Washington University in St. Louis
Washington University Open Scholarship

Arts & Sciences Electronic Theses and Dissertations

Arts & Sciences

Winter 12-15-2022

Politics During and After Democratic Backsliding

Benjamin Rieth Schneider
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

 Part of the Political Science Commons

Recommended Citation

Schneider, Benjamin Rieth, "Politics During and After Democratic Backsliding" (2022). *Arts & Sciences Electronic Theses and Dissertations*. 2753.
https://openscholarship.wustl.edu/art_sci_etds/2753

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST.LOUIS

School of Arts & Science
Department of Political Science

Dissertation Examination Committee:
Randall Calvert, Chair
Dino Christenson
Brian Crisp
Justin Fox
Andrea Katz

Politics During and After Democratic Backsliding

by

Benjamin R. Schneider

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2022
St. Louis, Missouri

© 2022, Benjamin R. Schneider

Table of Contents

List of Figures

List of Tables

Acknowledgments

It is hardly original to say that writing a dissertation is a difficult process. In the early stages of working on a topic in the spring of 2020, much of my graduate school experience was turned upside down and stayed mostly upside-down through the remainder of my time working on this project. I have several people to thank for their either direct or indirect role in the culmination of this project.

First, I want to thank Randy Calvert and Brian Crisp for their role in advising me on this project over its life. I know advising on something not entirely in one's wheelhouse is not always comfortable, but I am thankful for the help and effort both placed in learning a lot about democratic backsliding as I developed my dissertation. I also want to thank my other committee members for their help and comments throughout the process: Justin Fox, Dino Christenson, and Andrea Katz. Also thanks to everyone else who gave comments throughout this process: Margit Tavits, Dan Butler, Dominique Lockett, Bryant Moy, Jaerin Kim, and I am sure others. Lastly, thanks to all of the professors and graduate students in the WashU Political Science Department who taught me along the way as I progressed from someone who knew not very much to someone who still does not know a lot, but has a very niche understanding of some mathematical and political concepts: Matt Gabel, Sarah Brierley, Randy Calvert, Ryden Butler, Guillermo Rosas, Miguel Pereira, Keith Schnakenberg, JB Duck-Mayr, Jacob Montgomery, Min Hee Seo, Sunita Parikh, Margit Tavits, Julia Park,

Chris Lucas, Erin Rossiter, Brian Crisp, Michael Bechtel, Justin Fox, David Carter, William O'Brochta, and Taylor Carlson. Lastly, a major thanks to my mom reading through my dissertation and helping me edit even despite her aversion to mathematical notation.

Next, I want to thank my family for their support throughout this process. My parents Scott and Lori Schneider and my sisters Delaney and Alexandra. Also a thanks to my family dog Gibbs and my own dog Tally for being there when I desperately needed some distance from my dissertation.

Lastly, I want to thank my friends and support system. My close friends in the department have helped me immensely: Hyun Woo Lim, Dominique Lockett, Bryant Moy, and Patrick Rickert. Also my friends outside of graduate school who helped me get my mind off of my dissertation and coursework at necessary times: Lauren and Mason Bo, Maggie Potter, Mathew Ferguson, Mariel Brown, Lauryn Schmelzer, Shubha Soman, Isabella Spinozzi, and David Lukner. A major thanks to my career advisor Arlene Taich and the rest of the career group who helped me navigate the job market while working on my dissertation: Hossein Foroushani and Sam Smith. And lastly, the Dissertation Support Group at WashU for help keeping me accountable: Dr. Karolyn Senter, Zhaoyi Ling, Vasudha Nukala, Niu Xiangfeng, Jane Li, Prasanthi Ramakrishnan, Josh Landman, Aditi Vashist, and Rasha Alsugair.

I am profoundly grateful to the help and support I have received the last 5 years and am glad to have shared an, at times isolating, process with some fantastic people.

Benjamin R. Schneider

Washington University in Saint Louis

August 2022

ABSTRACT OF THE DISSERTATION

Politics During and After Democratic Backsliding

by

Benjamin R. Schneider

Doctor of Philosophy in Political Science

Washington University in St. Louis, 2022

Professor Randall Calvert, Chair

Why do voters support backsliding incumbents? Under what conditions are voters more permissive to backsliding? In recent years, scholars have been interested in what leads to voters to support backsliding incumbents where we should expect them to be sanctioned and generally these answers have been focused on partisanship and polarization. I build on this literature by arguing that societal threat reduces the negative evaluation of backsliding actions and highlights the need for competent leaders to protect against future crisis. I develop an original formal model that considers the different strategies available to backsliding incumbents and shows that societal threat benefits backsliders most when they have a competence advantage. Otherwise, societal threat can work against them, and given the ability to increase societal threat salience, they will not do so. I empirically verify a conclusion of the model that higher societal threat leads to more democratic backsliding for some types of threat. This research has implications for the persistence of backsliding incumbents and the prospects of removal.

What does the opposition do after democratic backsliding? Much of the backsliding literature has been focused on the rise and persistence of backsliding governments, but less focus has been placed on what happens after these governments leave power. The democratization literature has some helpful expectations for successor behavior, but a newer

framework is needed because backsliding countries are different from autocracies. I use an original formal model to rigorously explore when a successor to an initial backslider either prioritizes restoring democratic institutions and norms, or uses weakened institutions to pursue policy priorities and maintain power. I show that three main factors drive this decision: the durability of restorative measures, the magnitude of the subverted institutions, and the electoral safety of a successor. When restorative measures are not durable, a successor never benefits from restoring institutions. When restoration is durable, more severely subverted institutions make maintaining backsliding less appealing to a successor. Electoral safety impacts different kinds of backsliding differently: restoring executive constraints is less advantageous as electoral safety rises, but the reverse is true for restoring free and fair elections. This study provides an understanding of both the choices of successors and the prospects for long term democratic resurgence after a single instance of backsliding.

How does backsliding affect voters' preference for democratic institutions? Most of the literature on voters' acquiescence to backsliding focuses on the initial infraction and electoral efforts of backsliding incumbents, but less is known about how the preferences among those who did not support the incumbent are affected. I consider how the severity of backsliding enacted by an incumbent of the opposite party affects voters' preferences for or against backsliding when deployed by their own party. I posit that voters consider the trade-off between backsliding from their *preferred* leaders and the policies enacted. When voters are more focused on policy, especially after more extreme backsliding actions, they are more likely to support their leaders retaliating against the initial backslider. This makes restoration unlikely because of incentives that produce a cycle of retaliation which can set the stage for democratic failure. I evaluate the theory using an original survey experiment in the United States, a country that has recently replaced a backsliding incumbent. I found that respondents were more likely to prefer a retaliating candidate when the previous administration enacted

more extreme backsliding. Expectations with respect to partisanship were not supported by the survey experiment. This study has implications for the prospects of either democratic resurgence or democratic failure in post-backsliding states.

Chapter 1

Democratic Backsliding and Why it Matters

1.1 Introduction

Democratic backsliding is a collection of actions that weaken democratic institutions in order to benefit a leader's political or personal goals (Bermeo 2016; Mechkova, Lührmann, and Lindberg 2017; Schedler 2010; Waldner and Lust 2018). Backsliding actions directly compromise a state's democracy by attacking institutional safeguards meant to limit the power of branches of government or by making it more difficult for voters to select the government. These attacks have severe ramifications for the future of a democratic state as they work to damage institutions and limit democratic participation (Waldner and Lust 2018). These actions are typically gradual and develop slowly throughout the tenure of on authoritarian-minded leader (e.g., Erdoğan's Turkey, Orbán's Hungary, Duda's Poland) who is usually the head of state, and aided by a strong party in the legislature (Driesen 2021).

These gradual declines in democratic quality contrast with quick and dramatic seizures of power by leaders, referred to as a self-coup or autogolpe (e.g., Fujimori's Peru), or by the military that used to be the major avenue of democratic collapse (Huq and Ginsburg 2018; Levitsky and Ziblatt 2018).

Backsliding events have become more common in the past couple of decades (Lührmann and Lindberg 2019). Cases range from more established backsliding regimes in Hungary, Poland, Turkey, and Venezuela to newer governments with early backsliding such as Brazil and the recently removed, but still relevant, Trump administration in the United States. Other cases like India, Israel and Mexico have also had more recent turns away from democracy that have received less attention. Beyond these cases, dozens more have experienced some level of democratic decline in recent years. Among the most recent cases, in Tunisia, President Kais Saied put forward a new constitution for referendum on July 25th of 2022 that was overwhelmingly approved by voters (Sands and Murphy 2022). This new constitution helps Saied concentrate powers of the presidency and allows him to declare emergencies in order to dissolve parliament and extend his term in office (Parker 2022). The number and severity of these cases makes the study of democratic backsliding increasingly noteworthy.

In this project, I focus on two stages of democratic backsliding: during and after. Considerable excellent work has been devoted to the precursors of backsliding and the ascent of would-be backsliders, so the "before" will not receive as much attention. In Chapter 2, I consider how voters' opinions of a backsliding incumbent develop and under what conditions they prefer to retain or replace the incumbent. I focus my answer for this question in terms of societal threat and how it can help backsliding incumbents justify their backsliding actions to voters by inducing voter support for a competent leader. I use both a formal model to explore this mechanism and empirical application to explore a key conclusion from my model: higher societal threat allows for additional democratic backsliding.

In Chapters 3 and 4, I explore what we should expect when backsliders are removed. This is a frontier of the literature due to the small number of cases, but is crucial to understand since democratic backsliding is a multi-step process and those steps do not have to occur under one backslider or even one party. In Chapter 3, I rigorously explore the preferences of the successors to initial backsliders through a formal model and outline three key factors that drive the decision to restore institutions or maintain subverted institutions: the durability of restorative measures, electoral safety, and the magnitude of the subverted institutions. In Chapter 4, I focus on voters and under what conditions they are more or less likely to support further backsliding. I evaluate this through a survey experiment in the United States, where a backslider has been recently removed from office. I specifically expect that former out-party voters prefer a more aggressive retaliatory candidate the more severe the backsliding perpetrated by the incumbent.

Ultimately, this work helps push the study of backsliding forward in two key ways. The first is by providing an additional factor that determine whether a voter chooses to retain a backsliding incumbent — societal threat — and how this is moderated through already established considerations of partisanship, polarization, and leader competence. The second is pushing forward the understanding of retaliatory backsliding actions and when the alternation of parties can provide a balm or an irritant to the democratic health of a country. By rigorously exploring the incentives of successors through formal theoretical work, and the preferences of voters through experimental work, I provide initial expectations for this frontier in the backsliding literature. In the rest of this chapter, I first detail the existing work on democratic backsliding. Then, I use observational data to detail the number of recent cases of backsliding and the prospects for democracy after those infractions. Lastly, I distinguish the backsliding actions I am specifically evaluating that will feature throughout the remainder of the dissertation.

1.2 Democratic Backsliding Background

The trend of the recent decline in democratic quality in countries throughout the world has gone by several names: democratic recession (Diamond 2015), democratic deconsolidation (Foa and Mounk 2017*b*), constitutional retrogression (Huq and Ginsburg 2018), constitutional rot (Balkin 2017), and the term used in this inquiry — democratic backsliding (Bermeo 2016). Regardless of the term used and the exact definition prescribed, the general picture is the same, that several constitutional democracies throughout the world have experienced recent movements away from democracy. With respect to the "waves" of democratization[1] discerned by Huntington (1991), this new movement concluded the post Cold-War third wave of democratization and has ushered in a new period of authoritarian reversion (Chu et al. 2020; Foa and Mounk 2017*a*; Haggard and Kaufman 2016; Lührmann and Lindberg 2019). This literature is relatively new and heavily theoretical, but builds on a historical understanding of movement to and from democracy. This newer phenomenon has some distinguishing characteristics from past democratic reversion, and a wealth of new empirical applications and measures that have pushed the study of democracy further than previously possible. In this section, I introduce the wider literature by discussing how democratic backsliding has been measured and theorized.

[1]Generally these are divided into three waves of democratization and three waves of reversion (two plus this most recent trend of backsliding). The first wave of democratization or the "long wave" began in the 1820's and died off as the first reversion began in the 1920s. This first reversion (1922-1945) peaked during World War II and after which lead to the second wave of democratization from 1945 to 1960. After, a second reversion occurred from 1960 to 1975. The most recent "third wave", which was the result of the decline of the Soviet Union and other democratic movements world wide began as the second reversion concluded. When Huntington wrote about this wave, he theorized that a third reversion may eventually come, and since then, scholars generally view the rise of democratic backsliding starting in 1998 in Venezuela up until today has marked this third reversion (Lührmann and Lindberg 2019). See section 1.3 for a chart mapping these three cycles using measures of democratic quality.

1.2.1 Measurement

The most recent wave of democratic backsliding began sometime between the late 1990s and mid 2000s. Ever since this time, on average, several indicators of democratic quality like Freedom House and the Varieties of Democracy Project (V-Dem) have shown, on average, a consistent degradation of democratic performance. Backsliding is defined as a degradation of the qualities of democratic governance through a series of incremental actions (Waldner and Lust 2018). This is distinct from the more prevalent military coups or other dramatic seizures of power that used to be much more common (Svolik 2018). The less dramatic, progressively deteriorating actions that are emblematic of democratic backsliding are more subtle, and sometimes are so gradual, that voters sometimes fail to recognize any wrongdoing (Bermeo 2016).

Democratic backsliding is a somewhat opaque concept, and consequently, it is difficult to measure these changes in a uniform way across time and space. Changes in measures of regime type and quality are typically used, but because backsliding is a gradual process, an effective measurement that can adequately capture changes is difficult to derive and even more difficult to be agreed upon. It is for this reason that a considerable amount of work on democratic backsliding is either theoretical, formal theoretical, or qualitative. Empirical work tends to focus on single-country studies or a selection of geographically close countries (like western Europe or Latin America). It is for this reason that uniform cross-national theories of democratization and democratic backsliding rarely receive universal support. That said, as measures improve, the ability to quantify the changes in democratic quality will open up the literature of democratic backsliding to new empirical applications.

Generally, a backsliding action or event requires degradation in two of three characteristics: competition, participation, and accountability (Waldner and Lust 2018). Several indicators

of regime quality are commonly used to measure these changes, but they vary considerably in their sensitivity to change. Freedom House (Repucci and Slipowitz 2022), Polity IV (Marshall, Gurr, and Jaggers 2017), and V-Dem (Coppedge et al. 2020) are three of the major cooperative projects that measure regime quality and the changes in these measures are commonly used to assess backsliding. Other measures from Skaaning, Gerring, and Bartusevičius (2015) and Wahman, Teorell, and Hadenius (2013), which aggregate Freedom House and Polity IV, have similarly measured these concepts. Ultimately, the lack of a single gold standard indicator makes using a variety of measures attractive to assess several factors that describe democratic backsliding as the best approach. In Chapter 2, I rely on three "high-level" indices from V-Dem each of which aggregates several fine-grained measures to assess institutional change that can capture democratic backsliding. I detail these measures further in section 1.3 where I demonstrate the prevalence of backsliding events in since 1990.

1.2.2 Precursors

As democratic backsliding is a difficult concept to effectively pin down, a considerable amount of work has been undertaken in recent years to better explain the phenomenon. While I focus on two specific instances within this literature — the choice of voters to remove a backsliding incumbent, and the preferences of out-party voters and successors to initial backsliders when they return to power — there is a wide array of other work that I do not contribute to, but is foundational. This dissertation primarily focuses on the "during" and "after" phases of democratic backsliding, but the precursors to backsliding are important in shaping the path of what comes during and after. This work does not heavily feature in the remainder of this project, but certain concepts are influential when considering a backslider's tenure and what occurs after their removal, and I note when this is the case. I divide the precursors into four

groups: leader-driven explanations, voter-driven explanations, institutional explanations, and other explanations.

With respect to the leader-driven work, this literature is focused on the political actors that are backsliding-minded office-seekers and when they can rise to power. Political actors are a significant driving force for whether backsliding measures are enacted or not. Studies of Latin America showed that the prevalence of moderate leaders has been a key factor in staving off backsliding (Mainwaring and Pérez-Liñán 2013*a*). After the election of a more extreme leader, like Brazilian President Jair Bolsanaro in 2018, this dynamic may be challenged because a backsliding-minded leader can take advantage of a "superpresidential" system (Lapper 2021). Further, the increase in the frequency of political outsiders entering office as (mostly right-wing) populists is important as backsliding is a leader-driven process. Ultimately, the presence of a mostly non-democratic, disloyal opposition that accedes to power is a significant factor in traditional democratic collapse (Linz and Stepan 1978). While these historical dramatic seizures via coups have given way to gradual backsliding actions, the rise of group hostile to democracy is crucial to both the beginning of backsliding and its persistence. Factors such as partisanship, polarization, and societal threat also help drive voter preferences (the demand-side) and enables an increase of the authoritarian-minded politicians who become backsliders.

While leader-driven explanations are attractive, authoritarian-minded politicians do not enter office on their own. The role of voters responding to changing societies, especially in response to factors that I classify as societal threat[2], has been connected with the rise of populist politicians, especially those from the far-right (Baccini and Sattler 2021; Ballard-Rosa et al. 2021; Bustikova 2014; Colantone and Stanig 2018; Inglehart and Norris 2017; Lucassen and Lubbers 2012). Populists tend to care less about constitutional norms that

[2]See Chapter 2 for more.

make democratic persistence possible (Issacharoff 2018) and when these preferences are mixed with cultural grievances like rising immigration (Aleinikoff 2018), as is commonly the case with right-wing populism, the movement to disenfranchise the voices of ethnic minorities rises (King and Smith 2018). Backsliding actions that work to disenfranchise are a key focus of this project, and I discuss this group of strategies further in section 1.4. Increased voter dissatisfaction with the party system is an important factor that helps populists gain power. Populists gained power due to rising anti-system sentiments from the right in Hungary and from the left in Venezuela (Scheppele 2018; Tushnet 2018).

A key area where the leader-driven and voter-driven explanations work in tandem is with the role of partisanship and polarization. These concepts are influential because when voters are most concerned with policy, they are willing to overlook democratic subversion when it is leveraged in service of their preferred outcomes (Balkin 2017; Nalepa, Vanberg, and Chiopris 2018; Graham and Svolik 2020; Grillo and Prato 2020; Braley et al. 2021; Svolik 2018, 2019). The presence and growth of anti-system parties or factions is a historical component of the literature (Linz and Stepan 1978), so the growth of "populist" parties, adversarial to democracy, is an important explanation for both the onset of democratic backsliding and the continuation in office of backsliders. Partisanship and polarization are also important factors in the persistence of democracy and will be featured in each chapter.

Theories of a country's vulnerability to democratic decline based on institutional alignments are a traditional focus in study of democratic backsliding. Most of this literature is based on the literature around horizontal and vertical accountability (O'Donnell 1998; Smulovitz and Peruzzotti 2000). Theoretical work has focused on preferences for some institution or coalition configurations: a focus on using grand-coalitions to forestall democratic decline by making the persistence of democracy a shared fate (Lijphart 1977), better performance of proportional representation which avoid malapportionment of representatives

over majoritarian representation (Reynolds 2010), and better performance of parliamentary systems which diffuse power from a single individual over presidentialism (Linz 1990; Maeda 2010). The endogeneity of these institutional configurations remains a significant concern in this literature though, because the same factors that determine the institutions in a country also impact democratic performance and quality (Cheibub et al. 2007; Mainwaring and Shugart 1997). For example, the proliferation of presidential systems in Latin America is also a direct result of factors that impact persistence of democracy, like the history of military dictatorships and executive hegemony (Pérez-Liñán, Schmidt, and Vairo 2019). Lastly, work on specific institutions and their role in staving off democratic backsliding, namely an independent judiciary, are also featured in the backsliding literature (Gibler and Randazzo 2011; Kelemen and Blauberger 2017). Ultimately, institutional explanations are some of the most hotly debated theories and this has driven the greater focus on formal theoretic work and single-country studies in recent years.

Finally, there are some other factors that are commonly cited in the literature, but do not fit cleanly into the categories above. Many of these explanations focus on economic factors and have mixed results (Bermeo 2009; Boix 2010; Haggard and Kaufman 2012, 2016; Sitaraman 2018; Slater, Smith, and Nair 2014). The rise of inequality has been cited as factor that contributes to the rise of populism, but the robustness of effect is up for debate (Waldner and Lust 2018). Other work has highlighted the work of crisis and how temporary powers become more permanent (Driesen 2021; Gross 2018). Crisis and its relationship to societal threat is central to the theory of Chapter 2 where voters' fear of crisis is a key feature in the persistence of backsliding. Finally, returning to the waves of democratization and reversion, Huntington (1991) outlined factors that preceded the past two global reversions to authoritarianism and some of these factors are relevant today. Weakness of democratic values amongst elites and general public, economic setbacks that intensified social conflict,

social and political polarization, and movements to exclude political groups were among the precursors that remain salient today.

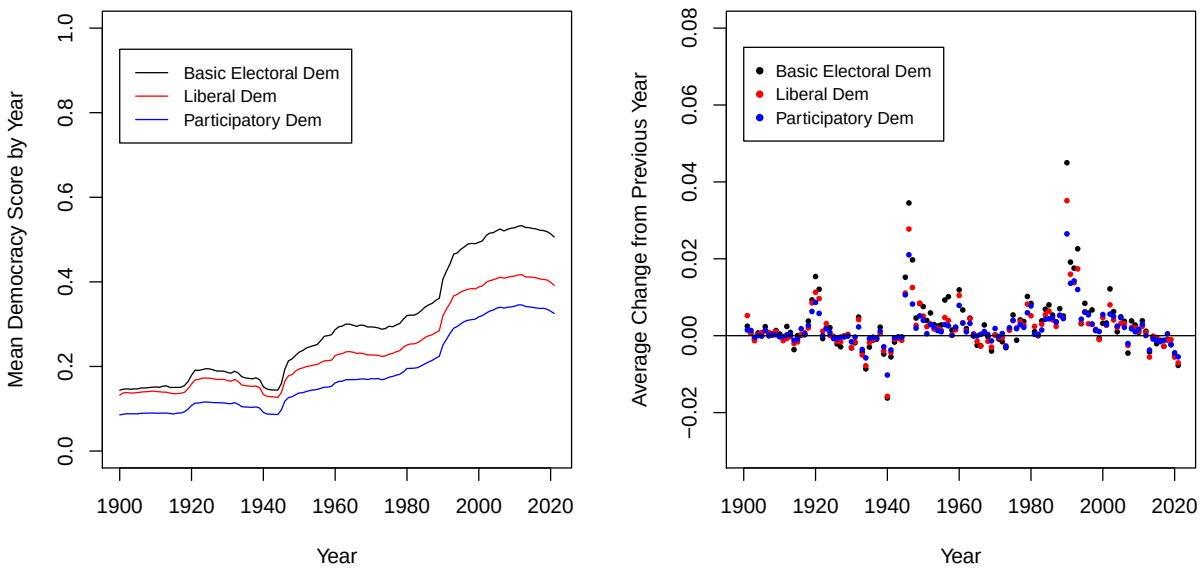
Ultimately, I am contributing to two select areas within the literature, described in more depth in the subsequent chapters: a voter's choice to retain or replace a backslider, and the preferences of former out-party voters and leaders after a backslider is removed. In each of these instances there is room for theoretical innovation and targeted inquiry to assess the key factors that affect backsliding and its persistence.

1.3 How Prevalent is Democratic Backsliding?

As discussed, precisely measuring democratic backsliding is a significant challenge, but I rely on measures from V-Dem to measure change at the year level (Coppedge et al. 2020). I use three different indicators of democratic quality. The first is a measure of basic electoral democracy in line with Dahl (1971) polyarchy to assess the extent of suffrage and whether elections are free and fair. The second is a rigorous measure of participatory democracy to serve as a high bar for voter access to elections. Last, is a measure for liberal democracy which helps assess the roles of constraints within and on government. These three measures help illustrate changes in quality of different institutions that are impacted when backsliding takes place. In table 2.2 of Chapter 2, I further elaborate on the measures I use for the empirical application.

To better display these measures in context of the waves of democratization and reversion (Huntington 1991), I show both the average score of each measure across all countries in the sample in each year since 1900 and also the average change from the previous year in figure 1.1. This figure helps illustrate the waves of democratization and the waves of democratic

Figure 1.1: Average Yearly Democracy Scores and Average Year-Change.



reversion including the most recent trend of democratic backsliding[3]. While the measures are heavily correlated on average, there is some useful variation at the year level that makes using all three attractive to assess change in democratic quality.

In order to explore recent democratic backsliding, I use this data to identify the number of cases of backsliding since 1990. This helps capture the most recent wave away from democracy (Lührmann and Lindberg 2019). I aggregate the country-year data from V-Dem to the country-government level to assess how democratic institutions changed during the tenure of an incumbent between elections. For the governments that met a baseline electoral democracy score greater than 0.5 (around the average score throughout this time period), I isolated those that had either a small or large decline in one of the three democracy indices from V-Dem. Of these observations, I distinguish between governments that were retained or

[3]The exact years of these waves is up for debate. Also, while considerable regional variation exists within each of these time periods, this overarching trend is helpful for detailing how democratic quality has changed over time. For a more fine-grained treatment of democratic waves, see Gunitsky (2018).

replaced at the end of the electoral term. Electoral retention and replacement of backsliders is a key focus of each of the three substantive chapters and the prospect of democratic improvement after replacement are of significant focus in Chapters 3 and 4.

If a backslider is replaced by a successor, there are two options that I propose the successor can choose between: *retaliation* or *restoration*. Under retaliation, the successor uses weakened institutions to pass preferred policies and maintain power with greater certainty. Alternatively, the new government can prioritize restoration by repairing institutional constraints and norms, but at the expense of fully enacting policy goals or retaining office with higher certainty. I use the observational data to identify cases either when a backslider was retained and either reformed or continued backsliding, or a backslider was replaced and the successor either restored institutions or retaliated. I divided these changes into one of four categories: all indices were worse after the election, the indices were mixed (some better, some worse), all three improved, or the observation is the most recently elected government in the sample for that country. In table 1.1, I show the number of cases by each category and the percent of the cases per column.

This demonstrates that when significant democratic backsliding is enacted, the retention of the incumbent is much higher than when any small democratic deviation is recorded. This gives some evidence that those who engage in backsliding are more likely to retain office, in line with expectations that some backsliding actions allow backsliders to retain power with greater certainty. Also, the prospects for democracy seem slightly better under a successor government than when the incumbent retains office. A voter's choice to retain or replace is the focus of Chapter 2. The choice of the successor government to retaliate or restore institutions is the focus on Chapter 3, and the preferences of out-party voters for their successors is the focus of Chapter 4.

Table 1.1: Prospects for Democracy After a Backslider Given Retention or Replacement.

	Any decline in an indicator (426 Cases)		Severe decline in an indicator (78 Cases)	
	Incumbent Retained 248 Cases (58.2% of Total Cases)	Incumbent Replaced 178 Cases (41.8%)	Incumbent Retained 53 Cases (67.9%)	Incumbent Replaced 25 Cases (32.1%)
All Worse After	95 (38.3% of Column)	56 (31.5%)	27 (51.0%)	11 (44.0%)
Mixed Scores After	52 (21.0%)	43 (24.2%)	3 (5.7%)	1 (4.0%)
All Better After	56 (22.6%)	55 (30.9%)	14 (26.4%)	7 (28.0%)
Most Recent Obs.	45 (18.1%)	24 (13.5%)	9 (17.0%)	6 (24.0%)

Note: Country-Government cases since 1990 (i.e., "Poland 2015-2020" as a single observation). I subset to those with a Polyarchy score greater than 0.5 so each case was marginally considered democracies at the beginning of the term in office. I identified both minor cases (any drop in one of the three indices) and major cases (a drop of 0.05 or greater in one of the three indices). These are not mutually exclusive groups. Cells show the number of country-government cases in that category and the percentage of that category per column.

When looking at the decisions by successors to backsliders, the choice to restore institutions has some mixed examples. The governments preceding Bolsanaro's rise in Brazil in 2018 and Orban's rise in Hungary in 2010 showed considerable democratic decline, and the choice of these known backsliders continued that pattern. Meanwhile, after backsliding actions in the United States from 2016 to 2020, the successor government improved democratic institutions. Many of the other instances of restoration after backsliding were in countries that experienced dramatic democratic collapse, such as Estonia after 1992, Thailand after 2007, and Mali after 2013.

Among those that retained power, the further decline of democratic institutions occurred in commonly cited cases like Venezuela from 1999 to 2005, Turkey from 2007 to 2015, and Poland from 2015 to 2020. Cases where democracy improved after retention include Ecuador after 2013, the Dominican Republic after 2016, and Romania after 2019. Retained incumbents that continued to attack institutions over multiple terms in office receive the most attention in the literature, but single term blips that are corrected are important as well. While I focus on successors, understanding when backsliders restore institutions in a future term is another

interesting avenue of research. As measurement of these concepts improves, it will be easier to assess these changes both within country and cross-sectionally.

With these observations guiding the understanding of the number of backsliding events since 1990, we can see that backsliding is a significant phenomenon both on average and with respect to the number of cases. In the next section, I detail two specific kinds of backsliding actions that I focus on throughout this project and provide examples of these backsliding actions.

1.4 Varieties of Democratic Backsliding Actions

There are many different kinds of democratic backsliding, but for the purpose of this project I focus on gradual backsliding strategies that attack democratic institutions. These gradual actions help erode institutions in favor of an incumbent, but they are not the only kind of actions scholars have identified as antagonistic to democracy. There is considerable literature of the influx of religiosity in politics, especially in settings where secularism had been a defining characteristic, such as Turkey (Varol 2014) and India (Mate 2018). Other actions focus on limiting basic rights such as freedoms of speech and press, such as the case of Turkish President Erdoğan's strategy of suing and resultantly silencing journalists who criticize him (Varol 2014). While these manifestations of backsliding are important, my focus on institutional attacks is intended to highlight a set of actions with clear goals and results either in terms of policy making or electoral performance.

There are two gradual kinds of backsliding that incumbents use: *executive aggrandizement* and the *strategic manipulation of elections* (Bermeo 2009). The two approaches serve authoritarian-minded incumbents jointly: first to make their preferred policies easier to enact by eliminating challenges to these policies within government, and second to maintain power

from election-to-election by attacking opposition candidates or limiting the voting rights of supporters of the opposition. When democratic institutions are damaged, it can lead to lower levels of freedom, higher inequality, worse government responsiveness, and lower governmental accountability (Diamond and Morlino 2004). These strategies are rarely assessed in tandem, though they are commonly used together by backsliding governments. I detail both sets of strategies and state some expectations about why they should work differently in practice, both with respect to values around democratic institutions and how their actual mechanical differences are distinct.

In some existing models, backsliding events are modeled as increasing the likelihood that voters will punish the incumbent because of executive overreach, but do not directly affect electoral institutions (Graham and Svolik 2020; Grillo and Prato 2020; Nalepa, Vanberg, and Chiopris 2018; Svolik 2018). I suggest that these models are addressing *executive aggrandizement*, which Bermeo (2016) defines as, "when elected executives weaken checks on executive power one by one, undertaking a series of institutional changes that hamper the power of opposition forces to challenge executive preference." Executive aggrandizement is a tool for incumbents to take over or to make existing institutions more favorable to their goals. These backsliding measures are likely to concern voters who strongly value certain democratic norms (e.g., role of the opposition in government, independence of the judiciary, freedom of the press, etc.), but supporters may not be as concerned with these actions.

Incumbents implement executive aggrandizement when their ability to govern is constrained by other government actors (e.g., other branches of government, or the civil service) and they want to act in defiance of these constraints. Attacks against institutions that can limit the executive, such as the judiciary, are common examples of such backsliding (Gibler and Randazzo 2011; Howell, Shepsle, and Wolton 2019). For example, in Poland, the ruling PiS saw the Constitutional Tribunal as a major obstacle to implementing their preferred

policies. The attack on the Constitutional Tribunal was accomplished by proactively packing the court, appointing judges to fill existing seats and pushing existing adversarial judges out. After the court was successfully captured, the PiS was able to leverage the now sympathetic court in order to partake in further backsliding actions that could no longer be constrained (Sadurski 2018).

Other work has modeled backsliding decisions as making re-election more certain for the incumbent (Luo and Przeworski 2019). Bermeo (2016) calls these backsliding activities the *strategic manipulation of elections* which:

> "denote a range of actions aimed at tilting the electoral playing field in favor of incumbents... including hampering media access, using government funds for incumbent campaigns, keeping opposition candidates off the ballot, hampering voter registration, packing electoral commissions, changing electoral rules to favor incumbents, and harassing opponents — but all done in such a way that the elections themselves do not appear fraudulent."

Distinct from executive aggrandizement, the strategic manipulation of elections alters institutions to reduce the ability of citizens to remove an incumbent from power. This should be an action that is more unnerving to voters because it directly constrains the most important role they have in a democracy: voting for and against politicians. Incumbents likely use these strategies when they are uncertain about either the likelihood of re-election or elections relevant to their cause. However, these actions can serve to activate voters against them, especially if voters are invested in democratic norms such as the freeness and fairness of elections. We should expect voters to have stronger negative reactions to strategic manipulation of elections than executive aggrandizement for this reason.

Incumbents strategically manipulate elections by attacking infrastructure and candidate entry in elections. These actions help create an uneven playing field to limit the chances or choices of the opposition while also refraining from engaging in outright fraud (Levitsky and Way 2010). For example, the 1999 Constituent Assembly elections in Venezuela featured new electoral institutions as a result of Chávez's new constitution approved via referendum earlier that year after failing to pass in the legislature. In the election, Chávez's party won 121 of the 131 seats with just under 66% of the vote (Landau 2018). This was made possible by aggressively drawing districts to limit the voice of supporters of the opposition and prevent equitable representation. Also consider the treatment of the Turkish political party, the HDP, which primarily serves as a party for the Kurdish ethnic minority. The HDP has had several leaders and members of the Grand National Assembly jailed on various charges including its leader Selahattin Demirtaş after increasingly better performances against Erdoğan's AKP. The Turkish government spent the last few years trying to ban the party altogether (BBC 2021). These examples illustrate how the strategic manipulation of elections tilts the electoral playing field and entrenches incumbents. I summarize the specific tools for both kinds of backsliding in table 1.2 (Bermeo 2016; Schedler 2010).

While these two groups of strategies — executive aggrandizement and the strategic manipulation of elections — often occur in tandem, they are distinct and serve different goals: the first is to make governing easier so that incumbents can enact their agenda with reduced difficulty, and the second is to ensure that they and their allies remain in office to continue this practice. In order to better explore and explain voter support of backsliders, my empirical models in Chapters 2 and 3 will consider both kinds of backsliding to first explore the role of societal threat and how it affects the support of voters by justifying undesirable backsliding actions, and then explore the choice of successors to initial backsliders. As stated, we should not expect voters to view these two strategies with the same severity, and because

Table 1.2: Two Kinds of Backsliding and Their Respective Tools (Bermeo 2016; Schedler 2010)

Backsliding Type	Tool	Definition	Example
Executive Aggrandizement (weakening checks from within government)	Disempowerment	formally removing powers, constraining exercise of powers	Changing scope of the judiciary, seizing legislating powers
	Agent Control	installing favorable actors, create incentives for actors to not challenge executive	Inserting loyalists on courts, judiciary, etc. structure incentives to maintain loyalty
	Fragmentation	divide an institution's responsibilities, keep institutions from acting with a unified voice	Split a court into parts, splinter the role of the legislature
	Insulation	limit an institution's access to external networks (esp courts)	Remove legal network from courts
Strategic Manipulation of Elections (weakening checks outside government)	Market Restrictions	opposition parties fragmented or excluded, specific citizens excluded	Ban parties, disenfranchisement or suppression (targeted)
	Preference Distortions	voters prevented from expressing true preferences	Voter Intimidation, vote buying
	Vote Distortions	targeting laws before elections that limit opposition parties	Change rules for ballots that disadvantage a specific group

their outcomes are different, a theoretical model should take that into consideration. In Chapter 2, I allow voters to have two sets of democratic preferences for each set of backsliding strategies for that reason.

1.5 Looking Forward

Democratic backsliding is an increasingly important area of study in political science, both because of the rising prevalence of these events and the normative impact of weakened democracy. This dissertation project contributes to the literature of democratic backsliding in two key ways.

The first is to provide a greater understanding of the choice that voters make to retain or replace their backsliding incumbent (Chapter 2). I propose that societal threat is an additional important component that combines with past explanations of partisanship and polarization to show why backsliding incumbents are so commonly retained. When voters feel

threatened they are more likely to support politicians who they deem as competent, which is a characteristic that many backsliders project through their forceful governing style.

For backsliding to yield to democratic resurgence, the removal of the authoritarian-minded officials is likely *necessary*, but whether this removal is *sufficient* is where the remainder of the dissertation comes into focus. Accordingly, the dissertation's second main contribution is to assess the prospects for democracy *after* democratic backsliding (Chapter 3 for successor governments and Chapter 4 for voters). Democratization has historically been cyclical, and after a long period of democratization, a trend toward reversion soon followed. This current trend of democratic backsliding may be the response to the "third wave" of democratization, but what comes after? This current moment of democratic backsliding is distinct from past democratic reversions because of the lack of major seizures of power through a military coup, popular uprising, or an autogolpe. Modern democratic backsliding involved the gradual capture of institutions that both constrain an executive within government and outside of government by damaging checks on the executive and electoral institutions respectively. The incentives of successors to repair institutions or instead to retaliate against initial backsliders may say a lot about how long the current backsliding moment will last. Simultaneously, the preferences of former out-party voters for their new leaders also affect this path.

In Chapter 2, I find that societal threat is related to new democratic backsliding actions through an original formal model and an empirical test of one of the implications of the model. In the model, I show the relationships between higher societal threat and a voter's willingness to retain an incumbent. I derive expectations for the kinds of backsliding we should observe more when societal threat is high, and the parameters that impact when retention should be expected or not. As backsliding incumbents can project societal threat, I also expect that competent incumbents will manipulate societal threat to engage in more democratic backsliding. In the empirical application, I show a relationship between changes

in factors that help measure societal threat and new backsliding actions. I specifically show a relationship between social stability and external threats with new backsliding actions.

In Chapter 3, I use an original formal model to show that when restorative actions of successors are not durable, or a future backslider can easily undo them, we should never expect democratic institutions to be restored. The extent to which institutions are subverted also impacts these preferences where as backsliding measures are more severe, restoration is more attractive. Electoral safety of a successor in the absence of subverted institutions increases incentives to restore depending on whether executive aggrandizement or the strategic manipulation of elections is considered. Higher electoral safety makes restoring executive constraints less likely, but makes restoring free and fair elections more likely.

Finally, in Chapter 4, I report an original survey experiment fielded in the United States, showing that out-party voters prefer their leaders to retaliate *more* as the backsliding incumbent engages in more extreme backsliding. This finding conflicts on the surface with the implication from Chapter 3, that extreme backsliding should lead to greater restoration preferences. This conflict between short-term and long-term incentives is crucial for the prospects of democratic resurgence.

Ultimately, this work serves to move the backsliding literature forward toward further consideration of the mechanisms through which democracy is repaired or destroyed. The prospects for democracy is an area that will continue to be important in the coming decades, and understanding the sometimes competing incentives that drive preferences both for leaders and voters is crucial.

Chapter 2

Threat Induced Voter Support For Democratic Backsliding

2.1 Introduction

Why do voters support backsliding incumbents? Under what conditions are voters more permissive of backsliding and can they be manipulated by incumbents? Backsliding events have become more common in the past couple of decades (Lührmann and Lindberg 2019) and generally the literature has focused on within-government factors such as the desire to more quickly pass policy or societal factors like polarization. Voters also play a major role because they can use their vote to punish incumbents that engage in backsliding or to deter backsliding actions. Despite these tools, backsliding incumbents frequently maintain power, so the study of how voter sanctioning of a backslider breaks down is important as an additional explanation of how consolidated democracies backslide.

Current explanations point to the role of polarized societies and partisanship to explain why voters support incumbents that subvert democracy (Beaulieu 2014; Graham and Svolik 2020; Svolik 2018). Work has also shown that when voters are most concerned with competence, competent incumbents can consolidate their power quickly and make removal difficult (Luo and Przeworski 2019). These factors can be complicated when the intention of the incumbent is uncertain (Helmke, Kroeger, and Paine 2019), and when backsliding actions are incentivized (Grillo and Prato 2020; Nalepa, Vanberg, and Chiopris 2018). These explanations are helpful, but miss a key area of focus of the campaigns of backsliders that interacts with polarization, partisanship, and competence: societal threat.

I argue that voters can be especially motivated by societal threat — a perceived precursor to a crisis in which there is an expected loss of status quo benefits — and alleviating the potential consequences of a crisis. New, salient threats in the lead up to elections can be a boon for the chances of backsliding incumbents, because the fear of crisis leads voters to overlook subversion of democracy in favor of stability in the face of uncertainty. This can happen even when voters generally support democratic norms and institutions. Societal threat salience is not exogenous and can be mediated through the messages released by the government, which can augment the salience of existing threats. This manipulation of societal threat level can be used by backsliding incumbents to meet their goals, especially when they believe voters may be swayed to support them with higher levels of societal threat. Electoral campaigns are a key area where voters can punish the incumbent, therefore the focus on the windows when voters can remove backsliding incumbents is necessary.

I construct a formal model that explores the voters' choice to support backsliding incumbents, specifically focusing on two kinds of backsliding: *executive aggrandizement* or the weakening of constraints on policy making for the executive and the *strategic manipulation of elections* or actions targeted at tilting the electoral playing field in an incumbent's favor. I

explore how these two different kinds of backsliding actions are made permissible to voters via societal threat, along with other factors including: the level of democratic support of voters, and the competence gap between the incumbent and challenger. I also consider when incumbents can manipulate societal threat to benefit themselves.

I find that higher societal threat helps allow incumbents to engage in more democratic backsliding. This is because societal threat justifies backsliding by making competent incumbents more desirable. This gives competent incumbents the additional ability to engage in backsliding. In the model, incumbents only engage in the strategic manipulation of elections when they are not as competent as the challenger, so societal threat only benefits competent incumbents and allows them to engage in further executive aggrandizement. Lastly, these incentives make it so that when competent incumbents can manipulate and increase threat (e.g., in Turkey, Erdoğan's focus on the G$\acute{}$ulen movement), they will do so in order to engage in more backsliding. This gives some background as to why re-election campaigns for backsliding incumbents are frequently focused on societal threat.

While some factors in the formal model are difficult to assess empirically, I focus my empirical application on the relationship between societal threat allowing more democratic backsliding. I use observational data from the Varieties of Democracy Project (V-Dem) (Coppedge et al. 2020) and the country-risk firm Fitch Connect to establish a relationship between newer, salient societal threats and backsliding actions. By using a measure like those from Fitch, I am also able to estimate *real* societal threat instead of threat manipulated by a backsliding incumbent. I show in this empirical application that societal threats concerning social stability and security with respect to external threats are significantly related to the introduction of additional backsliding measures. This finding, paired with more nuanced findings from the formal model, allow for a greater understanding of the importance of societal threat in both enabling backsliders and improving their chances of retaining office.

This work is important because it explores why backsliders are routinely re-elected, and how these conditions lead to further democratic decline. The level of societal threat plays a significant role in differential level of support for politicians. Increased focus on the conditions within backsliding countries are important because the longer backsliders have control, the more likely full democratic failure is to occur. Understanding the conditions when backsliding incumbents are vulnerable to removal is crucial to identify when democratic resurgence may be possible in the future.

2.2 Background

Misbehaving incumbents are ideally punished, but incumbents who engage in democratic backsliding are commonly supported for re-election. This support is puzzling because it challenges what we assume about democratic accountability. Further, even when a citizenry generally supports democratic institutions that safeguard against overreach by bad actors, these actions can be condoned by the re-election of a backsliding incumbent.

Much of the literature around democratic backsliding involves this breakdown of the sanctioning process. I summarize past answers and discuss the relevant factors that are central to my theory. Specifically, I focus on voters' preferences for democratic norms and institutions, and formal work that posits partisanship, polarization, uncertainty, and incumbent competence as explanations for the support of democratic backsliders. Lastly, I detail *societal threat* and how it has been used in other contexts to explain support for authoritarianism and its relevance to democratic backsliding.

2.2.1 Past Explanations

There is debate whether voters' opinions of democratic norms and institutions are sufficiently high for the persistence of democracy. More simply, whether democracy is only a "good" thing when a voter's preferred party is in power. While some have theorized that support for democracy has fallen and this explains backsliding (Foa and Mounk 2016, 2017*a*,*b*), others question the decline of democratic values and believe that though generational changes have made preferences for democracy change over time, there should still be a sufficient level of support for democracy to persist (Alexander and Welzel 2017; Norris 2017; Voeten 2016). A complementary explanation is that voters simply have majoritarian preferences and that actions of a rightly elected incumbent are by nature democratic (Grossman et al. 2022). Whether democratic values are robust or not, the declining preference for democratic norms and institutions was a feature of backsliding movements in the twentieth century (Huntington 1991). For those in power, there are two key values that erode in backsliding regimes: the *mutual toleration* of opposing parties who see competitors as legitimate opponents and *forbearance*, where minority rule is valued by the majority in expectation of being a future minority (Levitsky and Ziblatt 2018). These same values can permeate to voters and drive selective preferences for democracy.

Even when preferences for democratic institutions and norms are high, backsliding actions can manifest. Actions that approach backsliding, but do not fully cross the threshold, are sometimes incentivized in democratic systems in the form of "constitutional hardball", where pushing constitutional limits within government is done in order to get more favorable outcomes while in or out of power (Fishkin and Pozen 2018; Helmke, Kroeger, and Paine 2019; Pozen 2018; Shugerman 2019). When this behavior is incentivized, authoritarian-minded

politicians take advantage of the benefits of being able to engage in backsliding-adjacent actions (Grillo and Prato 2020).

For example, in the United States, Senators engaged in "hardball" behaviors with respect to judicial appointments from 2010 to 2020. Early in President Obama's administration after losing a filibuster-proof majority in the Senate, Republicans began to increasingly filibuster judicial nominees. In 2013, then Senate Democratic Majority Leader Harry Reid ushered in the "nuclear option" to change rules and remove the filibuster for appointments other than for those to the Supreme Court. Upon gaining a majority in the Senate after the 2014 mid-term elections, Republicans increasingly rejected President Obama's judicial appointees culminating in refusing to vote on Merrick Garland, a replacement for the Supreme Court in 2016. After President Trump's election in 2016 and the extension of the "nuclear option" to Supreme court nominees, Republicans were able to successful pack both the Supreme Court and other Federal Courts playing by much different rules than a decade before. These actions paid off for Republicans because the Supreme Court has since produced favorable rulings on policy priorities such as gun rights and abortion. A notable feature of constitutional hardball is that none of these actions were illegal, and most were the result of rule changes or norm-breaking behavior to overcome or introduce gridlock. Gridlock in government is a key component that incentivizes constitutional hardball even among democratically-minded politicians, but it is the authoritarian-minded counterparts that use these initial measures to springboard into more explicit backsliding actions (Grillo and Prato 2020). This is important because it shows that system incentives can overcome values for democratic institutions meant to safeguard against backsliding. This pairs well with the existing literature around "constitutional qualms" (Christenson and Kriner 2017*a*) where voters may support constitutional hardball actions for the sake of overcoming obstacles and enacting or maintaining preferred policies.

Beyond the distinction between hardball and backsliding, multiple explanations of voter support of backsliders focus on political polarization and partisanship to overcome democratic preferences. Generally, voters would support a challenger to a backsliding incumbent, but only if the challenger's policies are comparatively attractive. If the challenger is not sufficiently moderate, then the voter is willing to overlook the incumbent's backsliding because of the incumbent's more attractive ideological position compared to the challenger (Graham and Svolik 2020; Svolik 2018). This similarly holds for voters with extreme preferences. Policy extremism is a common feature of backsliding regimes (Mainwaring and Pérez-Liñán 2013*b*), but under explanations that focus under partisanship, this would require moderation that is not common. Because backsliders are generally ideologically extreme and must also appeal to a sufficiently large portion of the electorate to maintain power, this has highlighted the role of polarization to explain where backsliding proliferates (Svolik 2018).

Polarization's impact on citizen behavior can also be complicated by uncertainty of the incumbent's intentions (Nalepa, Vanberg, and Chiopris 2018). When the true intent of the incumbent — preference for or against future backsliding — is unknown, it creates uncertainty amongst voters. Polarization makes removing backsliding incumbents difficult because supporters of the incumbent may be less willing to sanction borderline or minor backsliding behavior because the alternative may be ideologically much less appealing. When borderline behaviors go unrecognized, they may escalate into future backsliding actions. Backsliding is a gradual process and this plays a key role because the desire to remove a misbehaving incumbent may not be high enough to counteract polarization. That said, many backsliders are more forthcoming about their intentions to change the system and thus their motives may not require much divining beyond listening to the incumbent.

Distinct, but complementary to partisanship and polarization, other explanations have focused on the role of competence and how especially competent incumbents use their

popularity to consolidate power (e.g., Chávez in Venezuela and Erdoğan in Turkey). In this explanation, voters are more likely to support an especially competent incumbent and will only tolerate backsliding behavior up to a certain point. The location of this point is determined by the preferences of voters (e.g., technocratic, institutionalist, or strict policy preferences). However, after that point, it becomes difficult for voters to remove the incumbent because of the backsliding advantage that was already enacted (Luo and Przeworski 2019)). Generally, models do not consider when backsliding can also impact the ability to remove incumbents. Therefore, this is an important factor to consider as it is a main goal of one of the two main kinds of backsliding (the strategic manipulation of elections). In my theory, I consider the role of competence more as it pertains to addressing crisis.

To summarize, existing work varies in the key causes that lead to backsliding, but generally agree that a voter's choice to remove a backsliding incumbent is often complex and uncertain. As this complexity and uncertainty increases, it becomes more likely backsliding goes unpunished, even when a voter supports democracy. This means that even though we expect voters to observe and punish obvious instances of backsliding, they often do not. From the literature, the factors that are worthy of focus are democratic preferences of voters with respect to norms and institutions, the difference between the competence or ideological position of the incumbent and challenger, and the different goals and outcomes of different kinds backsliding actions. In addition to these, I propose that societal threat plays a key role in the breakdown of sanctioning of backsliders: justifying backsliding actions by inducing support for a competent incumbent.

2.2.2 Societal Threat and Authoritarianism

A societal threat is any factor that is perceived as a *precursor* to a crisis in which there is an expected loss of status quo benefits. These crises can take several forms: terrorist

attacks, violent crime, losing a job or savings, loss of social status, natural disasters, and so on. Societal threat is a concept from the political psychology literature and has been connected to support of right-wing authoritarianism and social dominance preferences (Doty, Peterson, and Winter 1991; Duckitt and Fisher 2003; Lavine, Lodge, and Freitas 2005). I extend the use of the concept as a factor to induce support for backsliding.

Any given country has some level of societal threat, but backsliding regimes commonly have elevated levels of societal threat that affects politics. Societal threats may not be *actual* precursors to crisis, but they must be seen that way for a voter. Alternatively, they may be presented that way to a voter. Consider that in the United States, the areas with the lowest proportion of immigrants tend to have the highest opposition to immigration (Luca et al. 2022). The central motivator of societal threat is that voters have an intangible fear of a crisis in which something will happen that will make them worse off. When voters are focused on societal threat, negative outcomes can occur such as lower trust in government, lower engagement in politics, and support for less democratic actors. Table 2.1 provides examples of societal threats and crises that may arise as a result of these threats.

Table 2.1: Examples of Societal Threats and Potential Crises.

Societal Threat Examples	Perceived Potential Crises
Rising Immigration	Personal economic hardship, impending loss of status, crime
Climate Change	Personal experience with natural disasters, loss of housing or resources
Rising Crime	Being a victim of violent crime, living in an unsafe neighborhood
Economic Stagnation	Loss of a job, housing, or savings
COVID-19	Getting sick or dying, a loved one getting sick or dying
Rising Violence/Terrorism	Personal impact of mass-shooting or terrorist attack

Note: Societal threats do not need to be actually related to the potential perceived crisis. In some of the cases, the reverse relationship is true (i.e., in the U.S. immigrants are less likely to commit crimes). Societal threat operates at the psychological level, so the perception of a potential crisis is what is needed, not a direct empirical connection.

The literature tying societal threat and authoritarianism (or the authoritarian personality) is well established in political psychology and has focused on how societal threats can activate

latent authoritarian preferences in order to address potential crises (Adorno et al. 1950; Sales 1973). The presence of high societal threat has been tied with several authoritarian outcomes such as: support for the death penalty in sentencing (McCann 2008), social dominance (Duckitt and Fisher 2003), hate crimes (Doty, Peterson, and Winter 1991), and more support for hawkish military involvement and less support for civil liberties during the "War on Terror" (Hetherington and Suhay 2011). While threat is generally thought to work specifically on those high on the right-wing authoritarian (RWA) personality trait (Lavine, Lodge, and Freitas 2005), people low on this scale have also been shown to be susceptible to societal threat (Hetherington and Suhay 2011), so it is an important consideration of all voters in how it impacts support for backsliders.

When considering backsliding more directly, societal threat has been studied under the context of support for far-right populists, or the radical right which has been a group specifically focused on backsliding in recent years (i.e. Hungary, Poland, and the United States). Specifically, the salience of identity issues and the success of minorities in society (minorities' growing success is perceived as taking away from the majority) has been shown to be a significant motivator of far-right party success (Bustikova 2014; Lucassen and Lubbers 2012). The idea of cultural backlash contributing to insecurities and status threat has also been tied to the rise of right-wing authoritarian-minded leaders who are thought to be especially susceptible to backsliding (Inglehart and Norris 2017; Mutz 2018). The immigration of ethnic minorities is a common cultural threat used by right-wing backsliders for these reasons. In addition to culturally defined threats, economic issues in the form of import shocks from rising globalization have been tied to the growth of support for far-right and populist parties (Baccini and Sattler 2021; Ballard-Rosa et al. 2021; Colantone and Stanig 2018). While not all backsliding regimes are marked by a far-right or generically extreme ideology, the role of societal threat in the rise of eventual backsliders has been well documented.

Lastly, under the threat of terrorism, voters have been shown to overlook the reputation of right-wing incumbents while punishing left-wing incumbents (Di Lonardo 2019). Consider that violence from the PKK (a Kurdish separatist group) in Turkey has commonly been used my Erdoğan to justify the crackdown on the HDP (the primarily Kurdish political party) and to disenfranchise Kurdish voters with support of nationalist Turks (BBC 2015). It is unsurprising then, that the same role that societal threat played in raising the position of extreme voices would also be instrumental in how support for backsliders develops with time.

Outside of support for the radical right, which is not the sole perpetrator of backsliding (e.g., Chavéz in Venezuela), research has also shown that societal threat and crisis can precede democratic erosion. Terrorism has been shown to endanger democracy in the target country (Merolla and Zechmeister 2009), widespread discrimination has been tied to democratic instability (Goldstone et al. 2010), and economic inequality can also be a driver of eroding democratic quality (Bermeo 2009; Slater, Smith, and Nair 2014). Societal threat is an important component that can determine vote choice by targeting the security of voters.

I theorize that the role of societal threat with respect to democratic backsliding helps offset the "sting" of backsliding by preferring a more competent leader. More specifically, societal threat offsets negative opinions of backsliding actions because it induces support for a competent incumbent that can alleviate the negative impacts of a potential crisis. This is especially the case when a voter sees a challenger as especially incompetent or even a threat itself. Because societal threat is a mostly psychological concept, determined at a community or country level and interpreted at the individual level, it is an especially useful tool for backsliders to use to their advantage.

2.3 Societal Threat's Impact on Backsliding

In regimes where incumbents backslide, elections and re-election campaigns can be particularly contentious. Voters have the opportunity to sanction and remove the incumbent, but this choice can be made more difficult depending on the conditions within the country. Despite the ability to remove backsliding incumbents, voters commonly vote to affirm undemocratic leaders. A key reason for this breakdown is the role that societal threat plays in justifying backsliding actions by inducing support for a competent leader. This dynamic produces additional incentives for an incumbent to highlight the level of societal threat within the country and can lead to a tumultuous environment when the choice of voters to sanction backsliders is crucial for the persistence of democratic institutions.

Backsliding has been described as a reactionary process to elements within or outside of government (Dresden and Howard 2016) and past research on societal threat has shown that voters are also reactive when societal threat is high (Hetherington and Suhay 2011; Sales 1973). I propose that higher societal threat levels play a key role in voter support of democratic backsliders. When societal threat is high and a strong leader is needed, voters shift from sanctioning the incumbent's undemocratic performance to selecting a competent leader who can deal with a potential crisis. This general process has been documented in the literature of democratic accountability for other misbehaving incumbents (Fearon 1999; Przeworski et al. 1999; Smulovitz and Peruzzotti 2000) and especially relevant with respect to backsliding. All voters have some general support level of democratic institutions and norms instilled through the "civic culture" (Almond and Verba 1963), but when voters feel threatened they are willing to bend some from their values, especially when it means choosing a competent leader. I adopt the concept of incumbent *competence* (Luo and Przeworski 2019) and extrapolate it to dealing with potential crisis.

A competent incumbent is able to effectively act during a future crisis in a voter-approved manner. Competence subsumes some other factors such as partisanship, but is distinct because it also relies on experience and ability to minimize the impacts of a future crisis. For example, if rising crime is seen as a salient societal threat, a competent leader for one voter may heavily focus on funding police in a "tough-on-crime" approach, whereas another may prefer a focus on transformative justice and reforming the justice system to fix "root causes." Either of these approaches may convey competence to a voter, but is also interpreted in terms of other traditionally important factors like partisanship. The focus on competence in this theoretical approach is especially useful, because backsliders are usually seen as strong leaders due to their forceful governing approach, so they can project competence. In the model, I assume both the societal threat level and the competence of both the incumbent and challenger are exogenous, but both are subject to manipulation and other endogeneity. While I consider the manipulation of societal threat later, I do not do the same for competence, but this is an important future direction. Ultimately, societal threat can help incumbents get away with backsliding, and this has implications for how societal threat plays into voters' preferences.

2.3.1 Additional Considerations of Societal Threat

When incumbents are competent and they expect minor backsliding actions will be punished, they are incentivized to increase the salience of societal threats to justify their past actions and take advantage of voter insecurity. By being able to make the decision murky and manipulate voters' ability to judge their record, incumbents can reduce the potential negative impacts of democratic accountability via their removal from office (Maravall 1999). For supporters of leaders that border demagoguery, a focus on the fears of voters can help skirt accountability (Mercieca 2020). In the absence of high societal threat, an incumbent may choose to manufacture or augment societal threats and potential crises to meet their needs.

In this case, a backsliding incumbent may latch onto something minor so he can raise the societal threat salience around an issue or invent a threat all together (i.e. the Trump administration's focus on migrant caravans in the lead up to the 2018 and 2020 elections). The threat, augmented or artificial, usually highlights an issue area where incumbents or their parties have some issue ownership, in order to draw distinction with the opposition and signal competence and the ability to deal with crisis. In some ways, this is counter-intuitive, because incumbents should be highlighting their own successes, and focusing on the unstable nature of the country may not seem to achieve this. However, in the setting of an electoral campaign with an untested and/or more democratic challenger, this strategy can be effective and take advantage of the insecurity of voters.

Despite the role that societal threat can play in a voter's decisions to support backsliding incumbents, the presence of societal threat does not necessarily mean that it will induce incumbent support. Voters who do not view the incumbent as particularly competent would not be expected to react to threat by supporting the incumbent. In fact, it can have the opposite effect of increasing the support for the challenger. Societal threat is ultimately a psychological concept and its impact happens at the individual level, so the threat of certain crises may move voters very little.

Lastly, we should expect voters to respond to threat differently depending on the kind of backsliding measures employed. I distinguish between two key gradual kinds of backsliding that incumbents use: weakening other political institutions that can challenge the executive or *executive aggrandizement* and weakening the ability of the political opposition to win elections or the *strategic manipulation of elections*. While executive aggrandizement is likely to be a negative act for voters concerned with democratic institutions, this group of voters may not be large enough to prevent backsliders from maintaining office. Especially since these voters would need to be sufficiently knowledgeable about the separation of powers and

the administrative state. When incumbents strategically manipulate elections, voters may react more strongly. This is because the main input voters have in the democratic process is voting and when this is impacted, voters may have strong negative feelings that change their behavior.

To illustrate that different backsliding actions may activate voters differently, consider the example of Istanbul's 2019 mayoral race. The election was re-run after questionable calls of fraud from President Erdoğan, after his party (AKP) lost the seat after holding it since 1994.[4] In the follow-up election, the opposition party increased its margin of victory from 13,000 votes to 800,000 votes (or from 0.16 to 9.22 percentage points), and turnout rose. This was the one of the first instances of Erdoğan alleging voter fraud in his nearly 17 year tenure, and it was met with a groundswell against the ruling party. There is reason to believe that the strategic manipulation of elections is seen differently by voters in Turkey, especially because the AKP had engaged in wide-reaching executive aggrandizement for much of the last decade with little electoral punishment. Societal threat also likely played a role in this result, but in a way that hurt the AKP. A significant financial crisis was the main backdrop to the election which most Turks saw as the responsibility of the government's ruling party. This was in contrast to previous societal threats utilized to justify backsliding such as violence from the PKK (a Kurdish militant group) used to justify cracking down on the HDP and the 2016 coup attempt leading to the creation of a new constitution and purging the civil service. Different societal threats likely serve different ends, and may not always be positive developments for a government if it can not project competence.

Ultimately, we should expect a few things about the nature of elections with backsliding incumbents. First, the level of societal threat matters in voters' decision making, especially

[4]The AKP has held the seat since 2004. Before that, the seat was held by the FP, the predecessor to the AKP, since Erdoğan's election to the seat in 1994. This personal attachment to the seat was seen as an additional reason Erdoğan fought so hard to re-run the election after the narrow loss.

when there is a disparity between the competence of the incumbent and challenger. Second, even when democratic values should indicate that voters sanction backsliding incumbents, societal threat can help a competent incumbent justify backsliding actions. Third, when incumbents know that they can benefit from a higher-threat environment than their current environment, they will act to increase threat salience to improve their chances of re-election, but this may backfire under some conditions. And lastly, the two different kinds of backsliding operate differently in this framework because they have structural differences in practice and voters view them with different levels of severity. The focus on societal threat and the different backsliding tools available pairs well with, and builds on, past work to explain how sanctioning of backsliding incumbents often breaks down. This explanation also provides room for democratic actors to focus on limiting the role of threat messaging that can dominate election time in backsliding states.

2.4 Model Set-Up

There is some known state of the world $\omega \in (0, 1)$, where ω represents societal threat — the probability that a crisis will occur in the next period after an election. I focus on two actors: an incumbent (I) and a pivotal voter (V). There is also a challenger (C) which is present as an alternative to the incumbent, but has no strategic behavior. I first detail the the actors, their choice sets, and payoff functions. After, I summarize their strategies and beliefs.

2.4.1 Actors and Information

The incumbent (I) and challenger (C) have a fixed level of competence known at the beginning of the game. I define competence as the level of skill and expertise a politician has in dealing with crisis as preferred by the voter (V). This value includes factors such as experience and

partisanship. Importantly, competence is not a stand in for partisanship, but partisanship can moderate how competence is assessed by a voter. Consider that different politicians may have different responses to a crisis that are influenced by ideology, but I combine these factors into one measure that the voter uses to evaluate her choices. Ultimately, the voter is the most interested in selecting the most competent politician who can minimize the impact of a crisis should it occur. I use Ω to indicate whether a crisis occurs where $\Omega \in \{0, 1\}$. When $\Omega = 1$, a crisis occurs after the election and when $\Omega = 0$, no crisis occurs. Societal threat (ω) is the probability that a crisis will occur or: $\mathbb{P}(\Omega = 1) = \omega$. The competence of the incumbent is represented as ϕ_I and the competence of the challenger, ϕ_C. The difference between these two ($\phi_I - \phi_C$) is the competence gap: $\tilde{\phi}$. Generally, I assume $\tilde{\phi} > 0$, or the incumbent is more competent, but also consider the reverse case.

The Incumbent

At the beginning of the game, the incumbent assigns effort over two different kinds of backsliding: executive aggrandizement and the strategic manipulation of elections. The incumbent assigns effort based on the choice of e where e_A corresponds to the amount of effort placed into executive aggrandizement and e_M is the amount of effort placed into the strategic manipulation of elections where $e_A \in [0, 1]$ and $e_M \in [0, 1]$. When $e_A = 1$ the incumbent maximally engages in executive aggrandizement. When $e_M = 1$, the incumbent maximally manipulates elections. When effort is equal to zero for a given action, that kind of backsliding action is not enacted. After observing the incumbent's choices (e_A and e_M), the voter will either choose to re-elect the incumbent ($R = 1$), or elect the challenger ($R = 0$). If re-elected, the incumbent receives a benefit of office for another term in the form of $b > 0$.

Executive aggrandizement allows the incumbent to remove institutional constraints against governing in both the present and the future. If $e_A > 0$, the incumbent will receive

an additional benefit of office $b_A > 0$ for being able to institute policy more easily both in the current period, and in the future if re-elected. This benefit is *moderated* by the effort assigned to executive aggrandizement ($e_A * b_A$). Lastly, retaining office is always more preferred to executive aggrandizement: $b > b_A$.

The strategic manipulation of elections allows incumbents to entrench themselves and make it more difficult for voters to remove them from office. If $e_M > 0$, and the voter chooses to remove the incumbent ($R = 0$), the voter will fail to remove the incumbent with some probability: $e_M * f$ where $f \in (0, 1)$. When election subversion is successful, even though the true preference of the voter was to remove the incumbent, the strategic manipulation of elections allowed the incumbent to maintain power. Future payoffs are subject to a discount factor: $\delta \in (0, 1)$. Ultimately, the incumbent is interested in maintaining office and engaging in executive aggrandizement when possible. The incumbent's utility can be expressed as follows:

$$U_I(R, e_A, e_M) = \begin{cases} e_A b_A + \delta(b + e_A b_A) & R = 1 \\ e_A b_A + e_M f \delta(b + e_A b_A) & R = 0 \end{cases}$$

The Voter

The voter begins the game by viewing the societal threat level ω and the competence of the incumbent, challenger, and this gap (ϕ_I, ϕ_C, and $\tilde{\phi}$). In the event that a crisis occurs ($\Omega = 1$), the voter will receive a negative payoff, $t > 0$. The value of ϕ of the victorious candidate can lessen the impact of this threat where the actual payoff is $-(t - \phi)$ where $\phi_I < t$ and $\phi_C < t$. In the absence of a crisis ($\Omega = 0$), there is no penalty.

The voter also has preferences for democratic norms and institutions. These preferences are represented by $\beta < 0$, where negative values indicate that a voter is opposed to backsliding actions and supports democratic institutions and norms. The voter has specific opinions on both kinds of backsliding: β_A for executive aggrandizement and β_M for the strategic manipulation of elections. These values can be thought of as the negative cost of having a backslider for another term in office and are weighted by the incumbent's backsliding efforts: $e_A * \beta_A$ and $e_M * \beta_M$.

The voter uses all of the above information to make the decision of whether to choose the incumbent or the challenger. Absent backsliding, the voter is only interested in selecting the more competent leader[5]. The voter only has payoffs after the election, so I do not discount these payoffs. The voter is concerned both with the negative payoff of having a backslider in office in the future as well as the competence of the winner, because this can directly limit her negative exposure to a crisis should it occur. The general utility function for the voter is determined by R, e_A, e_M, and Ω:

$$U_V(R, e_A, e_M, \Omega) = \begin{cases} e_A\beta_A + e_M\beta_M - \Omega[(t - \phi_I)] & R = 1 \\ e_M f(e_A\beta_A + e_M\beta_M) - \Omega[e_M f(t - \phi_I) - (1 - e_M f)(t - \phi_C)] & R = 0 \end{cases}$$

Sequence of Play

Before the game begins, fixed values of the competence of the incumbent ϕ_I and challenger ϕ_C, societal threat ω, and the voters opinions of both kinds of backsliding β_A, β_M are known to all players. Then the game begins:

[5]This same assumption is used in Luo and Przeworski (2019).

1. The incumbent chooses effort values e_A and e_M.

2. Based on I actions, the voter V either re-elects (R=1) the incumbent I or elects (R=0) the challenger C.

3. If $R = 0$ and $e_M > 0$, incumbent maintains power with probability $e_M * f$. Otherwise, the voter's choice is respected.

4. After the election, a crisis occurs $\Omega = 1$ with probability ω.

5. Payoffs realized by players.

2.4.2 Strategies

I begin with the voter's strategy because it has a direct impact on the choices of the incumbent. I then break down the incumbent's preferred actions given the behavior of the voter.

The Voter

The voter maximizes the expected utility of retaining or removing the incumbent given the effort levels e_A and e_M, the level of societal threat ω, the magnitude of negative cost should a crisis occur t, the revised penalty voters place on another term under a backslider β_A and β_M, and the competence parameters of the incumbent ϕ_I and challenger ϕ_C. Given the parameters, the expected value of the voter's payoff given the choice to re-elect or not is:

$$\mathbb{E}U_V(R, e_A, e_M) = \begin{cases} e_A\beta_A + e_M\beta_M - \omega[(t - \phi_I)] & R = 1 \\ e_M f(e_A\beta_A + e_M\beta_M) - \omega[e_M f(t - \phi_I) + (1 - e_M f)(t - \phi_C)] & R = 0 \end{cases}$$

When the voter chooses a backsliding incumbent, the negative payoff of diminished democracy is combined with the term that dictates the probability of crisis after the election.

Should a crisis occur, the competence of the incumbent is important to the voter because it diminishes the negative payoff of the threat. When the voter chooses to remove the incumbent and there is not electoral subversion ($e_M = 0$), the voter does not incur a negative payoff from backsliding and her payoff relies on the competence of the challenger in the event of a crisis. When subversion does occur ($e_M > 0$), an incumbent that strategically manipulates the election may retain office if the voter chooses to remove with probability $e_M f$. Therefore, depending on the incumbent's effort allocation e_A and e_M the voter must assess whether $\mathbb{E}U_V(R = 1) \geq \mathbb{E}U_V(R = 0)$ and if so, the voter will choose to re-elect. The incumbent knowing this information can then consider what action to choose given the parameters.

The Incumbent

The incumbent knows what the voter will do depending on the values of the parameters listed above and makes decisions based on the expectation of retaining office or not and the benefit of engaging in executive aggrandizement (b_A). I will explain these strategies in terms of $\mathbb{E}U_V(R = 1) \geq \mathbb{E}U_V(R = 0)$ for the voter, because the incumbent is acting with this in mind, and the incumbent's own payoffs when applicable.

I first consider the much more likely case when the incumbent is more competent: $\tilde{\phi} > 0$. Backsliding incumbents are generally seen as strong leaders due to their forceful governing style and should be able to project a high level of competence. The incumbent's chief goal is to be re-elected. Absent backsliding, the incumbent will always be re-elected if more competent than the challenger. In this case, societal threat only impacts the voter's preferences with respect to competence and alleviating potential crisis.

The incumbent's secondary goal is to engage in as much executive aggrandizement as possible. If the incumbent backslides and the voter's negative opinions of backsliding outweigh

the advantage from being more competent than the challenger, then we should not expect backsliding. In this case, the best choice for the other effort value is $e_M = 0$. This is both because engaging in the strategic manipulation of elections serves no benefit for incumbent in terms of payoffs; and because it may lead the voter to choose $R = 0$. The following identifies the maximum possible value of e_A where retaining office is still certain and also demonstrates the mechanism through which higher competence helps offset more democratic backsliding:

$$\begin{aligned} e_A\beta_A - \omega(t - \phi_I) &\geq -\omega(t - \phi_C) \\ e_A\beta_A &\geq \omega(t - \phi_I) - \omega(t - \phi_C) \\ e_A\beta_A &\geq -\omega\tilde{\phi} \\ e_A &\leq -\frac{\omega\tilde{\phi}}{\beta_A} \end{aligned} \tag{2.1}$$

If inequality 2.1 holds for $e_A = 1$, then the voter will re-elect ($R = 1$) for any amount of executive aggrandizement up to the maximum amount. Also, because β_A is negative, there is a sign flip. To refocus this cut-off in terms of societal threat, I restate inequality 2.1 in terms of ω. I substitute $e_A = 1$ so this is the level of societal threat where the incumbent can engage in the max executive aggrandizement and still retain office with certainty:

$$\begin{aligned} -\frac{\omega\tilde{\phi}}{\beta_A} &\geq 1 \\ \omega &\geq -\frac{\beta_A}{\tilde{\phi}} \end{aligned} \tag{2.2}$$

If inequality 2.2 does not hold, but inequality 2.1 holds for some value of $e_A = e_A^* = -\omega\tilde{\phi}/\beta_A$ and the righthand side is less than 1, then the incumbent must decide whether this value is preferable to more democratic backsliding. By choosing this effort value, the incumbent would receive the payoff: $U_I(R = 1, e_A = e_A^*, e_M = 0)$. As stated previously, when

$R = 1$, that the incumbent has no incentive to set $e_M > 0$ because it has no payoff benefit. The alternative, where the incumbent chooses some effort for executive aggrandizement $e_A > e_A^*$ will also lead the voter to select $R = 0$. After this threshold is passed, when $R = 1$, all further increases in e_A and e_M increase U_I, so the best choice when $R = 0$ is $e_A = 1$ and $e_M = 1$. This produces the following payoff: $U_I(R = 0, e_A = 1, e_M = 1)$. So, the incumbent decides whether the more minor executive aggrandizement under certain re-election is better than giving up election certainty for maximum subversion or $U_I(R = 1, e_A = e_A^*, e_M = 0) \geq U_I(R = 0, e_A = 1, e_M = 1)$. I set up this inequality in terms of the incumbent payoffs given these actions and solve for e_A^*. On the right-hand side, because $e_A = 1$ and $e_M = 1$, the effort values drop out. After solving for e_A^*, I plug in $-\omega\tilde{\phi}/\beta_A$ and solve for ω to represent this threshold in terms of societal threat:

$$\begin{aligned} e_A^* b_A + \delta(b + e_A^* b_A) &\geq b_A + f\delta(b + b_A) \\ e_A^* &\geq \frac{b_A(1 + \delta f) + \delta b(f - 1)}{b_A(1 + \delta)} \\ -\frac{\omega\tilde{\phi}}{\beta_A} &\geq \frac{b_A(1 + \delta f) + \delta b(f - 1)}{b_A(1 + \delta)} \\ \omega &\geq -\frac{\beta_A}{\tilde{\phi}} \left[\frac{b_A(1 + \delta f) + \delta b(f - 1)}{b_A(1 + \delta)}\right] \end{aligned} \tag{2.3}$$

If inequality 2.3 holds, the societal threat level is sufficient for the incumbent to prefer a lower level of executive aggrandizement e_A^* to maximal subversion under an uncertain election result. Generally, I expect the incumbent to prefer e_A^* unless f is very high.[6] In this case, the role of the voter is mostly irrelevant because elections can be subverted to an extreme degree. For a backsliding incumbent up for re-election for the first time, f should be much closer to 0.

[6]Consider the case of f=0. This produces the following: $\omega \geq -\frac{\beta_A}{\tilde{\phi}}\left[\frac{b_A - \delta b}{b_A}\right]$, which is negative when $b_A < \delta b$. Because $b_A < b$, then for sufficiently high values of $\delta > \frac{b_A}{b}$ when f is low, there will never be an incentive to maximally subvert. As f rises, the maximal subversion option becomes more attractive.

Finally, I consider when the incumbent is less competent: $\tilde{\phi} < 0$. In this case, the voter always chooses $R = 0$. As stated earlier, in order to maximize payoffs and the chance to retain office under subverted institutions, the incumbent chooses $e_A = 1$ and $e_M = 1$ and expects $U_I(R = 0, e_A = 1, e_M = 1)$. This is one case where societal threat is not relevant because its role with respect to competence hurts the incumbent regardless. I generally do not expect this case because I expect an incumbent to project competence, but it gives some expectations for a comparatively incompetent incumbent attempting to maintain power. I expand on the incumbent's choices more later when the incumbent can manipulate threat.

2.5 Properties of Equilibrium

With the actors and strategies set up, I can now set up the equilibria and explore the factors that impact them. I first focus on how the voter responds to societal threat. Specifically, the ways different parameters and the incumbent's effort levels impact the level of societal threat when the incumbent is either retained or replaced. Then I establish the incumbent's best actions in terms of societal threat and under which effort levels the incumbent can engage in backsliding and be retained in terms of societal threat thresholds. Lastly, I use these thresholds to consider when the incumbent can augment societal threat, and show under which scenarios this is expected and how it impacts the incumbent's effort levels.

2.5.1 Voter's Reaction to Societal threat

I focus on the preferences of the voter and how changes to parameters and the incumbent's effort levels affect whether the voter re-elects a backslider. I consider preferences in terms of societal threat and expand the expression $\mathbb{E}U_V(R = 1|e_A, e_M) \geq \mathbb{E}U_V(R = 0|e_A, e_M)$:

$$e_A\beta_A + e_M\beta_M - \omega(t - \phi_I) \geq e_M f(e_A\beta_A + e_M\beta_M) - \omega[e_M f(t - \phi_I) + (1 - e_M f)(t - \phi_C)]$$

I simplify the inequality, and solve for ω to identify the critical threshold that societal threat must rise above for the voter to prefer retaining the incumbent when the incumbent backslides. I substitute $\tilde{\phi}$ for $\phi_I - \phi_C$ to make the competence gap a single parameter. This process yields the equation in lemma 1.

Lemma 1 (The Threat Cutoff for a Voter to Support a Backsliding Incumbent) *The societal threat threshold $\omega \geq \omega^*$ at or above which the voter supports the incumbent is:*

$$\omega^* = \frac{(e_M f - 1)(e_A \beta_A + e_M \beta_M)}{\tilde{\phi}(1 - e_M f)}$$

Lemma 1 identifies the critical value of societal threat level ω^* that induces incumbent support from the voter given other parameters. This equation is true for any values of the effort parameters where retention is possible, including when backsliding is not present. When the voter chooses to retain the incumbent even after backsliding, I call this a *sanctioning breakdown*. I take the partial derivative of ω^* with respect to the parameters of interest to get comparative statics for parameters of interest. These are general results to show how various parameters behave in the model. The main choices established for the incumbent are the most important component and I will detail those after.

Proposition 1 (Comparative Statics for Sanctioning Breakdown) *The critical threat level needed for a voter to retain a backsliding incumbent is:*

i. Decreasing in $\phi_I - \phi_C \equiv \tilde{\phi}$.

ii. Decreasing in β_A and β_M $\tilde{\phi} > 0$.[7]

iii. Increasing in e_A and e_M when $\tilde{\phi} > 0$.

[7] β_A and β_M are negative because they are the *penalty* for retaining a backsliding incumbent, so this result can also be interpreted as increasing in democratic preferences.

Proposition 1 shows how the parameters of interest impact the critical societal threat threshold needed for a voter to prefer retaining the incumbent given a set of parameters. In practice, this means that as the incumbent increases the competence advantage over the challenger (larger $\tilde{\phi}$), the level of societal threat needed for democratic sanctioning to fail is lower. As the voter has less negative opinions of backsliding actions (higher β_A or β_M), the societal threat level needed for incumbent retention decreases. As effort values increase, the societal threshold rises. Most of these results are reliant on the incumbent being more competent. Without a competence advantage, the statement $\mathbb{E}U_V(R = 1|e_A, e_M) \geq \mathbb{E}U_V(R = 0|e_A, e_M)$ can never be true.

Now that I have analysis for how different parameters interact with societal threat with respect to the voter's preferences, I can conclude the model by considering when an incumbent can benefit by raising the salience of societal threat.

2.5.2 When Incumbents Can Manipulate Societal Threat

Proposition 1 established how societal threat and other parameters impact the voter's choice to retain or remove the incumbent. I now consider an incumbent who is aware of the voter's behavior and is able to manipulate the salience of societal threat. Incumbents are focused primarily on maintaining office and secondarily being able to more easily enact their preferred policies by engaging in executive aggrandizement. When incumbents are unsure of their re-election prospects, they may rely on the strategic manipulation of elections to increase their chances of re-election. When I detailed the incumbent's strategies earlier, I discussed their best responses knowing the voter's preferences. I now restate those preferences with respect to the societal threat level needed in order for certain backsliding actions to be preferable or not.

Lemma 2 (Incumbent Actions In Basic Model) *The incumbent's best choices given the following conditions are:*

i. When $\tilde{\phi} > 0$ *and* $\omega \geq -\frac{\beta_A}{\tilde{\phi}}$, *I chooses* $e_A = 1$ *and* $e_M = 0$.

ii. When $\tilde{\phi} > 0$ *and* $-\frac{\beta_A}{\tilde{\phi}} > \omega \geq -\frac{\beta_A}{\tilde{\phi}}\left[\frac{b_A(1+\delta f)+\delta b(f-1)}{b_A(1+\delta)}\right]$, *I chooses* $e_A = -\frac{\omega\tilde{\phi}}{\beta_A}$ *and* $e_M = 0$.

iii. When $\tilde{\phi} > 0$ *and* $\omega < -\frac{\beta_A}{\tilde{\phi}}\left[\frac{b_A(1+\delta f)+\delta b(f-1)}{b_A(1+\delta)}\right]$, *I chooses* $e_A = 1$ *and* $e_M = 1$.

iv. When $\tilde{\phi} < 0$, *I chooses* $e_A = 1$ *and* $e_M = 1$.

Lemma 2 shows the incumbent's best responses given the competence gap, the level of threat, and the payoffs. I incorporate the inequalities from section 2.4.2 to construct the thresholds of ω where different actions are preferred and resultantly the four parts of lemma 2. Inequality 2.2 is the threat threshold where at or above, the incumbent can assign effort $e_A = 1$ (part i). Below this level, the incumbent needs to assign some value reflected in inequality 2.1 and for this to be preferred, the societal threshold established in inequality 2.3 must be true (part ii). Below this threshold, full subversion ($e_A = 1$ and $e_M = 1$) will be preferred (part iii). When the incumbent has no competence advantage, there are no relevant threat thresholds and full subversion will be chosen (part iv). This approach helps the incumbent both get the maximum payoff and maximize the probability of retention under subverted electoral institutions.

Using this information, we can now consider the ability of the incumbent to manipulate the threat level. Suppose that the incumbent can increase the level of societal threat for some cost c to a value $\bar{\omega}$ where $1 > \bar{\omega} > \omega$. This ability to raise societal threat impacts the actions of the incumbent. In practice, we are looking to incumbents detailed in lemma 2 who with higher threat can engage in more backsliding than previously possible.

Proposition 2 (Incumbent Actions with Threat Manipulation) *When the incumbent can manipulate threat, more executive aggrandizement is possible when:*

(i) When $\tilde{\phi} > 0$, $b_A > c$, *and* $\bar{\omega} > -\frac{\beta_A}{\tilde{\phi}} > \omega$, *I chooses* $e_A = 1$.

(ii) When $\tilde{\phi} > 0$, $e_A b_A > c$, *and* $-\frac{\beta_A}{\tilde{\phi}} > \bar{\omega} > -\frac{\beta_A}{\tilde{\phi}}\left[\frac{b_A(1+\delta f)+\delta b(f-1)}{b_A(1+\delta)}\right]$, *I chooses* $e_A = -\frac{\bar{\omega}\tilde{\phi}}{\beta_A}$.

In both scenarios in Proposition 2, a higher, manipulated level of societal threat allows incumbents to engage in more executive aggrandizement than otherwise would be possible. For incumbents in (i), this allows increase effort from $e_A = -\frac{\omega\tilde{\phi}}{\beta_A}$ to $e_A = 1$. This increased effort corresponds to incumbents going from part ii of lemma 2 to part i because of the increased societal threat level. For (ii), a higher level of e_A is now possible than before: $-\frac{\bar{\omega}\tilde{\phi}}{\beta_A} > -\frac{\omega\tilde{\phi}}{\beta_A}$. This higher effort level corresponds to a change within part ii of lemma 2. I do not consider the case of an incumbent moving from part iii to ii because this is a case where less backsliding occurs, and I expect part iii to be rare. This result shows that when incumbents can manipulate the threat level, this creates more situations where incumbents engage in executive aggrandizement while maintaining power. This does not impact more strategic manipulation of elections, because only less competent incumbents engage in it (except for part iii in Lemma 2). The ability to exaggerate existing threats can be a powerful tool for competent incumbents to use in order to maximize their ability to engage in backsliding without being removed from office. Higher threat does not benefit less competent incumbents because voters will not prefer them for any level of societal threat. This gives greater context to threat-focused campaigns of backsliders such as the focus on immigration in Hungary, Poland, and the United States or the focus on violence from the PKK in Turkey.

There are three main takeaways from the formal model. The first, and the focus of the empirical application in the next section, is that higher societal threat helps allow incumbents

to engage in more democratic backsliding. This is because societal threat justifies backsliding because voters are willing to overlook actions to retain competent incumbents, and higher threat allows the incumbent to get away with more backsliding. Second, higher societal threat allows competent incumbents to benefit from executive aggrandizement. This is because only incumbents that are less competent or close to the competence of the challenger will be expected to strategically manipulate elections. Lastly, incumbents that manipulate or raise the salience of societal threat are more likely to be engaged in executive aggrandizement. While some of these conclusions rely on concepts like voters' preferences for democracy or a competence gap between incumbent and challenger, I believe the finding that additional threat allows for additional backsliding is an important testable implication of this model. In the next section, I show a connection between new backsliding actions and changes in some measures of societal threat from year to year.

2.6 Data for Empirical Application

While some elements of the formal model are difficult to adequately measure such as incumbent and challenger competence or aggregated voter preferences for democratic norms and institutions, there are two concepts that can be estimated to empirically evaluate an important outcome of the model: increased societal threat allows for additional democratic backsliding actions. In order to evaluate this empirically, I need appropriate measures for both change in democracy and societal threat at the country-year level. I use data from V-Dem (Coppedge et al. 2020) and Fitch Connect to measure these concepts respectively.

First, the Varieties of Democracy of Project (V-Dem) has a considerable historical database that categorizes and evaluates institutions at the country-year level. V-Dem provides hundreds of different indicators to classify governments, the quality of democracy,

Table 2.2: Summary of High Level Indices from V-Dem

Indicator	Motivating Question	Clarification/Components
Liberal Democracy	To what extent is the ideal of liberal democracy achieved?	constitutionally protected civil liberties; strong rule of law; an independent judiciary; executive checks and balances that, together, limit the exercise of executive power
Polyarchy/ Electoral Democracy	To what extent is the ideal of electoral democracy in its fullest sense achieved?	suffrage is extensive; political and civil society organizations can operate freely; elections are clean and not marred by fraud or systematic irregularities; elections affect the composition of the chief executive of the country; freedom of expression and an independent media between elections
Participatory Democracy	To what extent is the ideal of participatory democracy achieved?	The participatory principle of democracy emphasizes active participation by citizens in all political processes, electoral and non-electoral

Note: Each of these measures is bounded 0-1 where the higher the value, the more complete democracy as defined is realized.

and several other important concepts. I work with three different high-level indices from V-Dem that aggregate several of the more granular measures to assess macro-level considerations of democratic quality. One of the measures best approximates the kind of democracy attacked through executive aggrandizement: liberal democracy. This value falls where there are less checks and balances between government institutions. For example, an executive engaging in executive capture of the judiciary would produce a lower score in this measure. The other two capture institutions attacked by the strategic manipulation of elections: basic or electoral democracy which captures the concept of "polyarchy" from Dahl (1971) and participatory democracy. The Polyarchy measure helps ascertain the more basic electoral institutions, while the participatory democracy measure is a more rigorous standard that evaluates voters' access to the system. These two indexes help give a more complete understanding of institutions attacked through the strategic manipulation of elections. While in the formal model only executive aggrandizement had a positive relationship with societal threat, this was largely driven by the modeling choices where the result of the election is known with certainty based on the values of parameters and the preferences of the pivotal voter. In practice, that level of certainty is not possible, so I include measures of both kinds of democratic backsliding. The description from V-Dem for each measure is in table 2.2.

To assess threat, I use a country-risk dataset from Fitch Connect. While societal threat can manifest in several different ways, I believe this data allows for assessing some key societal threat concepts that I summarize in table 2.3. Crisis can manifest in several different ways and these measures help assess several different kinds of societal threats that can be used to justify backsliding. Not all threats are beneficial for a backsliding incumbent though, and we should not expect all of the threat measures to impact backsliding to the same degree. Through these various measures, I can evaluate economic, social, and security/physical threats. I anticipate that some of these threats will be more useful to justify backsliding than others, especially those related to social threats and physical/security threats. Lastly, these measures are useful for estimating *actual* societal threat, and not exaggerated claims of backsliding incumbents.

Table 2.3: Societal Threat Measures from Fitch Connect

Measure	Definition	Kinds of Threats Estimated
Social Stability Index	Reflects unemployment, inflation, and public unrest	Economic and Social Threats
Security/External Threats Index	Reflects security threats, regional profile and international constraints	Physical/Security Threats
Short Term Economic Risk Index	Growth risk; monetary risk; fiscal risk; external risk; and financial risk.	Economic Threats
Crime and Security Index	Conditions with respect to interstate conflict risk, terrorism and crime, including cyber crime and organized crime	Physical/Security Threats
GDP Growth	Year change in GDP as percent	Economic Threats

Note: Each of these measures, except for % GDP, is bounded 0-100 where higher values indicate less threat in each of the categories.

The empirical strategy I use is OLS regression to assess the year change of each of the democracy measures from V-Dem. I use the change from the previous year for each of the three high-level V-Dem indexes as the dependent variables because certain threats may not help backsliding incumbents engage in backsliding equally. The change in each threat measure from the previous year is used as the independent variables because it is the increase

of societal threat that provides the cover for new backsliding actions. I control for logged GDP because wealth plays a substantial role in democratization (Bunce 2000; Geddes et al. 1999; Lipset 1959; Milner and Mukherjee 2009; Przeworski and Limongi 1997). I also control for each of the three democracy index scores from the given country-year observation to make sure the results are not driven by regime quality. Lastly, I use year fixed effects.

This empirical strategy helps estimate the year-to-year impact of societal threat changes on democratic backsliding the most effectively given the availability of data. In the model, I transform the V-Dem measures by multiplying them by 100, so an increase of a societal threat measure (which indicates societal threat decreased) is on the same scale as a change in the democracy measures.

2.7 Empirical Results

The goal of this empirical application is to clarify whether factors that we might identify as societal threats actually produce the effects from the formal model. By including several measures of societal threat that may affect voters, I can identify which factors are especially important in relation to new democratic backsliding measures. Societal threat is a macro-level concept interpreted at the individual level, and different societal threats may provide more cover for democratic backsliding than others.

The results from the empirical strategy are in table 2.4. In this case, each model corresponds to a different dependent variable for the various democracy indices from V-Dem. For the purpose of interpretation, a positive coefficient for a threat indicator means that a year decrease in the societal threat index (higher threat) is associated with a decrease in the democracy measure (additional democratic backsliding).

Table 2.4: Results for the Impact of the Change of Societal Threat on the Change of Democratic Indicators

	Δ Liberal Dem	Δ Polyarchy	Δ Participatory Dem
Δ Social Stability	0.048***	0.040**	0.017
	(0.015)	(0.017)	(0.013)
Δ Security and External Threats	0.072***	0.075***	0.039*
	(0.025)	(0.028)	(0.020)
Δ Short Term Economic Risk	0.011	0.028	0.012
	(0.019)	(0.022)	(0.016)
Δ Crime and Security	0.0005	−0.009	−0.001
	(0.021)	(0.023)	(0.017)
% GDP Growth	−0.002	0.0004	0.014
	(0.020)	(0.023)	(0.017)
Intercept	0.792	−0.700	−0.135
	(0.993)	(1.110)	(0.811)
Observations	1,006	1,006	1,006
R^2	0.031	0.031	0.017

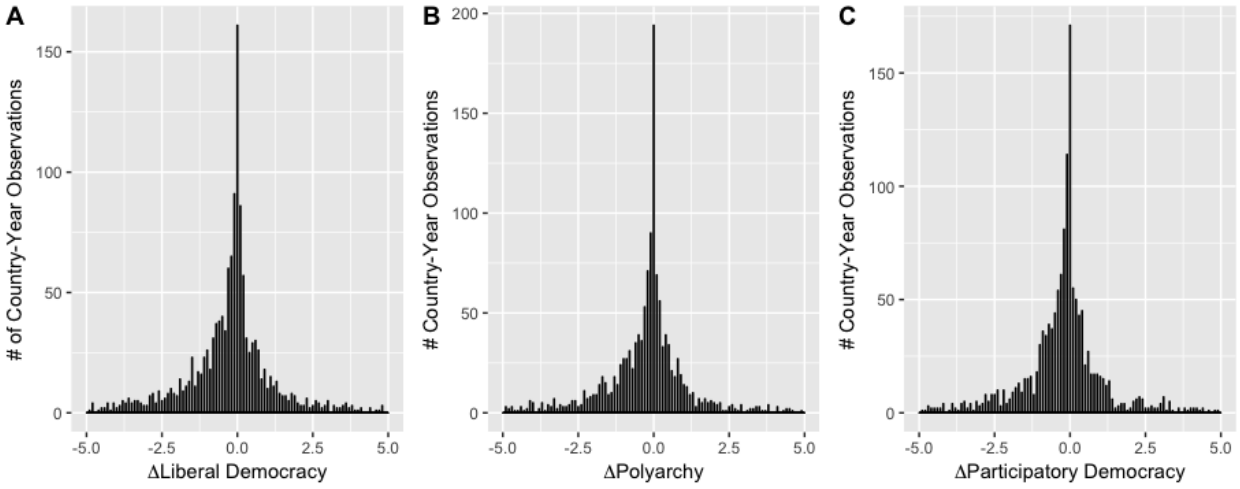
Note: *p<0.1; **p<0.05; ***p<0.01

These results provide some insight as to how the change in threat environment in a given year allows for more or less backsliding from an incumbent. In this case, both the social stability measure and crime/external threat measures provided positive and significant coefficients for two of the democracy measures. This would show that these factors may be threat types that especially allow for new backsliding actions, which is how societal threats behave in the model. Whereas the others do not capture the kinds of societal threats with the theorized effect. The coefficients are small, but each of the dependent variables are tightly distributed around a mode of zero and this impacts the substantive significance.

To better show the magnitude of these findings relative to the dispersion of the data, I show the distribution of the year-change of each of the dependent variables in figure 2.1. This figure shows that small changes in democratic quality are particularly common from

year to year. When these are aggregated over the course of a backslider's term in office or over multiple terms, these small changes become increasingly more important.

Figure 2.1: Distribution of Change in Democracy Measures.



Note: Because most of the observations are concentrated around zero, I bounded these figures to better show the distribution around zero. This omits 81 observations in panel A, 104 in panel B, and 58 in panel C that were either greater than 5 or less than -5.

To put these results into greater context, consider an example where the social stability measure falls by 10 in a given year. This would be associated with a 0.48 reduction in the liberal democracy measure and 0.4 reduction in the Polyarchy measure when controlling for the previously mentioned factors and the other threat indicators. While this seems like a relatively small change, because of the nature of the measures, this is still noteworthy and could indicate some additionally damaged institutions. For example, in the United States in 2018, there was a 0.5 decline in the Polyarchy measure and a 0.6 decline in the liberal democracy measure. This was the year of the start of the highly controversial family separation policy at the southern border and the protest response to these actions[8]. While the changes may seem small, because they are at the year level, the impact over several years is considerable.

[8]These specific examples were sources from the Democratic Erosion Event Database or DEED.

Ultimately, the societal threats measures with respect to social stability and security or external threats proved to behave in line with the formal model. Other measures that help assess societal threats, but did not behave as societal threat did in the formal model include short-term economic risk, crime and security, and percent GDP growth. These differences make sense, because some examples of societal threat may not have sufficiently looming crises that induce voter support for democratic backsliding and resultantly make democratic backsliding possible for incumbents. Future work can include additional examples of societal threat, or engage in individual-level analysis of the relationship of societal threat and backsliding support in a survey experiment. This may allow for assessing some of the more nuanced results from the model such as the ability to engage in more democratic backsliding while being able to manipulate societal threat.

Democratic backsliding is a gradual process over the course of several years, so small changes from year to year can add up over the course of a backslider's term in office, and establishing that threat has an impact in these changes is noteworthy. Further, because each of these measures is just one example of a way that societal threat may manifest, a societal threat related to social stability may provide cover to subvert institutions one year, while a more salient security or external threat may justify a backsliding action in a subsequent year. At the aggregate, these strategies can be helpful to successfully maintain power while engaging in backsliding actions. Lastly, because societal threat can only work against a backslider if they are incompetent in the face of crisis, different kinds of threats may give the necessary cover to overcome shortcomings in one area.

2.8 Conclusion

How democratic backsliders maintain power is a question that is increasingly more important in the study of democratic backsliding. While the normative hope is that leaders who subvert democratic norms and institutions are punished and removed, this is commonly not the case. Existing work has explored the role of partisanship and polarization and how these factors allow backsliders to maintain power because voters care more about policy than democracy. I posited that societal threat also plays an important role in how these factors relate to voter support of democratic backsliding.

I argued that societal threat impacts a voter's evaluation of an incumbent's tenure by raising the importance of a voter selecting a competent leader. I also considered how different kinds of backsliding that can be employed impact these dynamics. These factors were featured in my formal model where I concluded that additional societal threat provided room for more democratic backsliding. I showed additional results to indicate that the kinds of backsliding employed are also important. Specifically, I showed that executive aggrandizement should be particularly aided by societal threat and was the backsliding that competent incumbents would be most likely to engage in. These findings, paired with existing understanding of polarization and partisanship, help give an additional dimension to the understanding of why voters may vote to retain backsliding incumbents.

I also empirically tested a main conclusion of my formal model, that additional societal threat allows more democratic backsliding. I evaluated this theoretical implication by using observational data from the Varieties of Democracy Project to measure democratic quality and its changes as backsliding, and societal threat indicators from a country risk dataset from Fitch Connect. I showed that changes in societal threat from one year to another were related to changes in democratic quality. These results specifically showed that additional

societal threat from social stability and external threats were associated with backsliding at the country-year level. When these results are aggregated to the government or term level, societal threat can be an important component in providing cover for incumbents to engage in, and to get away with democratic backsliding.

The expectations from the formal model and results from the empirical give additional insight as to why backsliders are typically retained despite their misdeeds. Societal threat is a potent force in justifying backsliding misdeeds by leading voters to support an especially competent leader, who is usually the incumbent. This would mean that only under circumstances like the Istanbul Mayoral Race of 2019 where societal threat worked against the ruling government should we expect societal threat to lead to punishment. Lastly, even if societal threat is not high, incumbents can manipulate its salience to benefit themselves.

This theoretical framework pairs well with existing expectations and provides an explanation for why incumbents may not be retained due to a competence disadvantage. In the next two chapters, I consider when incumbents are able to be removed from power and whether the prospects for democracy are promising given the removal of a backsliding incumbent.

Chapter 3

What Does the Opposition do After Democratic Backsliding?

3.1 Introduction

What does the opposition do after democratic backsliding? The study of democratic backsliding has generally focused on precursors to and the early stages of an authoritarian-minded democratically elected leader seizing power. But what happens after that backslider is out of power? Successors[9] have three main choices they can make: retaliate and escalate an undemocratic cycle, focus on repairing institutions to stave off democratic deconsolidation at the expense of policy goals, or take a path somewhere in between. The importance of this choice is considerable, because a successor can either aggressively respond to an initial backslider and pave the way for a cycle of backsliding that can culminate in full democratic

[9]This paper focuses on opposition successors who were at odds with the previous backslider. I expect successors to backsliders of the same party (e.g., in Venezuela Nicolás Maduro succeeding Hugo Chávez after his death, or PRI in 20th century Mexico) to behave in a similar way to their predecessors.

failure, or leave the misdeeds of an initial backslider as an aberration by restoring democratic institutions. We can better understand the long-term consequences of this decision by exploring the incentives that shape that choice.

While a successor to an initial backslider's choice to restore institutions or not has not been studied, there is existing work that informs expectations for this choice and how it is made. Most of these expectations come from the more established democratization literature which tracks how a country moves from autocracy to democracy. Democratization is distinct from backsliding because states with recently removed backsliders have intact, albeit damaged, democratic institutions and their voters have existing democratic preferences and expectations that are not present in autocracies that move toward democracy. Despite these important differences, the choices of an autocrat looking to reform and a successor considering whether to restore norms and institutions have useful parallels. Work on elite-driven democratization provides some useful explanations of the conditions under which autocrats reform institutions and make them more democratic (Dahl 1971; Gonzalez 2008; Haggard and Kaufman 1999, 2016). Specifically, the way leaders focus on maximizing security and minimizing risk is central to the choice of when and how to institute reforms (Riedl et al. 2020; Schedler 2002). For autocrats, the concern is creating enough legitimacy to stave off a popular uprising while maintaining sufficient control. Successors to backsliders should behave similarly by weighing the restoration of legitimate institutions, which constrain both future opponents and themselves, with short-term policy goals and the desire to keep the initial backslider out of office.

I theorize that three factors shape the decision of the successor whether to *restore* by repairing norms and institutions to their previous state and refusing to engage in further norm-breaking behavior, or *retaliate* by continuing the use of damaged institutions and norm-breaking behavior to the successor's benefit and against opponents. First is the *durability of*

the restoration or whether restorative actions pursued by a successor are likely to constrain a future undemocratic actor. If a future backslider can unravel restorative measures with relative ease, then the incentives for the successor to constrain their actions, only to have an opponent not play by the same rules, are not present for restoration to be attractive. Second, is the *magnitude of the subverted institutions*. In other words, this is the difference between the outcomes provided by damaged institutions when in power and out of power. If damaged institutions allow marginally better results for an incumbent, then the risk of eventually being out of power is not enough of a deterrent to maintain damaged institutions. The final factor is *electoral safety* or how confident the successor feels in the ability to win future office. Electoral safety plays a considerable role in determining whether policy is prioritized at the expense of democratic restoration because short-term gains and long-term strategies may conflict.

I develop an original formal model in which a newly elected government has the option to restore institutions or retaliate after two kinds of democratic backsliding: *executive aggrandizement* where constraints on the executive have been weakened so that policies are easier to pass, and the *strategic manipulation of elections* where elections had been made less free and fair and thus tilted towards the incumbent. When successors restore institutions, it impacts their ability to enact their preferred policies and makes election outcomes more fair. When successors retaliate, they can enact their preferred policies more easily and win re-election with greater certainty. Restoration is *durable* when restorative actions are binding for future leaders, and if those are not binding, restoration is not durable. For the *magnitude* of the subverted institutions, I consider both kinds of backsliding and how the institutions they impact drive advantages with respect to either policy making or electoral performance. Lastly, I conceptualize *electoral safety* as the baseline probability that the successor gets elected absent electoral manipulation.

I find that the most important factor to determine preference for restoration is durability. I explored durability by considering two extreme conditions: an optimistic case where restoration is completely durable and can never be reversed, and a pessimistic case where restoration is never durable and will always be reversed a successor loses re-election. When restorative actions are not durable, restoration is never preferred, but under durable restoration, there are some conditions where restoration is preferred. I also find that, under durable restoration, more aggressively subverted institutions for both kinds of backsliding make maintaining backsliding less attractive for the successor. Additionally, under durable reforms, electoral safety has different effects on maintaining the two kinds of backsliding. As electoral safety increases, successors prefer restoring executive constraints less, but prefer restoring free and fair elections more. These results are driven by the nature of the institutions attacked and have important implications for successor preferences. Finally, by assessing durability using two extreme cases — completely durable and completely non-durable restoration — I can assume that restoration can be preferable, under some conditions, even when restorative measures are not completely durable. Restorative measures are likely to fall somewhere between fully durable and never durable, so this is an important and noteworthy finding.

As backsliding regimes in countries like Turkey, Hungary, and Poland are likely to eventually yield to the opposition, a greater understanding of what happens after this is novel and important. When conditions are favorable for backsliding to continue, a long term series of tit-for-tat actions can snowball into far worse outcomes that can culminate in full democratic failure. However, with some attractive conditions, such as sufficiently large electoral safety, and strong consequences for maintaining backsliding in the future, restoration can be not only a normatively beneficial path, but the optimal path. The rigorous interrogation of these incentives in the formal model allow for important initial work on this important window of opportunity to counteract backsliding.

3.2 Background

The democratization literature provides a useful starting point for thinking about how successors in a post-backsliding state will behave. Democratization is defined in terms of a transition from autocracy to democracy, but because of the broadness of this definition, it is difficult to establish a uniform theory (Rustow 1970). Some historically relevant democratization theories are less useful to the backsliding case such as: modernization theory (Lipset 1959) and its addendums (Bunce 2000; Geddes et al. 1999; Milner and Mukherjee 2009; Przeworski and Limongi 1997), theories of consociational democracy (Andeweg 2000; Lijphart 1977), the authoritarian legacy of the state (O'Donnell, Schmitter, and Whitehead 2013), and the early process of institutionalization (Huntington 1968). These theories mostly outline how states develop democracy out of nothing, which is not the baseline condition of backsliding states looking to reform. I instead focus on work that has outlined the uncertainty in democratization, work on reform-centered explanations that track under what conditions autocrats gradually institute democratic institutions, and work on fledgling democracies in the wake of authoritarianism. While the context of this literature is both fledgling and emerging democracies, focus on fragile democracies on the verge of repair is a useful extension of these theories. I will summarize this literature and connect it to backsliding to show what we can learn.

In formerly authoritarian regimes, there is uncertainty when transitioning to democracy. The state goes from a certain authoritarian state to an uncertain "something else" (O'Donnell, Schmitter, and Whitehead 2013). This general approach is helpful for recently backsliding countries because of the uncertain choice of the new government. The choice of the next government — or current, in most democratization literature — is whether to democratize or not. Much of this work also focuses on the role of liberalization by increasing individual rights

and institutionalization of civil society, elections, and political parties or pacts. While these concepts are notable features of more traditional transitions, under damaged democracies these factors are less important. Most damaged democracies need to repair institutional arrangements like electoral rules and restore the civil society, but the creation of new institutions from the ground up is not necessary. The same goes for individual rights, which I assume are mostly present in order for a country to be considered a democracy in the first place. These sorts of choices will not be a significant focus of this study; instead I focus on the institutional democratic restoration a new executive can undertake in relation to existing institutions after a backslider yields. The uncertain choice is between the current subverted state or an uncertain arrangement of some or all restored institutions.

Beyond uncertainty, a major question in the democratization literature is why autocrats introduce democratic reforms that may eventually lead to their removal. This top-down approach to democratization is distinct from instances of popular rebellion and the early choices of a new government to democratize because it is often gradual and calculated so that the autocrat can safely maintain power. Reform centered explanations show that franchise expansion (Dahl 1971; Haggard and Kaufman 1999, 2016) and the break up of political monopolies (Gonzalez 2008) are sometimes employed by the authoritarians, albeit under conditions where it is safe to do so (Riedl et al. 2020; Schedler 2002). This cautious approach to expanding democratic institutions and norms is usually to add legitimacy to the autocrat and prevent the popular rebellions that would result in a bottom-up approach to democratization. Successors of initial backsliders inherit damaged institutions that they can either wield for their own purpose, or restore. Similar to reforming autocrats, restoration is more likely when the successor feels safe doing so. While autocrats reform to give greater legitimacy and prevent violent uprising, successors may restore democratic institutions to constrain themselves, and also future backsliding-minded leaders. For this to be possible, the

new rules would need to be either sufficiently beneficial, or be able to sufficiently constrain opponents in the future.

Other than top-down reform in autocracies, there are also historical instances of fledgling democracies with recent authoritarian legacies that are useful to look toward. In states with successors to autocrats, there is a similar trade-off for successors to backsliders: the desire to develop new democratic institutions in a bottom-up approach, and the concern of keeping the former autocrat or like-minded autocrats at bay during this fragile period. In a fragile setting like post-Fujimori Peru, where a personalist autocracy yielded to an unstable party setting, initial democratic hesitancy in order to ward off any potential autocratic actors set the groundwork for a meaningful democratic system accountable to voters (Kelly 2003; Levitsky 1999; Levitsky and Cameron 2003). This sort of democratization with initial guardrails has relevance to successors of initial backsliders. While post-backsliding countries start with fleshed-out institutions and some history of democratic competition, these considerations are important for successors to weigh in their decisions to restore institutions, or do so more carefully to avoid a future backsliding threat. Past understanding shows that the former opposition may attempt to aggressively institute measures to limit the role of past authoritarians in government (Linz and Stepan 1978). While this is probably normatively the most beneficial route to prevent autocratic resurgence, it may also lead to a cycle of autocratic actions if pursued beyond the short-term. This same dynamic is important to evaluate in terms of backsliders: immediate reform may be much more detrimental than a more gradual reform process.

Using this historical understanding of democratization and modernizing it to fit a context where democratic institutions, albeit weak ones, are in place, lends to some expectations of how we should expect the successors of initial backsliders to behave when they enter office. Ultimately, the most useful parts of the democratization literature for thinking about

post-backsliding choices are the considerations of security and risk. I believe security impacts choice in two ways: how secure a successor feels in office and whether restorative actions will be respected in the future. Meanwhile, risk is dependent on the nature of the damaged institutions themselves. Taking over powerful institutions today may be attractive, but if those powers can also be used against the successor in the future, restoring them may be the best course. I explore these concepts more in the next section where I detail the options available and how different conditions may lead the former opposition to behave once they regain power.

3.3 When Will Successors Restore Institutions?

After democratic backsliding occurs and an opposition party comes into power, there are several options that the successor can take. The successor can use weakened institutions to pass preferred policies and maintain power with greater certainty. This option may be attractive for successors, but it can set the stage for future resurgence of backsliders and perpetuate an undemocratic cycle that may lead to democratic failure. Conversely, the new government can prioritize repairing and restoring institutional constraints and norms, to the detriment of policy goals and the certainty that comes from more favorable electoral institutions. Under this option, a successor constrains actions today in hopes that future actions by an opponent are also constrained. Lastly, the successor may pursue a hybrid approach, blending the need for undoing existing policies, preventing future resurgence of backsliders, and restoring select democratic institutions. I distinguish between two groups of backsliding strategies that can be enacted: executive aggrandizement and the strategic manipulation of elections [10]. I generally assume that by the time a backslider is removed

[10]See table 1.2 in Chapter 1 for a more detailed description of these strategies and examples of actions that fall under both.

from power, both kinds of backsliding have been enacted and the successor has two different sets of damaged institutions that they can either repair or maintain.

Under executive aggrandizement, within-government constraints on the executive have been attacked. When these constraints are damaged, it is easier for the executive to exert its will over other elements of government that it does not traditionally control. In practice, this allows the ruling party to more easily implement their preferred policy program. Consider the case of Poland where the PiS aggressively reshaped judicial institutions and the composition of constitutional courts. A successor to the PiS may retaliate by executing similar measures in order to reshape courts in favor of the successor to be able to more easily pass policies, but adding to the volatility of the institution. Meanwhile, the successor could restore the judicial institutions by reforming court appointments in order to de-politicize the institutions and make the courts more balanced. In both cases, these measures would require significant action, and produce very different results, especially depending on the expected behavior of a future backslider.

Under the strategic manipulation of elections, the electoral institutions have been manipulated in ways to make removal of the incumbents or entry of opponents more difficult. These actions help make the ruling party more entrenched and able to maintain power for a longer time than otherwise possible by requiring a larger majority or plurality of voters to remove the ruling party than under completely free and fair elections. Consider the 2022 Hungarian Parliamentary elections where, under a mixed-member parliament, the constituency districts were drawn such that the ruling party won 87 of 106 districts with 52.5% of the vote[11]. A successor to Fidesz, should they overcome this hurdle and enter office, could retaliate by re-drawing lines to benefit the new governing party and disadvantage the

[11]This amounted to 82% of seats available in that tier. Under the proportional party list, the ruling coalition won 48 of the 93 seats (about 52% of the seats) with 54% of the vote in that tier.

initial backslider's party. Conversely, the successor can restore electoral institutions either by requiring more balanced districts, or only using the proportional tier so that different parties can win with a majority of the votes and no party is entrenched[12].

3.3.1 Factors that Feature in the Formal Model

Using this understanding of both sets of backsliding actions and how restorative and retaliation measures produce different incentives for successors to initial backsliders, I can detail some expectations that I will expand upon more rigorously in the formal model. I argue that there are three key factors that shape the choice of a successor to reform institutions or not: the *durability* of the restorative measures, and the *magnitude* of the subverted institutions, and *electoral safety*.

The durability of restoration is the most important component for the successor. If successors believe that restorative measures cannot be easily undone if the backslider's party regains power, then restoration is durable. Absent this durability, the incentives for successors to constrain themselves, only to have their opponents not play by the same rules, are not sufficient for restoration to be attractive. Consider a case where a successor chooses to restore an executive constraint by limiting extent to which executive actions bypass the legislature, as was the norm before the initial backslider's tenure. This action may prove to be detrimental if a future backslider does not mind breaking that norm for a second time. However, if the norm is expected to be durable in the future, it may prove to be a necessary and beneficial sacrifice.

[12]Empirically distinguishing between restoration and retaliation may be a slightly more complex endeavor considering that institutions are changing in favor of the successor regardless of the path. I assert that a democratic indicator should be able to capture the difference as one path is a notably more democratic development than the other. To other actors such as supporters of the initial backslider, there may be little distinction, so the difference between impressions of democratic quality and objective democratic quality, to the extent that it can be achieved, is an important area for empirical consideration in future work on this topic.

The magnitude of both kinds of backsliding actions — or the extent of the subversion of executive constraints and electoral institutions — is also an important determinant for the successor's decision. Small subversive measures can produce different incentives than wide-ranging attacks on democratic norms and institutions. This concept is also dependent on the kinds of institutions that have been damaged. For executive aggrandizement, the range of policies that can be enacted without executive constraints is larger than the range of policies available when constraints are in place. In the formal model, I use a parameter for the level of *polarization* of policy positions absent executive constraints to capture this concept. I consider polarization as the ideological distance between parties. In other words, if two polarized parties are competing against each other, a successor may be more inclined to constrain institutions because the alternative when out of power is significantly worse compared to the benefit in power.

With respect to the strategic manipulation of elections, I refer to the severity with which electoral institutions were attacked as the *subversion advantage*. In this case, assuming that every time one of two parties enters office, the incumbent can entrench itself in office to some degree, the difference between these two outcomes is the subversion advantage. In other words, the level of entrenchment when a backslider is in power minus the level of entrenchment when a successor is in power. A moderate advantage relative to the baseline today may not be worth it if a severe disadvantage is possible under a future backslider. I expect more subverted institutions are less appealing for a successor to maintain.

A successor has electoral safety when they believe that, absent any electoral manipulation, they have a sufficiently high probability of being re-elected. Electoral safety is an important component in the choice to restore institutions because it may provide the reforming successor more time in power should they choose to eschew subverted electoral institutions. This expectation impacts preferences for restoration. If a successor enters office and has a narrow

opportunity to act before yielding again, there may be different consequences for different institutions. For example, if a successor has a low level of electoral safety there may be more incentives to restore executive constraints in expectation that their opponent will spend more time in power later. Simultaneously, they may prefer the boost from damaged electoral institutions when in power and choose not to restore that set of institutions. I expect electoral safety to not have the same impact on both sets of backsliding actions for this reason.

Ultimately, the successor's choice is crucial for the future of a democracy. If a successor is policy and power motivated, then the desire to restore institutions may be low, and this can continue a cycle of backsliding that culminates in democratic failure. Even if a successor is interested in restorative measures, there are numerous institutional factors that can make restoration advantageous or not. The durability of restorative measures, the magnitude of the damaged institutions, and electoral safety all can shape the choice: sometimes in competing ways. The decision of a successor is important to fully understand because the removal of a backsliding incumbent may not be sufficient to stem the tide of backsliding.

3.4 The Formal Model

Consider two parties, B and D (a backslider and a democratic successor), with policy positions on a number-line where party D has a bliss point at $X_D = 0$, and party B at $X_B = c$ where $c > 0$. At the beginning of the game, party D beat incumbent party B which had instituted backsliding measures. In the wake of backsliding under party B there are two different ways institutions have been attacked. There were attacks on constraints against the executive or executive aggrandizement, and attacks on the electoral system or the strategic manipulation of elections.

In the absence of executive constraints, parties implement their bliss point, but when executive constraints are in effect, there is some inner range of acceptable policies that can be enacted: $[a, b]$ where $0 < a < b < c$ [13]. Under the backslider, the executive constraints were removed. Executive constraints are either in effect, or not, but if they are in effect, the magnitude of the constraints is impacted by the range of a and b. The utility of each party is the distance between their bliss point and the existing policy X_t at time t, subject to a quadratic penalty. The game begins at time $t = 0$. Payoffs are subject to a common discount factor δ. The basic payoffs per period are:

$$u_D(X_t) = -(X_t)^2$$

$$u_B(X_t) = -(X_t - c)^2$$

I conceptualize c as the *magnitude* of executive aggrandizement because it is the range of acceptable policies, whereas under executive constraints, the comparative value is $b - a$. An alternative way to interpret c is the magnitude of *polarization*. When executive constraints are in place, the parties select policies from the range of $[a, b]$, but absent constraints, the region increases to $[0, c]$. As c increases, the policy penalty for D when B is in office becomes higher, so the constrained policies may become more appealing.

Typically there is uncertainty around elections due to robust democratic electoral institutions. The winner of free and fair elections at time t is determined by a probability, where under normal conditions B wins with probability $p_N \in (0, 1)$ without loss of generalizability. I use $1 - p_N$ to represent D's *electoral safety*. As p_N increases, electoral safety decreases. Electoral institutions have been subverted by the backslider, so that the office holder has a higher probability of re-election than they would under a free and fair election. Under D, the

[13] This configuration of policies is inspired by Alesina (1988).

probability of B's victory is p_D and under B it is p_B, where $0 < p_D < p_N < p_B < 1$. The magnitude of subverted elections is conceptualized as the *subversion advantage* or $p_B - p_D$. The subversion advantage helps represent the magnitude of the difference of electoral fortunes depending on the party in power[14].

At the beginning of the new government, both of these sets of institutions, previously used against D, can now be leveraged in its favor. Party D has separate choices to restore or retaliate on one or both of executive aggrandizement and the strategic manipulation of elections. When restorative measures are *durable*, D's choices remain in place for the rest of the game, even when B returns to power. Conversely, when restorative measures are *not durable*, D's choices remain in place only until B returns to power for the first time. After, B attacks restorative measures and institutions revert to their subverted states for the remainder of the game.

In the analysis, I only analyze the restoration decision of D, but this requires taking into account the future actions of B when it regains office. I achieve this by constructing a discounted infinite-horizon continuation payoff for D's decisions under the condition of durable or non-durable restoration. In other words, I examine an optimistic and a pessimistic equilibrium in an infinite-horizon game respectively, but instead of mapping B's payoffs, I proceed as though its optimal action takes one of two extreme forms. In the case of *durable* reform, B finds it rational to maintain D's restoration of whichever institutions it chooses. This occurs if B's payoffs and expectations made B willing to pursue a tit-for-tat or grim trigger pattern of future reforms and backsliding: as long as D maintains either or both types of restoration, B will respond in kind. In the case of *non-durable* reform, B is not cooperative with restorative measures, so it will return to backsliding whenever it returns to office and

[14]If manipulation is equivalent to gerrymandered districts, then p_D represented the prospects of D being able to re-draw its own tilted electoral districts. If manipulation is control of election judges or other infrastructure, then p_D helps estimate those same institutions captured by D and party D partisans.

regardless of D's restorative measures. After this, I assume D will also give up on restoration when it regains office for a second time.

Finally, I use an infinite horizon game for two key reasons. First, D must consider its actions both in terms of present utility and in the event that B returns to power. It is the risk of B re-entering power that drives the desire of D to restore institutions. Second, while institutions are only subject to D's choice and durability in the model, backsliding is a multi-step process that occurs over the course of multiple terms in office. Considering the future horizon of institutions gives greater realism to the model and provides useful functionality for the sake of solving. The events proceed as follows:

1. At time $t = 0$, party D wins the election and inherits subverted institutions.
2. Party D chooses to restore one, both, or neither institution. If restorative measures are durable, these choices are fixed for the remainder of the game. If measures are not durable, then restored institutions will be subverted permanently after B's first victory.
3. If executive constraints are in effect, party D implements $X_0 = a$. Otherwise $X_0 = 0$.
4. If elections are free and fair, $\mathbb{P}(\text{B wins}) = p_N$. Otherwise p_D. Leader for $t = 1$ chosen.
5. Under executive constraints, if D wins it enacts $X_t = a$, and if B wins it enacts $X_t = b$. Absent executive constraints, D enacts $X_t = 0$ and B enacts $X_t = c$.
6. Payoff from X_t discounted by δ^t.
7. Under free and fair elections, the winner for $t = t + 1$ is determined by p_N. Absent free and fair elections, the winner of the next period is determined by p_D if D holds office and p_B if B holds office.
8. Steps 5-7 are repeated infinitely.

3.4.1 Continuation Values and Durability

The choice to restore institutions is based on D's continuation value (forward-looking payoff for the remainder of the game) given its actions and durability. I detail the continuation values first for durable restorative measures and then for non-durable restorative measures.

There are four different potential combinations of restored or not institutions and each has its own set of continuation values. I used the symbol V for when both institutions are restored, W for when only electoral institutions are restored, Y for when only executive constraints are restored, and Z for when neither set of institutions is restored. For each of these symbols, there is a different result depending on which party is in power. For example, V^D is the continuation value for party D when it is in power and V^B when B is in power and all institutions are restored. The continuation values are listed in terms of each other such that V^D is in terms of itself and V^B. Continuation values for each combination of institutions V, W, Y, and Z are in table 3.1

Table 3.1: D's Continuation Values for Each Combination of Institutions (Durable)

Notation	Interpretation	CV given the current office-holder
V	The payoff when both sets of institutions are restored	$V^D = u_D(a) + \delta[(1-p_N)V^D + p_N V^B]$
		$V^B = u_D(b) + \delta[(1-p_N)V^D + p_N V^B]$
W	The payoff when only electoral institutions are restored	$W^D = u_D(0) + \delta[(1-p_N)W^D + p_N W^B]$
		$W^B = u_D(c) + \delta[(1-p_N)W^D + p_N W^B]$
Y	The payoff when only executive constraints are restored	$Y^D = u_D(a) + \delta[(1-p_D)Y^D + p_D Y^B]$
		$Y^B = u_D(b) + \delta[(1-p_B)Y^D + p_B Y^B]$
Z	The payoff when neither set of institutions are restored	$Z^D = u_D(0) + \delta[(1-p_D)Z^D + p_D Z^B]$
		$Z^B = u_D(c) + \delta[(1-p_B)Z^D + p_B Z^B]$

When restorative measures are *not durable*, the continuation values are not in terms of themselves the same way because only D plays by the new institutional rules. After B's first

return to power, any restored institutions revert to the subverted state. These continuation values are in table 3.2. For an example of how these differ, consider V^D in table 3.1 and V^{D*} in table 3.2. V^D is represented in terms of itself and V^B or the terms where both institutions are restored depending on the office holder. When restorative measures are not durable, V^B will never occur because B will immediately subvert institutions when it returns to office. Therefore, V^{D*} is represented in terms of itself as long as it maintains power uninterrupted, but after B wins for the first time, is in terms of Z^B when B is in power and Z^D when D returns to power.

Table 3.2: D's Continuation Values for Each Combination of Institutions (Non-Durable)

Notation	Interpretation	CV given the current office-holder
V	The payoff when both sets of institutions are restored	$V^{D*} = u_D(a) + \delta[(1-p_N)V^{D*} + p_N Z^B]$
W	The payoff when only electoral institutions are restored	$W^{D*} = u_D(0) + \delta[(1-p_N)W^{D*} + p_N Z^B]$
Y	The payoff when only executive constraints are restored	$Y^{D*} = u_D(a) + \delta[(1-p_D)Y^{D*} + p_D Z^B]$
Z	The payoff when neither set of institutions are restored	$Z^D = u_D(0) + \delta[(1-p_D)Z^D + p_D Z^B]$
		$Z^B = u_D(c) + \delta[(1-p_B)Z^D + p_B Z^B]$

I use these two sets of continuation values to derive the conditions for equilibrium first for non-durable restoration, and then for durable restoration. I introduced these in the reverse order because the in durable the institutions do not change, so it is easier to set up. Deriving the conditions for equilibrium for the durable case is much more involved though, so I begin with the results for non-durable restorative measures. The ways in which V^{D*}, W^{D*} and Y^{D*} differ from V^D, W^D and Y^D show the impact of the durability of restorative measures and how it impacts the willingness of a successor to restore institutions.

3.5 When Restoration is Not Durable

I begin by solving the more simple case where restorative measures are not durable. In this case, given any choice of D for restorative measures, B will always undo them when it regains power. After this occurs, D reverts to the damaged institutions should it regain office as a sort of grim trigger response. This scenario is the most pessimistic approach to restorative choices. In table 3.2, I detailed continuation values absent durable restoration. To evaluate when restoration is appealing, I compare the continuation values when D restores one or both sets of institutions, V^{D*}, W^{D*} and Y^{D*}, to the continuation value when D does not restore any: Z^D.

In each of the cases where D restores institutions, it gives up either or both the policy 0 for a and the probability that B wins of p_D for p_N for as long as D maintains power, and after which point, the values revert to the less democratic alignment. I show that it is *never* advantageous for D to restore institutions when they are not durable.

I first prove why reforming electoral institutions is never preferable by comparing the continuation value W^{D*}, where D enacts non-durable electoral restoration, and Z^D, where no institutions are restored. For W^{D*}, let T be the random variable giving the number of periods D stays in office when pursuing electoral restoration. In each period $1, ..., T$ the policy choice will be 0, and D's payoff will be $u_D(0)$. T has a negative binomial distribution with parameter $1 - p_N$. For Z^D, let T' be the number of periods D stays in office if D does not pursue any institutional restoration. In each period $1, ..., T'$, the policy choice will be 0, and D's payoff will be $u_D(0)$. T' has a negative binomial distribution with parameter $1 - p_D > 1 - p_N$.

D's expected payoff given T, when D restores electoral institutions, is

$$\begin{aligned}\mathbb{E}(W^{D*}|T) &= \mathbb{E}\sum_{t=0}^{T-1}\delta^t u_D(a) + (\mathbb{E}\delta^T)Z^B \\ &= u_D(0)\,\mathbb{E}\left[\frac{1-\delta^{T-1}}{1-\delta}\right] + (\mathbb{E}\delta^T)Z^B.\end{aligned}$$

Similarly, when D maintains subverted institutions given T', the expected payoff is

$$\mathbb{E}(Z^D|T') = u_D(0)\,\mathbb{E}\left[\frac{1-\delta^{T'-1}}{1-\delta}\right] + (\mathbb{E}\delta^{T'})Z^B.$$

Each of these expressions is a weighted sum between the $u_D(0)$ term and the common constant Z^B. I then need to compare the expectations of the terms $\mathbb{E}\delta^T$ and $\mathbb{E}\delta^{T'}$. Both T and T' are distributed negative binomial such that: $T \sim NB(r=1, p=1-p_N)$ and $T' \sim NB(r=1, p=1-p_D)$. The general PMF for a negative binomial variable k where k is the number of successes and $r=1$ is the number of failures until the institutions revert to the subverted state is:

$$\mathbb{P}(K=k) = \binom{k+r-1}{k}(1-p)^r p^k.$$

Substituting the parameter values for T:

$$\mathbb{P}(T=k) = \binom{k}{k}(p_N)(1-p_N)^k = (p_N)(1-p_N)^k.$$

Thus the expectation of δ^T is

$$\mathbb{E}\delta^T = \sum_{k=0}^{\infty}\delta^k(p_N)(1-p_N)^k = \frac{p_N}{1-\delta(1-p_N)}, \tag{3.1}$$

and similarly

$$\mathbb{E}\delta^{T'} = \frac{p_D}{1-\delta(1-p_D)}. \tag{3.2}$$

I set (4) greater that (5) to verify which expectation is greater:

$$\frac{p_N}{1-\delta(1-p_N)} > \frac{p_D}{1-\delta(1-p_D)}$$

To verify this, I cross multiply and cancel common terms:

$$p_N(1-\delta(1-p_D)) > p_D(1-\delta(1-p_N))$$

$$p_N > p_D.$$

Therefore $\mathbb{E}\delta^{T'} < \mathbb{E}\delta^{T}$, and since the weight on the more preferred $u_D(0)$ term — which is better than Z^B because $u_D(0) = 0$ and $Z^B < 0$ — is higher when $\mathbb{E}\delta^{T'}$ is lower, the expected payoff for Z^D is higher than W^{D*}. That is, for all parameter values, $W^{D*} < Z^D$, so in the absence of durability, institutional restoration of elections is never preferable. Now, to prove restoring executive constraints is never preferable. I look at Y^{D*}, which benefits from T', but implements the policy a:

$$\mathbb{E}(Y^{D*}|T') = u_D(a)\,\mathbb{E}\left[\frac{1-\delta^{T'-1}}{1-\delta}\right] + (\mathbb{E}\delta^{T'})Z^B.$$

In this case, the weight on the $u_D(\cdot)$ term is the same between the two cases Z^D and Y^{D*}, but the $u_D(\cdot)$ term is higher in Z^D than in Y^{D*} and the same weights are on the shared, negative Z^B term. Because $u_D(0) > u_D(a)$ because $u_D(0) = 0$ and $u_D(a) < 0$, the expected payoff for Z^D is higher than Y^{D*}. So, for all parameter values, $Y^{D*} < Z^D$, and institutional restoration of executive constraints is never preferable. Finally, when combining

both executive and electoral restoration in V^{D*}, D is worse off for both sets of institutions compared to Z^D:

$$\mathbb{E}(V^{D*}|T) = u_D(a)\,\mathbb{E}\left[\frac{1-\delta^{T-1}}{1-\delta}\right] + (\mathbb{E}\delta^T)Z^B$$

In this case, D is giving up both the advantage in winning probability that comes from T' and the better policy outcome that comes from $u_D(0)$. This means that both the $u_D(\cdot)$ term is better in Z^D, but also the weight on the term is higher because $\mathbb{E}\delta^{T'} < \mathbb{E}\delta^T$. So, the expected payoff for Z^D is higher than V^{D*}. So finally, for all parameter values, $V^{D*} < Z^D$, so in the absence of durability, institutional restoration of elections and executive constraints is never preferable. This means that for any combination of restored institutions, D is worse off than it would be maintaining subverted institutions.

Proposition 3 (Reforming when restoration is never durable) *When restoration is not durable, there is no combination of parameters where D prefers to restore any institutions.*

Proposition 3 states that for any combination of non-durable restorative measures, D is strictly worse off engaging in restorative measures. This result for non-durable institutions is useful as a fringe case, but while it is likely that B will not want to abide by the new restorative measures D can passes, it is unlikely that trampling existing institutions for a second time would go without consequences. To complement the non-durable case, in the next section I rigorously explore when restorative measures are completely durable and the different parameter configurations that produce preferences for or against restoration.

3.6 When Restoration is Durable

When restorative measures are durable, there is no guarantee that restoration is *always* disadvantageous. Under durable restorative measures, B can never revert to backsliding so it plays by the rules set by D in perpetuity. This scenario is the most optimistic setting where any restorative measures are certain to hold and should be a "high water mark" for the prospects of restoration. In order to assess choice under durable restorative measures, I first need to re-write the continuation values in terms of the parameters in order to compare across the different institutional combinations. I do this using the values from table 3.1 and I solve for each continuation value as follows: for V^D, (1) solve V^D in terms of V^B, (2) solve V^B in terms of V^D, (3) substitute solution of (2) into (1) and solve for V^D. This produces a solution for the continuation value V^D using only parameters:

$$V^D = \frac{u_D(a)(1-\delta p_N) + \delta p_N u_D(b)}{1-\delta}$$

A similar process for the other continuation values yields:

$$W^D = \frac{u_D(0)(1-\delta p_N) + \delta p_N u_D(c)}{1-\delta}$$

$$Y^D = \frac{u_D(a)(1-\delta p_B) + \delta p_D u_D(b)}{(1-\delta)(1-\delta p_B + \delta p_D)}$$

$$Z^D = \frac{u_D(0)(1-\delta p_B) + \delta p_D u_D(c)}{(1-\delta)(1-\delta p_B + \delta p_D)}$$

Using these values, I can analyze the factors that would determine the best choice for the successor. The comparisons that I focus my analysis on are: $V^D > W^D$, $V^D > Y^D$,

$W^D > Z^D$, and $Y^D > Z^D$. I do not consider $V^D > Z^D$ because it is too complicated to solve directly, but I can leverage the transitive application of conditions to get at this. Consider that if $V^D > W^D$ and $W^D > Z^D$ or $V^D > Y^D$ and $Y^D > Z^D$ are true, then we can state $V^D > Z^D$, though these are only sufficient conditions. I show the substantive interpretation for each of these comparisons in table 3.3.

Table 3.3: Conditions to Solve for Nash Reversion and Interpretation

Notation	Interpretation
$V^D > W^D$	When only executive institutions have been compromised, D is better off restoring executive constraints than not.
$V^D > Y^D$	When only electoral institutions have been compromised, D is better off restoring free and fair elections than not.
$W^D > Z^D$	When both sets of institutions have been compromised, D is better off only restoring free and fair elections than not.
$Y^D > Z^D$	When both sets of institutions have been compromised, D is better off only restoring executive constraints than not.

I then solve each of these expressions for a common parameter: the discount factor δ. For each set of inequalities, I simplified when possible, and solved for delta. For example, in the case of $W^D > Z^D$, I began with the following inequality:

$$\frac{u_D(0)(1-\delta p_N)+\delta p_N u_D(c)}{1-\delta} > \frac{u_D(0)(1-\delta p_B)+\delta p_D u_D(c)}{(1-\delta)(1-\delta p_B+\delta p_D)}$$

On the left hand side is the continuation value when only executive constraints were restored, and on the right hand side the continuation value when neither set of institutions was restored. Significant algebra is needed to solve for δ and simplify, but I repeated this process for the other inequalities. I specifically solve for δ because it is present in all of the inequalities of interest and it can be solved for for the purpose of doing comparative statics. No other parameter had both of these properties. An alternative way to interpret δ is the

level of *urgency* at the beginning of D's term, where a low level of δ shows a high level of focus on short-term goals as opposed to long-term impacts.

I approach the equilibrium in terms of Nash reversion, or the best choice for D looking forward regardless of the history of the game. In the case of $V^D > W^D$ for instance, electoral restoration took place previously, but that history is not important in D's decision. I simplified these expressions to create cut-offs and plugged-in the parameters a, b, and c into the $u_D(X)$ equation to solve for a critical discount factor: δ^*, where when the discount factor is at or above this level, restoration is preferred. As δ^* increases, the corresponding restoration condition is harder to satisfy.

Lemma 3 (Successor Cutoffs for Preferring Restoring Institutions) *The critical discount factor threshold, $\delta \geq \delta^*$, in which successors prefer restoring institutions, for each choice of interest is:*

i. $V^D > W^D$: $\quad \delta^* = \frac{a^2}{p_N(a^2-b^2+c^2)}$

ii. $V^D > Y^D$: $\quad \delta^* = \frac{p_N-p_D}{p_N(p_B-p_D)}$

iii. $W^D > Z^D$: $\quad \delta^* = \frac{p_N-p_D}{p_N(p_B-p_D)}$

iv. $Y^D > Z^D$: $\quad \delta^* = \frac{a^2}{p_B a^2+p_D(c^2-b^2)}$

I use the thresholds from lemma 3 to analyze how different parameters impact the position of the threshold, and identify when it is within the range of $(0, 1)$ so restoration can be the preferred path. For each set of results, I take the partial derivatives of each of δ^* with respect to my parameter of interest. In both $V^D > W^D$ and $Y^D > Z^D$, D is considering whether or not to restore executive constraints, but in one of these cases electoral institutions

remain subverted ($Y^D > Z^D$). Unlike the decision to restore executive constraints, δ^* is the same for both $V^D > Y^D$ and $Y^D > Z^D$, because the policies $(0, a, b, c)$ cancel out. I consider how these equations differ in each of the sets of results. These results can be interpreted in terms of the parameter regions in which restoration is or is not preferred and how these parameters regions grow or shrink depending on δ and the parameters of interest.

The parameters of interest for each of the three sets of results are divided into the magnitude of the subverted institutions for both kinds of backsliding, and finally electoral safety. The magnitude of subverted executive constraints is measured by c or the policy that B implements absent executive constraints. I refer to this as the polarization parameter in my results. The more polarized the two parties, the greater the potential policy penalty when out of power. I assume that without executive constraints, a party implements that bliss point. As that bliss point becomes more extreme, I expect that it will require additional subversion of executive constraints. The extent of electoral manipulation is measured by the *subversion advantage*, $p_B - p_D$, or the difference between B's probability of victory when in power minus the probability of victory for B when D is in power. Finally, I assess electoral safety, represented by p_N or the probability of B's victory absent electoral subversion regardless of the office holder. For the results for each of these parameters of interest, I graphically represent the various δ^* curves. In these accompanying figures to the main results, I alter various parameters to provide additional understanding of how various factors interact and impact preferences for restoration.

3.6.1 Polarization and Restoration

When D considers whether or not to restore executive constraints, the parameter of interest is c or the polarization between the two parties when the executive is unconstrained. In this case, the magnitude of executive subversion is reliant on whether policy will bounce between

an inner constrained range of $[a, b]$ or the unconstrained $[0, c]$. When holding the inner range fixed, the larger the second range, the greater risk for maintaining the damaged institutions. In order to assess the role of the size of executive subversion on restoration preferences, I take the δ^* from Lemma 3 for both $V^D > W^D$ and $Y^D > Z^D$. In $V^D > W^D$ elections are free and fair regardless, and in $Y^D > Z^D$ elections are subverted regardless. I take the partial derivative of these δ^* with respect to c to get the results for Proposition 4.

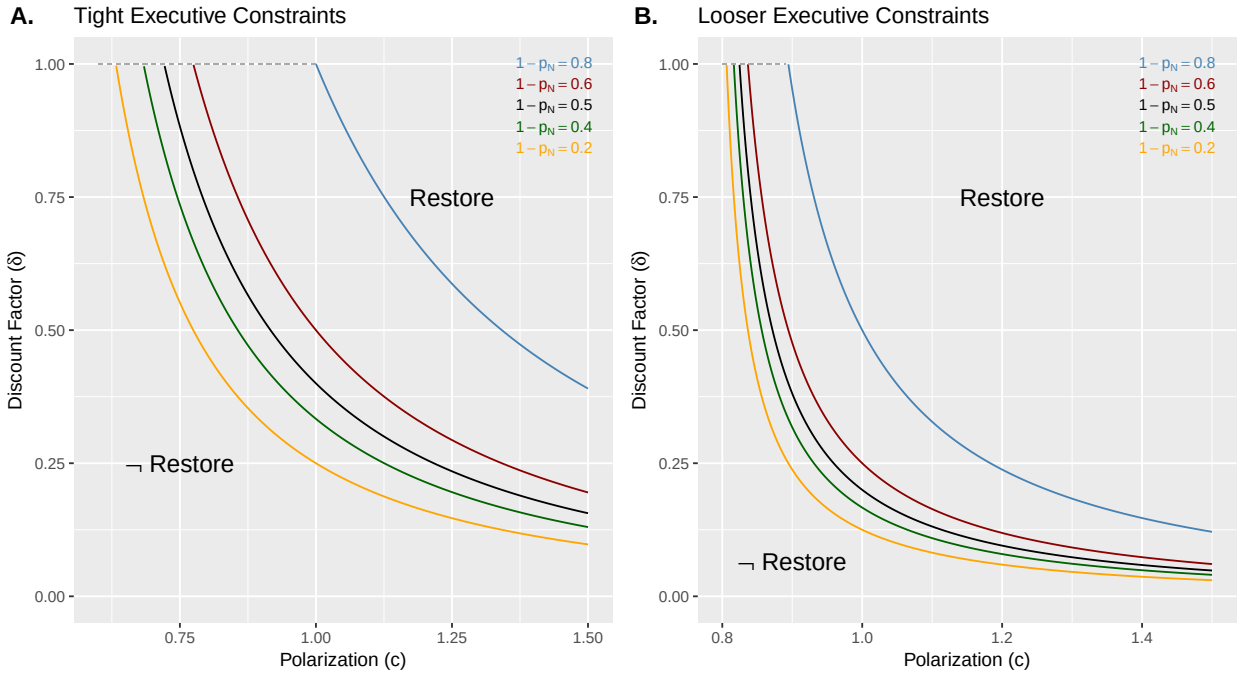
Proposition 4 (Polarization and Restoring Executive Constraints) *The critical discount factor needed for a successor to prefer restoring executive constraints is decreasing in the polarization parameter, c.*

In other words, higher levels of polarization require a lower discount factor to induce the democratic restoration of executive constraints. This result is true both when election are restored or not. In terms of the parameter space, a higher level of polarization produces a larger region where restoration is preferred. This means that higher polarization makes it less worthwhile to maintain executive aggrandizement and this is true regardless of the size and position of inner range of b and c, though these values will effect the position the threshold itself. The risk of more polarized outcomes can be an effective deterrent to maintaining executive aggrandizement when the choice of D is durable.

While proposition 4 is helpful for evaluating polarization, there are other parameters that can affect this relationship. I am specifically interested in the impact of the electoral institutions or p_N when elections are free and fair and p_B and p_D when elections are subverted. In each of those cases, I am also interested in the inner range of policy when the executive is constrained $[a, b]$, and its position relative to c. I first consider D's preferences when elections are not subverted ($V^D > W^D$) and then consider the choice when elections are subverted

$(Y^D > Z^D)$. I plot the δ^* curves and alter the parameters of interest to show how the different parameters varying together impacts preferences for restoration.

Figure 3.1: The Effect of Electoral Safety $(1 - p_N)$ on the δ and c Combinations Where D Favors Restoration.

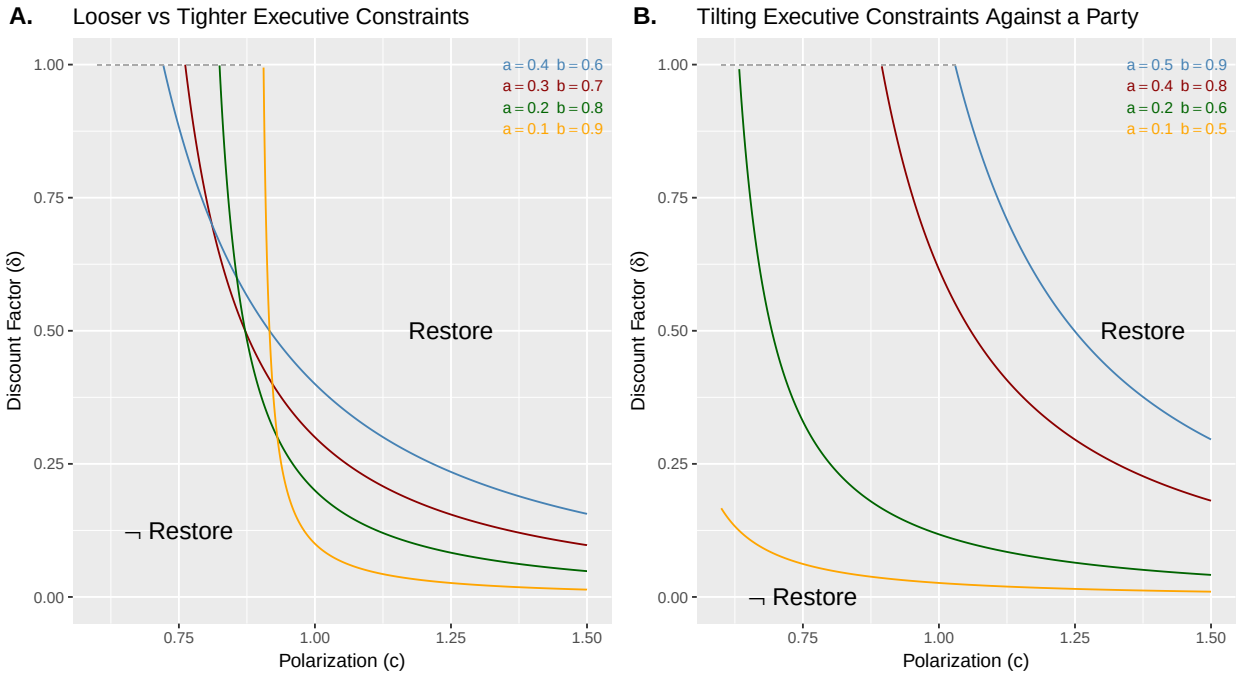


Note: δ^* curves for D's choice to restore executive constraints when electoral subversion has already been restored $(V^D > W^D)$. In both panels I manipulate the electoral safety $(1 - p_N)$ of D. In this case, the higher safety numbers mean a lower probability of D winning election. In panel A , the inner range of policy under a constrained executive is smaller ($a = 0.4$ and $b = 0.6$). In panel B, I expand this inner range ($a = 0.2$ and $b = 0.8$).

In figure 3.1 I show the impact of electoral safety $(1 - p_N)$ on polarization and the choice to restore executive constraints. As p_N increases, D's electoral prospects are worse. In panel A, I show the impact of electoral safety with a narrow range of constrained policies if executive constraints are restored (small range of a and b) and in panel B I show how those same values of $1 - p_N$ impact the curves with a larger range of constrained policies. As electoral safety decreases, the region where restoration is preferred grows larger, and this is especially the case when the executive constraints allow a wider range of policies. In both

panels, when $c = 1$, the constrained range of policies — $[a, b]$ — is not biased in favor of either party relative to their bliss points. However, when $c > 1$, executive constraints are biased in favor of D, and when $c < 1$, executive constraints are biased in favor of B. Because it is unlikely that B would have removed executive constraints that benefited it, and rising polarization is a defining characteristic of the backsliding literature (Graham and Svolik 2020), it is more useful reference the portion of the figure where $c > 1$.

Figure 3.2: The Effect of Executive Constraints (a and b) on the δ and c Combinations Where D Favors Restoration.

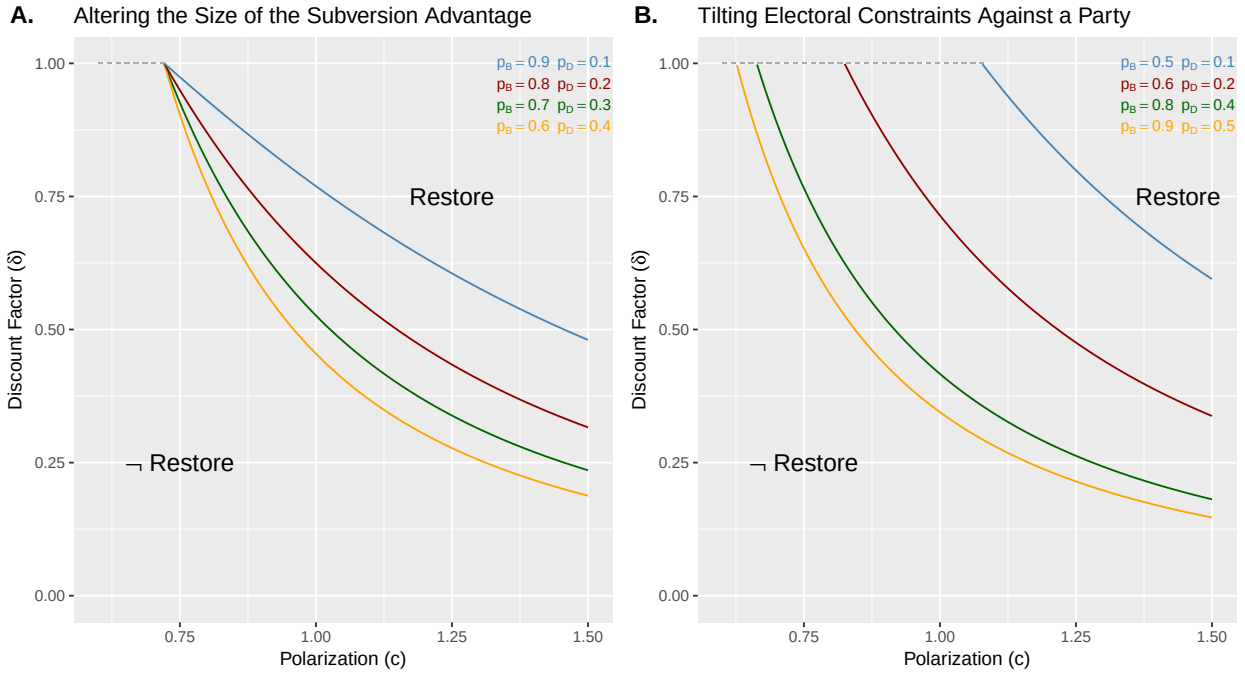


Note: δ^* curves for D's choice to restore executive constraints when electoral subversion has already been restored ($V^D > W^D$). In both panels I explore the inner range of constraints against the executive (a and b). In panel A I show the impact of increasing the size of this internal range. In Panel b I show the impact of "tilting" the range in favor of either D or B. In both panels I fix $p_N = 0.5$.

In figure 3.2 I show the impact of altering the range of acceptable policies (a and b) when executive constraints are restored. In panel A, I progressively expand the range of acceptable policies when the executive is constrained. Similar to above, when $c = 1$, the

range is not biased in favor of either party relative to their bliss points. As this range grows larger, so does the region where restoration is preferred and this is most observable when $c > 1$. In panel B, I explore when the inner range is "tilted" for or against one of D or B. In this case, the more tilted the constrained institutions are in favor of D, the larger the region where D prefers restoring. The converse is true for when institutions are "tilted" in favor of B. When I consider the impact of executive constraints on the choice to restore executive constraints when electoral institutions are subverted regardless ($Y^D > Z^D$), the results are largely similar[15].

Figure 3.3: The Effect of Damaged Electoral Institutions (p_D and p_B) on the δ and c Combinations where D Favors Restoration.



Note: δ^* curves for D's choice to restore executive constraints when electoral subversion is present ($Y^D > Z^D$). In panel A I show the impact of increasing the size of the subversion advantage ($p_B - p_D$) when symmetrical around 0.5. In panel B I show the impact of "tilting" damaged electoral institutions in favor or against D. In both panels I fix $a = 0.4$ and $b = 0.6$.

[15]There are minor differences driven by the short-term benefit D has in retaining office with probability $p_D - 1$ instead of $p_N - 1$, but because these figures are mostly the same, I exclude them here.

Lastly, I consider the choice to restore executive constraints when electoral institutions are subverted ($Y^D > Z^D$). In figure 3.3, I assess the impact of damaged electoral institutions (p_D and p_B) on the choice to restore executive constraints. In panel A, I show the impact of the size of the subversion advantage ($p_B - p_D$) on the preference to restore democratic institutions. The larger the subversion advantage, the larger the difference in outcomes when in power and out of power. As the subversion advantage decreases, the region where restoration of executive institutions is preferred grows larger. This is mostly driven by the probability that B can win while D is in power (p_D). When this value is very low, as it is the larger the subversion advantage, the the risk of B winning and instituting an extreme policy (c) is also low. Conversely, as the subversion advantage shrinks, the threat of the extreme policy rises and the region where restoration is preferred expands.

In panel B, I hold the subversion advantage constant and consider "tilted" electoral institutions. In this case, the more tilted the institutions are to D, the smaller the region where restoration is preferred. In practice, this unlikely because B would not have an incentive to damage electoral institutions in a way to benefit D. Instead, it is more helpful to consider the δ^* curves where the damaged electoral institutions are tilted in favor of B. Even when the inner range of policies is skewed in favor of B, as it is when $c < 1$, there are incentives for D to restore executive constraints when electoral institutions are biased against it.

The effect of polarization on the preference to restore executive constraints can be altered by electoral rules or the constraints themselves, but is a positive relationship regardless. A successor can even settle for unappealing constrained executive institutions if the initial backslider is especially extreme. Consider two sets of competing parties: one highly polarized set, and another set with moderate polarization. Restoration is more likely to benefit a successor in the highly polarized case because even a sub-optimal set of executive constraints would be much more preferred than the alternative of wild swings in policy depending on

the leader. Under more moderate polarization, the risk is not as great, so the executive constraints may be there to stay. These results are important for the policy considerations of a successor and restoring executive constraints, because retaining damaged executive institutions can be a disadvantageous choice under some conditions. Next, I consider the restoration of electoral institutions.

3.6.2 Electoral Institutions and the Subversion Advantage

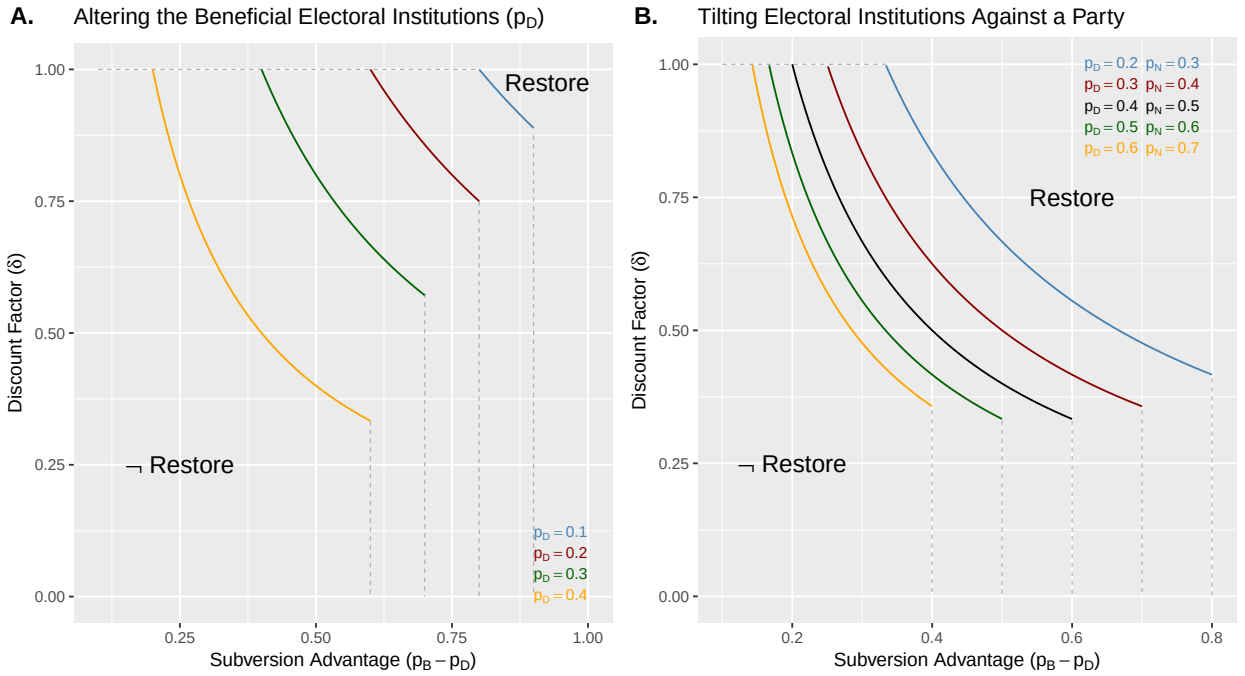
Next, I evaluate the choice to restore free and fair elections and how the subversion advantage $(p_B - p_D)$ impacts this choice. When the subversion advantage is high, it means that while in power, D benefits from especially favorable electoral institutions (p_D), but when out of power, the punishment is more severe (p_B). This additional level of risk is important to consider when the alternative is the fixed value p_N that exists within the range of (p_D, p_B) and is reliable and more moderate. Unlike the previous proposition, the two δ^* equations from lemma 3 where D considers restoring fair electoral institutions are identical. Also unlike the results from the previous proposition, where the subversion advantage interacted with polarization to increase preferences for restoration of executive constraints, polarization has no impact on restoring electoral institutions. I take the partial derivative of this shared δ^* with respect to $p_B - p_D$ to get the results for the subversion advantage.

Proposition 5 (The Subversion Advantage and Restoring Electoral Institutions) *The critical discount factor needed for a successor to prefer restoring free and fair elections is decreasing in the subversion advantage, $p_B - p_D$.*

Similar to the polarization result, the larger the magnitude of electoral subversion, the lower the discount factor needed to prefer restoring free and fair elections. In other words, as the subversion advantage increases, the region where restoration is preferred grows larger.

When the magnitude of the subversion advantage is higher, the potential risk is also higher and the successor increasingly prefers to restore free and fair elections. Next, I consider how the component parts of the subversion advantage (p_D and p_B) and its alternative (p_N) impact the δ^* curves and the regions where D prefers to restore electoral institutions.

Figure 3.4: The Effect of Electoral Institutions (p_D and p_N) on the δ and $p_B - p_D$ Combinations Where D Favors Restoration.



Note: δ^* curves for D's choice to restore electoral institutions regardless of executive constraint restoration ($V^D > Y^D$ and $W^D > Z^D$). In panel A I show the impact of increasing the size of the subversion advantage ($p_B - p_D$) when symmetrical around 0.5. In panel B I show the impact of "tilting" damaged electoral institutions in favor or against D while maintaining $p_N - p_D = 0.1$. In both panels I fix $a = 0.4$ and $b = 0.6$. In both panels, beyond the right-bound of each curve, there are no parameter space for that given curve.

In figure 3.4, I consider the impact of the probability B wins when D is in power (p_D) and the probability B wins under free and fair elections (p_N)[16]. In panel A, I fix electoral safety to a "fair" level ($p_N = 0.5$) and explore how progressively better odds for D's re-election

[16]Because I vary the subversion advantage in one of the axes and fix the value of p_D, the value of p_B changes depending on the size of the subversion advantage.

when in office (p_D) and subversion advantage ($p_B - p_D$) impact the preference to restore electoral institutions. The length and position of the δ^* curves are different because as p_D decreases, the potential *range* of the subversion advantage decreases. These curves show that the better the odds for D when in control of subverted institutions (smaller p_D), the smaller the region where restoration is preferred.

The left and right bound of each of the δ^* curves as they appear are also important reference points to interpret the results. First, the left bound, where the curve intersects with $\delta = 1$, is the value of of the subversion advantage when the subverted institutions are not biased in favor of either party. Beyond this point, lower subversion advantages mean that the subverted institutions are biased in D's favor, so there is no scenario where D restores electoral institutions. These values of the subversion advantage are unlikely, because the institutions were subverted by and are likely skewed in favor of B. Second, the right bound of the curve is the highest possible value of the subversion advantage given p_D, or when p_B approaches 1. In this case, even if D knows that it will never win again after the first time B gains power, the short term benefit of maintaining power can be high enough that D will not restore. Because this is an extreme case, this is where D is most likely to restore. However, with a sufficiently low level p_D, perhaps in a country where the government has significant controls of elections, restoration may not be the best choice unless D is focused more on long-term impacts as opposed to short-term gains.

In panel B, I vary the restored electoral institutions (p_N) and the subverted institutions when D is in power (p_D) to show the impact of tilted institutions when electoral safety varies. The black curve in panel B is identical to the yellow-orange curve in panel A, and helps show the impact altering the two sets of institutions together on the choice of restoring electoral institutions. As both subverted and restored institutions are more biased against D, the region where restoration is preferred grows. In each δ^* curve, when the subversion advantage

is equal to 0.2, the subverted institutions are not biased to either party relative to the baseline probability under restored institutions. In the two δ^* curves where the subverted electoral institutions are biased towards D (the red and blue curves), there is no possible value of δ for restoration to be preferred. Therefore, both the degree of tilt of the subversion advantage relative to electoral safety (p_N) and the bias of electoral safety itself impacts preferences for restoration of electoral institutions.

When combining the results on restoring electoral institutions together, the impact of damaged and undamaged institutions is just as important as any biases they contain. Election subversion likely will never favor the successor and this substantially impacts preferences to restore these institutions. Consider the cases of electoral manipulation where parties or candidates are banned or groups of people are disenfranchised like the case of Kurdish voters in Turkey. It would be difficult for a successor to symmetrically wield those powers against the other side, so subverted institutions will very likely be tilted in favor of the initial backslider. Therefore, in cases like these, successors should be very likely to restore electoral institutions when those actions cannot be easily undone. Next, I conclude the results with the findings for electoral safety.

3.6.3 Electoral Safety and Restoration

Lastly, I explore the impact of the electoral safety under free and fair elections ($1 - p_N$), on D's incentive to restore both sets of institutions. I represent these results in terms of p_N, where when p_N is lower, D has a higher level of electoral safety. I look to three of the δ^* equations from lemma 3 to assess the impact of electoral safety on restoration preferences. For restoring executive constraints, I select the choice to restore executive constrains when

elections are free and fair ($V^D > W^D$)[17]. For restoring subverted electoral institutions, I look to both instances where electoral institutions can be restored ($V^D > Y^D$ and $W^D > Z^D$).

Proposition 6 (Electoral Safety and Restoration Preferences) *The critical discount factor needed for a successor to prefer restoring democratic institutions is:*

i. decreasing in p_N for executive constraints ($V^D > W^D$)

ii. increasing in p_N for electoral institutions ($V^D > Y^D$ and $W^D > Z^D$)[18]

In other words, as the successor has less electoral safety (higher p_N and lower $1 - p_N$), the discount factor required to prefer restoration is decreasing for executive constraints, but increasing for electoral institutions. In other words, as safety becomes lower, the region where restoring executive constraints is preferred expands, but the region where restoring free and fair elections is preferred shrinks. This naturally follows the understanding of the structure of the two kinds of backsliding. For the first part, when D increasingly does not anticipate winning elections in the future, it knows that placing an upward bound on its opponent's policy when they are out of power is more attractive. This is a policy motivated decision that leads to the preference to restore executive constraints. Alternatively, when subverted elections are in play, as electoral safety drops, D has increasingly less to lose and would prefer a moderate bump in their chances when in power (p_D) over a comparatively more minimal disadvantage when out of power (p_B).

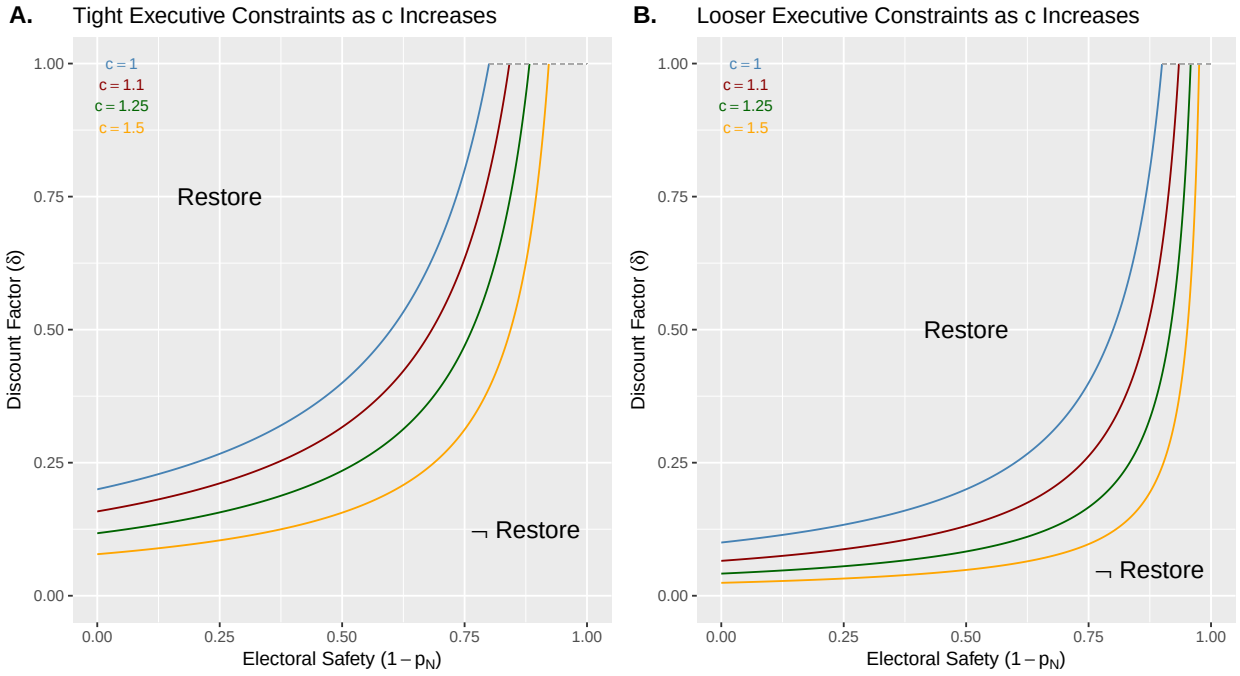
Similar to the previous two propositions, I graphically represent how the parameters explored so far, namely the magnitude of subverted institutions bias for each set of backsliding

[17] I do not look at the remaining case ($Y^D > Z^D$) because there are never free and fair elections so p_N is not a relevant parameter.

[18] Alternatively, when restating for electoral safety ($1-p_N$): (1) Increasing in $1-p_N$ for executive constraints ($V^D > W^D$) and (2) Decreasing in $1 - p_N$ for electoral institutions ($V^D > Y^D$ and $W^D > Z^D$).

actions, impact these results. For the choice to restore executive constraints, I explore altering first polarization and then the constrained executive institutions. For the choice to restore electoral institutions I show the impact of the bias and size of the subversion advantage.

Figure 3.5: The Effect of Polarization (c) on the δ and $1 - p_N$ Combinations Where D Favors Restoration.

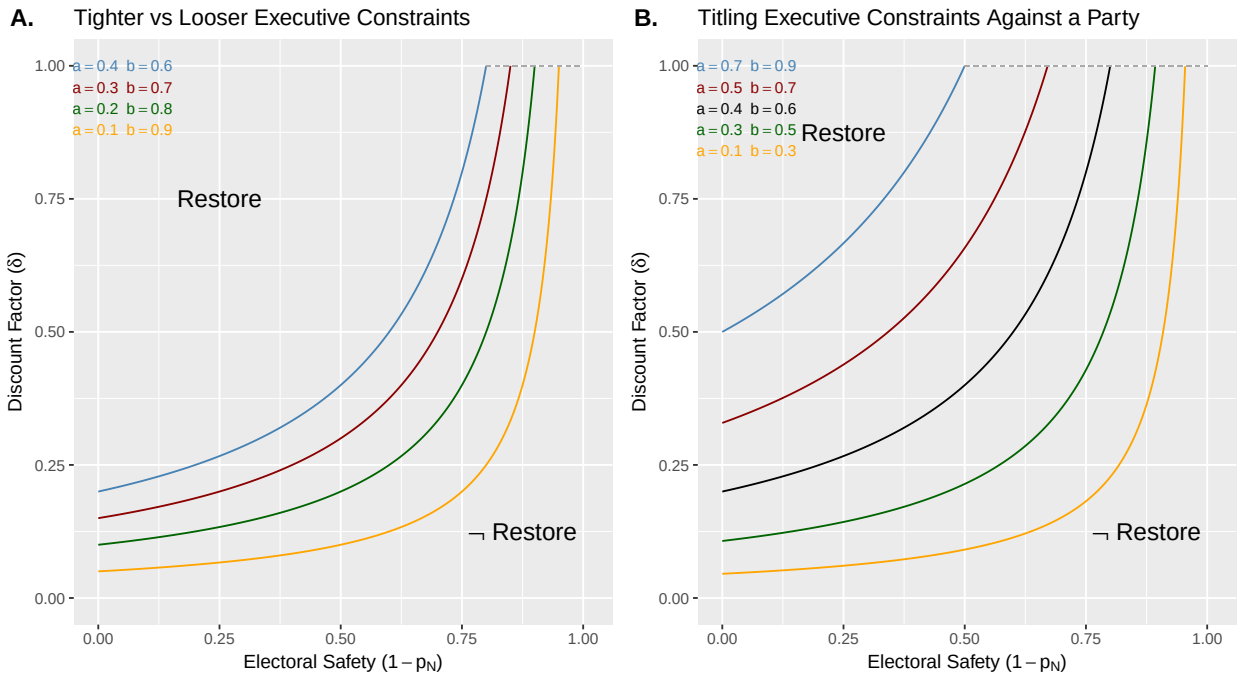


Note: δ^* curves for D's choice to restore executive constraints when electoral subversion has already been restored ($V^D > W^D$). In panels A and B I show the impact of increasing the polarization parameter c. In panel A, I set a narrow range of constrained policies $a = 0.4$ and $b = 0.6$. In Panel B, I set a wider range of constrained policies $a = 0.2$ and $b = 0.8$.

Figure 3.5 shows how polarization and electoral safety $(1 - p_N)$ jointly affect preferences for restoring executive constraints. In panels A and B, I show that as polarization increases, the region where restoration is preferred grows. In panel A, I set a narrow range of constrained policies while in panel B I set a wider range. As polarization rises, the constrained range of policies is more biased in favor of D, but this effect is more influential for the δ^* curves is when

electoral safety is high, otherwise, as electoral safety decreases, the impact of polarization is smaller because the region where restoration is preferred is already large.

Figure 3.6: The Effect of Executive Constraints (a and b) on the δ and $1 - p_N$ Combinations Where D Favors Restoration.

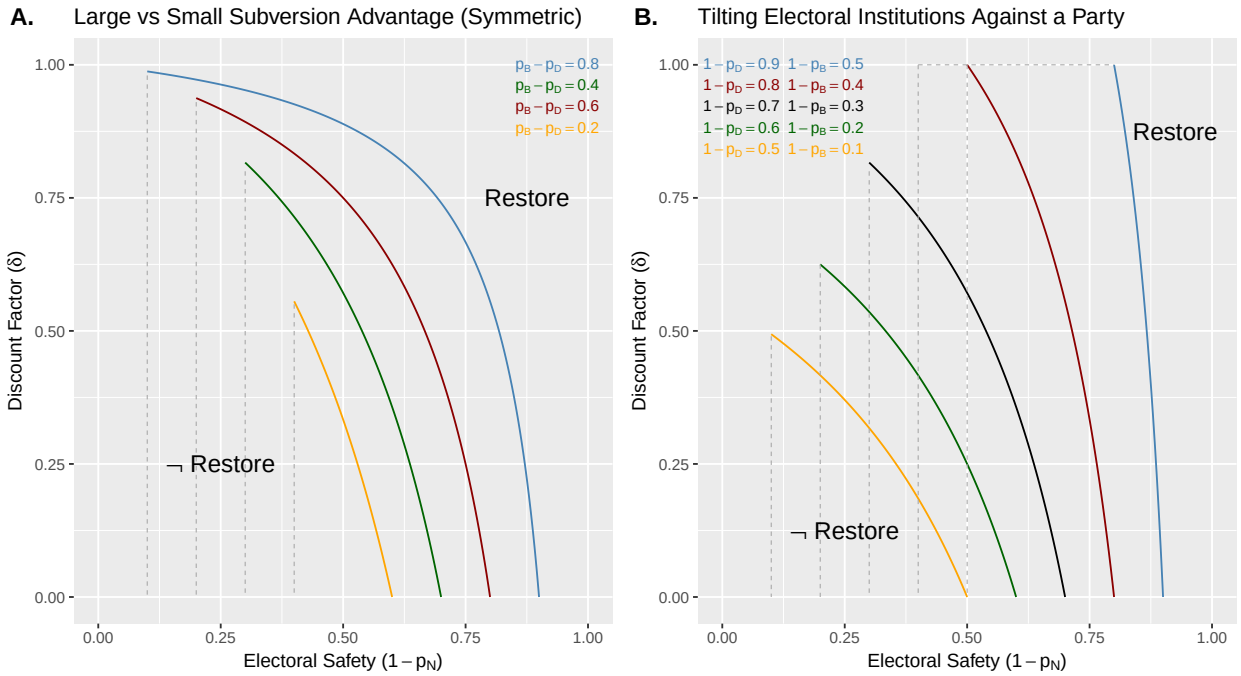


Note: δ^* curves for D's choice to restore executive constraints when electoral subversion has already been restored ($V^D > W^D$). In panel A I show the impact of increasing the constrained range of policies symmetric around $c/2$ where $c = 1$. In panel B I explore a tilted range of constrained institutions with a fixed range of $b - a = 0.4$ and where polarization is fixed to $c = 1$.

In figure 3.6 I consider the size and bias of the executive constraints as they relate to electoral safety and the preference for restoring those constraints. In panel A, the size of the executive constraints are altered or the range of a and b is increased with fixed polarization. As this range expands, so does the region where restoration is preferred. The magnitude of the executive constraints is most influential when electoral safety is high, because the risk of a higher of value of B's constrained policy is much lower than the benefit of a lower value of D's constrained policy. In panel B, the size of the executive constraints is fixed, but its

bias is altered. In this case, the blue line in panel A is the same as the black line in panel B. As executive constraints are more biased to D, the region where restoration is preferred is unsurprisingly larger. Meanwhile, D will prefer restoring less when executive constraints are biased to B, which is an unlikely world because B likely would not have removed executive constraints in the first place. Ultimately, these results are interesting because more modest restorative measures are much more preferred than more severe restorative measures.

Figure 3.7: The Effect of the Subversion Advantage ($p_B - p_D$) on the δ and $1-p_N$ Combinations Where D Favors Restoration.



Note: δ^* curves for D's choice to restore electoral institutions regardless of executive constraint restoration ($V^D > Y^D$ and $W^D > Z^D$). In panel A I show the impact of increasing the size of the subversion advantage ($p_B - p_D$) when symmetrical around $1 - p_N = 0.5$. In panel B I show the impact of "tilting" damaged electoral institutions in favor or against D. In both panels I fix $a = 0.4$ and $b = 0.6$.

Lastly, in figure 3.7, I show the impact of the size and position of the subversion advantage ($p_B - p_D$) on electoral safety ($1 - p_N$) and the preference to restore electoral institutions. In panel A, I alter the size of the subversion advantage so the range of possible $1 - p_N$ increases

with the subversion advantage[19]. As the subversion advantage increases, so does the region where restoration of electoral institutions is not preferred. In other words, when D has especially favorable institutions today (low p_D or high $1 - p_D$), it is less likely to restore, especially when electoral safety is not high. In each δ^* curve, as electoral safety increases, the preference for restoration increases. In panel B, the subversion advantage is fixed, and I shift the subverted institutions to be biased for or against D. In this panel, $1 - p_N$ can only take a value between $1 - p_D$ and $1 - p_B$, so the curves begin and end in different locations. The more biased the institutions are against D (lower $1 - p_D$ and $1 - p_B$), the larger the region where restoration is preferred. Also, as institutions are more biased in favor D, the more $\delta^* > 1$ which δ can never meet, so restoration is never preferred in those cases.

While the results for electoral safety are driven by the nature of the two sets of institutions, they provide some interesting competing incentives for successors that consider whether to restore. If greater electoral safety makes restoration more likely for one set of institutions than another, then it is easier to imagine a scenario where one set of institutions is restored for one purpose and another is not restored for a different reason. These results are interesting in light of the findings from the previous two propositions as well because they help detail the case where sometimes restoration of one set of institutions is highly preferred, which in other cases, restoration may not produce beneficial outcomes. While these results are in light of durable restorative measures, that also provide useful context to the more likely case where durability is not ironclad. I discuss this along with the main implications from the formal model in the next section.

[19]Where p_B and p_D are symmetric around $1 - p_N$ when $1 - p_N = 0.5$.

3.7 Semi-Durability: Limitations and Discussion

Now that I have systematically explored how the three main factors that I theorize impact a successor's choice to restore elections — durability, the magnitude of subversion, and electoral safety — and how they relate to each other and combine to produce similar or competing incentives, I will discuss two of the main limitations of the model. First, it is unlikely that restorative measures are either completely durable or completely not durable. But, by analyzing these two fringe cases — the optimistic case of completely durable restorative measures and the pessimistic case of completely non-durable restoration — I can say more about the likely world where restored institutions are "sticky," but not ironclad. The second is that even if restoration is preferred it may not always be possible. If not all institutions *can* be fully restored, more modest instances of restoration can be made to stem the tide of backsliding. Before concluding, I discuss these two limitations in terms of the model and the implications of this research on our understanding of what happens after a backslider is removed in the real world.

In terms of the model, when considering some level of durability between the two extreme cases of completely durable and completely non-durable, I would expect that the δ^* curves would shift up as restorative measures become less durable. In any of the figures, suppose that the backslider B could re-damage institutions with some probability π when B re-gains power. As π increases, I expect that the δ^*'s would shift higher within the parameter space of each figure. As this occurs, there will be less of the δ^*'s in a position where a value of δ can satisfy a preference for restoration. To address the limitation that restorative measures may be costly, consider that there is some cost to restore institutions for D. I assume that given some cost for restorative measures, this would impact δ^* in a similar way, by shifting the curves up and reducing the possible number of values of δ where D prefers restoration. I expect

that extreme subversion, there should still be some values of δ possible where restoration is preferred, even if those actions are costly.

While a general concern of restoring institutions is constraining one's actions only to have the other side not live by those same constraints, institutional changes should not be so fragile. Assuming some fragility, there are still cases where restoration is preferable. Consider out-parties in Turkey, Hungary and Poland. In Turkey for instance, the opposition party (CHP) is heavily constrained by the size of the governing party (AKP). Should it gain power and attempt to walk back some of the backsliding measures of Erdoğan, it likely will enter office with low electoral safety, probably having won in a close election with factors dramatically in its favor. In this case, constraining the executive may be beneficial even if the CHP expects the AKP may undo the restorative measures after re-taking power. In Poland and Hungary, where electoral institutions are more clearly biased in favor of the ruling parties, even if opposition parties gained power they likely could not wield the institutions significantly to their benefit, so similar incentives may exist for restoring electoral institutions. In the recent case of the United States, President Biden, who won despite institutions biased against his party, has shown resistance in engaging in the same executive over-reach of his predecessor and resisting calls to retaliate (Beals 2022), but has explored reform of elections, judges, and the civil service. In this case, the other limitation is especially noteworthy.

Unlike in the model, some restorative measures are not always possible. In the American case, even if Democrats wanted to restore electoral institutions, most of the bias of these institutions is enacted at the state level and not the federal level. Gerrymandering in drawing of districts for the House of Representatives, strict voter ID laws, and alarming new efforts to propose alternative slates of electors after Presidential elections are three examples of efforts at the *state* level to tip the scales at the national level and are not able to be neatly fixed by a new administration of Democrats. Instead, restorative measures may take time and happen

over several terms in office. Despite *existing* institutional durability, the fundamentals of the model remain relevant for these choices because even if restorative measures are not absolute, they can have more modest benefits that follow the general relationships in the result section.

Ultimately, while the model has some limitations for the real world, such as the nature of institutional restoration being not durable enough or existing institutions being too durable, it can tell us a lot about expectations for restoration and the conditions under which we should most and least expect to see it. Even expectations of more minor restorative actions or retaliatory measures are useful theoretical innovations that can inform and shape this new frontier in the literature.

3.8 Conclusion

Understanding both how backsliding either persists or is reversed is a natural, yet unstudied extension of the backsliding literature. While existing democratization literature provides several useful points of departure such as the focus of risk aversion of new leaders, backsliding regimes are notably different, because they still possess democratic norms and institutions, even if they were left in a state of disrepair.

I argued that the main factors that impact whether successors to backsliders pursue restoration of democratic institutions are the durability of restorative measures, the magnitude of subverted institutions, and electoral safety. Durability is the most important factor for restoration because absent durability, there are no incentives to restore if the other side will not follow the same constrained rules. Given some level of durability though, there are situations where restorative measures are preferred even if the confidence in a future backslider playing by the new rules is not certain. The magnitude of subversion both institutions is also important, and when this grows larger, so do the preferences for restoration for each

set of institutions. Lastly, electoral safety impacts preferences for restoration differently depending on the nature of the institutions being reformed. While higher electoral safety leads to higher preferences for restoring electoral institutions, it leads to lower preferences for restoring executive constraints. These main results, along with the systematic analysis of how these various factors interact, provide a full model detailing preferences for restoration.

Future work can build on this theoretical contribution by exploring some of the smaller conclusions using observational data. While many factors of the model are difficult to adequately estimate, analyzing some macro-level considerations for successors to backsliders and observing the actions they take may provide some useful findings for how this theoretical model holds up under real world conditions. The focus on establishing this formal model and the intensive theoretical analysis in the results sections provided several interesting conclusions about preferences for restoration such as restoration being preferred in situations where a successor may not immediately benefit from the new rules due to biased institutions. These results, both major and minor, are some of the first contributions to answer a question that will only continue to be more important in the backsliding literature and political science as a whole.

This study shows that democracy may still be imperiled even when a more democratic government succeeds a backsliding one. This has implications for both when the opposition to either the AKP in Turkey, the PiS in Poland, or Fidesz in Hungary returns to power, and also the choice of the current Biden administration in the wake of backsliding under the Trump administration. If the succeeding government prioritizes policy by retaliating against the initial backslider and neglects to meaningfully restore and repair democratic norms and institutions, then a tit-for-tat cycle of backsliding may lead to complete democratic failure. This study is important because when voters opt for democratic alternatives to their backsliding leaders, there may be a narrow window of time to repair democratic institutions

to serve as a bulwark against future backsliding. Understanding this new set of incentives will allow political scientists and policy makers to better understand the most productive priorities for the long-term democratic health of a country. In the next chapter I shift my focus to voters and their preferences for their leaders should they succeed a backsliding incumbent, and under what conditions they prefer a restoring successor or a retaliator.

Chapter 4

Retaliation or Reformation: Voters' Democratic Preferences After Democratic Backsliding

4.1 Introduction

How does backsliding affect voters' preference for democratic institutions? Are voters more likely to support backsliding from co-partisans if the opposition engaged in democratic backsliding first? The literature around voters' support of leaders engaging in democratic backsliding is relatively new, but less focus has been placed on voters' preferences for maintaining democracy. While past explanations of voter support have focused on partisanship, polarization, and the role of societal threat (Chapter 2), democratic preferences themselves are often taken as given and are understudied. The focus on supporters of initial backsliders is useful, but considering how detractors will act when they regain power is an important

extension and can give greater insight about the long-term democratic health of a country. If the opposition prefers to retaliate against the initial backslider instead of repairing institutions to prevent future democratic subversion, then a long term pattern of tit-for-tat retaliation can eventually yield a full authoritarian reversion.

When a backsliding incumbent yields to a successor formerly in the opposition, there is uncertainty about the path forward: democratic restoration or continued democratic backsliding either by maintaining damaged institutions or damaging them further. While the focus of the backsliding literature on voters is commonly focused on their tolerance of backsliding, considering what happens when a successor is chosen is noteworthy. In Chapter 3, I explored when the former opposition regains control and the choice to either continue to use the damaged institutions to retaliate or to repair them to prevent future misuse. It is additionally important to explore how voters view these choices and how that impacts the prospects for democracy after the backslider is removed. In the American context, the literature on constitutional qualms (Christenson and Kriner 2017*a*; Reeves and Rogowski 2018) provides some competing expectations of the permissiveness of co-partisans to actions that approach backsliding. In the literature on constitutional hardball (Fishkin and Pozen 2018; Helmke, Kroeger, and Paine 2019; Pozen 2018; Shugerman 2019), incentives to push democratic limits can lead to a breakdown in the mutual deterrence that maintains democracy. The practice of retaliatory hardball is a key area where the retaliation or restoration choice can be made, and a democratic breakdown can occur after too many retaliatory actions.

I argue that the trade-off between policy and democratic preferences is fundamental for voters' overall choice between retaliation or restoration. Generally, voters may or may not react adversely to democratic backsliding if their own party is responsible. I expect that for supporters of the out-party, if their leaders are able to push forward policy at the expense of democratic norms and institutions when the party retakes power, the voters will generally

approve, *especially after backsliding*. I also consider the counterfactual condition where an incumbent did not engage in backsliding to compare with backsliding incumbents to show how this trade-off changes. If voters only care about policy and not democratic governance, we should expect little differentiation. If voters care about democratic norms at the expense of policy, then they will prefer their successor behave more democratically to the detriment of their policy goals. Additionally, if voters are sensitive to democratic norms but value policy, they will prefer the more democratic path after a non-backsliding or minor backsliding incumbent, but prefer retaliation after more severe backsliding. I expect these voters to be sensitive to the severity of the backsliding employed.

To test the theory, I look to the United States: a country that both recently removed an incumbent that engaged in backsliding and has a history of tit-for-tat constitutional hardball. Further, the United States is a laboratory for exploring voter preferences for a successor to an initial backslider as the choice between retaliation and restoration has dominated the early days of the Biden Administration. The administration has signaled a desire to improve weakened institutions, in line with restoration-minded preferences (Wise and Walsh 2021). Simultaneously, some provisions of the robust policy program the administration campaigned on were passed into law, but other components required executive action — an example of minor backsliding because executive orders generally require some executive overreach on a policy area that was not amenable under other means — outside of the normal legislative process to enact key provisions due to considerable gridlock (Vazquez, Sullivan, and Klein 2022). Executive actions alone are also insufficient to enact the policy program because of the prevalence of adversarial courts that can strike down statutes (Breuninger 2022; Liptak 2022). This would indicate that in order to pass the outlined agenda, there would either need to be some backsliding action to capture the courts or eliminate gridlock. However, some of the measures to counteract these forces remain divisive (Freeman 2021; Yokley 2022). These

competing goals can provide an array of choices to politicians, but also directly involve voters. If the base of the Democratic party prioritizes eliminating gridlock or changing courts via backsliding measures to pursue policy change while they have unified control of government, then this may create incentives for their leaders to pursue these goals beyond the incentives established in the previous chapter. This recent history, along with a more established history of constitutional hardball, make the United States an appropriate setting to gauge voter preferences in a survey experiment.

In the survey experiment, I evaluate voters' preferences after hypothetical backsliding by the opposite-party. Specifically, I examine their preference between one of two same-party candidates: a *restorer* who prefers a more democratic approach to governance at the expense of full policy goals, or a *retaliator* who prefers pursuing policy reversals at the expense of democratic institutions. Respondents make this choice after random assignment to an incumbent who engaged in no backsliding, minor backsliding, or severe backsliding. I expect that voters prefer the restorer over the retaliator more when the incumbent does not backslide or engages in minor backsliding actions than when the incumbent enacts major backsliding. Beyond the theoretical framework, I also test recent findings that Republicans are more supportive of backsliding (Fishkin and Pozen 2018; Grumbach 2022; Mann and Ornstein 2016; Shugerman 2019). I expect, because of this existing literature, that Republican respondents across all categories to prefer retaliators when compared to Democratic respondents.

I find that the severity of backsliding impacts voters' preferences for the retaliating candidate. More extreme backsliding measures were met with higher retaliator support than milder backsliding cases. Meanwhile, the control condition, where no backsliding was enacted, produced unexpected results. These results may indicate that the choice for voters after backsliding is distinct from initial support for backsliding, but not in the expected direction. Respondents who were randomly assigned to an incumbent that did not backslide

supported the retaliator *more* than those who were randomly assigned to the minor backsliding incumbent, but *less* than those who were assigned the severe backslider. Lastly, contrary to expectations, I found that Republicans were slightly less likely to support the retaliator.

This research is important to explore whether the removal of a backsliding incumbent can help slow or stop the cycle of democratic backsliding, or if the success of one backslider can create a series of tit-for-tat actions between rivaling factions that devolves into retaliatory hardball. If one instance of backsliding leads to alternating parties operating on progressively more fragile institutions, a full democratic breakdown can occur. This study isolates one instance of the choice of a successor in this chain, but a series of retaliatory backsliding actions can increase the threat to democracy because of the continued erosion of democratic institutions and trust. Within this snapshot, I focus on voters' behavior specifically because of how politicians and elites get cues on what to prioritize, and vice versa. Therefore, this chapter combined with Chapter 3 can provide greater understanding of the importance of the narrow window where democratic resurgence after backsliding can occur. If voters are not concerned about restoring democratic institutions after they are degraded, then this can be another step along the path of retaliatory hardball and democratic backsliding. Consequently, if leaders are incentivized by their supporters to subvert democracy in order to advance their policy programs, and voters only care about democratic conduct when their politicians are out of power, then these incentives may not bode well for electoral democracies experiencing backsliding.

4.2 Background

The study of voter support for backsliding has predominately focused on when and why voters support undemocratic candidates and backsliding incumbents. Little focus has been placed

on how these opinions change after a backsliding government yields to a successor. This is understandable given the limited instances of this change, but is normatively important to consider when current backsliders are replaced. These events can be a crucial window to restrain backsliding and stave off future back-and-forth hardball to prevent democratic failure. I summarize existing arguments around voter support, and more narrowly, voters' value of democratic norms and institutions and how they are willing to bend given characteristics like partisanship, polarization, and societal threat.

As I detailed in chapter 2, there is considerable debate around whether voters preferences for democratic norms and institutions have fallen or not in recent years (Alexander and Welzel 2017; Foa and Mounk 2016, 2017*a*,*b*; Norris 2017; Voeten 2016). Whereas other work has approached these values alternatively by framing preferences as majoritarian or not (Grossman et al. 2022). While the understanding of these preferences at the aggregate is difficult to assess, it remains important to attempt to understand voters' motivation as falling support for democracy amongst voters makes the emergence of a backsliding actor more probable (Huntington 1991).

Other work has taken these preferences as given, and explored the factors that makes voters weight these preference more or less. Arguments around partisanship and polarization (Graham and Svolik 2020; Mazepus and Toshkov 2022; Singer 2021; Svolik 2018) and the role of societal threat that raises the stakes of elections and makes voters and elites risk averse (Chapter 2) have served to show how even democratically-minded voters are willing to support those who subvert democracy. In this work, voters are generally concerned about policy chiefly, and preferences against backsliding actions secondarily. In this chapter, I explore the role of an important policy goal and how it impacts an out-party voter's willingness to support a backsliding-minded candidate depending on whether the opposite-party incumbent engaged in severe backsliding or not.

In the American context, these same concepts, namely partisanship and polarization, have also been found to provide support for unilateral actions of the executive (Christenson and Kriner 2017*a*; Reeves and Rogowski 2015, 2016). When there are policy motivations for stretching the range of acceptable actions, voters are more willing to approve of these actions and look past their "constitutional qualms". The evidence of voters' willingness to overlook constitutional misdeeds in favor of policy is a bit mixed (Reeves et al. 2017; Reeves and Rogowski 2018), and I continue the exploration of that trade-off. While most of this work is in the context of unilateral action of American presidents, I believe these same concepts are relevant for backsliding actions more generally.

This same concept has been discussed in literature on "constitutional hardball" (Helmke, Kroeger, and Paine 2019) with respect to the choice of elected leaders and their incentives to push against constitutional boundaries, but not to break them. Constitutional hardball generally speaks to challenging democracy at the margins, or minor backsliding actions used for policy gains. These initial infractions lead to a series of retaliatory hardball measures that have been a key feature in recent American politics (Fishkin and Pozen 2018; Pozen 2018; Shugerman 2019). As these hardball actions get progressively more egregious, the likelihood of more explicit backsliding actions rises. Constitutional hardball has been increasingly associated with the Republican Party at the national (Fishkin and Pozen 2018; Shugerman 2019) and state level (Grumbach 2022). Meanwhile, at the voter level, there is evidence in the United States that voters overestimate the willingness of members of the opposite party to break norms and this drives support for norm breaking by their own party (Braley et al. 2021).

The existing literature provides useful context for understanding voters' democratic norm and institution preferences. Isolating how those preferences shift in the face of backsliding by the opposite party is a useful extension to consider in the chain of retaliatory backsliding. I

consider this more in the next section where I outline the theoretical framework for a voter's preferences for or against a candidate that pursues democratic subversion to enact policy goals. Within this choice, I consider how the presence of existing backsliding and its severity impact voters' eventual choice. I explore the importance of the choice during the window after democratic backsliding by isolating the events after a backslider's removal. Voters may prioritize policy and signal their leaders to continue a series of back-and-forth retaliatory backsliding. Alternatively, they may prefer that some policy goals are sacrificed to enact restorative measures to repair democracy. The contours of this choice at the voter level, paired with the successor level (Chapter 3) can provide collective and competing expectations for when backsliders are removed.

4.3 Theory

If a backslider leaves office and is replaced by the former opposition, the officeholder responds with a key choice.[20] To simplify for the purpose of the voter, the choice can be divided between two general paths: *restoration* where the opposition focuses on repairing damaged institutions, or *retaliation* where recently subverted norms and institutions are maintained or damaged further. These choices also have direct consequences for the ability of the former opposition to implement its preferred policies.

Under restoration, a democratic successor focuses on restoring the democratic norms and institutions that were targeted by the backsliding predecessor. Examples of restoration include introducing constraints on the executive, restoring civil service protections, and making elections more free and fair[21]. This path is normatively beneficial for the quality of

[20]The focus of Chapter 3

[21]Initial backsliders commonly attack and damage many of these institutions as counter-majoritarian, and use that as justification to seize power. By doing this, these leaders eliminate constitutional constraints and are able to later misuse those new powers in unadvertised ways (e.g., the 1999 constitution in Venezuela was meant to streamline the legislative process but dramatically entrenched the new party, thus creating

democracy, but has an important trade-off with policy. I assume that pursuing a restoration route, especially restoring constraints on the executive, would come at the expense of the new leader instituting a comprehensive policy program. This means that a successor may be able to undo some policies of the backslider, but will not be able to leverage a recently constrained government to completely enact a full set of policies.

The alternative to restoration is retaliation, where a successor either uses damaged institutions or further damages them to both more effectively change policy and to tilt the electoral playing field in an advantageous way. In practice, retaliators may take advantage of the weakened executive constraints to pursue their goals more easily, use the control of the bureaucracy to their benefit, or tilt the electoral playing field to their benefit. Retaliation can also manifest as symmetrical backsliding actions if the existing institutional alignment disproportionately benefits the initial backslider's party (i.e., if the initial backslider packed courts, then the successor would have to further pack courts or remove judges). Regardless, retaliation means that the successor aims to use backsliding to pass policies more easily and manipulate the election infrastructure. This can be achieved by maintaining the recently worsened status quo or through tit-for-tat actions. For the purpose of this project, I do not distinguish between these, but they are generally subject to the nature of institutions. For example, for backsliding that led to advantages in the composition of constitutional courts, retaliation would require some level of symmetrical backsliding to either add sympathetic judges or remove adversarial judges. Meanwhile, backsliding that created new executive powers for a given policy area can likely be similarly leveraged by a successor and would not require new backsliding to retaliate.

a permanent majority). If a restorative measure helps push back by re-introducing these institutional constraints, it is not necessarily to limit the majority, but to appropriately limit the exercise of executive power.

I incorporate these competing paths (i.e., restore vs. retaliate) in the context of work (Christenson and Kriner 2017*a*) that explores a trade-off between the *means* and *ends* of policy change. The *means* are the governance changes or actions that make policy possible. The *ends* are the policy outcomes that result from the means used. I classify backsliding actions as one of the means to an end that is usually a policy outcome or an advantage in the electoral sphere. Voters with strong preferences for democratic norms and institutions are concerned with the means employed, whereas voters who are more focused on policy outcomes will overlook the means in pursuit of the ends. This produces a trade-off where a change in the severity of the means (how much backsliding) is positively related to the ends (the comprehensiveness of the policy enacted). In the long run, parties exchange power; therefore, the relative policy gain today may not be worth the potentially worse policy enacted by the other party after a future election. Because of this, I expect this framework to operate in the short-term, because in the long term, the means and ends will be more balanced out by more extreme policy outcomes both in favor of and against the voter's interests. The importance of the long-term incentives for the successors is explored in depth in Chapter 3.

4.3.1 Means and Ends Preferences of Voters

I distinguish between three ideal types of voters depending on the value they place on the trade-off described above: ends voters, means voters, and balance voters. In table 4.1, I summarize schematically the preference of each of these three voters for a successor to replace an out-party incumbent that engaged in either no backsliding, minor backsliding, or major backsliding. The first column shows expectations for voter preference when no backsliding has occurred. This condition is the counterfactual for this study, but is the traditional focus in the backsliding literature that identifies which voters prefer a backslider without provocation.

The second two columns show preferences after backsliding has occurred and are the focus of this paper.

Table 4.1: The Preferred Successor Type for Each of Three Voter Types Given an Incumbent Action.

	Incumbent's Action with Respect to Democratic Institutions		
	No Backsliding	Minor Backsliding	Major Backsliding
Ends voter	Initial Backslider (≈ Retaliator)	Retaliator	Retaliator
Balance voter	Status Quo (≈ Restorer)	Restorer or Retaliator	Retaliator
Means voter	Status Quo (≈ Restorer)	Restorer	Restorer

Two ideal types of voters — means voters and ends voters — do not respond to the variation of the incumbent's action. For means voters, they are concerned about process and going about policy change in the "correct way." Therefore, regardless of the actions of the incumbent, they will always prefer a restorer who will restore democratic institutions, even at the expense of policy. Ends voters, conversely, always prefer a retaliator because they are only concerned with policy outcomes and if a preferred outcome requires backsliding, they will support it. These voters are focused on short-term gains only, even if in the long-run democratic subversion makes them worse off because of the alternation of parties in office. These voters are typically important in other work on backsliding because they are a key coalition that enables backsliders to remain in power since they care less about backsliding when policy successes are achieved.

Unlike the other two ideal types of voters, balance voters exist on a continuum because they respond to variation in the severity of backsliding and are generally context-driven. In the absence of backsliding from the other side, they are skeptical of candidates that propose engaging in backsliding to achieve policy goals, and instead prefer a candidate more committed to the status quo (a restorer in this case who has nothing to restore). After severe and unprecedented backsliding — and I assume also considerable policy change — they prefer

a retaliator since the means have already been damaged and the policy would be expected to be more distant from their preferred position. Under a milder version of backsliding, there is some cutoff where the balance voter would prefer a restorer under that threshold, and a retaliator above that threshold. The location of this theoretical cut-off occurs on a continuum where minor agitation can cause retaliator support for some, while others have a more significant threshold before they support democratic subversion to fight back. For the purpose of this paper, I consider the policy enacted by the former backslider as given. In reality, more extreme backsliding can be used to pass more extreme policy, but this is not always the case. Regardless of the actual policy, I argue that it is the trade-off between policy and backsliding severity that drives voters' responses, and the magnitude of the policy and severity of the backsliding determines the cut-off. To summarize, all three of these groups of voters are important for the selection of a successor, and determining the path that the successor takes.

This framework highlights a crucial element of the future of democratic stability: whether voters treat democracy as being "worth it." When supporters of would-be backsliders see democracy as a necessary inconvenience then they will be deterred from engaging in backsliding. In contrast, when voters prefer their leaders to be unencumbered by rules they see as unnecessary, the path to democratic collapse is much clearer. Erdoğan once said that "[d]emocracy is like a tram. You ride it until you arrive at your destination, then you step off" (Çakır 2012). And more explicitly, "[d]emocracy is a means, not an end"(Çakır 2012)[22]. While elected and non-elected elites are crucial in signaling these paths to voters, sometimes voters can actively cheer on democratic subversion and push their leaders on this path. Democratic subversion is normatively bad and there are additional negative outcomes such as lower levels of personal freedoms, higher inequality, hampered government responsiveness, and lower

[22]In Turkish: "Demokrasi bir tramvaydır, gittiğimiz yere kadar gider, orada ineriz" and "Demokrasi amaç değil araçtır" respectively.

governmental accountability (Diamond and Morlino 2004). These negative outcomes are the key reason why we should know how preferences against democracy develop. We can speak to voters' role in either contributing to, or working against democratic stability by understanding their preferences within this trade-off more thoroughly.

This framework provides two important measurments that can help us understand the preferences of voters and how they decide whether to support a restorer or a retaliator. The first measurement is the proportion of the different voter types in an electorate. It is unlikely that the three types are evenly represented, and existing work has been divided about what voters value at the aggregate. Further, this typology is useful for disaggregating voter preferences and can help explain different results in the literature that may be direct consequences of experimental manipulation (i.e., a too severe or not severe enough manipulation). The second measurement is the location of the means-ends cut-off for the balance voters (above which they prefer the retaliator and below the restorer). If voters in the aggregate respond similarly across the three incumbent actions in table 4.1, it would indicate a dearth of balance voters. However, if there are meaningful aggregate differences in the response to the incumbent's actions, then the balance voters are likely driving these differences. Also, depending on experimental manipulation, we can assess where different kinds of backsliding actions are in relation to this threshold and provide useful information about the preferences of voters after backsliding.

To put these concepts in context, I return to the American case to detail some specific expectations based on the theory and the existing understanding of the party system in the United States. Constitutional hardball — a concept similar to several retaliatory minor backsliding actions — has been a key feature in American Politics, both within the executive branch and the legislature, for much of the last century. Under Franklin D. Roosevelt, efforts to capture the courts were undertaken to execute a significant policy program in the face of

the Great Depression. In the 1970s and 1980s, Nixon and Reagan helps spearhead the capture of the bureaucracy along with other scandals that involved overreach (Watergate and Iran-Contra respectively). In the 1990s, Newt Gingrich initiated a series of constitutional hardball actions that persists today. Recently, Mitch McConnell has been a significant instigator of constitutional hardball within the Senate. Outside of the legislature, under each the three most recent former presidents, executive action and overreach has been undertaken, almost always controversially and in service of partisan goals, to exercise undue authority in military conflict, protection for undocumented immigrants, and banning visitors or immigrants from dubiously selected countries, to name a few. A defining characteristic of this literature is the asymmetric nature of the execution of hardball and the reliance that Republicans have taken to create gridlock, pack courts, and aggressively undermine the executive branch when they do not control it (Fishkin and Pozen 2018; Grumbach 2022; Mann and Ornstein 2016; Shugerman 2019). For this reason, I expect that Republican respondents are more likely to prefer retaliators than Democrats regardless of the action of the incumbent.

The cues that voters send their leaders after a backsliding incumbent is removed are an important component leaders use to determine the path ahead, and can combine with or contradict the expectations of successor behavior from Chapter 3. If moderation and restoration are not valued by voters who put politicians into office, then it is unlikely that those approaches will be enacted by elected leaders. This not only affects the initial backsliders, but also their successors. In settings like Venezuela where a dysfunctional "partarchy" dominated from the 1961 constitution to Chavéz's election in 1998 (Landau 2018), the path from democratic backsliding to democratic failure was a multi-step process that gradually damaged norms and institutions until an actor could execute a full authoritarian reversion. This theory helps isolate a step in that process and provides a window to show how preferences and

incentives at the voter level can either perpetuate that process or stall the cycle in favor of democratic restoration.

4.4 Experimental Design

I evaluate this theoretical framework through a survey experiment administered to respondents in the United States. Since the United States recently removed a backsliding incumbent in a democratic election, this recent history paired with the longer history of constitutional hardball makes voters in the United States a useful sample to evaluate the theoretical framework. The structure for the experiment is summarized in table 4.2. Two iterations of the experiment were fielded during the month of June 2022.

Table 4.2: Experiment Structure

Pre-Treatment	-Collect covariates for respondents including partisanship -Identify policy area important to respondent (R) for use in experimental design: a favored and disfavored policy identified (see table 4.3) -Assess opinions on American democracy
Treatment (Random Assignment Blocked by Partisanship)	-R given vignette in which the incumbent is from the opposite party -Policy area identified above is incorporated into this vignette
	Incumbent experimental conditions (see figure 4.1) in the vignette: **1. Control: Incumbent did not engage in backsliding: passed disfavored policy into law using status quo measures** **2. Treatment 1: Incumbent engaged in moderate backsliding: passed disfavored policy into law using executive powers** **3. Treatment 2: Incumbent engaged in major backsliding: passed disfavored policy into law by removing SC justices**
	-R chooses between two own-party candidates (with same policy) in a congressional primary except for one distinction (see figure 4.2): *1. A restorer who plans to restore or maintain democracy with lower probability of favored policy success* *2. A retaliator who plans to retaliate or engage in backsliding with higher probability of favored policy success*
Post-Treatment	-Re-assess opinions on American democracy -Assess R's interest in politics

Note: The randomized groups are bolded. The main choice for the respondent is italicized.

Table 4.3: Policy Proposals Used in Vignettes.

Issue Area	Democratic Proposal	Republican Proposal
Abortion	Eliminate legal restrictions on abortion before viability	Institute additional legal restrictions on abortion
Healthcare	Create of a public option available to anyone without private health insurance	Eliminate the Affordable Care Act (also known as Obamacare)
Guns	Require background checks and ban semi-automatic weapons	Eliminate legal restriction on the right to bear arms
Taxes	Raise taxes on corporations and highest earners	Reduce high taxes on job creators
Immigration	Create a pathway to citizenship for undocumented immigrants that arrived as children (also known as dreamers)	Reject asylum seekers and combat illegal immigration
Climate Change	Create new taxes on polluting industries and use revenue to fund alternative sources of power	Make effective use of fossil fuel resources and eliminate regulations on businesses

Note: For the issue area that the respondent selects, the disfavored policy is plugged into the incumbent vignette (see figure 4.1 for example text for a Democratic respondent) and the favored policy is plugged into the congressional candidate vignettes (see figure 4.2 for example text for a Republican respondent).

In the pre-treatment portion of the experiment, respondents were shown a table of six issue areas (abortion, guns, healthcare, taxes, immigration, and climate change) with proposed policies from both Democrats and Republicans within that issue area. Policies for both Democrats and Republicans are presented in table 4.3. Respondents were asked to choose one of six issue areas where they think the difference between the two parties is important. After this decision, the respondents selected which policy is closest to their preferred policy[23]. These policy choices were used in the vignettes respondents evaluated. In the main vignette, an opposite-party incumbent government passed the disfavored policy or the policy not selected by the respondent. In the two own-party candidate profiles that the respondent must choose between, both candidates plan to enact the respondents' selected

[23]These issue areas and policies where taken from the 2020 Democratic Party platform and the 2016/2020 Republican Party platform.

policy. These policies were used for the purpose of providing the respondent with an example that employed an issue area that concerned them[24].

Figure 4.1: Incumbent Vignettes for Democratic Respondents

Same for all groups: Suppose that the government is controlled by the Republicans and the Democrats are in the minority. In the upcoming election, your congressperson, a Republican, is up for re-election. Since the previous election, the Republicans have passed legislation including a high profile law to pass DISFAVORED POLICY.

Control 1: (No additional text)

Control 2: This law was made possible by the majority support it received in Congress.

Treatment 1: The passage of this law was made possible by Congress giving the executive branch controversial new powers. This change was criticized as unconstitutional, and allowed the Democrats to pass laws that they otherwise could not.

Treatment 2: The passage of this law was made possible by the addition of new justices to the Supreme Court to combat a conservative majority that would have struck down the legislation. This action was widely condemned as unconstitutional, and allowed the Democrats to pass laws that they otherwise could not.

Note: Each group receives the first three sentences detailing who is in power and what was accomplished. DISFAVORED POLICY refers to the policy in the issue area that is the opposite of their preferred choice. For a Democratic respondent, this would correspond to the Republican proposal column in table 4.3. Treatment conditions correspond to the three incumbent actions in table 4.1.

In the experiment, respondents were presented with an incumbent government controlled by the party opposite that of the respondent. The respondent learned that their disfavored policy in their selected issue area was put into effect by the government. Figure 4.1 contains the text for each incumbent vignette for a respondent who selected the Democratic party's policy proposal. In the control groups, the incumbent administration passed the law either with unstated means or with majority support. In the first iteration of the experiment, no statement was made about how the law was passed. In a second iteration, two control groups

[24]While some policies were selected at different frequencies for Democrats and Republicans (i.e., a higher proportion of Republicans selected immigration), the different issue areas were not statistically different from each other with respect to the dependent variable.

were included, the first group was the same and the second group contained a statement that the law was passed with a majority of the Congress. In the first treatment group, the administration engages in a moderate instance of backsliding where the legislature cedes powers to the executive so it can implement policies that were not possible through legislation. In the second treatment group, in a more extreme action, the legislature adds additional Supreme Court Justices, without whom the policy would not have been possible.

Figure 4.2: Congressional Candidate Vignettes for Republican Respondents

CONTROLS 1 AND 2:
Restorer: Plans to work with a new governing majority of Republicans and following existing rules in government to reverse Democratic legislation on ISSUE AREA and passing a law to FAVORED POLICY. Experts believe these goals are not likely to pass into law.
Retaliator: Plans to give a Republican President controversial new powers to delegate lawmaking on ISSUE AREA to reverse Democratic legislation on ISSUE AREA and passing a law FAVORED POLICY. Experts criticize the controversial new powers as unconstitutional, but believe these goals are more likely to pass into law.

TREATMENT 1:
Restorer: Plans to work with a new governing majority of Republicans to remove the new powers for the President and restore previous government practices damaged by Democrats. Also plans to reverse Democratic legislation on ISSUE AREA and pass a law in FAVORED POLICY. Experts praised the plan to restore constitutional lawmaking, but believe these goals are not likely to pass into law.
Retaliator: Plans to give additional controversial powers to a Republican President to combat the changes made by Democrats. Also plans to reverse Democratic legislation on ISSUE AREA and pass a law to FAVORED POLICY. Experts criticize the additional powers as unconstitutional, but believe these goals are more likely to pass into law.

TREATMENT 2:
Restorer: Plans to work with a new governing majority of Republicans and not significantly alter the legal framework left by Democrats. Also plans to reverse Democratic legislation on ISSUE AREA and pass a law to FAVORED POLICY. Experts praised the plan to not retaliate against Democrats on the Supreme Court, but believe these goals are not likely to pass.
Retaliator: Plans to work with a new governing majority of Republicans and add additional conservative justices to counteract the justices Democrats confirmed. Also plans to reverse Democratic legislation on ISSUE AREA and pass a law to FAVORED POLICY. Experts criticize adding more judges as unconstitutional, but believe these goals are more likely to pass into law.

Note: FAVORED POLICY refers to the policy in the issue area that is the preferred choice of the respondent. For a Republican respondent, this would correspond to the Republican proposal column in table 4.3. ISSUE AREA refers to the selected issue area from the Issue Area column in the same table.

After evaluating the incumbent, the respondent was presented two own-party congressional candidates — a restorer and a retaliator — who could run against their opposite party congressperson. The only difference between the candidate profiles between groups is the kind of backsliding of which they are responding. Figure 4.2 contains the text for each pair of candidates for each of the treatment groups for a respondent who selected a preferred policy from the Republicans. The restoration candidate plans to either maintain or repair institutions regardless of the incumbent's action. The retaliation candidate plans to engage in backsliding regardless of the previous vignette. In the control group, the retaliator plans to initiate backsliding (identical to the backsliding in treatment 1) and in the two treatment groups plans to retaliate by either maintaining new backsliding or engaging in symmetrical backsliding. The restoration candidate's policy goals are identical to the retaliator, but are presented as less likely to pass compared to the retaliator. The respondent is informed that both candidates are about as likely to win both the primary election and the general election to deter a strategic voting choice.

The respondents evaluated the two challenger candidates on three questions. The first is the main dependent variable: a forced choice for which candidate gets the respondent's vote. The second and third questions are which candidate they think would be better for the country, and which they would be more likely to support for higher office. These questions allowed the respondent to consider the candidates as a whole and determine whether they see a distinction between them on other subjective measures in addition to the forced choice. These binary choices, along with minor differences between the vignettes makes estimating the effect of backsliding on preferences clear.

4.5 Expectations

The experimental structure above makes the comparison across groups possible in order to evaluate the theoretical structure. In this section I briefly summarize how the experiment connects to the theory section and discuss three groups of expectations for the experimental results.

First, there are two treatment groups to determine whether the severity of backsliding actions is important for a voter determining whether to support a candidate who prioritizes democratic restoration after backsliding. The proportion of respondents that select the restorer in the second treatment group (the more extreme backsliding) provides an estimate for the proportion of means voters in the sample. Meanwhile, a difference between the two treatment groups would provide evidence that there is a contingent of balance voter respondents who weigh the trade-off between policy and backsliding and have some threshold above which a retaliator is preferred and below which a restorer is preferred. If there is no difference between the two treatments, it means that voters are either not discerning backsliding severity, or do not care and intend to retaliate regardless of how the rules were broken. A future study can be designed to provide a wider array of backsliding actions with different levels of severity to assess an individual and aggregate cut-off for balance voter respondents.

Second, the control condition should provide a useful baseline from which to compare the two experimental conditions. These groups directly correspond to the three columns for the incumbent actions in table 4.1. The absence of backsliding in the control is useful for assessing initial support for backsliding without provocation and should allow for the detection of ends voters who are concerned about policy outputs and not concerned with the dissolution of democracy. The proportion of respondents that select the retaliator in

this group then can be compared to the proportions of the two treatments to determine whether the actions of the incumbent are important in determining the preferred actions of the successor. If there are meaningful differences in support for retaliators in the two treatment conditions when compared to the control, this indicates that respondents are more likely to prefer retaliators when backsliding has occurred and provides evidence for the presence of balance and means voters. If there is no clear distinction between retaliator support between the control and treatment groups, it may show that democratic preferences are not as influential, or the choices are not as comparable as expected. For instance, a voter choosing between a status quo and initial backsliding candidate without the incumbent backsliding may not be sufficiently analogous to a choice between a retaliator or restorer after backsliding.

Lastly, there is reason to expect some heterogeneity depending on the partisan identity of the respondent. Historically, Republicans have been more associated with democratic backsliding at the state level (Grumbach 2022) and work on asymmetric constitutional hardball has asserted Republicans have been more likely to engage in democratic subversion (Fishkin and Pozen 2018; Shugerman 2019). If elected Republicans are responding to an electorate that has a higher proportion of ends voters than the populace as a whole, this may indicate important differences between Democrats and Republicans with respect to backsliding in practice. A key strength of the design is that I make sure that regardless of the partisan identification of the respondent, they are always responding to an out-party incumbent so I can explore two key decisions from past and future American elections. First, I can simulate both Democratic voters' preferences in a Democratic primary under a Republican incumbent backslider such as in 2020, and also current Biden supporters' preferences for and against backsliding in the current administration. Additionally, I can simulate the choice of current Republicans and their preferences for the kinds of candidates they will support in the

2022 midterm elections and in the 2024 Republican primary for President. If Democrats and Republicans have similar levels of support for retaliators across the experimental groups, it would indicate that voters of the two parties do not have dramatically different preferences, and the differences between the two parties in government (Mann and Ornstein 2016) may be driven by factors beyond their voters.

Each of these three expectations from the survey experiment will provide novel understanding of how previous backsliding impacts support for or against additional backsliding. This design, and the results derived from it, can provide additional understanding of the stability of democratic institutions and the extent to which they can be robust against the preferences of their own citizens. Candidate selection in these elections will be an important component in either the way back to, or the continued departure from, democracy and is a unique strength of this experimental design. If democratic failure is a multi-step process of retaliatory actions, this survey experiment provides a realistic way to look at the key decisions in one step of that process.

4.6 Results

For each of the three groups of potential results, evidence for the expected relationships was only found with respect to the first expectation: voters respond to severity of backsliding by expressing greater support for a retaliator. For the second expectation, the control group did not appear to function as intended. Because of this, I re-ran the survey experiment with an additional control group, but observed similar results. Lastly, with respect to partisanship, Republicans were not more supportive of retaliators than Democrats. If anything, they were slightly less likely to support the retaliator. This is would indicate that Republicans' tendency to be the more intractable party in government (Mann and Ornstein 2016) is not driven by

overwhelming preferences of their voters (at least within the confines of this study). I discuss each of these three groups of expected results and interpret the findings for each of the two iterations of the experiment: study 1 and study 2.

4.6.1 Study 1

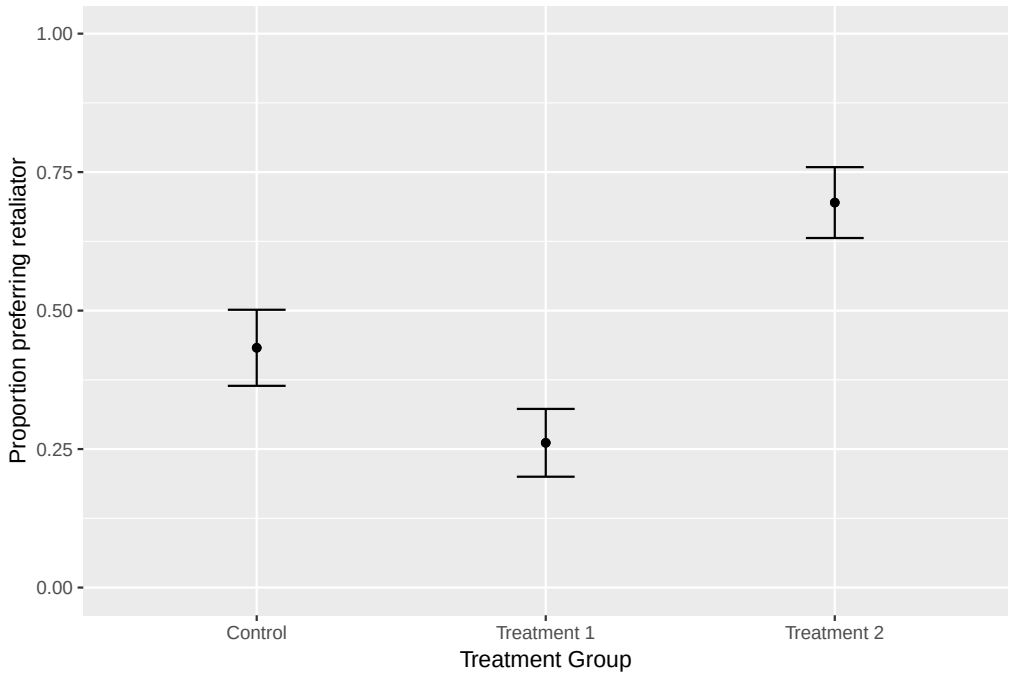
The first survey experiment was fielded on Prolific, a survey platform, between June 16 and June 19 of 2022. It was distributed to a representative sample of 600.[25] The basic results by treatment group are shown in figure 4.3. In each figure, I include the point estimate which is the proportion of respondents in a given group that preferred the retaliator. In addition to the point estimates, I include 95% confidence intervals. I only include results with respect to the respondent's choice between the two candidates for the election[26]. To assess results with respect to partisanship, I display results for each treatment group divided by party.

In figure 4.3, I display the basic results by treatment group. Each group was statistically different from the others, but not in the expected direction with respect to the control group. In line with expectations, I found that support for the retaliator is dramatically higher in the second treatment group compared to the first treatment group. This shows that respondents were more likely to support the retaliator the more extreme the backsliding action. Retaliator support was 0.434 higher in the treatment 2 than in treatment 1. Results with respect to the control group were less expected, where more supported the retaliator than in treatment 1 (0.172 higher in the control), but less than in treatment 2 (0.262 lower in the control).

[25]Prolific provides a service to acquire a representative sample. While the sample is not balanced by partisanship, as are most survey platforms, demographic balance of gender, race, and age are in line with estimates from the 2020 U.S. Census. Around 71% of the respondents selected the Democratic policy as their preferred policy.

[26]Results for which candidate the respondent believed was better for the country and which candidate they were most likely to support in the future produced similar results. This means that the respondents' evaluations of the candidates were in line with their main forced-choice, despite one being explicitly less democratic.

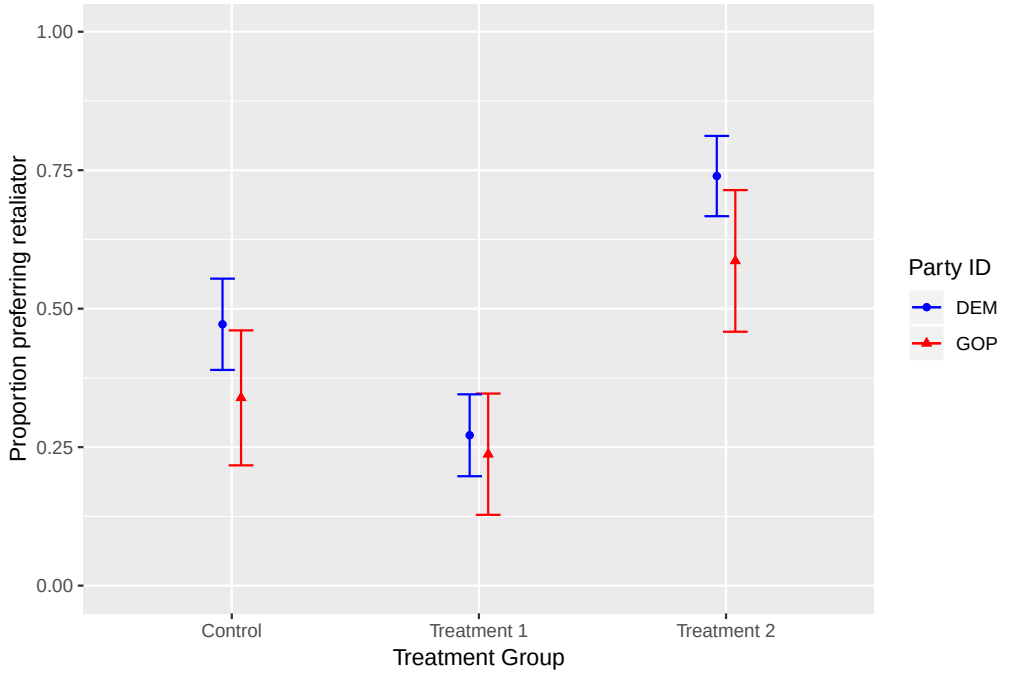
Figure 4.3: Study 1 Results by Treatment Group (Point Estimates with 95% Confidence Intervals).



One reason for the unexpected control group results may be the nature of the vignette. The control group only outlined that a disfavored law had been passed and no information was provided regarding its passage. In American politics, there has been a trend of increased constitutional hardball that has bordered on backsliding (Fishkin and Pozen 2018). In this case, voters may have presumed some malfeasance worse than the condition in treatment 1. Additionally, the choice between a retaliator and a restorer in the two treatment groups may be distinct from the choice between a status quo candidate and a backsliding candidate in the control. These factors may have led to some of the differences. Regardless, about 43% of respondents in the control group preferred the backsliding-minded candidate over the status quo candidate. Within the theoretical structure, and considering that some backsliding may be implied, I would expect that value represents the ceiling of ends voters in this sample, but because the level in treatment 1 is lower, the proportion of ends voters may be as well. Meanwhile, treatment 2 helps assess the proportion of means voters, who select the restorer

despite severe democratic subversion. Due to these unexpected results with the control group, I re-ran the survey experiment with an additional control group with an explicit statement that the law was made possible due to its majority support. I report these results in Study 2.

Figure 4.4: Study 1 Results by Party and Treatment Group.



The final set of expected results was regarding partisanship. I expected that Republicans would be more likely to prefer the retaliator, but contrary to this expectation, when aggregating the results across the treatment groups, Republicans across all groups tended to prefer the retaliator a statistically significant 0.109 *less*. In figure 4.4, I disaggregate the results by party to show how partisanship impacted each of the treatment groups. Within the three groups, there was only a significant difference in treatment group 2[27] This may have been driven by the fact that the vignette about the Supreme Court for Democrats, where Republicans

[27]At the $p < 0.05$ level. The control group had a statistically significant difference at the $p < 0.1$ level.

engaged in court packing, was closer to reality than the reverse case for Republicans. I discuss the importance of these results more in study 2.

Ultimately, I found that more extreme backsliding actions were met by voters with higher preferences for retaliators. This shows that voters respond to the severity of backsliding and provides evidence for a significant proportion of balance voters, though the proportion in this study is difficult to assess because the control group did not function as intended. I re-ran the survey experiment and included an additional control group to see if these results were driven by the control group phrasing and if the results were replicable. Lastly, in contrast with expectations, Republicans supported the retaliator less than Democrats.

4.6.2 Study 2

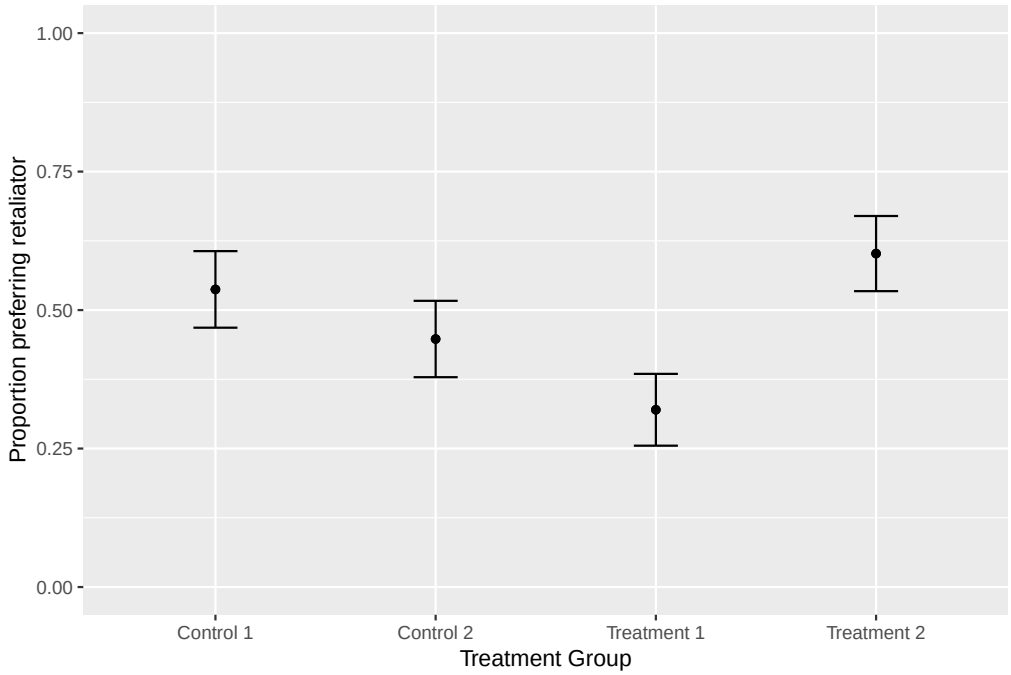
The second survey experiment was fielded on Prolific, between June 22 and June 24 of 2022 [28]. The survey was fielded to a representative sample of 800[29]. I re-ran the survey with an additional control group to see if the results in study 1 were driven by an unintended implication of backsliding or malfeasance in the control group.

In figure 4.5, I show the by-group results where the results mirror those in study 1. In this iteration of the survey experiment, the main finding from study 1 — the higher preference for the retaliator in treatments 2 over treatment 1 — remained significant in study 2. In this case, support for the retaliator was 0.345 higher in treatment group 2 than in treatment

[28] Around 95% of the responses were recorded prior to the release of the Supreme Court's decision on Dobbs v. Jackson Women's Health Organization on the morning of July 23rd. While the issues of abortion and the Supreme Court were important in the survey design, I do not believe the salience of the decision significantly impacts these results because of the number of respondents that may have been affected by it was quite small. 183 respondents selected abortion as their most important issue area (the second most selected issue after healthcare which was selected by 208) of these, about 70% of respondents preferred the policy from the Democrats.

[29] Similar to Study 1, the results were balanced on gender, age, and race/ethnicity. Around 77% of respondents in this study selected the Democratic policy as their preferred policy, so this sample is slightly more liberal than in Study 1.

Figure 4.5: Study 2 Results by Treatment Group.



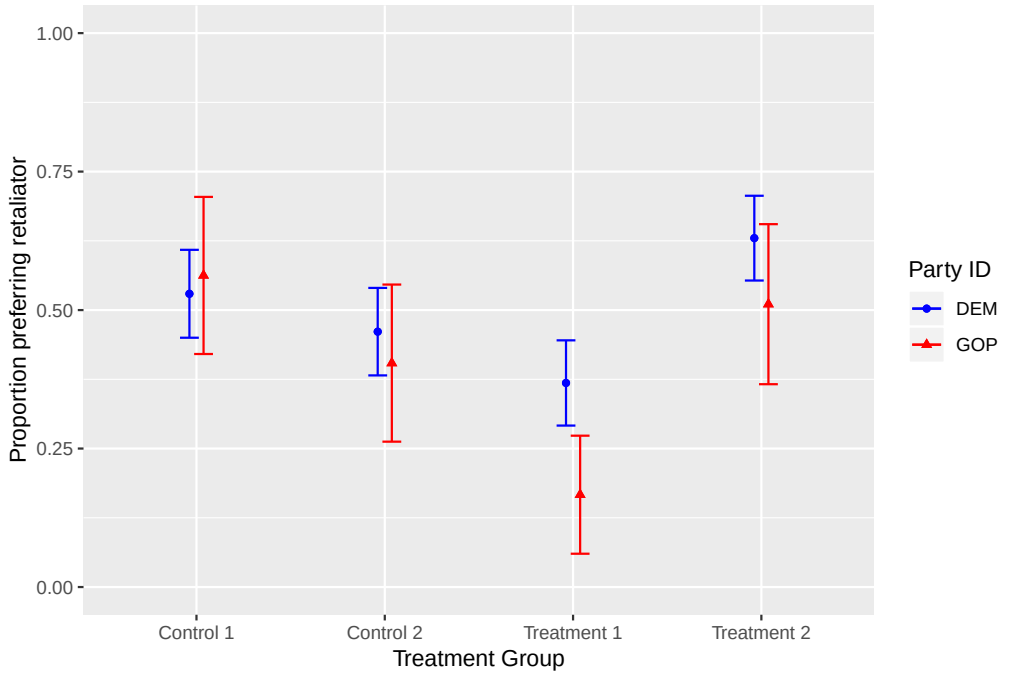
1. This gave additional evidence that voters are more likely to prefer a backsliding-minded candidate the more backsliding the previous government perpetrated.

This result is noteworthy because it is in direct contrast with the incentives for the successor in Chapter 3. In the theoretical model from Chapter 3, the more extreme the backsliding of the initial incumbent, the more a successor should prefer *restoration*, whereas in both study 1 and 2, the comparison between treatments 1 and 2 show that voters prefer *retaliation* more after severe backsliding. This shows that there is conflict between the long-term best choice of incoming successors and the more short-term preferences of the voters who elect them. These competing incentives likely lead to restorers being less preferred when they actually provide the long-term best option for a voter. The means-ends framework in this chapter, helps contrast with the long-term approach of Chapter 3 to help detail why the prospects for restoration may not be so simple.

Despite the addition of a second control group, results regarding the treatment groups relative to the control group produced similar results. The first control group (the same as study 1) was only statistically distinguishable from treatment group 1, where the voters preferred a retaliator after minor backsliding 0.312 less than when no information was provided. Meanwhile, control 2, where respondents were informed that the disfavored law passed with majority support, was statistically distinguishable from both treatment groups, but in the same direction as study 1. When compared to the second control group, support for the retaliator was 0.14 higher in the treatment 2 and 0.205 lower in treatment 1. Even though the control groups did not produce the expected results, the two control groups produced slightly different results from each other. There was slightly lower support (0.107) for the retaliator in control 2 than control 1 (though only significant at the $p < 0.1$ level). This difference provided some evidence that the initial control group may be capturing more with respect to bias against the opposite-party incumbent government than initially intended. These results, paired with the previous iteration, indicate that the control condition is not as comparable to the treatment groups as expected. Whether those unexpected findings were driven by built-in assumptions of backsliding or some other factor remains unclear. Future work can restructure the control group to provide an additional reference point.

Lastly, the the second experiment produced roughly the same results with respect to partisanship at the aggregate, but different results by group. In figure 4.6, I break down the study 2 results by partisanship. Republicans, when aggregating across all groups, preferred the retaliator about 0.08 less, and this was a significant difference aggregated across all groups. Similar to how the results within treatment 2 in study 1 drove results, this result seems to be driven by the difference in treatment 1. While the result across all groups was the same in both studies, in control groups, Republicans and Democrats were statistically undistinguishable, and there are no consistent results for either treatment groups across the

Figure 4.6: Study 2 Results by Party and Treatment Group



two studies. Lastly, because the sample was not balanced by partisanship, a greater number of Republican respondents may impact the results. Future work can further explore the role of partisanship, but in this inquiry, the results are mixed and sometimes in the unexpected direction.

4.7 Implications

Combining the results from both iterations of the survey experiment, one of the expectations outlined in section 4.5 was supported while the others produced unexpected results. I discuss the results relative to the expectations and the implications of these findings for the understanding of voters' democratic preferences.

First, when comparing the two treatment groups, the level of backsliding perpetrated by an initial backslider is an important factor for voters to determine whether to support a

restorer or a retaliator. In line with expectations, when backsliding actions are more extreme, voters are less likely to be focused on the means (democratic governance) and more focused on policy results that a retaliator provides. This finding is in contrast to the incentives for the successor to a backsliding incumbent from Chapter 3, where the more extreme the backsliding, the more appealing it is to restore institutions. With respect to voters, the long-term prospects for policy and democracy may not be as important. When voters care most about policy and retaliation, the prospects for democracy are the most fragile, especially when choosing a successor to an already effective backslider. This dynamic is an important implication of both Chapters 3 and 4 and provides expectations for the persistence of a cycle of retaliatory backsliding.

Second, results comparing the control groups to the treatment groups were unexpected. Retaliator support in the control conditions fell between the two treatment conditions instead of being the expected baseline from which retaliator support would increase. This could be driven by the fact that the two candidates in the control were not as comparable to those in the treatment conditions. The matchup between an initial backslider versus a status quo candidate in the control groups, and a retaliator versus a restorer in the treatment groups may be too distinct from each other to cleanly compare. Additionally, respondents may have inferred unsavory behavior in the absence of more information of how their disfavored policy was passed. Given the proliferation of constitutional hardball in recent years, it may be difficult for respondents to imagine the opposite party behaving with the utmost deference to norms and institutions. The slight difference between the two control groups in study 2 provided some very modest support that this may be a factor at play, but was not enough to produce the anticipated results. In the future, establishing a more appropriate control is an important area to extend and replicate this research. Lastly, treatment group 1 may not have adequately represented backsliding to voters if they believed the provision of additional

powers by congress was legitimate (Christenson and Kriner 2017*b*). Some other example of executive overreach may have provided more expected results. Regardless, the control groups relative to treatment 2 produced the expected results.

Third and finally, with respect to openness to backsliding actions, Republican politicians in the real world and Republican respondents from these studies were less aligned than anticipated [30]. This unexpected difference could be the result of current events influencing Democrats and Republicans and impacting the way respondents interpreted the vignettes. Democrats currently have unified control of government, so Republicans may prefer retaliation less because they know their leaders cannot currently leverage executive overreach. That said, the experiment simulates a response to an opposite party administration, with which Republicans can relate. I believe the current "moment" in American politics, where Democrats are focused on responding to the Trump administration and Republicans look to future elections and their response to the Biden Administration, helps provide some internal validity to these choices, so this explanation is less convincing. However, built in advantages for Republicans in backsliding already enacted may make the vignettes for Democratic respondents activate real-life preferences more than for Republicans, especially when such actions, like gerrymandering, have gone unpunished. These issues are a consequence of survey experimental research in general, and there may eventually be important real-world cases to design a study around that produce results with greater external validity. For the purpose of this inquiry, the design was an important first step in evaluating the important decision to retaliate or restore institutions.

While some of the expectations of this project were not observed in the data, the research has important implications for the prospects of democratic resurgence. In a setting where a

[30]Disaggregating these preferences further may show a distinction between different kinds of Republicans. Such as strong Trump supporters and more weaker supporters, or another important dimension like evangelical identification.

backsliding incumbent was removed in a democratic election, the more severe the backsliding incumbent, the more opposite-party voters prefer a candidate that retaliates more over a candidate who tries to improve democratic quality and reduce tension. This preference is in line with the concept of retaliatory hardball. This result is driven by how voters weight their policy preferences relative to their democratic preferences: the means and ends trade-off. Under more modest transgressions, democratic resurgence may be more possible than in the wake of considerable democratic subversion. This provides a grim picture for countries like Hungary, Poland, and Turkey which are tightly controlled by backsliders who levied multiple significant backsliding actions.

Future work can consider the different kinds of democratic backsliding featured in Chapters 2 and 3, where I expect actions targeting electoral institutions would be especially alarming to voters. For the sake of simplicity in this chapter, I focused on executive aggrandizement, but the strategic manipulation of elections may produce more extreme results. Further, alternative settings, such as the subnational level, where Grumbach's "laboratories against democracy" are most prolific may be useful settings to assess voter preferences. Lastly, in the comparative context, other countries with significantly different institutions may produce different results driven by institutional structure or the nature of partisanship and polarization. In a setting like Turkey, where polarization exists more on a secular-Islamism dimension as opposed to a left-right distinction, important differences may arise. There is considerable work that still needs to be done to help refine theoretical expectations.

4.8 Conclusion

When backsliding incumbents are eventually removed, the preferences of the successors and their voters will be important to determine the prospects for democratic resurgence, or further decline. The work in this chapter, combined with the theoretical work in Chapter 3 provided some useful, albeit competing results regarding the preferences for restoration or retaliation.

I argued, in coordination with existing work, that voters consider a trade-off between the means employed and the ends realized when supporting a backsliding-minded politician. I built on this work by looking at the vote choice of out-party voters and the potential way back to democratic governance after backsliding. I also detailed a typology of voters and the level of their value on democratic governance or policy goals in leading to their choice to support retaliatory or restoration-minded politicians. Using this theoretical structure, I executed a survey experiment in the United States to evaluate the expectations derived from the theory.

The results from two iterations of this survey experiment showed that voters more strongly value a retaliator when the initial backslider enacts more extreme backsliding measures. Results relative to a situation where the incumbent did not backslide provided mixed results, but interesting points of departure for future research. Additionally, expectations that Republicans would be more likely to support retaliators were not found. If anything, Republicans supported restorers more. Re-running this experiment after the election of a unified Republican government may provide interesting complimentary results.

This research helped us learn how preferences for democratic institutions and norms may be situational depending on the actions of the previous government, and how this context impacts democratic stability. By evaluating this theory empirically in a setting where

backsliding has yielded to an opposition party, I was able to shed light on how we should expect democracy to rebound or stagnate after backsliding. The prospect for democracy after democratic backsliding is highly contingent on the preference of voters. If after backsliding, voters want their leaders to engage in a tit-for-tat approach where democracy continues to be eroded and likely eventually degrade completely, then the likelihood of positive developments in states recently exiting backsliding regimes like the United States is not high. This is extended to the longer term prospects in countries like Poland, Hungary, and Turkey with current legacies of backsliding and many more countries that have experienced democratic recession in recent years.

References

Adorno, Theodor W, Else Frenkel-Brunswik, Daniel J Levinson, and R Nevitt Sanford. 1950. "The Authoritarian Personality." *Social Studies Series: Publication No III, New York: Harper and Row*: Forthcoming.

Aleinikoff, T Alexander. 2018. "Inherent Instability: Immigration and Constitutional Democracies.". In *Constitutional Democracy in Crisis*. Oxford University Press.

Alesina, Alberto. 1988. "Credibility and policy convergence in a two-party system with rational voters." *The American Economic Review* 78(4): 796–805.

Alexander, Amy C, and Christian Welzel. 2017. The myth of deconsolidation: Rising liberalism and the populist reaction. Technical report ILE Working Paper Series.

Almond, Gabriel A, and Sidney Verba. 1963. *An approach to political culture*. JSTOR.

Andeweg, Rudy B. 2000. "Consociational democracy." *Annual Review of Political Science* 3(1): 509–536.

Baccini, Leonardo, and Thomas Sattler. 2021. "Austerity, economic vulnerability, and populism." *Available at SSRN 3766022*: Forthcoming.

Balkin, Jack M. 2017. "Constitutional crisis and constitutional rot." *Md. L. Rev.* 77: 147.

Ballard-Rosa, Cameron, Mashail A Malik, Stephanie J Rickard, and Kenneth Scheve. 2021. "The economic origins of authoritarian values: evidence from local trade shocks in the United Kingdom." *Comparative political studies* 54(13): 2321–2353.

BBC. 2015. "Turkey's Pro-Kurdish HDP Party Offices Attacked in Ankara." BBC.

BBC. 2021. "Turkey moves to ban pro-Kurdish HDP opposition party." BBC.

Beals, Monique. 2022. "Biden: No plans for Executive Order on Guns.".

Beaulieu, Emily. 2014. "From voter ID to party ID: How political parties affect perceptions of election fraud in the US." *Electoral Studies* 35: 24–32.

Bermeo, Nancy. 2009. "Poverty, Inequality, and Democracy (II): Does Electoral Democracy Boost Economic Equality?" *Journal of Democracy* 20(4): 21–35.

Bermeo, Nancy. 2016. "On democratic backsliding." *Journal of Democracy* 27(1): 5–19.

Boix, Carles. 2010. "Origins and persistence of economic inequality." *Annual Review of Political Science* 13(1): 489–516.

Braley, Alia, Gabriel S Lenz, Dhaval Adjodah, Hossein Rahnama, and Alex Pentland. 2021. "The Subversion Dilemma: Why Voters Who Cherish Democracy Vote it Away." *Available at http s://osf. io/my987*: Forthcoming.

Breuninger, Kevin. 2022. "U.S. judge Blocks Biden vaccine mandate for federal workers in latest blow to White House covid agenda." CNBC.

Bunce, Valerie. 2000. "Comparative democratization: Big and bounded generalizations." *Comparative political studies* 33(6-7): 703–734.

Bustikova, Lenka. 2014. "Revenge of the radical right." *Comparative Political Studies* 47(12): 1738–1765.

Çakır, Ruşen. 2012. "Erdoğan: Demokrasi Tramvayında Ihtiraslı Bir Yolcu." Gazete Vatan.

Cheibub, José Antonio et al. 2007. *Presidentialism, parliamentarism, and democracy*. Cambridge University Press.

Christenson, Dino P, and Douglas L Kriner. 2017*a*. "Constitutional qualms or politics as usual? The factors shaping public support for unilateral action." *American Journal of Political Science* 61(2): 335–349.

Christenson, Dino P, and Douglas L Kriner. 2017*b*. "Mobilizing the public against the president: Congress and the political costs of unilateral action." *American Journal of Political Science* 61(4): 769–785.

Chu, Yun-han, Kai-Ping Huang, Marta Lagos, and Robert Mattes. 2020. "A lost decade for third-wave democracies?" *Journal of Democracy* 31(2): 166–181.

Colantone, Italo, and Piero Stanig. 2018. "The trade origins of economic nationalism: Import competition and voting behavior in Western Europe." *American Journal of Political Science* 62(4): 936–953.

Coppedge, Michael, John Gerring, Adam Glynn, Carl Henrik Knutsen, Staffan I Lindberg, Daniel Pemstein, Brigitte Seim, Svend-Erik Skaaning, and Jan Teorell. 2020. "Varieties of Democracy." : Forthcoming.

Dahl, Robert A. 1971. "Polyarchy: Participation and Opposition. New Haven: Yale University Press." : Forthcoming.

Di Lonardo, Livio. 2019. "The partisan politics of counterterrorism: reputations, policy transparency, and electoral outcomes." *Political Science Research and Methods* 7(2): 253–269.

Diamond, Larry. 2015. "Facing up to the democratic recession." *Journal of Democracy* 26(1): 141–155.

Diamond, Larry, and Leonardo Morlino. 2004. "The quality of democracy: An overview." *Journal of democracy* 15(4): 20–31.

Doty, Richard M, Bill E Peterson, and David G Winter. 1991. "Threat and authoritarianism in the United States, 1978–1987." *Journal of personality and social psychology* 61(4): 629.

Dresden, Jennifer Raymond, and Marc Morjé Howard. 2016. "Authoritarian backsliding and the concentration of political power." *Democratization* 23(7): 1122–1143.

Driesen, David M. 2021. *The Specter of Dictatorship: Judicial Enabling of Presidential Power*. Stanford University Press.

Duckitt, John, and Kirstin Fisher. 2003. "The impact of social threat on worldview and ideological attitudes." *Political Psychology* 24(1): 199–222.

Fearon, James D. 1999. "Electoral accountability and the control of politicians: selecting good types versus sanctioning poor performance." *Democracy, accountability, and representation* 55: 61.

Fishkin, Joseph, and David E Pozen. 2018. "Asymmetric constitutional hardball." *Columbia Law Review* 118(3): 915–982.

Foa, Roberto Stefan, and Yascha Mounk. 2016. "The danger of deconsolidation: The democratic disconnect." *Journal of democracy* 27(3): 5–17.

Foa, Roberto Stefan, and Yascha Mounk. 2017*a*. "The end of the consolidation paradigm." *Journal of Democracy*: Forthcoming.

Foa, Roberto Stefan, and Yascha Mounk. 2017*b*. "The signs of deconsolidation." *Journal of democracy* 28(1): 5–15.

Freeman, James. 2021. "Opinion | mason-dixon poll: Americans still don't like court-packing." Dow Jones amp; Company.

Geddes, Barbara et al. 1999. "What do we know about democratization after twenty years?" *Annual review of political science* 2(1): 115–144.

Gibler, Douglas M, and Kirk A Randazzo. 2011. "Testing the effects of independent judiciaries on the likelihood of democratic backsliding." *American Journal of Political Science* 55(3): 696–709.

Goldstone, Jack A, Robert H Bates, David L Epstein, Ted Robert Gurr, Michael B Lustik, Monty G Marshall, Jay Ulfelder, and Mark Woodward. 2010. "A global model for forecasting political instability." *American Journal of Political Science* 54(1): 190–208.

Gonzalez, Francisco E. 2008. *Dual transitions from authoritarian rule: Institutionalized regimes in Chile and Mexico, 1970–2000*. JHU Press.

Graham, Matthew H, and Milan W Svolik. 2020. "Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States." *American Political Science Review* 114(2): 392–409.

Grillo, Edoardo, and Carlo Prato. 2020. "Reference points and democratic backsliding." *American Journal of Political Science*: Forthcoming.

Gross, Oren. 2018. "The normal exception." *Constitutional Democracy in Crisis*: Forthcoming.

Grossman, Guy, Dorothy Kronick, Matthew Levendusky, and Marc Meredith. 2022. "The Majoritarian threat to liberal democracy." *Journal of Experimental Political Science* 9(1): 36–45.

Grumbach, Jacob. 2022. *Laboratories against democracy: How national parties transformed state politics*. Vol. 182 Princeton University Press.

Gunitsky, Seva. 2018. "Democratic waves in historical perspective." *Perspectives on Politics* 16(3): 634–651.

Haggard, Stephan, and Robert R Kaufman. 1999. "4. The Political Economy of Democratic Transitions." In *Transitions to democracy*. Columbia University Press , 72–96.

Haggard, Stephan, and Robert R Kaufman. 2012. "Inequality and regime change: Democratic transitions and the stability of democratic rule." *American Political Science Review* 106(3): 495–516.

Haggard, Stephan, and Robert R Kaufman. 2016. "Democratization during the third wave." *Annual Review of Political Science* 19: 125–144.

Helmke, Gretchen, Mary Kroeger, and Jack Paine. 2019. "Exploiting asymmetries: A theory of democratic constitutional hardball." *Unpublished manuscript, University of Rochester*: Forthcoming.

Hetherington, Marc, and Elizabeth Suhay. 2011. "Authoritarianism, threat, and Americans' support for the war on terror." *American Journal of Political Science* 55(3): 546–560.

Howell, William G, Kenneth Shepsle, and Stephane Wolton. 2019. "Executive absolutism: the dynamics of authority acquisition in a system of separated powers." *Available at SSRN 3440604*: Forthcoming.

Huntington, Samuel P. 1968. *Political order in changing societies*. Yale university press.

Huntington, Samuel P. 1991. "Democracy's third wave." *Journal of democracy* 2(2): 12–34.

Huq, Aziz, and Tom Ginsburg. 2018. "How to lose a constitutional democracy." *UCLA L. Rev.* 65: 78.

Inglehart, Ronald, and Pippa Norris. 2017. "Trump and the populist authoritarian parties: the silent revolution in reverse." *Perspectives on Politics* 15(2): 443–454.

Issacharoff, Samuel. 2018. "Populism versus democratic governance." *GRABER, Mark A.; LEVINSON, Sanford; TUSHNET, Mark. Constitutional democracy in crisis*: Forthcoming.

Kelemen, R Daniel, and Michael Blauberger. 2017. "Introducing the debate: European Union safeguards against member states' democratic backsliding." *Journal of European Public Policy* 24(3): 317–320.

Kelly, Jana Morgan. 2003. "Counting on the past or investing in the future? Economic and political accountability in Fujimori's Peru." *The Journal of Politics* 65(3): 864–880.

King, Desmond, and Rogers M Smith. 2018. "Populism, Racism, and the Rule of Law in Constitutional Democracies Today." *Constitutional Democracy in Crisis*: Forthcoming.

Landau, David. 2018. "Constitution-making and authoritarianism in Venezuela: the first time as tragedy, the second as farce." *Constitutional democracy in crisis*: Forthcoming.

Lapper, Richard. 2021. "Beef, Bible and Bullets: Brazil in the age of Bolsonaro." In *Beef, Bible and Bullets*. Manchester University Press.

Lavine, Howard, Milton Lodge, and Kate Freitas. 2005. "Threat, authoritarianism, and selective exposure to information." *Political psychology* 26(2): 219–244.

Levitsky, Steven. 1999. "Latin America's Imperiled Progress: Fujimori and Post-Party Politics in Peru." *Journal of Democracy* 10(3): 78–92.

Levitsky, Steven, and Daniel Ziblatt. 2018. *How democracies die.* Broadway Books.

Levitsky, Steven, and Lucan A Way. 2010. "Democracy's past and future: Why democracy needs a level playing field." *Journal of Democracy* 21(1): 57–68.

Levitsky, Steven, and Maxwell A Cameron. 2003. "Democracy without parties? Political parties and regime change in Fujimori's Peru." *Latin American Politics and Society* 45(3): 1–33.

Lijphart, Arend. 1977. *Democracy in plural societies: A comparative exploration.* Yale University Press.

Linz, Juan J. 1990. "The perils of presidentialism." *Journal of democracy* 1(1): 51–69.

Linz, Juan J, and Alfred Stepan. 1978. *The breakdown of democratic regimes (Vol. 1).* Johns Hopkins University Press.

Lipset, Seymour Martin. 1959. "Some social requisites of democracy: Economic development and political legitimacy1." *American political science review* 53(1): 69–105.

Liptak, Adam. 2022. "Supreme Court refuses for now to restore Biden Plan on immigration enforcement." The New York Times.

Luca, Davide, Javier Terrero-Davila, Jonas Stein, and Neil Lee. 2022. "Progressive Cities: Urban-rural polarisation of social values and economic development around the world." : Forthcoming.

Lucassen, Geertje, and Marcel Lubbers. 2012. "Who fears what? Explaining far-right-wing preference in Europe by distinguishing perceived cultural and economic ethnic threats." *Comparative Political Studies* 45(5): 547–574.

Lührmann, Anna, and Staffan I Lindberg. 2019. "A third wave of autocratization is here: what is new about it?" *Democratization* 26(7): 1095–1113.

Luo, Zhaotian, and Adam Przeworski. 2019. "Democracy and its vulnerabilities: Dynamics of democratic backsliding." *Available at SSRN 3469373*: Forthcoming.

Maeda, Ko. 2010. "Two modes of democratic breakdown: A competing risks analysis of democratic durability." *The Journal of Politics* 72(4): 1129–1143.

Mainwaring, Scott, and Aníbal Pérez-Liñán. 2013*a*. *Democracies and dictatorships in Latin America: emergence, survival, and fall.* Cambridge University Press.

Mainwaring, Scott, and Aníbal Pérez-Liñán. 2013*b*. "Lessons from Latin America: Democratic breakdown and survival." *Journal of Democracy* 24(2): 123–137.

Mainwaring, Scott, and Matthew S Shugart. 1997. "Juan Linz, presidentialism, and democracy: a critical appraisal." *Comparative Politics*: Forthcoming.

Mann, Thomas E, and Norman J Ornstein. 2016. *It's even worse than it looks: How the American constitutional system collided with the new politics of extremism*. Basic Books.

Maravall, Jose Maria. 1999. "Accountability and Manipulation." *Democracy, accountability, and representation*: Forthcoming.

Marshall, Monty G, Ted Robert Gurr, and Keith Jaggers. 2017. "Center for Systemic Peace." *Polity IV Project*: Forthcoming.

Mate, Manoj. 2018. "Constitutional Erosion and the Challenge to Secular Democracy in India." *Constitutional Democracy in Crisis*: Forthcoming.

Mazepus, Honorata, and Dimiter Toshkov. 2022. "Standing up for Democracy? Explaining Citizens' Support for Democratic Checks and Balances." *Comparative Political Studies* 55(8): 1271–1297.

McCann, Stewart JH. 2008. "Societal threat, authoritarianism, conservatism, and US state death penalty sentencing (1977-2004)." *Journal of Personality and Social Psychology* 94(5): 913.

Mechkova, Valeriya, Anna Lührmann, and Staffan I Lindberg. 2017. "How much democratic backsliding?" *Journal of Democracy* 28(4): 162–169.

Mercieca, Jennifer. 2020. *Demagogue for president: The rhetorical genius of Donald Trump*. Texas A&M University Press.

Merolla, Jennifer L, and Elizabeth J Zechmeister. 2009. "Democracy at risk." In *Democracy at Risk*. University of Chicago Press.

Milner, Helen V, and Bumba Mukherjee. 2009. "Democratization and economic globalization." *Annual Review of Political Science* 12(1): 163–181.

Mutz, Diana C. 2018. "Status threat, not economic hardship, explains the 2016 presidential vote." *Proceedings of the National Academy of Sciences* 115(19): E4330–E4339.

Nalepa, Monika, Georg Vanberg, and Caterina Chiopris. 2018. "Authoritarian backsliding." *Unpublished manuscript, University of Chicago and Duke University*: Forthcoming.

Norris, Pippa. 2017. "Is Western democracy backsliding? Diagnosing the risks." *Forthcoming, The Journal of Democracy, April*: Forthcoming.

O'Donnell, Guillermo A. 1998. "Horizontal accountability in new democracies." *Journal of democracy* 9(3): 112–126.

O'Donnell, Guillermo, Philippe C Schmitter, and Laurence Whitehead. 2013. *Transitions from authoritarian rule: Tentative conclusions about uncertain democracies*. JHU Press.

Parker, Claire. 2022. "Resistance builds in Tunisia as populist leader seeks more power." WP Company.

Pérez-Liñán, Aníbal, Nicolás Schmidt, and Daniela Vairo. 2019. "Presidential hegemony and democratic backsliding in Latin America, 1925–2016." *Democratization* 26(4): 606–625.

Pozen, David E. 2018. "Hardball and/as Anti-hardball." *NYUJ Legis. & Pub. Pol'y* 21: 949.

Przeworski, Adam, and Fernando Limongi. 1997. "Modernization: Theories and facts." *World politics* 49(2): 155–183.

Przeworski, Adam, Susan Carol Stokes Stokes, Susan C Stokes, and Bernard Manin. 1999. *Democracy, accountability, and representation*. Vol. 2 Cambridge University Press.

Reeves, Andrew, and Jon C Rogowski. 2015. "Public opinion toward presidential power." *Presidential Studies Quarterly* 45(4): 742–759.

Reeves, Andrew, and Jon C Rogowski. 2016. "Unilateral powers, public opinion, and the presidency." *The Journal of Politics* 78(1): 137–151.

Reeves, Andrew, and Jon C Rogowski. 2018. "The public cost of unilateral action." *American Journal of Political Science* 62(2): 424–440.

Reeves, Andrew, Jon C Rogowski, Min Hee Seo, and Andrew R Stone. 2017. "The contextual determinants of support for unilateral action." *Presidential Studies Quarterly* 47(3): 448–470.

Repucci, Sarah, and Amy Slipowitz. 2022. "Freedom in the world 2022." : Forthcoming.

Reynolds, Andrew. 2010. *Designing democracy in a dangerous world*. OUP Oxford.

Riedl, Rachel Beatty, Dan Slater, Joseph Wong, and Daniel Ziblatt. 2020. "Authoritarian-led democratization." *Annual Review of Political Science* 23(1): 315–32.

Rustow, Dankwart A. 1970. "Transitions to Democracy." *Comparative Politics* 2(3): 337–363.

Sadurski, Wojciech. 2018. "Constitutional Crisis in Poland." *Constitutional Democracy in Crisis*: Forthcoming.

Sales, Stephen M. 1973. "Threat as a factor in authoritarianism: an analysis of archival data." *Journal of personality and social psychology* 28(1): 44.

Sands, Leo, and Matt Murphy. 2022. "Tunisia referendum: Voters give president near unchecked power." BBC.

Schedler, Andreas. 2002. "The nested game of democratization by elections." *International Political Science Review* 23(1): 103–122.

Schedler, Andreas. 2010. "Democracy's past and future: Authoritarianism's last line of defense." *Journal of democracy* 21(1): 69–80.

Scheppele, Kim Lane. 2018. "The Party's Over.". In *Constitutional Democracy in Crisis*. Oxford University Press.

Shugerman, Jed Handelsman. 2019. "Hardball vs. Beanball." *Columbia Law Review* 119(3): 85–122.

Singer, Matthew M. 2021. "Fiddling while Democracy Burns: Partisan Reactions to Weak Democracy in Latin America." *Perspectives on Politics*: Forthcoming.

Sitaraman, Ganesh. 2018. "Economic Inequality and Constitutional Democracy." *Constitutional democracy in crisis*: Forthcoming.

Skaaning, Svend-Erik, John Gerring, and Henrikas Bartusevičius. 2015. "A lexical index of electoral democracy." *Comparative Political Studies* 48(12): 1491–1525.

Slater, Dan, Benjamin Smith, and Gautam Nair. 2014. "Economic origins of democratic breakdown? The redistributive model and the postcolonial state." *Perspectives on Politics* 12(2): 353–374.

Smulovitz, Catalina, and Enrique Peruzzotti. 2000. "Societal Accountability in Latin America." *Journal of democracy* 11(4): 147–158.

Svolik, Milan. 2018. "When polarization trumps civic virtue: Partisan conflict and the subversion of democracy by incumbents." *Available at SSRN 3243470*: Forthcoming.

Svolik, Milan W. 2019. "Polarization versus democracy." *Journal of Democracy* 30(3): 20–32.

Tushnet, Mark. 2018. "Comparing right-wing and left-wing populism." *Constitutional Democracy in Crisis*: Forthcoming.

Varol, Ozan O. 2014. "Stealth authoritarianism." *Iowa L. Rev.* 100: 1673.

Vazquez, Maegan, Kate Sullivan, and Betsy Klein. 2022. "5 significant bills and 5 executive orders Biden signed in his first year – and his 5 biggest priorities for 2022." Cable News Network.

Voeten, Erik. 2016. "Are people really turning away from democracy?" *Available at SSRN 2882878*: Forthcoming.

Wahman, Michael, Jan Teorell, and Axel Hadenius. 2013. "Authoritarian regime types revisited: updated data in comparative perspective." *Contemporary Politics* 19(1): 19–34.

Waldner, David, and Ellen Lust. 2018. "Unwelcome change: Coming to terms with democratic backsliding." *Annual Review of Political Science* 21(1): 93–113.

Wise, Alana, and Deirdre Walsh. 2021. "House approves Major Election Reform and Voting Rights bill.".

Yokley, Eli. 2022. "Voters are divided over changing the Senate's rules for voting rights, but most back changes to the electoral count act.".