), identification depends on there being no regressor (including the constant) that appears in all the payoffs for any given actor. To satisfy this condition, the group's status quo payoff, S_R , is specified without a constant term and the government's victory payoff when the group backs down, V_G , is normalized to zero. These are standard identification restrictions with this model (e.g. Crisman-Cox and Gibilisco 2019; Whang, McLean and Kuberski 2013).

The status quo is specified using covariates that directly affect the decision to initiate a dispute. Here, I focus on factors that might make peace more or less attractive. Specifically, I look at whether the state is newly formed, and logged population from the National Materials Capabilities (NMC) dataset provided by the Correlates of War (COW) project. New states provide a period of unrest/transition that may make the status quo more or less palatable, while population is included here as it is frequently one of the strongest predictors of civil dispute onset, and as Bartusevičius and Gleditsch (2018) argue, it is of particular relevance at the initial decision step. In fact, they go so far as to argue that population *only* matters at this stage in the selection process when they use it to satisfy the exclusion restriction in their Heckman selection model. As such, I include it here as a predictor of whether the

Table 1: Specification of the main model

Variable	Utilities	Description	Source
Fixed to 0	V_G	Identification restriction	–
Constant	$V_R, C_G, \bar{W}_i, \bar{a}$	Omitted from S_R for identification	–
New state	S_R	Is the state less than 10 years old?	COW
Population	S_R	logged total population	COW
Ethnic fract.	C_G	0-1 measure of ethnic heterogeneity	Fearon and Laitin (2003)
Oil	C_G, \bar{W}_R	Is more than 1/3 of G 's export revenue is from oil?	Fearon and Laitin (2003) & World Bank
Democracy	C_G, \bar{a}	G 's polity2 score and polity2 squared	Polity IV
Separatist dispute	V_R, C_G, \bar{a}	Is R a (potential) separatist group? (0/1)	Graduate research assistant coding and MAR
Mount. terrain	\bar{W}_i	Percentage of country that is mountainous (logged)	Fearon and Laitin (2003)
GDP per capita	\bar{W}_i	Per capita income (logged)	Penn World Table
Military size	\bar{W}_i	Military personnel per capita (logged)	COW
Other groups	\bar{a}	A count of the total number of actual and potential challengers (logged)	CONIAS & MAR

group initiates a civil dispute/crisis.

The payoffs for outcome CL (the government gives into a group's demand) are based on covariates that affect the value of the good and the reputational concerns associated with giving into the group. For the group, this outcome results in their victory-without-fighting payoff (V_R) and it is specified based on whether they seek territorial separation. I use this variable to capture the idea that different types of goals might be valued differently across

conflict types. The group's goal are hand-coded by graduate student research assistants for the CONIAS conflicts and based on the MAR territorial concentration variable for remaining cases.¹¹

For the government, this outcome results in their concession payoff (C_G). Here I specify it with whether the group has territorial separation as a goal, with the understanding that different types of disputes may be particularly bitter or difficult to resolve. Additionally, I include democracy and democracy squared measured using the state's polity2 score. Finally, I control for ethnic fractionalization to account for possible reputation concerns among governments. Specifically, the finding by Walter (2006) that states are less willing to offer concessions when there are lots of other potential challengers. I account for this by including the measure of ethnic fractionalization from Fearon and Laitin (2003).

For both sides' civil conflict payoff (\bar{W}_R, \bar{W}_G), I include the percent of mountainous terrain (from Fearon and Laitin 2003), the country's per capita GDP (from the Penn World Table), and the military personnel per capita (from the NMC data). These three variables reflect the economic and military capabilities within the conflict, with mountainous terrain serving as a pre-conflict proxy for would-be rebel's conflict ability (Lyal 2010). Additionally, for the group's armed conflict payoff, I include an indicator for whether the state is an oil exporter. This variable is based on Fearon and Laitin (2003) and records if more than 1/3 of their export revenue is based on oil exports using World Bank Development Indicators. It is included to capture the level of natural resources that may be available to a potential rebel group.

Finally, when groups back down they face a possible punishment (\bar{a}). Here I include democracy and democracy squared as a proxy for the level of external punishment imposed by the state, which may vary by regime type. Additionally, I use the number of other actual and potential groups within the country to proxy for the strategic context. The idea here is that with more groups there will be enhanced internal pressure for the group to not back

¹¹The logic here being that territorial concentration makes it easy for groups to coordinate on territory as a goal and as such this becomes the closest proxy in the MAR data.

down so that it can stand out within the competitive environment, which is a standard argument in theories of outbidding (e.g., Conrad and Greene 2015). Additional and more expansive specifications are considered in the online appendix. Ideally, a pre-treatment measure of group-level organizational ability could be added however it is not clear that such a measure exists that would describe all the potential and actual groups in the data. As a workaround, I also consider a model in Appendix C that specifies \bar{a} as a group-specific fixed effect. This fixed-effects specification has the advantage of capturing all the group-level differences, including organizational capacity and goals, that affect the punishment payoff. The main result still emerge under the fixed-effects specification, and as such, we stick to the simpler specification for the main presentation. The government’s victory payoff when the group back down (V_G) is normalized to zero which makes it a reference category for comparing its other payoffs. The main specification is presented in Table 1.

4 Results

The main results of the NPL estimation are presented in Table 2. The five columns in the table denote the associated outcome, payoff, variable, estimate and standard error for each the covariates in the model. The χ^2 statistic tests this model against an all-constants version of the model; the large value of the test statistic provides support for this model over the null model to test for whether the covariates offer any explanatory power over the outcomes. As additional measures of model fit I also consider the two “sensible” payoff restrictions mentioned in Schultz and Lewis (2005). The restrictions are that: the government prefers receiving the good without a fight to conceding the good ($V_G = 0 > \hat{C}_G = X_{C_G}\hat{\beta}_{C_G}$), and groups prefer receiving the good without resistance to the status quo ($\hat{V}_R = X_{V_R}\hat{\beta}_{V_R} > \hat{S}_R = X_{S_R}\hat{\beta}_{S_R}$); I find that these restrictions hold in every observation. Further the group’s victory payoff is also larger than the civil conflict and backing down payoffs for every observation, while the government also prefers that the group to back down rather than fight in very observation. The fact that nearly all of the results satisfy these sensible payoff orderings

provides some face validity to these estimates.

Table 2: Structural estimates

Outcome	Payoff	Variable	Estimate	Std. Error
Status Quo (<i>SQ</i>)	S_R	New state	-0.80	0.12
		Population	-0.07	0.03
Gov't Conciliates (<i>CL</i>)	V_R	Constant	-0.03	0.66
		Separatist dispute	-0.10	0.37
	C_G	Constant	-3.13	1.39
		Democracy	0.03	0.02
		Democracy squared	-0.00	0.00
		Separatist dispute	0.08	0.13
Civil Conflict (<i>CC</i>)	\bar{W}_R	Ethnic fract.	-0.50	0.29
		Oil Exporter	-0.18	0.20
		Constant	-2.01	0.23
		Percent mount. terrain	0.04	0.04
		GDP per captia	-0.13	0.08
	\bar{W}_G	Military personnel per captia	-0.06	0.04
		Oil Exporter	0.21	0.12
		Constant	-2.39	1.66
		Percent mount. terrain	0.16	0.05
		GDP per captia	-0.43	0.10
Group backs down (<i>BD</i>)	\bar{a}	Military personnel per captia	-0.06	0.05
		Constant	-2.83	0.33
		Democracy	0.02	0.01
		Democracy squared	-0.00	0.00
		Other groups	-0.14	0.08

Log-likelihood = -864.51

Observations = 2794

$\chi^2 = 83.30$

Turning our attention to the estimates, the first thing to note is that these are the direct effects that each variable has on that actor's utility for one particular outcome. As a result, the direct interpretation of any one parameter is very limited as the individual choices and outcomes are generated by the highly nonlinear choice probabilities. With that caveat in mind, there are still a few points of interest that can be discerned from the table. First, the common results that civil peace is fragile in new states and states with a large population appear here as both reduce the group's status quo value. Second, concessions are costly for governments. Notably there is a large negative constant term in C_G . Concessions are less

desirable in more ethnically heterogeneous states. This result is very intuitive and supports ideas from Walter (2006) and others who find that in more diverse states, governments are less willing to make concessions as it may encourage other potential challengers to organize. Additionally, we see an upside-down U shape between regime type and punishment where autocratic and democratic regimes have a lower utility for offering concessions than anocracies. One way to interpret this result is that democratic and autocratic states have strong institutions that are capable of punishing leaders for backing down during a dispute with a group as this may be seen as harming the state’s reputation against other prospective challengers. In contrast, Anocratic regimes have some of the weakest internal institutions and as such appear to face the lowest level of punishment for granting concessions to groups. Other notable results are that the constants on the conflict payoffs are strongly negative for both actors as well, demonstrating that violent conflict is costly.

The main point of interest here, however, is on the punishment parameter \bar{a} and its relation to the status quo outcome S_R . Specifically, to test Hypothesis 1 we need to know if the estimated values of \bar{a} are less than the estimated values of S_R . The first thing we note is that for every observation in the data the payoff for backing down is less than the status quo. The second thing we do is test the linear hypothesis for each observation, where we find that for every observation we can reject the null hypothesis that $\bar{a} - S_R = 0$ at the 5% level. As such we find support in favor of hypothesis one: in every observation considered, would-be rebel groups face the dilemma wherein making a demand and backing down is strictly worse than not making a demand at all. Additionally, this is not just the external costs of belligerence. In every observation we also note that \bar{a} is worse for groups (more negative) than the going into civil conflict, and we reject the hypothesis that $\bar{a} - \bar{W}_R = 0$ in 97% of these cases. This provides direct evidence that potential rebel groups perceive backing down as a costly proposition (more costly than escalation to conflict), this novel finding extends results from Prorok (2016) into the selection stage of the conflict process.

Finding that backing down from threats is bad for groups provides some insight into

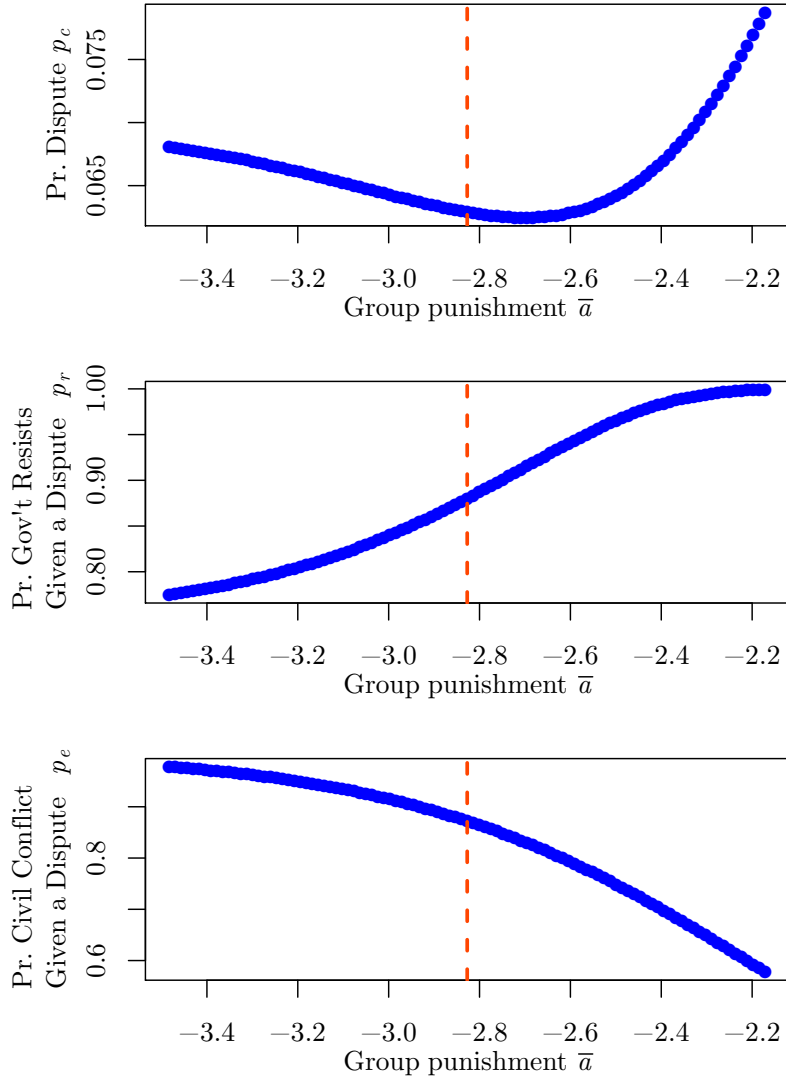
the dispute escalation process and offers a new explanation for why some intrastate disputes escalate to armed conflict while others do not. Specifically, the existence of these punishments suggests that groups engage in risky behavior when they initiate a dispute. As such, when the government does not back down, groups may find themselves escalating to costly civil conflict because they expect conflict to be less costly than the expected punishment for backing down. While this escalation process is familiar to interstate conflict scholars, the above results provide some of the first direct empirical evidence for these types of punishment concerns exist within the intrastate conflict process.

In Table 2, the only significant predictor of punishment is the number of other groups ($p < 0.10$). On average, as the number groups challenging the state increases, punishments for backing down become more severe. This result suggests a kind of outbidding dynamic may be at work, where groups are conscious of the fact that they may be competing with each other for resources or exposure Conrad and Greene (2015). Under this outbidding logic, we would expect that backing down would become increasingly costly to a group's reputation as the number of other groups increases, raising the incentive for internal punishment. I explore this result further in Appendix D. I find that across several specifications, the coefficient on the number of other challengers remains significant at conventional levels (5 or 10%), and that this remains the only significant predictor of expected punishment size. Future work should focus on expanding the data we have on violent challengers, non-violent challengers, and non-challengers to allow for a more thorough exploration of punishments.

4.1 The prospect of punishment and civil dispute escalation

Having established that groups, on average, expect to be punished for backing down in disputes with the state, we are now interested in finding out how the prospect of punishment affects both prospective rebel and government strategies. Specifically we are interested in whether these expected punishments have a coercive or deterring effect during this stage of the conflict process. The coercive effect, if present, manifests in two ways; as the size of the expected punishment increases, potential rebels should initiate more disputes and

Figure 3: Effects of punishment on strategies in an average conflict



governments should resist less. In contrast, if the deterrence effect is present we expect that increasing the size of the expected punishment would result in an increase in civil peace (i.e., groups are less likely to initiate a dispute).

To look for these effects, I first consider the effects that appear in an “average” interaction, where all covariates are fixed to either a mean or median value. With this particular data profile in hand, I then find all the equilibria for this average observation (only one in this case) and then calculate the comparative static for changing the size of the constant term in the punishment parameter \bar{a} . All the other payoffs are fixed to their estimated values.

One advantage of taking this comparative statics is it implicitly holds the baseline costs associated with initiating a dispute (i.e., belligerence costs) fixed. To see this consider that belligerence costs would enter all of the group’s non-status quo payoffs, by fixing V_R and \bar{W}_R and varying \bar{a} we are fixing the belligerence costs and isolating the effect that is strictly punishment. As a result, it allows me to identify the effects of punishment separately from the costs of initiation, which is historically difficult to do within an experimental or reduced form research design (Kertzer and Brutger 2016).

Figure 3 presents the comparative statics that the prospect of group punishment has on each actor’s strategy. Here we adjust the constant term in \bar{a} by raising and lowering it by two standard errors with the dashed line representing the estimated value. Adjusting the constant term allows us to take a comparative static that reflects how changes to these punishment parameters affect behavior in this average country-group dyad.

The top pane reflects the effect that prospective punishments have on the probability that a potential rebel group initiates a dispute (p_c). There are two things of note here. First, the overall probabilities are very small reflecting the rare events nature of civil disputes. Second, the relationship here is non-monotonic and takes a U -shape. On the side of the plot representing more severe punishments (closer to $-\infty$), the effect is positive, where an increased threat of punishment raises the prospect of a dispute (decreases the likelihood of civil peace). Moving approximately two standard errors towards negative infinity is associated with an almost 10% increase in the probability that a potential rebel group initiates a dispute. This result provides support for Hypothesis 2, in that it suggests that potential rebels are willing to exploit the credibility advantages of large punishments. Interestingly, a two standard error decrease in punishments (moving towards 0) is also associated with an increase (about 25%) in the probability of a dispute. While this latter effect offers some support for Hypothesis 3, it comes with a very specific caveat. When punishments are weaker than average, increasing them can result in deterrence, but only to a certain point. After this point, the coercion logic comes into play at this initial decision node.

The middle pane of Figure 3 considers the more direct aspect of the coercion argument by looking at how likely it is that a government resists a demand. Overall, there is a much clearer trend here where an increase in the average punishment increases the likelihood that the government offers a concession to would-be rebels. There are two implications of this trend. First, this is direct evidence in favor of the coercion hypothesis as governments are responding to increases in the punishment by giving concessions at a much higher rate. Second, it highlights the strategic aspects of the dispute escalation process: as punishments increase, coercion becomes an increasingly reliable strategy.

In the bottom-most panel we see an intuitive result where rebels are increasingly more likely to fight conditional on making a demand as the punishment for backing down increases. This combination of results, suggests that there is a definite strategic trade-off where would-be rebels risk a costly dispute in order to try and extract valuable concessions during the bargaining phase. Interestingly, in this “average” case there is only a very small effect on the unconditional probability of a civil conflict (not shown in the figure). However, the underlying conflict process is remarkably different despite the underlying risk of armed civil conflict being roughly unchanged. As punishments decrease (move closer to 0) disputes are more likely to end with the group making an empty threat and then backing down, while as punishments increase (toward $-\infty$) disputes are more likely to end with the government backing down and making a concession. Thus while armed conflict is avoided in both cases, the outcome and process are distinct.

Overall this average case analysis finds that the expectation of punishments can have either a deterring or a coercive effect on the decision to initiate a dispute. However, as punishments become more severe than average the coercive effect appears. This coercion story also appears when considering the likelihood that a government gives into a demand. As punishments increase, the government is increasingly less willing to risk a fight and is more likely to grant a concession to the rebel group to avoid armed conflict. While this average conflict setup provides an interesting overview, we might also be interested in how

Table 3: The marginal effects of punishment

	Marginal effect of $\bar{a} \rightarrow -\infty$
$\uparrow \text{Pr}(\text{Civil peace})$	45%
$\uparrow \text{Pr}(\text{Gov't gives in} \mid \text{Dispute})$	100%
$\uparrow \text{Pr}(\text{Civil Conflict})$	93%

Note: Percentages denote the proportion of observations where increasing the expected punishment (toward $-\infty$) increases the probability of the status quo, the probability that the government offers concessions, and the unconditional probability of a violent civil conflict.

these dynamics play out in the observed data.

One of the strengths of the statistical model is that under the estimated parameter values, different observations may have different comparative statics. Indeed, each observation can exhibit its own equilibrium behavior and its own marginal trend. This flexibility differs from standard reduced form analysis where, absent a flexible functional form, marginal effects will always be in the same direction across observations. Table 3 reports the percentage of government-group decades where the marginal effect of a more severe (more negative) punishment has a positive effect on a quantity of interest. Here I consider the probability of civil peace ($1 - p_c$), the probability that, given a dispute, the government grants a concession to avoid violence ($1 - p_r$), and the unconditional probability of armed conflict $p_c p_r p_e$.

In the first case, we see the same kind of split that appeared in Figure 3. In roughly 45% of observations increasing punishment increases the likelihood of civil peace. On the surface, this nearly even split looks like mixed evidence that supports both coercive and deterrence theories. However, an examination of marginal effects at plus and minus two standard errors reveals that in nearly 2/3 of observations, the U -shaped relationship between expected punishment size and the probability of a dispute appears.

Additionally, when we consider the effect on government concessions, we see that all the cases fall in line with the coercive theory's expectations. In every single observation, an increase in punishment results in an increased probability that the government gives into a demand. More severe punishments have a clear effect on government actions that appear

advantageous to potential challengers. On the whole, this result is strong evidence that the prospect of punishment greatly improves the bargaining position of would-be rebels despite the risk that they may have to pay it. As such, groups that want to maximize the likelihood of extracting concessions from a government should take lengths to have well structured punishment mechanisms that sanctions leaders who back down from challenges. While deterrence logic suggests that such a structure should discourage prospective challengers from initiating a dispute in the first place, the results here clearly suggest that the hand-tying benefit of punishment and the coercive strength that comes with is likely to lead to more threats and more policy concessions.

The final row in Table 3 provides an interesting result that highlights the importance of strategic model and the need to analyze the escalation steps within the conflict process. Here we see that increasing punishments makes civil conflict more likely in nearly all observations. On the surface, this may appear to be an unintuitive result. However, this result emerges from a kind of gambling dynamic. Specifically, as these punishments become more severe, governments become more likely to grant a concession; the allure of these concessions makes groups more willing to gamble, raising the probability of a dispute leading the group into conflict. As such, the unconditional risk of civil conflict rises as groups attempt to leverage the coercive effect of these punishments. That nearly all observations experience this increase in the risk of civil conflict as punishments increase, is further evidence that the emboldening effect of punishment dominates the deterrence effect.

5 Conclusion

In this paper, I asked do rebel groups face punishment for backing down from demands and how do these punishments effect the onset of both intrastate disputes and conflict? To address this question I used newly publicize data from the CONIAS project that records instances of internal disputes that do and do not escalate to violence from Bartusevičius and Gleditsch (2018). I fit these data to a dispute escalation game that tracks the bargaining

process from the decision to initiate a dispute through the decision to either back down or escalate to armed conflict.

Two main results emerged from this study. First, groups expect to be punished if they make a demand against the state and then back down. The prospect of this punishment is an important part of their strategic choice to initiate a dispute against the state. Second, depending on how strong these punishments are, they can have disparate effects on the likelihood of dispute onset. Specifically, I uncover a *U*-shaped relationship between the size of punishments and the likelihood that a group initiates an intrastate dispute.

When punishments are larger than average, factors that increase them make demands and threats more credible. Potential groups exploit this enhanced credibility by initiating more disputes. However, when punishments are weaker than average, increases to them make groups less likely to initiate a dispute. In this range, increasing punishments makes onset riskier for prospective rebels as the credibility benefits are less pronounced. Governments are very likely to resist demands from groups with small punishments, and groups are more likely to have to pay them.

Despite these competing effects at the initiation stage, there is clear evidence that the coercive effect of punishments dominates the deterrence effect when we turn to the government's decision. Specifically, increases in punishments lead to a decrease in the probability that the government resists demands. Governments understand that enhanced punishments tend to decrease the likelihood that a group will back down at its final node, and so governments back down instead rather than risk armed conflict. As such, governments are making a calculation that the risk of armed conflict is too much and are more willing to face their own reputational costs.

The primary contribution of this work is to provide direct evidence that would-be rebel groups face the prospect of punishment when deciding whether to initiate a dispute and that the effects of these punishments are non-monotonic. While recent work has identified that these punishments appear during armed conflict (Prorok 2016), the results presented

here show that punishment concerns appear at the earliest stages of the conflict process. Additionally, while there are good reasons to suspect that the threat of punishment might deter potential rebel groups, the evidence suggests that such punishments tend to have a more coercive effect that strengthen a prospective rebel's bargaining position and make governments more likely to back down before the armed conflict stage. Indeed these results provide a new take on why governments fight some rebels rather than others, wherein the size of the punishment faced by groups for backing down greatly effects the government's decision to either concede or risk a fight. An additional contribution is that these results provide new justification for the expansion of selectorate theories into the study of non-state actors (as in Cunningham and Sawyer 2019). Learning more about intragroup politics and structure provides us with a better sense as to where the risk of civil conflict may be highest and the results presented here suggest that selectorate theory may provide some of this insight.

Future work should focus on expanding the set of observables associated with non-violent CONIAS and other status quo cases. Notably, while recent data collection efforts have been very good at recording more observable characteristics of armed rebel groups (including their leader selection mechanisms), more should be done to record the observable characteristic for the non-violent groups in order to understand the escalation process. While such an undertaking will be undoubtedly costly, it will likely result in improved estimates of both the punishment parameters and in modeling the strategic selection into armed conflict. Another avenue of future work could also look to further disentangle the differences between external and internal punishment. While the structural model does some of that in the sense that it allows us to hold the belligerence costs fixed, a more direct separation of these forms of punishment may be interesting. Additionally, future scholars may also want to consider the benefits of a fitting a fully dynamic game of intrastate dispute escalation and de-escalation to data.

References

- Arena, Philip and Brian Hardt. 2014. "Incentives to Rebel, Bargaining, and Civil War." *International Interactions* 40(1):127–141.
- Bartusevičius, Henrikas and Kristian Skrede Gleditsch. 2018. "A Two-Stage Approach to Civil Conflict: Contested Incompatibilities and Armed Conflict." *International Organization* 73:225–248.
- Chatagnier, J. Tyson and Emanuele Castelli. 2019. "The Arc of Modernization: Economic Structure, Materialism, and the Onset of Civil Conflict." *Political Science Research and Methods* 7(2):233–252.
- Chiozza, Giacomo. 2002. "Is There a Clash of Civilizations? Evidence from Patterns of International Conflict Involvement, 1946-97." *Journal of Peace Research* 39(6):711–34.
- Conrad, Justin and Kevin Greene. 2015. "Competition, Differentiation, and the Severity of Terrorist Attacks." *Journal of Politics* 77(2):546–561.
- Crisman-Cox, Casey. 2020. "Democracy, Reputation for Resolve, and Civil Conflict." Unpublished manuscript.
- Crisman-Cox, Casey and Michael Gibilisco. 2018. "Audience Costs and the Dynamics of War and Peace." *American Journal of Political Science* 62(3):566–580.
- Crisman-Cox, Casey and Michael Gibilisco. 2019. "Estimating Crisis Signaling Games in International Relations: Problems and Solutions." Forthcoming at *Political Science Research and Methods*.
- Cunningham, Kathleen Gallagher. 2013. "Understanding strategic choice: The determinants of civil war and nonviolent campaign in self-determination disputes." *Journal of Peace Research* 50(3):291–304.

- Cunningham, Kathleen Gallagher and Katherine Sawyer. 2019. "Conflict negotiations and rebel leader selection." *Journal of Peace Research* 65(5):619–634.
- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3):577–592.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.
- Fearon, James D. 2004. "Why Do Some Civil Wars Last So Much Longer than Others?" *Journal of Peace Research* 41(3):275–301.
- Fearon, James D. and David D. Laitin. 2003. "Ethnicity, insurgency, and civil war." *American Political Science Review* 97(1):75–90.
- Fortna, Virginia Page. 2015. "Do Terrorists Win? Rebels' Use of Terrorism and Civil War Outcomes." *International Organization* 69(3):519–556.
- Huth, Paul K. 1996. *Standing your Ground: Territorial Disputes and International Conflict*. Ann Arbor: University of Michigan Press.
- Jo, Jinhee. 2011. "Nonuniqueness of the Equilibrium in Lewis and Schultz's Model." *Political Analysis* 19(3):351–362.
- Kertzer, Joshua D. and Ryan Brutger. 2016. "Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory." *American Journal of Political Science* 60(1):234–249.
- Kurizaki, Shuhei and Taehee Whang. 2015. "Detecting Audience Costs in International Disputes." *International Organization* 69(4):949–980.
- Lacy, Dean and Emerson MS Niou. 2004. "A theory of economic sanctions and issue linkage: The roles of preferences, information, and threats." *Journal of Politics* 66(1):25–42.

- Lapan, Harvey E. and Todd Sandler. 1993. "Terrorism and Signaling." *European Journal of Political Economy* 9:383–397.
- Lewis, Jeffrey B. and Kenneth A. Schultz. 2003. "Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information." *Political Analysis* 11(4):345–367.
- Lyall, Jason. 2010. "Do Democracies Make Inferior Counterinsurgents? Reassessing Democracy's Impact on War Outcomes and Duration." *International Organization* 64:167–92.
- Mahoney, Charles W. 2018. "Empty Threats: How Extremist Organizations Bluff in Terrorist Campaigns." *Studies in Conflict & Terrorism* Online first:1–21.
- Maoz, Zeev and Bruce M. Russett. 1993. "Normative and Structural Causes of Democratic Peace, 1946-1986." *American Political Science Review* 87(3):624–638.
- Nieman, Mark David. 2015. "Statistical Analysis of Strategic Interaction with Unobserved Player Actions: Introducing a Strategic Probit with Partial Observability." *Political Analysis* 23(3):429–448.
- Partell, Peter J. and Glenn Palmer. 1999. "Audience Costs and Interstate Crises: An Empirical Assessment of Fearon's Model of Dispute Outcomes." *International Studies Quarterly* 43(2):389–405.
- Poole, Keith T. and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29(2):357–384.
- Prorok, Alyssa K. 2016. "Leader Incentives and Civil War Outcomes." *American Journal of Political Science* 60(1):70–84.
- Prorok, Alyssa K. 2018. "Led Astray: Leaders and the Duration of Civil War." *Journal of Conflict Resolution* 62(6):1179–1204.

- Reed, William. 2000. "A unified statistical model of conflict onset and escalation." *American Journal of Political Science* pp. 84–93.
- Schultz, Kenneth A. 1998. "Domestic Opposition and Signaling in International Crises." *American Political Science Review* 92(4):829–844.
- Schultz, Kenneth A. 1999. "Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War." *International Organization* 53(2):233–266.
- Schultz, Kenneth A. and Jeffrey B. Lewis. 2005. "Learning About Learning." *Political Analysis* 14(1):121–129.
- Signorino, Curtis S. and Ahmer Tarar. 2006. "A Unified Theory and Test of Extended Immediate Deterrence." *American Journal of Political Science* 50(3):586–605.
- Trinn, Christoph, Thomas Wencker and Nicolas Schwank. 2016. "Introducing the Heidelberg Approach to Conflict Research." Available at https://www.uni-heidelberg.de/md/politik/personal/croissant/personal/trinn_et_al_-_introducing_the_heidelberg_approach.pdf. Last accessed: 13 February 2020.
- Walsh, Oonagh. 2002. *Ireland's Independence, 1880-1923*. New York: Routledge.
- Walter, Barbara F. 2006. "Building Reputation: Why Governments Fight Some Separatists but Not Others." *American Journal of Political Science* 50(2):313–330.
- Weeks, Jessica L. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization* 62(1):35–64.
- Weeks, Jessica L. 2012. "Strongmen and Straw Men: Authoritarian Regimes and the Initiation of International Conflict." *American Political Science Review* 106(2):326–347.
- Whang, Taehee, Elena V. McLean and Douglas W. Kuberski. 2013. "Coercion, Information, and the Success of Sanction Threats." *American Journal of Political Science* 57(1):65–81.