IGCC UC INSTITUTE ON GLOBAL CONFLICT AND COOPERATION

WORKING PAPER NO. 6
SEPTEMBER 2025

# The Backsliding Bloc? Selective Contestation and Alliance Realignment in the Liberal International Order

Anna M. Meyerrose and Irfan Nooruddin

## Abstract

How do democratic backsliders—states that retain aspects of democracy yet have become increasingly illiberal—behave in international institutions? Pressured between needing to preserve democratic legitimacy while minimizing external scrutiny of domestic illiberalism, these states engage in selective contestation: they continue to endorse the liberal international order (LIO), while systematically opposing its more intrusive institutions, particularly those promoting human rights and democratization. Using original data on voting and sponsorship patterns in the United Nations (UN) Human Rights Council (2006–23), we find backsliders align with autocracies, and against advanced Western democracies, on particularly sensitive and intrusive issues. These dynamics are confirmed by voting patterns in the UN General Assembly (UNGA). Text analysis of Universal Periodic Review reports reveals that backsliders' rhetoric is converging to that of other backsliders and autocracies over time. These patterns indicate backsliders are reshaping the international landscape—leaving the West isolated and undermining core facets of the LIO.

**Keywords:** Democratic backsliding, international organizations, multilateralism, human rights

## Authors

**Anna M. Meyerrose**
Assistant Professor, Arizona State University
Email: anna.meyerrose@asu.edu

**Irfan Nooruddin**
Professor, Georgetown University
Email: in62@georgetown.edu

## Suggested Citation

Meyerrose, Anna M. and Irfan Nooruddin. 2025. *The Backsliding Bloc? Selective Contestation and Alliance Realignment in the Liberal International Order*. IGCC Working Paper No 6. escholarship.org/uc/item/2w79t90g

## Introduction

After the Cold War ended, the United States and its closest allies built a powerful international coalition composed of both the core founders of the post-World War II liberal order and a larger group of advanced and new democracies. These democracies together worked through multilateral institutions to promote and protect Western liberal norms and values (Pevehouse, 2005; Donno, 2013). In recent years, consolidated autocracies—led mainly by China and Russia—have grown more brazen in their efforts to undermine this post-Cold War liberal international order (LIO) (Cottiero et al., 2025). Although more aggressive challenges by autocracies garner the most attention, the threat they pose would be less existential for Western hegemony and the stability of the LIO if its core advocates could reliably mobilize their broader coalition of democratic states to act as a counterweight against growing autocratic threat. We show here that this is no longer be the case. Today, the West must confront not just the rise of powerful autocracies, but also the consequences of an ongoing global democratic recession.

Democratic backsliding occurs when democratically elected officials (Bermeo, 2016; Bartels, 2023) intentionally weaken or erode one or more liberal democratic institutions (Haggard and Kaufman, 2021). By definition, all backsliders are democracies at the start of their transformation, and, indeed, most remain democratic even after their backsliding has plateaued, albeit much diminished and teetering on the brink of further erosion. Therefore, the consequence of the democratic recession for global governance is asymmetric, and a key international effect of backsliding is the splintering of the broad democratic coalition that the core founders of the LIO relied on in the post-Cold War era to maintain the international governance system they had constructed after World War II.

Anecdotal evidence abounds of backsliders shifting away from the West. In 2022, the United States and its allies drafted the first resolution in the United Nations (UN) Human Rights Council (UNHRC) targeting China for its ongoing human rights violations in Xinjiang. This resolution was defeated when a set of democratic backsliders joined consolidated autocracies in opposing it: Argentina, Brazil, and India abstained, while others, including Bolivia, Indonesia, and Senegal, voted against the resolution altogether.[1]

[1] "India along with Brazil abstain on UNHRC vote on Xinjiang," *The Economic Times*.

Similarly, in October 2024, for the second consecutive year, India abstained on a UNHRC resolution highlighting Russia's deteriorating human rights situation.[2] These trends extend beyond the UNHRC; in recent years, India has also joined China, Russia, and other consolidated autocracies in the UN General Assembly (UNGA) to block efforts to sanction human rights violations in Iran.[3]

Democratic backsliding has become increasingly common, yet we understand little about how these states behave on the international scene. Scholarly focus instead has been on geopolitical contestation by consolidated autocracies. Studies that have focused on backsliding have studied either its effects on the membership composition and functioning of international organizations (IOs) (Kelemen, 2020; von Borzyskowski and Vabulas, 2025; Winzen, 2025) or alternatively the ways in which backsliders leverage their votes and rhetoric within institutions to contest liberal norms and values (Lipps and Jacob, 2025; Meyerrose and Nooruddin, 2025). Knowing whether democratic backsliders are acting in concert with autocracies carries high stakes for our understanding of contemporary geopolitics. This paper tackles this challenge by seeking to identify if backsliders represent a new distinct bloc in international politics, and whether their emergence has altered alliance structures within the core institutions of the LIO.

Democratic backsliders in our conceptualization are democratic states that have shown measurable erosion in their commitments to the norms and practices considered foundational to democracy. Critically, the majority of these cases remain above minimal thresholds for democracy (see Figure 1 below)—that is, they are not autocracies and indeed remain closer on international democratic indices to consolidated democracies. But the process of undermining democratic checks and balances from within, of suppressing civil society and press freedoms to restrict criticism and accountability, we argue, should have observable effects on these states' external behaviors and preferences. These states have incentives to signal support for democratic norms to both their domestic and international audiences as a way of blunting accusations that they are fomenting autocratization. But, simultaneously, in an effort to undermine international actors that might scrutinize their behavior at home and validate domestic critics, they seek to weaken and delegitimize the oversight capacity of international institutions, especially those that are focused on human rights and democratization. In that aspect, if no other, their preferences converge with their autocratic counterparts.

---

[2] "India Expresses Guarded Concern Over Israel's Attack on UN Interim Force in Lebanon; Indian Nazi Propagandist in Hitler's Germany; Guj Journo Mahesh Langa Sent to Police Custody; Modi Lost His Mojo?" *The India Cable*, October 11, 2024. https://www.theindiacable.com/p/ india-expresses-guarded-concern-over?utm_source=publication-search.

[3] "India joins Russia, China in voting against resolution on Iran human rights at UN," *The Economic Times*, December 23, 2023, https://economictimes.indiatimes.com/news/india/ india-joins-russia-china-in-voting-against-resolution-on-iran-human-rights-at-un/ articleshow/106239146.cms?from=mdr.

These divergent impulses make predicting democratic backsliders' international behaviors complicated. They are undeniably still democratic, even if diminished by their own previous standards, and yet they are also particularly sensitive to criticism for policies and practices undermining liberal democracy at home, or, put differently, for creating the backsliding momentum. As such, they wish to be seen as members still of the international democratic "club," even as they work to undermine the very international organizations that club created to protect democracy worldwide. Therefore, we expect that, compared to consolidated autocracies, which oppose all components of the Western-backed LIO, backsliders will be more selective in their contestation. While these states might remain aligned with the West on less intrusive policies associated with the LIO (e.g., trade or currency policies), we expect them to oppose actively and systematically more intrusive international policies—such as those that promote human and minority rights or the other liberal institutions most commonly eroded in cases of backsliding—that risk validating domestic opposition critiques of their rule. By engaging in this selective contestation, they fundamentally alter the alliance structures within liberal international institutions, leaving the LIO's core founders and defenders increasingly isolated as their coalition of the willing is bled by the realignment of backsliders.

We test this argument using data from the UN human rights institutions. From their founding, these institutions have been central to the West's efforts to promote liberal international norms and values (Finnemore and Sikkink, 1998; Hafner-Burton, 2012; Voss, 2019). They are also among the most intrusive of international institutions as they subject states to regular evaluations—and public criticism—of their domestic human rights practices. Using an original data set on voting behavior and resolution sponsorship in the UNHRC between 2006–23, we show that over time backsliders have become significantly more likely to vote with consolidated autocracies and against the West on an important subset of resolutions: contentious and highly intrusive ones. We show further evidence of these dynamics in the context of the UNGA. In addition to voting patterns, we also analyze backslider states' rhetoric in the Universal Periodic Review (UPR). Using cosine similarity scores, we find that backsliders' rhetoric has become more cohesive over time as they have shifted away from the West and come to more closely resemble autocracies. However, important differences between backsliders and autocracies remain, providing evidence that the former are a distinct bloc with a unique set of interests and preferences. Taken together, these findings reveal the West is increasingly isolated at the international level as formerly democratically committed states are joining the group of backsliders, more actively and consistently opposing the most intrusive international policies and institutions, and, as a result, undermining core facets of the post-Cold War LIO.

# 1. The Authoritarian Turn in International Organizations

The LIO originated in the wake of World War I but expanded significantly after World War II. Western democracies created a new international system with multilateral institutions—most prominently, the UN, the Bretton Woods Institutions, and the General Agreement on Tariffs and Trade (GATT)—at its core (Börzel and Zürn, 2021). Through these, the West sought to institutionalize a new era of global governance that protected and furthered Western interests by promoting international economic exchange, encouraging international deliberation and policy coordination, and, more broadly, promoting liberal democratic norms and values that contrasted with Communist ideologies promulgated by the Soviet Union. Building on these foundational institutions, additional Western-led IOs proliferated after the Cold War, with politically oriented ones becoming inextricably linked to the construction of a liberal global system centered on democracy, elections, rule of law, and civil and political human rights (Barnett and Finnemore, 2021).

Liberal democratic values have been a core component of the LIO from its inception. The international human rights regime on which these values are based was first institutionalized by the Universal Declaration of Human Rights (UDHR) in 1948. The UDHR was initially drafted by Australia, Chile, China, France, Lebanon, the Soviet Union, the United Kingdom (UK), and the United States. We identify the advanced Western democracies in this group—Australia, France, the UK, and the United States—as the core founders of the LIO's human rights regime, referring to them throughout as the Human Rights Founders (HRF).[4] The liberal democratic values undergirding the LIO were first articulated by the HRF states, and after the Cold War ended, these states were at the heart of a global coalition of new and advanced democracies that continued to promote these values.

Today, there is growing concern that the HRF-backed LIO is under real threat for the first time since 1930 (Ikenberry, 2018). On the one hand, universalist multilateral institutions, including the UN, have always been plagued by internal dissent from illiberal regimes whose interests are not served or represented by the Western liberal agenda (Lake, Martin, and Risse, 2021).

[4] Other states can no doubt lay claim to sharing the parentage of the human rights regime. Our goal here is simply to identify a set of states whose participation in the drafting of the original documents is undisputed and whose interests over the next 75 years can clearly be contrasted with the authoritarian bloc. As discussed below, our results are robust to widening the set of countries included in the HRF category.

But as the relative power and influence of the United States has waned in the face of increasingly powerful and assertive China and other autocratic powers (Cooley and Nexon, 2020; Hyde, 2020), the LIO has been more overtly contested (Weiss and Wallace, 2021; Boyle, 2023; Pauselli, Urdínez, and Merke, 2023; Cottiero et al., 2025). This challenge against the Western-backed order has grown with the emergence of IOs composed primarily, if not entirely, of autocratic countries (Libman and Obydenkova, 2018; Kneuer et al., 2019; Cottiero and Haggard, 2023). Increasingly, illiberal regimes are using these autocratic IOs to contest long-established international liberal values (Cooley, 2016; Ginsburg, 2020*a*; Debre, 2025), reshape international legal standards (Ginsburg, 2020*b*), and justify their rule to international and domestic audiences (Bush, Cottiero, and Prather, 2024; Hafner-Burton, Pevehouse, and Schneider, 2025; Morrison et al., 2025).

However, the influence of these illiberal regimes is not limited to consolidated autocracies. Rather, as democratic backsliding has become increasingly common (Waldner and Lust, 2018; Ziblatt and Levitsky, 2018; Meyerrose, 2020; Haggard and Kaufman, 2021; Meyerrose, 2024, 2026), these backsliders—the regimes that emerge as a result of democratic weakening—are further unraveling long-established consensus within Western-backed IOs (Kelemen, 2020, 2024). New evidence suggests democratic backsliders are increasingly using their memberships in these organizations to insulate their domestic political power through both the rhetoric they adopt and their votes on contentious resolutions in multilateral fora (Baturo, 2023; Lipps and Jacob, 2025; Meyerrose and Nooruddin, 2025; Winzen, 2025). Nevertheless, compared to consolidated autocracies, we know relatively little about how democratic backsliders behave on the international stage, whether they represent a distinct bloc in today's geopolitics, and if their proliferation has altered alliance structures within the core institutions of the LIO. We turn to this question in the following section, outlining our theory of backsliders' preferences before explicating how these preferences should impact their international behavior and multilateral voting coalitions.

## 2. Democratic Backsliders on the International Stage

Democratic backsliding is an historically unprecedented phenomenon. Prior to the end of the Cold War, the primary threats democracies faced were more patently existential: democracies failed abruptly as the result of military or executive coups, and were replaced by fully autocratic regimes. However, since the end of the Cold War, total democratic collapse has become less common. Instead, attacks against democracies are more likely to come in the form of democratic backsliding (Bermeo, 2016; Geddes, Wright, and Frantz, 2018).

The starting point for any case of backsliding is a democracy. Democratic backsliding occurs when elected officials erode the liberal institutions that undergird democratic practice and executive constraints (Bartels, 2023). The specific institutions officials target vary from one state to the next, but typically include one or more of the following: the constitution, rule of law, civil and minority rights, judicial and media independence, and the separation of power within government, all of which facilitate executive aggrandizement at the cost of meaningful checks and balances (Waldner and Lust, 2018; Haggard and Kaufman, 2021; Meyerrose, 2026). Yet backsliding states, though increasingly illiberal in practice and norms, heed international and domestic incentives to appear democratic by continuing to hold regular and as least nominally free and fair elections. Therefore, the outcome of backsliding is a diminished democracy or, in the worst rare instances, a semi-autocratic regime (Luhrmann and Lindberg, 2019).

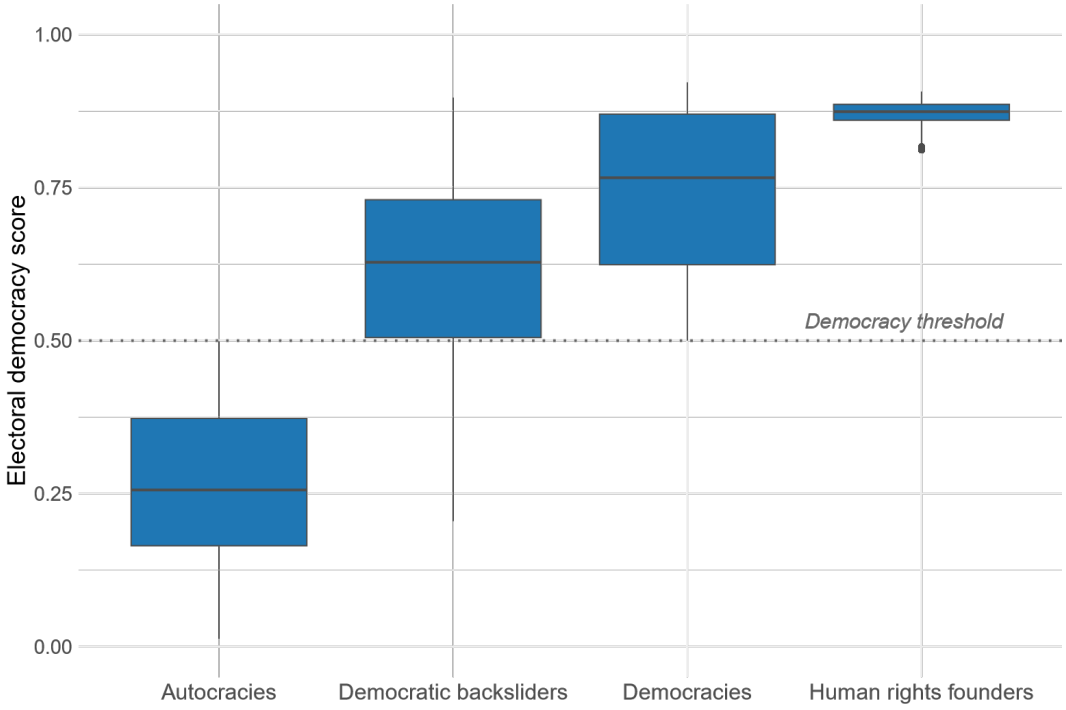
Similar to Levitsky and Way (2010)'s conceptualization of hybrid regimes, we argue democratic backsliders should be viewed as a distinct group of countries, rather than simply a transitory phase on the way to autocracy. Indeed, in terms of their overall regime status, backsliders are on average much closer to non-backsliding democracies than they are to consolidated autocracies, and the bulk of the observed variation in their electoral democracy scores lies above the commonly used 0.5 threshold for classifying democracies, *even while they are in a backsliding phase* (See Figure 1) (Coppedge et al., 2024).[5]

On the one hand, backsliders are an incredibly diverse set of states. Backsliding can begin in weak and consolidated, new or advanced, liberal or minimalist democracies. The political ideology of the elected officials that initiate backsliding can also vary; while attacks against democratic institutions have more commonly emanated from the right in Europe and North America, many cases of democratic erosion in Latin America have been initiated by the left.

[5] See Section 4 for a detailed discussion of how we define and identify these four groups of states.

Furthermore, the types of institutions that are eroded can vary from one case to the next; while some executives may focus primarily on undermining rule of law and judicial independence, others will instead prioritize attacks against the constitution and intragovernmental checks and balances, while yet others emphasize silencing civil society and the press.

**Figure 1.** While democratic backsliders have regressed in their levels of liberal democracy, they continue to hold free and fair elections in which all citizens are allowed to participate. Indeed, with few exceptions, they are notably more democratic than consolidated autocracies and thus should not be viewed simply as a subset of autocratic regimes.




*Data Source: Coppedge et al. (2024).*

Nevertheless, backsliders also share several core characteristics that both define them as a group while also distinguishing them from consolidated autocracies and from other democracies. These features inform their preferences and behavior at the international level. First, these countries have incentives to continue broadly to signal support for democratic norms. This allows them to maintain their international standing, shielding them from potential sanction for noncompliance with international values (Kelley, 2009). It also serves an important function domestically. Backsliders continue to hold largely free and fair elections and defend themselves as "true" democracies by returning power directly to "the people" rather than to unelected elites (Mudde and Rovira Kaltwasser, 2017).

Second, with several exceptions, most backsliders are developing economies and smaller states. Prior to their backsliding phase, as emerging democracies, they were largely aligned with and patronized by the West and continue to benefit from and depend on many of the Western-backed and financed international economic, political, and security structures.

Tempting as it is for opponents to label backsliders as autocrats, the truth is that such rhetoric is inaccurate. Backsliders adhere to the standards of elections as a means of determining popular will, but they diminish true democracy by insisting on crude majoritarianism at the expense of liberal protections of minority and dissenting voices. The populist insistence that they are the only legitimate representatives of "the people" and their attacks on perceived elite institutions that allegedly undermine the national interest is consistent with their rejection of any external criticism of how they rule domestically. Thus, even as democratic backsliders benefit from and endorse many aspects of the LIO, they reject its enforcement of political liberalism in the name of national sovereignty. Of course, their opponents at home are keen to glean support from international actors whose criticisms they can amplify domestically to bolster their case against the ruling regime. Consolidated autocrats—who make no pretension of being democratic—face no such balancing act. Their rejection of the LIO can be wholesale and unapologetic. As such, while democracies remain largely supportive of the LIO and its founding members (the HRF states), and autocracies oppose it in its entirety, we expect backsliders to be more selective in their contestation. While they will remain aligned with the West on the less intrusive international policy areas from which they benefit, they respond to incentives to oppose intrusive policies, particularly those promoting human rights and civil society participation, as such intrusions can highlight recent domestic regressions, fuel opposition criticism, and thereby serve as a threat to the regime.

Our analysis of the incentives facing backsliders has important implications for our understanding of how the growth of backsliders might affect multilateral institutions within the LIO. Our key insight is that the broad democratic coalition of nations that could be counted on to defend the LIO has splintered, especially in particular domains that represent the most intrusive and political ambitions of the LIO. Specifically, the human rights components of the LIO, which already pitted democracies against autocracies, rely today on a very fragile coalition of states since, as we show below, backsliding democracies increasingly deviate from their democratic counterparts in how they vote and speak on human rights topics. Unarrested, this has the potential to fundamentally alter the alliance structure within international institutions charged with upholding and protecting liberal norms and values. We apply this argument to the context of the UN human rights institutions, which epitomize the most ambitious and intrusive aspects of the LIO, and show that similar dynamics are also visible within the UN General Assembly on related issues.

## 3. The Consequences of Backsliding for International Human Rights Institutions

The contemporary UN human rights system consists of two distinct institutions: the UN Human Rights Council and the Universal Periodic Review. The UNHRC was founded in 2006 after its predecessor—the UN Commission on Human Rights—was dissolved (Hug and Lukács, 2014). The UNHRC has 47 member states that are elected by the UN General Assembly for three-year terms. In order for a state's human rights practices to be evaluated by the UNHRC, that state must first be identified as one of concern, and then one or more of the UNHRC member states must draft and put forth a resolution about that state. All members of the UNHRC then vote on that resolution. The UNHRC also votes on resolutions about more general human rights values such as support for women's rights, development, or freedom to protest; these resolutions are not targeted at any particular state. It is only for a certain subset of these resolutions that the UNHRC records how each individual state voted; others are adopted by voice votes.

The UNHRC and its predecessor represent the acme of postwar liberal ambition. Espousing a commitment to global enforcement of human rights standards championed—if not always practiced—by the West, it arguably is the most intrusive institution of the LIO. Granted, the UNHRC lacks real teeth, although its condemnations are not without material consequence (Lebovic and Voeten, 2006). But the real threat these institutions pose to human rights violators, especially states outside the orbit of protective Western patronage, is through naming and shaming (Finnemore and Sikkink, 1998; Hafner-Burton, 2005; Kelley and Simmons, 2015; Terman and Voeten, 2018). States targeted by the UNHRC might worry both about the material international consequences and also the possibility that having their human rights violations highlighted and condemned on the world stage would galvanize domestic opposition forces at home, leading to protests, rising instability, and even ouster.

The most cohesive and consistent bloc of supporters of the UNHRC's mandate has been the core founders of the post-World War II liberal international order (Burmester and Jankowski, 2014; Hug and Lukács, 2014; Burmester and Jankowski, 2018), the group of states we identify and refer to as the HRF states. As members of the UNHRC, these states consistently vote with each other more than 90 percent of the time (see Figure 2). More importantly for our investigation, these states have also used their economic, military, and ideological power to influence voting by other states in multilateral institutions, especially the UN; this influence was most pronounced on democratic states seeking alignment with the United States during the Cold War and even more so in recent years when the United States faced no real geopolitical rival until China emerged as a potential challenger (Dreher, Nunnenkamp, and Thiele, 2008; Vreeland and Dreher, 2014; Vreeland, 2019).

Against these HRF countries is arrayed a set of consolidated autocratic regimes, elected to the UNHRC from their world regions. Recent research shows that China engages in vote-buying efforts within the UN and other multilateral institutions (Brazys and Dukalskis, 2017; Kaya, Kilby, and Kay, 2021; Binder and Payton, 2022; Dreher et al., 2022; Lu, 2024; Steinert and Weyrauch, 2024), and both China and Russia increasingly rely on nascent blocs of likeminded states to challenge and constrain the UN human rights system (Inboden, 2021; Dukalskis, 2023). These consolidated autocracies predictably vote against the West on most resolutions, with the polarization between these two factions of the UNHRC being most pronounced when the target of the resolution is one of those consolidated autocratic states, a fact that we will exploit in our research design below.[6]

Between these two camps sit all other countries. These include stable democracies, ones with emerging-but-still-consolidating democratic institutions, and a newer group of democratic backsliders. Historically, the HRF's power to exert influence through the UNHRC rested on their internal cohesion and ability to attract support from these other states, and especially from those democracies more closely aligned with Western preferences (Voeten, 2000).

Over the last decade, this "liberal coalition" has grown more rickety. A primary cause, we argue, is widespread democratic backsliding. Democratic backsliding has two consequences for the UN human rights institutions: (a) global liberal values lose legitimacy and power in the resultant backslider states as populist leaders reject their universal pretensions by appealing to principles of state sovereignty and noninterference in domestic politics, and (b) these same leaders seek to protect themselves against intrusive international scrutiny that might embolden domestic opposition, increase their chances of international sanction, and thus threaten their grip on power.

Backsliders have incentives to behave strategically within international human rights institutions to protect themselves from intrusive policies and, relatedly, potential sanctions. Indeed, autocracies and states with poor human rights records (Hug and Lukács, 2014) or those dealing with domestic insurgents (Prasad and Nooruddin, 2024) have sought to defang the UNHRC before it turns the spotlight on them. Likewise, backsliders promote alternative conceptualization of human rights and target Western critics for their own alleged human rights deficiencies in an effort to deflect attention (Meyerrose and Nooruddin, 2025). We advance this prior scholarship by arguing that democratic backsliding internationally poses a grave—even existential—threat to the

[6] See Hug and Lukács (2014) for a related argument.

UN human rights institutions and, by extension, the ongoing promotion and protection of liberal norms and values by fundamentally undermining previous alliance structures within these institutions.

The liberal coalition of states faces growing contestation from a strong coalition of consolidated autocracies who use their membership in the UN human rights institutions to oppose LIO values and principles through the resolutions they sponsor, the votes they cast, and the rhetoric they use. What makes the current moment more threatening for the post-Cold War international system is the global phenomenon of democratic backsliding, which has fractured the pro-LIO coalition.[7] Like an ice floe detaching itself and floating away, democratic backsliders have moved away from the Western pole and perceive shared incentives to join the autocratic bloc to weaken the UNHRC and thereby to reduce the legitimacy of international criticisms their governments could face for growing repression of political and civil liberties at home. This framework yields a primary testable hypothesis:

> **Hypothesis 1:** *Democratic backsliders will be more likely to vote with the autocratic bloc on UNHRC resolutions.*

Yet, as noted above, we expect backsliders to be more selective than autocracies in their contestation of the LIO. It is only the policies that are the most intrusive, and that highlight directly the ways in which their regimes are out of sync with international democratic standards, that should be of concern to them. This differentiation exists in three ways in the UNHRC.

First, many of the resolutions voted on in the UNHRC are relatively nonintrusive and require states to voice their support or opposition for general human rights practices. For example, resolution A/HRC/RES/51/19 (2022) calls on states "to take measures to ensure the progressive realization of the human rights to safe drinking water and sanitation," while Resolution A/HRC/RES/39/4 (2018) advocates the "promotion of a democratic and equitable international order." As actors that benefit from the LIO, and seek to maintain their reputations as committed democrats, backsliders have few incentives to defect from the HRF and align with consolidated autocracies on these types of nonintrusive resolutions. Therefore, we further predict:

---

[7] Of course, U.S. actions on the international stage—and indeed even within the UN specifically—in the early months of the second Trump administration signal that the United States is no longer a reliable proponent and protector of the LIO; indeed, it may begin to behave in ways similar to the other backsliders that we focus on in this paper. However, since our analyses to follow end in 2023, and because for much of the post-World War II period the United States was undeniably one of the primary architects and defenders of the LIO, in what follows we treat the United States as such. We return to a discussion of the implications of ongoing changes within the United States for the LIO in more detail in the conclusion.

**Hypothesis 2:** *Democratic backsliders will be no more likely to vote with the autocratic bloc on general, nonintrusive UNHRC resolutions.*

Where we should see backsliders opposing the HRF, and aligning with autocracies, is on intrusive resolutions that validate the UNHRC's right to criticize domestic policies of sovereign actors. Such practice makes it more likely the spotlight will be placed on backsliders themselves in the future. In the context of the UNHRC, these more intrusive resolutions are the ones that, rather than simply stating general support for broad-based human rights norms, instead directly target (i.e., name and shame) specific states for their human rights violations. Therefore, we further predict:

**Hypothesis 3:** *Democratic backsliders will be more likely to vote with the autocratic bloc on UNHRC resolutions that target specific states.*

Even among these targeted resolutions, there exists important variation. In some cases, resolutions in the UNHRC are passed by broad consensus. Uncovering voting blocs, and determining if backsliders are indeed aligning with autocracies on issues that they view as threatening and intrusive, requires identifying more controversial votes that necessitate the explicit choosing of sides. We focus on resolutions on which a majority of consolidated autocratic states voted against a supermajority (75 percent) of HRF states. These resolutions have the virtue of unambiguously pitting the two geopolitical camps against each other, thereby forcing other UNHRC members to clarify their positions.[8] Further, by limiting our attention to those resolutions backed by a supermajority of HRF states, we can focus on the issues of greatest importance to the leaders of the LIO that are opposed by a majority of autocratic members of the UNHRC. If backsliders are moving away from the West, and toward the autocratic camp, as we have argued, these are the resolutions on which that movement should be most evident and most troubling for the LIO's future viability.

**Hypothesis 4:** *Democratic backsliders will vote with the autocratic bloc on contentious targeted UNHRC resolutions that pit consolidated autocracies against HRF states.*

The era of backsliding is widely dated as reaching its acme by 2013 (Nord et al., 2024). We argue the dynamics of the shifting allegiances of democratic backsliders should be particularly visible from 2013 onward, a period in which backsliding became largely complete and consolidated in many erstwhile stable democracies; and, when Western champions of democratic values were distracted by their own political crises at home. While the earlier years of the UNHRC (and its predecessor) might have been broadly

[8] This is the intuition behind the use of so-called "party unity" votes analyzed by scholars of the U.S. Congress (Desposato, 2005).

characterized as experiencing competition between democratic and autocratic blocs with all other states lying between these two poles, more recent experience suggests we need to consider backsliders as a potentially distinct "camp" unto themselves, separate from the non-HRF democracies of which they were once a part. We therefore expect that the realignment of democratic backsliders in the UNHRC to be more pronounced after 2013.

> **Hypothesis 5:** *Backslider states' shift away from the HRF states and toward autocracies in the UNHRC will be particularly pronounced after 2013, which marks the height of the global democratic recession.*

These shifts in backsliders' preferences should be evident not only in their voting behavior, but also in the rhetoric they use. The UPR process involves pairs of countries serving as reviewers for a given state. If backsliders are indeed moving away from the West on these core and highly sensitive issues, and increasingly toward autocracies, we would expect reports written by backsliders to be more similar both to those written by fellow backsliders, as well as to those written by autocracies; at the same time, they should diverge from reports composed by HRF states. Such rhetorical patterns would be evidence that backsliders as a group have developed a distinct set of preferences regarding international human rights norms, and how they are monitored. Therefore, we assess the degree of similarity in the language used by reviewers as a function of their regime type. Similar to the UNHRC voting, we expect these dynamics in UPR language to be particularly salient after 2013.

> **Hypothesis 6:** *The language democratic backsliders use in their UPR reports will be more similar to the language used by autocracies, and less similar to the language used by HRF states.*

> **Hypothesis 7:** *The shifts in language similarity for democratic backsliders in the UPR should be particularly prevalent after 2013.*

# 4. Identifying Voting Blocs in the UNHRC

To test Hypotheses 1 through 5, we explore the dynamics of voting in the UNHRC with data on recorded votes we collect directly from the UNHRC online library for all resolutions from 2006–23.[9] The resulting dataset contains information on how each member of the UNHRC voted on any given resolution for which votes were recorded, the state that is the subject of the resolution, if any, as well as details about which state(s) sponsored the resolution. Our dataset contains observations at the resolution-UNHRC member state level and consists of 22,911 individual state votes on 488 unique resolutions. These include both general and targeted resolutions. Among the latter, 22 states were targeted at least once between 2006–23.[10]

Membership in the UNHRC is based on an equitable geographic distribution of seats. When electing states to serve on the UNHRC, the UN General Assembly is asked to consider each candidate's commitment to human rights. Nevertheless, autocracies and other states with questionable human rights practices are regularly represented in the Council. In other words, both democracies and autocracies can be UNHRC members. Furthermore, as democratic backsliding has progressed and spread to a growing list of countries, the number of backsliders sitting on the Council has increased.

To identify the universe of democratic backsliders, we focus exclusively on states that began the period as democracies.[11] Next, since backsliding often involves attacks against liberal democratic institutions, we measure changes with respect to liberal, rather than electoral, democracy. Backsliding begins when a state experiences a negative 0.01 (or greater) annual change in its liberal democracy score, along with an overall decline of 0.1 or more. The episode continues if there is: i) an annual (negative) change in one of every five consecutive years; ii) no reverse (positive) annual change of 0.03 or more; and iii) no cumulative (positive) reversal of 0.1 over a five-year period.[12] Based on these criteria, we identify all episodes of democratic backsliding that occur within the time period covered by our data (2006–23); the full list of episodes is available in Appendix 1.

---

9 https://searchlibrary.ohchr.org/search?ln=en&cc=Voting. Last accessed July 30, 2024.

10 The states targeted in these resolutions, and the number of times they were targeted, are: Afghanistan (1), Belarus (17), Burundi (9), Colombia (1), Eritrea (5), Ethiopia (2), Georgia (6), Iran (12), Israel (78), Nicaragua (5), North Korea (5), Myanmar (6), Philippines (1), Republic of the Congo (2), Russia (4), South Sudan (2), Sri Lanka (6), Sudan (4), Syria (47), Ukraine (7), Venezuela (6), and Yemen (3).

11 That is, ones with a pre-backsliding electoral democracy score of 0.5 or higher.

12 We adopt these coding rules from the Varieties of Democracy's (V-Dem) Episodes of Regime Transformation (ERT) dataset (Maerz et al., 2021). While widely used to identify episodes of autocratization, the ERT dataset has nevertheless been criticized for its use of crisp cutoffs along continuous indices to distinguish one regime type from another, as well as for its use of arbitrary thresholds to determine whether or not a state is backsliding (see https://www.v-dem.net/documents/43/v-dem_dr2024_lowres.pdf, pp. 46-47). We take these critiques seriously and below show that our results are robust to a more stringent criteria for inclusion as a backsliding state: an overall reversal of 0.2, rather than 0.1, over a five-year period.

Theoretically, we are interested in whether backsliders are shifting their allegiances toward autocracies and away from HRF states. We begin by descriptively examining the levels of voting cohesion among four distinct and mutually exclusive groups of states within the UNHRC: consolidated autocracies, backslider states, HRF countries,[13] and all other democracies.[14] To measure voting cohesion, we follow Burmester and Jankowski (2018). For any given resolution for which individual votes were recorded, we first identify whether the majority vote for any given group was "yes" or not.[15] Next, for each individual state in that group, we code whether they voted with or against the average decision of their regime-type group. We then calculate the mean number of times that all members of a group voted with their group. These annual averages by group, as well as the number of members of each group represented in the UNHRC per year, are plotted in Figure 2.

As Figure 2 shows, HRF countries consistently vote together at high rates. But, to only a slightly lesser degree, so do consolidated autocracies, which is even more remarkable since the set of autocratic states is much larger numerically and much more diverse in virtually every other dimension than the four states in our HRF group. These mirrored levels of voting cohesion support our contention that the HRF states and consolidated autocracies represent two poles within the UNHRC. Backsliders and other democracies, by contrast, show less cohesion in their voting decisions, though, on average, the level of cohesion for democratic backsliders is similar to that of non-HRF stable democracies. This makes sense given that these backsliders were only recently part of that broader community of democracies.

Are these backsliders more likely to vote with the HRF states, or are they instead moving closer to the bloc of consolidated autocracies? On the one hand, the data suggest these backsliders are indeed increasingly breaking away from the HRF states. Figure 3 traces the annual average vote cohesion among all HRF countries and backsliders within the UNHRC. The overall trend has been for these two groups of states to become less cohesive over time; this shift becomes particularly evident after 2013, which, as we noted, marks the peak of the global democratic recession.

---

[13] For reasons discussed at the outset, we code Australia, France, the UK, and the United States as HRF founders, but our results are robust to expanding the set of states included in this category.

[14] This group includes all new and advanced democracies that are neither backsliders nor HRF states.

[15] The other possible votes in our dataset are "no," "abstain," or, in some cases, failure to participate in the vote. We concur with Morse and Coggins (2024) that abstentions signal a lack of support for a resolution, and therefore group them with "no" votes both here and in our analyses below. Morse and Coggins (2024) also argue that absences from votes in the UNGA can either signal a lack of capacity and resources, or can be used more strategically by weak states to counteract geopolitical pressures. Given this ambiguity, we exclude observations where a UNHRC member did not participate in a vote; these absences constitute only 0.006 percent of the observations in our dataset, and the findings remain the same if we include these observations in the "opposition" group of votes.

**Figure 2.** Voting Cohesion in the UNHRC

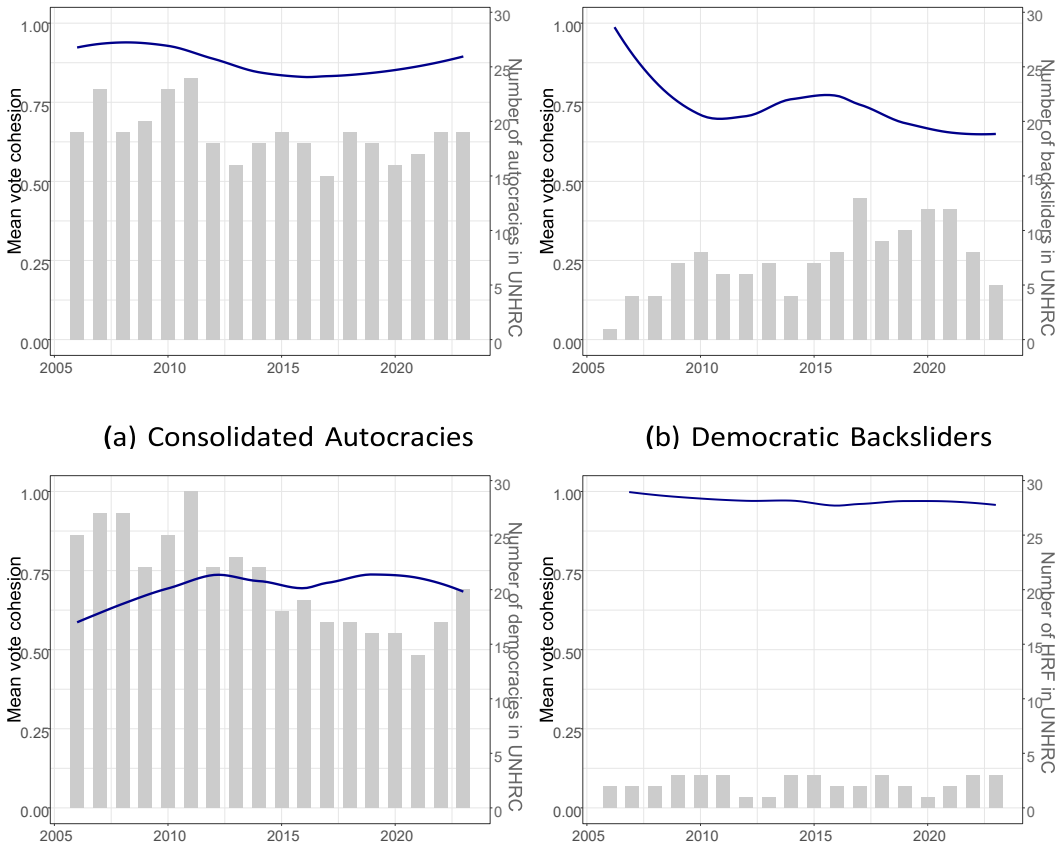
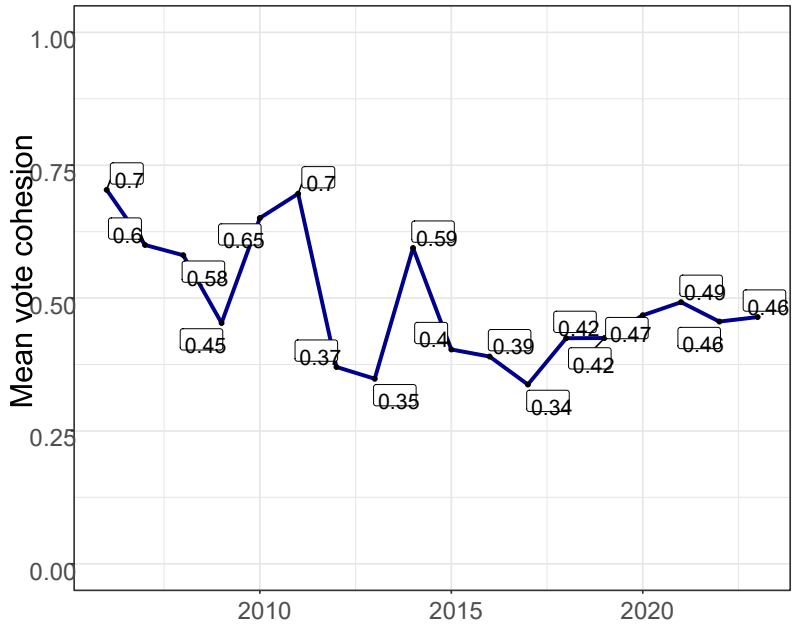



**Figure 3.** UNHRC Voting Cohesion: HRF Countries and Backslider States



We test these dynamics more systematically using data on all UNHRC resolutions for which votes were recorded from 2006 (the year the UNHRC in its current form was founded) through 2023. We calculate for each resolution the majority vote among all consolidated autocracies: either "yes," or against, which groups together "no" and "abstain" votes into a single category. Then, for each member of the UNHRC, we code whether they voted with or against the majority position of the autocratic bloc; this binary variable is our dependent variable and our unit of observation is at the UNHRC member-vote-resolution level of analysis. The main independent variables in these models are binary indicators for whether a state is a backslider state or a consolidated autocracy, with HRF states and democracies—i.e., all non-backsliding democracies—as the omitted reference category.[16]

We control for a range of factors that should impact voting decisions. First, we control for the voting state's liberal democracy score (Coppedge et al., 2024) with the expectation that higher levels of democracy will make states less likely to vote with the autocratic bloc. This also accounts for heterogeneity within the backslider group, as not all of these states start or end their backsliding episode at the same point on the liberal democracy index. Similarly, we control for the voting state's gross domestic product (GDP) per capita (World Bank, 2024), with the expectation that wealthier states will be less likely to vote with autocracies.[17] Finally, we also capture the voting state's voting affinity with the United States in the UNGA, since states that are more closely aligned with the United States should be less likely to vote with autocrats (Voeten, Strezhnev, and Bailey, 2009). We similarly control for target state characteristics, including liberal democracy score, GDP per capita, and agreement with the United States in the UNGA, as characteristics of the target state itself, rather than the nature of the human rights violations, may impact voting decisions.[18] We also control for whether or not the voting and target state are from the same geographic region (Coppedge et al., 2024).

[16] The coding rules for each category are as follows: (1) consolidated autocracies are states below the 0.5 threshold of V-Dem's electoral democracy score; (2) backsliders had a score above 0.5 on the electoral democracy index but then experience a cumulative decline on the liberal democracy index of at least 0.1 over a given period and meet the other criteria spelled out in Section 4 above; and (3) all other democracies include countries that score above 0.5 on the electoral democracy, including the four HRF countries we identify above. These groups of states are mutually exclusive in any given year. As we discuss below, our results are robust to alternative cut-offs for distinguishing democracies from autocracies, and for identifying backsliders.

[17] We log this variable to account for diminishing returns.

[18] We do not control for the voting or target state's human rights score (Cingranelli, Richards and Clay, 2014) because these are highly correlated with liberal democracy scores (0.87 and 0.77, respectively). Given the choice, we choose to include the liberal democracy scores instead to test directly whether being a backslider has any effect on voting patterns above and beyond their overall level of democracy.

Israel is the most frequent target of UNHRC resolutions in our data (78 resolutions). Research shows voting patterns against Israel in the UNHRC are distinct (Seligman, 2011). Therefore, we control for whether the resolution targets Israel. We do the same for Syria, since Syria is the second-most targeted state (47 resolutions) in this period. Finally, we calculate the percent of democratic sponsors of each resolution to capture the logic that resolutions sponsored mainly by democracies generate greater opposition from autocracies and backsliders.

We estimate four models. First, to test whether backslider states are more likely to align with autocracies in the UNHRC (Hypothesis 1), we estimate a model that includes observations for all resolutions—both general and targeted—for which votes were recorded. Next, to test Hypothesis 2, which predicts that backsliders should be no more likely to vote with the autocratic bloc on general, nonintrusive resolutions, our second model includes only observations for general resolutions that do *not* target a specific state. These are those resolutions that focus on general human rights such as access to clean drinking water.[19] Backsliders should not view these general resolutions as threatening, and in fact may use them as an opportunity to signal their commitment to democratic values. Third, Hypothesis 3 expects backsliders to be more likely to side with the autocratic bloc on resolutions that target a specific state, as these resolutions are significantly more intrusive. Therefore, model 3 includes only observations for votes on targeted resolutions. Finally, if backsliders are truly pushing back against the West on these domestically sensitive and intrusive policies, we should see this behavior especially when votes are contentious (Hypothesis 4). Therefore, the fourth model focuses exclusive on contentious targeted resolutions, which are those in which a supermajority (75 percent or greater) of HRF states voted one way, while the autocratic bloc voted in opposition.

Since our dependent variable is a binary variable, we estimate logistic regressions with year fixed effects.[20] In addition, because states vote on multiple resolutions, the observations in our data are not independent. Therefore, we also cluster standard errors by voting state. The results of these models are reported in Table 1.

[19] Since these resolutions do not target a specific state, this model and the first, which includes both targeted and general resolutions, necessarily do not control for target state characteristics.

[20] While some countries, such as the UK, Germany, and Cuba are represented across 15 years or more in our dataset, others, such as Iceland, Luxembourg, and Uzbekistan, were present for three years or fewer. Indeed, of the 124 countries that have served as members of the UNHRC since its inception in 2006, nearly half (60) are represented in six years or less. Because the number of years that a state serves in the UNHRC can vary significantly from one country to the next, and since small and unbalanced sample sizes can exacerbate bias in fixed effects models (Bell and Jones, 2015; Clark and Linzer, 2015), we do not include country fixed effects. Doing so would effectively remove those states that infrequently serve on the UNHRC from our analysis.

**Table 1.** Voting with the Autocratic Bloc in the UNHRC, 2006–2023

| | *Dependent Variable:* | | | |
|---|---|---|---|---|
| | Vote with autocratic bloc (1 = yes, 0 = no) | | | |
| | All | General | Targeted | Contentious targeted |
| | (1) | (2) | (3) | (4) |
| Backslider state | 0.12* | -0.06 | 0.29*** | 0.48*** |
| | (0.05) | (0.08) | (0.08) | (0.10) |
| Consolidated autocracy | 0.43*** | 0.33** | 0.58*** | 0.51*** |
| | (0.08) | (0.12) | (0.11) | (0.14) |
| Voting lib dem score | -0.63*** | -0.89*** | -0.29 | -0.93** |
| | (0.15) | (0.23) | (0.23) | (0.28) |
| Voting GDP pc (log) | -0.07*** | -0.13*** | -0.01 | -0.14*** |
| | (0.02) | (0.03) | (0.03) | (0.03) |
| Voting UNGA US agree | -9.40*** | -11.31*** | -8.46*** | -13.91*** |
| | (0.19) | (0.28) | (0.28) | (0.47) |
| Perc. demo sponsors | -0.01*** | -0.003*** | -0.002* | -0.002 |
| | (0.0005) | (0.001) | (0.001) | (0.001) |
| Target = Israel | | | 1.24*** | 1.90*** |
| | | | (0.23) | (0.28) |
| Target = Syria | | | -0.37*** | -0.52*** |
| | | | (0.09) | (0.12) |
| Target lib dem score | | | 0.82* | -0.24 |
| | | | (0.37) | (0.47) |
| Target GDP pc (log) | | | -0.07 | 0.08 |
| | | | (0.04) | (0.05) |
| Target UNGA US agree | | | -0.18 | 1.34** |
| | | | (0.37) | (0.47) |
| Voting-target UNGA affinity | | | -0.23 | 2.49*** |
| | | | (0.22) | (0.40) |
| Voting-target same region | | | -0.21 | -0.14 |
| | | | (0.12) | (0.14) |
| Year Fixed Effect | Yes | Yes | Yes | Yes |
| Clustered standard errors | Yes | Yes | Yes | Yes |
| Model type | Logit | Logit | Logit | Logit |
| Observations | 21,709 | 11,487 | 9,423 | 7,439 |
| Log Likelihood | -9,077.22 | -4,076.89 | -4,215.38 | -2,668.05 |
| Akaike Inf. Crit. | 18,202.44 | 8,201.79 | 8,492.76 | 5,398.10 |
| *Note:* | | | | *p<0.05; **p<0.01; ***p<0.001 |

We find compelling evidence in support of our theoretical framework. First, we find in Model 1 in Table 1 that, compared to other democracies and across all resolutions, backsliders are significantly more likely to vote with the autocratic bloc. However, we also find evidence of important variation across resolutions. When we focus exclusively on general, nontargeted resolutions in Model 2, the coefficient for backslider states is no longer statistically significant. However, once we consider exclusively targeted (Model 3) and contentious targeted (Model 4) resolutions, we find that backsliders are significantly more likely to vote with the autocratic bloc. These results suggest that backslider states align with autocratic states on resolutions that they view as potentially threatening to their regime persistence.

Is it just that backsliding states are becoming more autocratic and therefore beginning to align more closely with autocracies? Our answer is no, and we argue the results from Model 2 in particular provide evidence of this differentiation. Nevertheless, to further test if backsliders are in fact distinct from autocracies, or if instead they are simply the newest members of the autocratic bloc, in Table 2 we show that in three of the four models, backslider states are statistically distinct from autocracies. It is only in the fourth model—the one focused on contentious targeted votes that pit autocracies against the HRF group—that backsliders are indistinguishable from autocrats. This suggests backsliders are indeed a distinct group with their own emerging sets of interests, but behave like autocracies on the most invasive and threatening resolutions.

Backsliding can both begin and end at different levels of liberal democracy. As such, one concern may be that the effects of backsliding on voting decisions are conditional on a state's level of liberal democracy at the time of the vote. Our modeling approach accounts for this interactive effect. Logistic regressions such as those reported in Table 1 are fully interactive: even when no interaction terms are included in the model, the effect of one variable on the outcome of interest is conditional on the value of other covariates (Berry, DeMeritt, and Esarey, 2010). To further ensure the effects of backsliding are not limited to states at certain levels of liberal democracy, we also estimate the marginal effect of backsliding conditional on levels of liberal democracy, and find that both non-backsliding and backsliding democracies are less likely to vote with the autocratic bloc as their liberal democracy score increases. However, at all levels of liberal democracy, backsliding states are more likely to vote with autocrats than non-backsliding democracies.[21]

[21] See Appendix Figure A1.

**Table 2.** Pairwise differences between coefficient estimates for all UNHRC models, 2006–2023. "✓" indicates pairs of coefficients that are statistically different from one another ($p < 0.05$), while "✗" indicates no statistically significant difference.

| | General | All | Targeted | Contentious targeted |
|---|---|---|---|---|
| Backslider–autocracy | ✓ | ✓ | ✓ | ✗ |

These findings hold across alternative specifications. First, as noted above, V-Dem's ERT method for identifying backsliders has been critiqued for arbitrary thresholds, and coder-based measures more broadly may paint an overly grim picture of the state of global democracy (Little and Meng, 2024). Therefore, it is possible we have captured only minor changes in levels of democracy, rather than actual instances of democratic backsliding. To test this, we apply a stricter criterion: states must show a 0.2 (not 0.1) decline in liberal democracy over the episode to be coded as backsliders. Results remain substantively unchanged.[24] Similarly, some argue V-Dem's 0.5 cutoff for distinguishing democracies from autocracies may be too high, thereby misclassifying democracies and autocracies; Kasuya and Mori (2019) suggest 0.42 is more accurate. Therefore, we reestimate our models using this alternative threshold for all relevant variables, and again find substantively similar results.[25]

## 4.1 Shifts Over Time

Hypothesis 5 predicts that voting dynamics among backsliders in the UNHRC would be more pronounced after 2013, which marks an (admittedly arbitrary) inflection point when backsliding became particularly widespread and extensive (Nord et al., 2024). To test this prediction, we reestimate the same models from Table 1, disaggregating the observations into pre- and post-2013 samples.[22] The results for the main independent variables are reported in Figure 4.[23] We find evidence that backsliders have moved closer to autocracies on a wider range of resolutions over time. Prior to 2013, backslider states were more likely to vote with autocrats on only one type of resolution: contentious targeted ones. However, following 2013, they were more likely to vote with the autocratic bloc on all targeted, along with contentious targeted, resolutions. We also note the coefficient for backsliding states in the "All" model is significant at the 90 percent level in the post-2013 sample.

[24] See Appendix Table A1.

[25] See Appendix Table A2.

[22] We note that using alternative years as thresholds returns substantively similar results.

[23] These models include all control variables reported in Table 1, where relevant; full results are reported in Appendix Tables A3 and A4.

This suggests concerns about growing autocratic influence in IOs over the past few years is justified. But it also provides new evidence that the growing ability of autocrats to challenge the LIO is due in no small part to the global phenomenon of democratic backsliding that has considerably weakened the HRF's ability to mobilize a reliable coalition in international fora like the UNHRC. If one desires a silver lining, it is that this trend is relatively recent and so perhaps there remains time to reverse it through concerted action by the LIO's most ardent advocates.

### 4.2 Beyond Human Rights: Backslider States in the UNGA

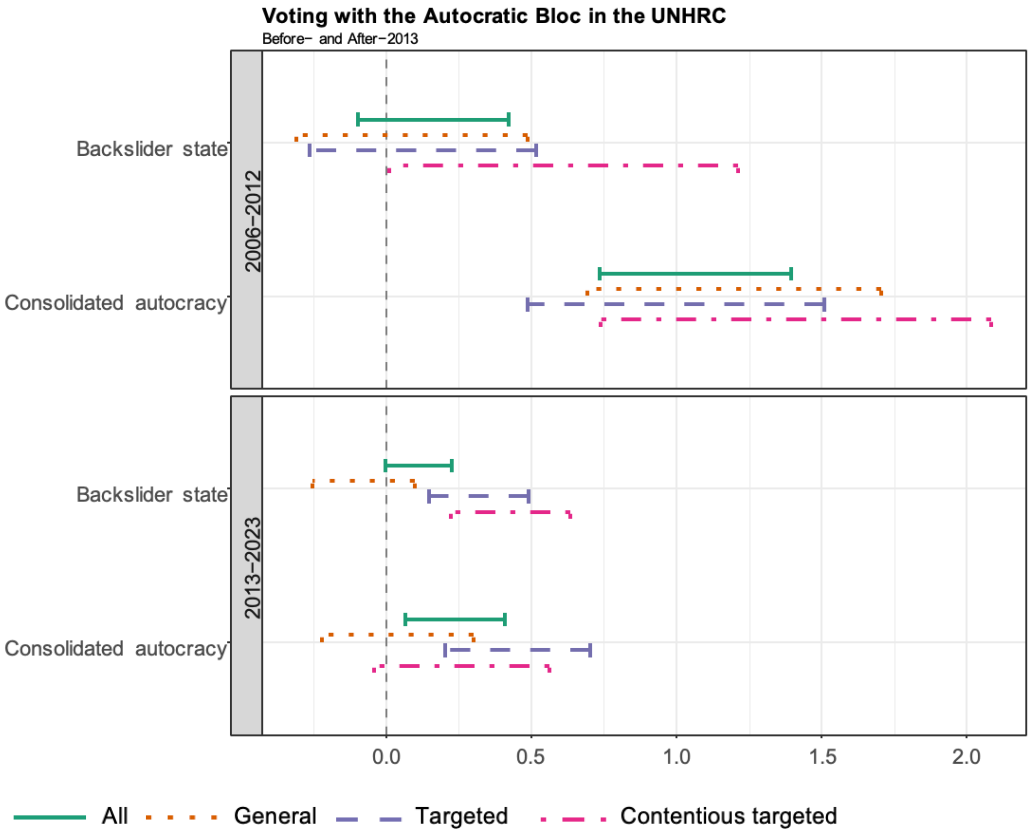
We find evidence that backslider states are more likely to vote with autocracies on targeted, contentious votes in the UNHRC, especially after 2013. However, there are several limitations to analyzing backsliders' behavior exclusively within the UNHRC context. First, the UNHRC was only founded in 2006, and thus limits our ability to observe historical voting trends. Second, not all states are represented in the UNHRC at all or at the same time, and the subset of states whose human rights practices are evaluated by the Council is limited; this may induce a selection effect in that we only observe the behavior of certain subsets of democracies, autocracies, and backslider states both in isolation and in reference to one another. Finally, the UNHRC is, by definition, focused on a single issue area—human rights—that, while the most intrusive and, therefore, we argue particularly relevant to backslider states' interests, is also especially likely to foster opposition from illiberal states, whether they are backsliders or consolidated autocracies. In other words, it is arguably a most likely case for us to see backslider states drifting toward autocracies. If our theory about the types of international policies that backsliding states contest is indeed correct, we would expect to see these states pushing back against issues of greatest importance to the West in contexts outside the UNHRC. These dynamics will be largely a post-Cold War phenomenon, and should be especially salient beginning in 2013.

Therefore, we analyze voting dynamics in the UNGA, an institution in which all UN members are represented and a wider range of issues are decided, as an additional test of our argument. Using UNGA voting data spanning 1946–2023 (Fjelstul, Hug, and Kilby, 2025), we estimate a series of models to approximate those reported in Table 1.[24] In the first model, we analyze votes for all resolutions (1946–2023), and in the second we focus exclusively on contentious resolutions (1946–2023).[25]

[24] These data include information for votes on both amendments and resolutions. Mirroring our approach in the UNHRC, we focus exclusively on resolutions for which roll call votes were taken.

[25] As before, contentious votes are ones where 75 percent or more of the HRF countries voted in one direction, while the majority of autocracies voted in opposition.

**Figure 4.** Prior to 2013, backslider states were more likely to vote with autocrats on only one type of resolution: contentious targeted ones. However, following 2013, backsliders were significantly more likely to vote with the autocratic bloc on both targeted and contentious targeted resolutions. The plot depicts 95 percent confidence intervals.




Above, we show backsliders vote on targeted UNHRC resolutions differently than they do general resolutions. Since UNGA resolutions rarely target specific states in the way that UNHRC resolutions often do, we use a proxy to capture "high stakes" resolutions. The third model includes observations for "important" votes. These are "votes on issues which directly affected United States interests and on which the United States lobbied extensively."[26] These important votes are identified by the U.S. State Department as being especially closely tied to its strategic goals. The State Department records how each state votes on these important resolutions, and uses these votes as a criterion for disbursing foreign aid (Carter and Stone, 2015). In other words, these are the issues of greatest importance to the United States. Given the resources the United States devotes to monitoring and lobbying these votes, their outcomes are arguably more consequential and intrusive for any specific countries that are targeted by these votes.

[26] https://2009-2017.state.gov/documents/organization/162417.pdf.

Therefore, backslider states should view these UNGA resolutions as particularly threatening.[27] Examples of topics identified as important have included resolutions related to Palestine, the U.S. embargo of Cuba, and human rights issues both broadly and in specific countries.

We also combine several of these resolution-level characteristics. In the fourth model, we analyze all important and contentious votes (1983–2022). Finally, in the fifth and sixth models, mirroring the results reported in Figure 4, we analyze temporal dynamics and divide the important, contentious votes into post-Cold War samples prior to (1990–2012) and during (2013–2022) the peak backsliding wave.

The dependent variable for the UNGA models captures whether a state voted with the autocratic bloc on a resolution, and our main independent variables identify backslider states and consolidated autocracies. We control for the voting state's liberal democracy score, GDP per capita, and agreement with the United States in the UNGA.[28] As before, each model includes year fixed effects, and we cluster standard errors by voting state. Figure 5 reports the coefficient for backslider states from all six models.[29]

The results confirm that backsliders are strategic in their contestation of international norms and policies. In models that include "all" and "contentious" resolutions—those that span from 1946–2023 and cover a range of issue areas—backsliders were significantly less likely to vote with the autocratic bloc. For "important" votes, we see a shift: for these resolutions, the coefficient for backslider states is no longer significant. However, once we identify important votes that were contentious—in other words, those that were not subject to broad-based international consensus—we find backsliders are significantly *more* likely to vote with autocracies in the post-Cold War era, and particularly beginning in 2013.

To summarize our analyses of voting data from the UNHRC and the UNGA, we have shown evidence that backslider states—whose governments are becoming increasingly illiberal yet still adhere to minimally democratic practices—are more likely to support the positions taken by consolidated autocracies on issues they view as particularly threatening and intrusive, even when this brings them in direct conflict with the preferences of the core founders of the post-Cold War LIO.

---

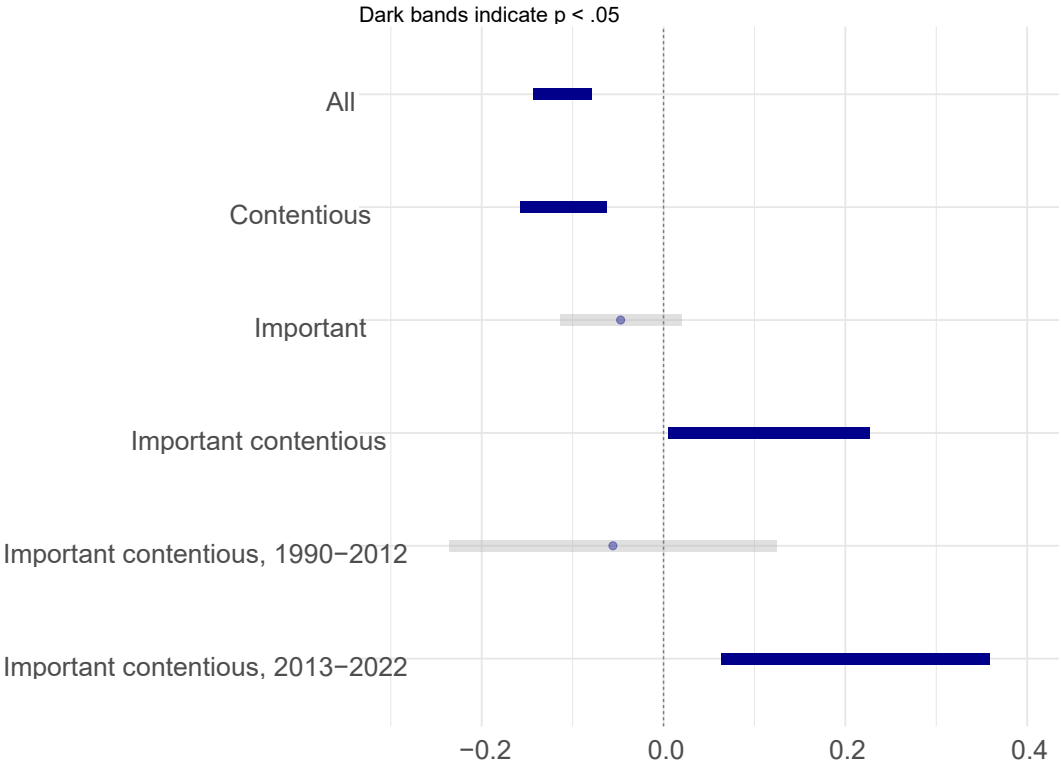
27 Coding of "important" votes by the U.S. State Department only began in 1983, and is available through 2022, so this model is restricted to observations during that time period.

28 Unlike in the UNHRC, we do not have data on resolution sponsorship, so cannot control for the percent of democratic sponsors. Furthermore, since UNGA resolutions rarely target specific states, there are no target-state level controls.

29 See Appendix Table A5 for full results.

These dynamics have been particularly pronounced since 2013, after which backsliding became a pervasive global phenomenon. However, voting behavior alone—which can be subject to vote buying or bloc dynamics—may not fully capture the dynamics of backslider states' preferences or methods of contestation. Furthermore, the types of resolutions for which votes are recorded constitute only a subset of all decisions made within these UN institutions (Hug, 2012). Therefore, we turn next to data from the Universal Periodic Review to test Hypotheses 6 and 7.

**Figure 5.** While overall backsliders are less likely to vote with autocrats in the UNGA, they align themselves with the autocratic bloc on contentious issues of greatest importance to the United States. The is particularly the case following the peak of the backsliding wave in 2013. The plot depicts 95 percent confidence intervals.



## 5. Rhetorical Shifts in the UPR

In 2006, the UN introduced the Universal Periodic Review as an additional mechanism for monitoring the human rights practices of its member states. Beginning in 2008, under the UPR all UN member states were subject to reviews of their domestic human rights practices every four-and-a-half years. In these review sessions, the human rights record of the state under review (SuR) is "peer reviewed" through an interactive dialogue between the SuR and all other UN member states. These reviews are conducted by the UPR Working Group, which is comprised of the 47 sitting members of the UNHRC; the UPR Working Group is assisted by a group of three states ("troikas"), all of whom are selected to serve as rapporteurs in the review process through a drawing of lots. All other states also have the opportunity to respond and evaluate the SuR. Indeed, and importantly for our purposes, participation in the UPR reviews frequently extends beyond the randomly selected troika states: in our dataset the mean number of countries that provided reviews is 62.6. These reviews are recorded in an outcome report (Cox, 2010; Terman and Byun, 2022), which is publicly available online.[30] We use these compiled reports to create a dataset that consists of 99,166 individual recommendations made for all UN member states via the UPR mechanism from 2008 through January 2022.

Specifically, we use the text content of these UPR reports to measure the semantic similarity of any two given reports written in the same session about the same state. Recent work in political science has used text-as-data methods to measure state preferences in the content of UN speeches (Kentikelenis and Voeten, 2021; Watanabe and Zhou, 2022), debates in the UN General Assembly (Baturo, Dasandi, and Mikhaylov, 2017), and UPR reports (Dai and Lu, 2024; Kim, 2024; Lu, 2024; Meyerrose and Nooruddin, 2025). The recommendations that states make in their UPR reports differ in content and severity of the critique of the SuR (Kim, 2024), and therefore serve as an informative way to capture states' positions both individually and in relation to one another.

Our dataset includes all potential dyads of state reviews written about the same SuR during the same UPR session. To measure the similarity between any two relevant reports, we use a common measure of semantic similarity: cosine similarity.[31] The cosine method captures the relative similarity between two text documents by counting word frequencies, rather than threads of words, to conduct sentiment analysis (Manning, 2008; Spirling, 2011); in other words, it uses a "bag of words" approach.

---

[30] https://upr-info-database.uwazi.io/en/. Last accessed June 10, 2025.

[31] For other applications of cosine similarity in political science, see: Garrett and Jansa (2015); Diodati, Marino, and Carlotti (2018); Jansa, Hansen, and Gray (2019); Hager and Hilbig (2020); Linder et al. (2020); Kim (2024); Lu (2024).

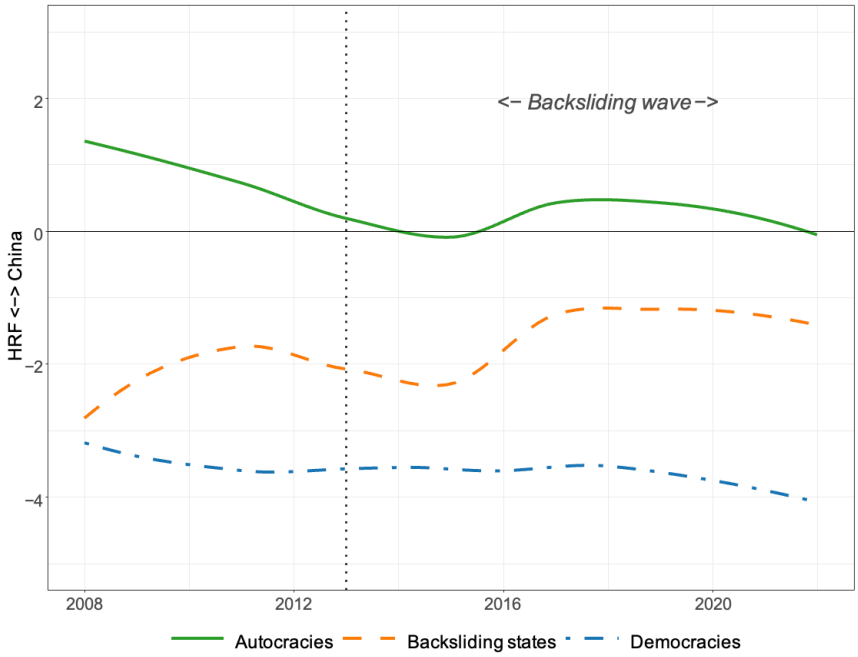
To estimate the cosine similarity for all relevant dyads in our data, we first pre-process the text of the UPR reports, removing numbers, punctuation, symbols, and stop words (e.g., "the" or "and"), and stemming and "lower casing" the text data. Next, we convert the text data to a matrix that records the number of times any given word appears in each report. We use this matrix to estimate the cosine similarity for all pairs of documents. Theoretically, cosine scores can range from 0, which indicates complete disagreement, to 1, which indicates complete agreement. In our dataset, the cosine scores have a mean value of 0.059 and a standard deviation of 0.069. We rescale these scores to take values from 0–100 for ease of interpretation. These 0–100 cosine similarity scores are the dependent variable in our analyses below.

We begin by exploring descriptive patterns in cosine similarity across regime type. We calculate autocracies', backslider states', and democracies' average annual cosine similarity scores first with the average for the HRF group, and then with China. Our expectation is that autocracies will be closer to China, democracies will be closer to the HRF, and backsliders will shift away from the HRF and toward autocracies over time. Figure 6 traces each group's average difference in cosine similarity with China and the HRF (average China similarity minus average HRF similarity). By this measure, positive values indicate a group is more similar on average to China, and less similar to the HRF countries. As Figure 6 shows, backsliders have on average begun to shift away from the HRF countries and toward China over time. However, these states nevertheless remain closer to the HRF on average than do consolidated autocracies, providing further evidence that backslider states' preferences are distinct from those of consolidated autocracies.

We also more formally analyze the cosine similarity of backsliders' UPR reports. To create our main independent variables, we again divide the reviewing states into four mutually exclusive groups: HRF countries, consolidated autocracies, backsliders, and democracies. Our main independent variables are each possible pairing between these four regime types,[32] with the HRF-HRF dyad as our omitted reference category. We choose to use these dyads as our baseline since these states constitute the most cohesive voting bloc in the UNHRC, as shown in Figure 2 and, as such, are the group whose UPR reports we would expect to be the most similar.

[32] Namely, backslider-backslider; backslider-HRF; backslider-autocracy; autocracy-autocracy; HRF-autocracy; democracy-democracy; democracy-backslider; democracy-autocracy; and democracy-HRF.

**Figure 6.** Over time, the content of backslider states' UPR reports has become more similar to that of China, and less similar to reports written by the HRF group.




Our models control for factors that might impact report similarity. At the recommending state level, we control for whether both recommending states are from the same geographic region; the difference between the recommending states' agreement scores with the United States in the UNGA; and whether either recommending state is also under review during the same UPR session. We also control for characteristics of the state under review, namely, that state's human rights score (Cingranelli, Richards, and Clay, 2014) and an indicator for whether the state under review was simultaneously a member of the UNHRC.

We estimate five ordinary least squares (OLS) models with four-way fixed effects for year, each recommending state, and the state under review, and we cluster the standard errors by dyad (of recommending states). The results of these models are reported in Table 3. In the first model, we include observations for all UPR reports written between 2008–22. In the second and third models, to explore the temporal dynamics of backsliding, we focus on different time periods: 2008–12 and 2013–22, respectively. In Models 4 and 5, we again pool the observations for all years, and instead explore whether the identity of the state under review impacts the similarity of reports. Specifically, in Model 4 we focus exclusively on reports written about any of the four HRF countries,[33] and in Model 5 we only include observations for reviews written about consolidated autocracies.

[33] The "SuR UNHRC member" control variable is dropped in Model 4 as it is collinear with year fixed effects.

**Table 3.** Cosine Similarity of UPR Reports, 2008–2022

| Dependent variable | *Cosine similarity* | | | | |
|---|---|---|---|---|---|
| | All | 2008-2012 | 2013-2022 | HRF | Autocracies |
| Model: | (1) | (2) | (3) | (4) | (5) |
| *Variables* | | | | | |
| DB-DB | 3.16*** | 2.45*** | 3.20*** | **1.62** | 4.43*** |
| | (0.58) | (0.93) | (0.66) | **(1.57)** | (0.782) |
| **DB-HRF** | **-0.219** | **-0.318** | **-0.350** | **-0.384** | **-0.186** |
| | **(0.37)** | **(0.58)** | **(0.38)** | **(0.74)** | **(0.50)** |
| **DB-Auto** | **2.18*** ** | **2.17** ** | **2.16*** ** | **-0.33** | **3.20*** ** |
| | **(0.57)** | **(0.95)** | **(0.66)** | **(1.57)** | **(0.77)** |
| Auto-Auto | 2.65*** | 4.08*** | 2.39*** | -1.37 | 3.89*** |
| | (0.58) | (1.05) | (0.67) | (1.60) | (0.78) |
| **HRF-Auto** | **-1.88*** ** | **-1.57*** ** | **-1.96*** ** | **-2.32*** ** | **-2.54*** ** |
| | **(0.22)** | **(0.51)** | **(0.23)** | **(0.45)** | **(0.30)** |
| **Demo-HRF** | **1.03*** ** | **0.83*** | **1.07*** ** | **-0.421** | **1.52*** ** |
| | **(0.23)** | **(0.44)** | **(0.24)** | **(0.45)** | **(0.32)** |
| Demo-Auto | 1.46** | 1.36 | 1.50** | -0.73 | 2.36*** |
| | (0.57) | (0.99) | (0.66) | (1.58) | (0.77) |
| Demo-DB | 3.01*** | 2.41*** | 3.10*** | 1.74 | 4.35*** |
| | (0.57) | (0.93) | (0.66) | (1.56) | (0.77) |
| Demo-Demo | 3.34*** | 2.86*** | 3.47*** | 1.64 | 4.95*** |
| | (0.57) | (0.97) | (0.66) | (1.58) | (0.77) |
| Both rec same region | 1.38*** | 1.18*** | 1.42*** | 1.61*** | 1.48*** |
| | (0.08) | (0.11) | (0.09) | (0.21) | (0.09) |
| US-agree diff between recs | -0.380** | -1.27*** | -0.861*** | 0.158 | -0.280 |
| | (0.18) | (0.41) | (0.22) | (0.99) | (0.23) |
| SuR HR score | -0.019*** | 0.001 | -0.015*** | 0.052 | -0.043*** |
| | (0.00) | (0.02) | (0.00) | (17,805.8) | (0.00) |
| SuR UNHRC member | -0.084*** | -0.992*** | -0.158*** | | -0.233*** |
| | (0.02) | (0.22) | (0.03) | | (0.03) |
| Either rec under review | -0.124*** | -0.088** | -0.151*** | -0.035 | -0.043** |
| | (0.02) | (0.04) | (0.02) | (0.09) | (0.02) |
| *Fixed-effects* | | | | | |
| Year | Yes | Yes | Yes | Yes | Yes |
| SuR | Yes | Yes | Yes | Yes | Yes |
| Rec 1 | Yes | Yes | Yes | Yes | Yes |
| Rec 2 | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 1,008,665 | 169,213 | 839,452 | 26,783 | 490,067 |
| $R_2$ | 0.06 | 0.09 | 0.07 | 0.09 | 0.09 |
| Mean of DV | 5.94 | 5.86 | 5.95 | 6.19 | 5.71 |
| SD of DV | 6.93 | 7.00 | 6.91 | 6.82 | 6.75 |

*Note: Clustered standard-errors by dyad in parentheses.* *p<0.05;**p<0.01;***p<0.001.
*DB = democratic backslider; HRF = human rights founder; Demo = democracy; Auto = autocracy.*

Several findings stand out. First, in four of the five models, reports written by pairs of backsliding states are significantly more similar than reports written by a pair of HRF countries (our baseline, omitted dyad). This suggests the existence of broadly shared interests across backslider states. We also see several interesting patterns among the mixed regime type dyads. On the one hand, we find no significant relationship between reports written by backsliders and HRF states in any of the five models. This suggests that, on average, across all reviews, backsliders' reports are relatively distinct from those written by the founding members of the global human rights regime, and that backsliders are shifting away from these states. While none of the backslider-HRF coefficients are significant, all of the coefficients for the HRF-autocracy dyads are negative and significant; this suggests backsliders are distinct from autocracies, occupying some middle ground between autocrats and HRF states.[34]

That said, backsliders often do align with autocrats in the UPR. The coefficients for the backslider-autocracy dyad are positive and significant in four of the five models, suggesting that, like in the UNHRC, backsliders are shifting away from the West to align more closely with autocrats. Tellingly, however, there is no significant relationship between backsliders and autocracies when an HRF state is under review. Therefore, although backsliders are increasingly pushing back against Western liberal values and intrusions that threaten their new regime, their evaluations of the LIO's core architects vary from those of consolidated autocracies, again highlighting their status as a distinct bloc with a unique set of interests.

In addition to being distinct from autocracies, the results also provide evidence that backsliders are separate from more stable democracies. Democracies write reports that align much more closely with those of HRF states: in four of the five models, the coefficient for the Demo-HRF dyad is positive and statistically significant, indicating they are more reliable partners for HRF countries. This is in stark contrast to the DB-HRF coefficients, none of which are significant; this is evidence that the HRF's core coalition is shrinking.

One important question is: how large are these estimated effect sizes? At the bottom of Table 3, we report the mean and standard deviation for the dependent variable for each relevant subset of the data. Among the significant coefficients for the main independent variables, the effect sizes range from a change of 0.11 (democracy-HRF, model 2) to 0.73 (democracy-democracy, model 5) standard deviations.[35]

---

[34] These results also validate conceptually our measure of text similarity: we would expect HRF states' reports to be less similar to ones written by autocracies when compared to a pair of two HRF countries.

[35] We note that while the coefficient sizes and the means and standard deviations for our dependent variables are seemingly quite small, they are in line with the magnitudes reported in other recent studies in political science that use cosine similarity to measure text similarity (e.g., Hager and Hilbig, 2020).

**Table 4.** Pairwise differences between coefficient estimates for UPR models from Table 3. "✓" indicates pairs of coefficients that are statistically different from one another ($p < 0.05$), while "✗" indicates no statistically significant difference.

| | All | 2008–2012 | 2013–2022 | HRF | Autocracies |
|---|---|---|---|---|---|
| DB-DB: Auto-Auto | ✓ | ✓ | ✓ | ✓ | ✓ |
| DB-DB: Demo-Demo | ✗ | ✗ | ✓ | ✗ | ✓ |
| Demo-Demo: Auto-Auto | ✓ | ✓ | ✓ | ✓ | ✓ |

*Note: DB = democratic backslider; HRF = human rights founder; Demo = democracy; Auto = autocracy.*

We also report in Table 4 whether the coefficients for each like-pair of states are significantly different from one another across all five models. Here we find that backsliders are significantly different from autocracies in all five models. This a critical point: while backsliders are becoming increasingly illiberal, they are a distinct type of regime, rather than the newest members of the autocratic bloc. That said, there may be concern that backslider states are therefore simply a subset of all non-HRF democracies. However, as Table 4 shows, this is also not the case: in three of the five models, a pair of backsliders is statistically distinct from a pair of democracies. While there was no evidence in Table 3 of shifts over time in backsliders' affinity with one another or with other regime types, the comparison between these two pairs in Table 4 is suggestive of the temporal dynamic predicted in Hypothesis 7: backsliders and democracies were not statistically distinct in the pre-2013 period (Model 2), but became so after 2013 (Model 3). This supports our overarching argument that backsliders were previously part of the broader international democratic coalition, but overtime they have broken away and become less reliable partners.

These findings are robust to a series of alternative specifications. As discussed in more detail in Section 4.1, we reestimate these models with cases of backsliding that meet the 0.2, rather than 0.1 threshold for overall decline,[36] and we use the 0.42, rather than 0.5, threshold, along the electoral democracy index to distinguish democracies from autocracies.[37] In both cases, our core findings remain the same.

[36] See Appendix Table A6.

[37] See Appendix Table A7.

Cosine similarity uses a bag-of-words approach to measure the similarity between two texts. The benefit of such an approach is that it is better able to capture deeper semantic similarities between two texts, including paraphrasing. Other measures of text similarity rely instead on syntactic similarities: they capture more exact word or phrase overlaps to identify commonalities. To ensure our results are not sensitive to the measure of similarity we chose, we also reestimate the models from Table 3 using a common syntactic measure of similarity—Jaccard similarity—as our dependent variable instead. Our key findings remain the same.[38]

Taken together, Tables 3 and 4 provide support for Hypotheses 6 and 7: backsliders are frequently aligned both with one another and with autocracies in their UPR reports. Nevertheless, important differences between backsliders and autocracies remain: backslider states' reports are not in direct opposition to those of HRF states in the way that autocracies are, and backslider dyads are statistically distinct from autocratic ones. Early on, backslider states were indistinguishable from democracies, yet over time, as backsliding has progressed, they have broken away from this group. This indicates that backsliders have emerged as a new and distinct bloc of states on the international stage.

## 6. Conclusion

Rising challenges from China, Russia, and other powerful autocratic states, coupled with political weaknesses among its founders and champions, call into question the future of the liberal international order. At the heart of the LIO's existential crisis is the question of whether the advanced democracies that created it are able to maintain and continue to expand their coalition of likeminded states when attempting to promote and protect Western liberal values, including human rights. While growing coordination and influence among autocratic states is undoubtedly a cause for concern, it would be less so if Western liberal democracies were able to maintain the support of a broader coalition of democratic states. However, we show that one consequence of the ongoing global democratic recession is that the West is losing members of its liberal coalition as a subset of democracies backsliders become increasingly illiberal and, as a result, engage in selective contestation of the most intrusive parts of the LIO. This shift pulls them away from the West as they drift closer toward established autocracies on a subset of particularly sensitive issues.

[38] See Appendix Table A8.

We explore these dynamics in the context of the United Nations, providing new evidence that backsliders are more likely to vote with autocratic states in the UNHRC on targeted and contentious resolutions, especially after 2013; we observe similar dynamics of selective contestation by backsliders in the post-Cold War period in the UNGA. Drawing on the text of reports from the UPR, we also show that backsliders are adopting rhetoric more similar to that espoused by autocracies as they drift away from the West. Nevertheless, these backsliders remain distinct both from more stable democracies and also from consolidated autocracies. This highlights the need to treat this group of a states as a unique bloc with a distinct set of preferences, rather than simply as transitional regimes with the same interests as autocrats.

Taken together, this evidence suggests that the core founding members of the LIO are becoming increasingly isolated in their efforts to promote and support liberal democratic values on the international stage, and will face additional barriers to holding illiberal regimes to account for their domestic human rights practices. The long-term consequence of this realignment of states in international institutions is further exacerbated by the United States' renewed isolationism under Donald Trump in his second term. As in his first administration, Trump has withdrawn the United States from the UNHRC. Indeed, his executive order doing so justified the decision by claiming that the UNHRC had become a forum for autocratic states to shield human rights offenders. In that, if nothing else, our research is in agreement. But, unchecked and unreversed, the United States' own democratic backsliding and refusal to defend the LIO against its detractors places the world in waters uncharted since the interwar era of the 1930s. Then, a fledgling effort at global governance failed spectacularly with devastating consequences. What happens next in our time is uncertain, but our research makes clear that the LIO, exemplified by the UNHRC and the international human rights regime it heads, finds itself with fewer champions than ever before, the sad culmination of a steady inexorable erosion of the Western democratic coalition over the past twenty years.

## References


Barnett, Michael and Martha Finnemore. 2021. *The Power of Liberal International Organizations*. Johns Hopkins University Press.

Bartels, L.M. 2023. *Democracy Erodes from the Top: Leaders, Citizens, and the Challenge of Populism in Europe*. Princeton University Press.

Baturo, Alexander. 2023. "Russia's Global Anti-Nazism Campaign: Seeking Support in International Organizations."

Baturo, Alexander, Niheer Dasandi, and Slava J. Mikhaylov. 2017. "Understanding state preferences with text as data: Introducing the UN General Debate corpus." *Research & Politics* 4(2):2053168017712821.

Bell, Andrew and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data." *Political Science Research and Methods* 3(1):133–153.

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

Berry, William D, Jacqueline H.R. DeMeritt, and Justin Esarey. 2010. "Testing for interaction in binary logit and probit models: is a product term essential?" *American Journal of Political Science* 54(1):248–266.

Binder, Martin and Autumn Lockwood Payton. 2022. "With Frenemies Like These: Rising Power Voting Behavior in the UN General Assembly." *British Journal of Political Science* 52(1):381–398.

Börzel, Tanja A and Michael Zürn. 2021. "Contestations of the liberal international order: From liberal multilateralism to postnational liberalism." *International Organization* 75(2):282–305.

Boyle, Michael J. 2023. The coming illiberal order. In *Survival 58.2*. Routledge pp. 35–65.

Brazys, Samuel and Alexander Dukalskis. 2017. "Canary in the coal mine? China, the UNGA, and the changing world order." *Review of International Studies* 43(4):742–764.

Burmester, Nicolas and Michael Jankowski. 2014. "Reassessing the European Union in the United Nations general assembly." *Journal of European Public Policy* 21(10): 1491–1508.

Burmester, Nicolas and Michael Jankowski. 2018. “One voice or different choice? Vote defection of European Union member states in the United Nations General Assembly.” *The British Journal of Politics and International Relations* 20(3):652–673.

Bush, Sarah Sunn, Christina Cottiero, and Lauren Prather. 2024. “Zombies ahead: Explaining the rise of low-quality election monitoring.” *The Review of International Organizations* 20:162–291.

Carter, David B. and Randall W. Stone. 2015. “Democracy and Multilateralism: The Case of Vote Buying in the UN General Assembly.” *International Organization* 69(1):1–33.

Cingranelli, David L., David L. Richards, and K. Chad Clay. 2014. “The CIRI Human Rights Dataset.” CIRI Human Rights Data Project, Version 2014.04.14.

Clark, Tom S. and Drew A. Linzer. 2015. “Should I Use Fixed or Random Effects?” *Political Science Research and Methods* 3(2):399–408.

Cooley, A. and D. Nexon. 2020. *Exit from Hegemony: The Unraveling of the American Global Order*. Oxford University Press.

Cooley, Alexander. 2016. Countering Democratic Norms. In *Authoritarianism Goes Global: The Challenge to Democracy*, ed. Larry Diamond, Marc F. Plattner and Scott Walker. Baltimore: Johns Hopkins University Press chapter 7, pp. 117–134.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Fabio Angiolillo, Michael Bernhard, Cecilia Borella, Agnes Cornell, M. Steven Fish, Linnea Fox, Lisa Gastaldi, Haakon Gjerlow, Adam Glynn, Ana Good God, Sandra Grahn, Allen Hicken, Katrin Kinzelbach, Joshua Krusell, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Natalia Natsika, Anja Neundorf, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Oskar Rydén, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundström, Eitan Tzelgov, Yi ting Wang, Tore Wig, Steven Wilson, and Daniel Ziblatt. 2024. “V-Dem [Country-Year/Country-Date] Dataset v14.”

Cottiero, Christina, Emilie M. Hafner-Burton, Stephan Haggard, Lauren Prather, and Christina J. Schneider. 2025. “Illiberal regimes and international organizations.” *The Review of International Organizations* 20:231–259.

Cottiero, Christina and Stephan Haggard. 2023. “Stabilizing Authoritarian Rule: The Role of International Organizations.” *International Studies Quarterly* 67(2):sqad031.

Cox, Eric. 2010. “State Interests and the Creation and Functioning of the United Nations Human Rights Council.” *Journal of International Law and International Relations* 6:87–120.

Dai, Xinyuan and Lucie Lu. 2024. “Beyond Liberal Narratives. China and the International Human Rights Order.” Working paper.

Debre, Maria J. 2025. *How Regional Organizations Sustain Authoritarian Rule: The Dictators’ Club*. Oxford University Press.

Desposato, Scott W. 2005. “Correcting for small group inflation of roll-call cohesion scores.” *British Journal of Political Science* 35(4):731–744.

Diodati, Nicola Martocchia, Bruno Marino, and Benedetta Carlotti. 2018. “Prime ministers unchained? Explaining prime minister policy autonomy in coalition governments.” *European Political Science Review* 10(4):515–536.

Donno, Daniela. 2013. *Defending Democratic Norms: International Actors and the Politics of Electoral Misconduct*. Oxford: Oxford University Press.

Dreher, Axel, Andreas Fuchs, Bradley Parks, Austin Strange, and Michael J. Tierney. 2022. *Banking on Beijing: The aims and impacts of China’s overseas development program*. Cambridge University Press.

Dreher, Axel, Peter Nunnenkamp, and Rainer Thiele. 2008. “Does US aid buy UN general assembly votes? A disaggregated analysis.” *Public Choice* 136:139–164.

Dukalskis, Alexander. 2023. “A fox in the henhouse: China, normative change, and the UN Human Rights Council.” *Journal of Human Rights* 22(3):334–350.

Finnemore, Martha and Kathryn Sikkink. 1998. “International Norm Dynamics and Political Change.” *International Organization* 52(4):887–917.

Fjelstul, Joshua, Simon Hug, and Christopher Kilby. 2025. “Decision-making in the United Nations General Assembly: A comprehensive database of resolution-related decisions.” *The Review of International Organizations* pp. 1–18.

Garrett, Kristin N. and Joshua M. Jansa. 2015. “Interest Group Influence in Policy Diffusion Networks.” *State Politics 38; Policy Quarterly* 15(3):387–417.

Geddes, Barbara, Joseph Wright, and Erica Frantz. 2018. *How Dictatorships Work: Power, Personalization, and Collapse*. Cambridge University Press.

Ginsburg, Tom. 2020*a*. "Authoritarian international law?" *American Journal of International Law* 114:221–260.

Ginsburg, Tom. 2020*b*. "How Authoritarians Use International Law." *Journal of Democracy* 31(4):44–58.

Hafner-Burton, Emilie, Jon C.W. Pevehouse, and Christina J. Schneider. 2025. "Good governance in autocratic international organizations." *The Review of International Organizations* 20:323–356.

Hafner-Burton, Emilie M. 2005. "Trading Human Rights: How Preferential Trade Agreements Influence Government Repression." *International Organization* 59:593–629.

Hafner-Burton, Emilie M. 2012. "International regimes for human rights." *Annual Review of Political Science* 15:265–286.

Hager, Anselm and Hanno Hilbig. 2020. "Does public opinion affect political speech?" *American Journal of Political Science* 64(4):921–937.

Haggard, Stephan and Robert Kaufman. 2021. *Backsliding: Democratic Regress in the Contemporary World*. Cambridge University Press.

Hug, Simon. 2012. What's in a Vote? In *APSA 2012 Annual Meeting Paper*.

Hug, Simon and Richard Lukács. 2014. "Preferences or blocs? Voting in the United Nations Human Rights Council." *The Review of International Organizations* 9:83–106.

Hyde, Susan D. 2020. "Democracy's backsliding in the international environment." *Science* 369:1192–1196.

Ikenberry, G. John. 2018. "The end of liberal international order?" *International Affairs* 94:7–23.

Inboden, Rana Siu. 2021. *China and the International Human Rights Regime*. Cambridge University Press.

Jansa, Joshua M., Eric R. Hansen, and Virginia H. Gray. 2019. "Copy and Paste Lawmaking: Legislative Professionalism and Policy Reinvention in the States." *American Politics Research* 47(4):739–767.

Kasuya, Yuko and Kota Mori. 2019. “Better regime cutoffs for continuous democracy measures.” *V-Dem Users Working Paper Series* 25:1–31.

Kaya, Ayse, Christopher Kilby, and Jonathan Kay. 2021. “Asian Infrastructure Investment Bank as an instrument for Chinese influence? Supplementary versus remedial multilateralism.” *World Development* 145:105531.

Kelemen, R. Daniel. 2020. “The European Union’s authoritarian equilibrium.” *Journal of European Public Policy* 27:481–499.

Kelemen, R Daniel. 2024. “Will the European Union escape its autocracy trap?” *Journal of European Public Policy* pp. 1–24.

Kelley, Judith. 2009. “D-minus elections: The politics and norms of international election observation.” *International Organization* 63(4):765–787.

Kelley, Judith G. and Beth A. Simmons. 2015. “Politics by number: Indicators as social pressure in international relations.” *American Journal of Political Science* 59(1):55–70.

Kentikelenis, Alexander and Erik Voeten. 2021. “Legitimacy challenges to the liberal world order: Evidence from United Nations speeches, 1970–2018.” *The Review of International Organizations* 16(4):721–754.

Kim, Yooneui. 2024. “Interests convergence in global human rights politics: Text analysis of Universal Periodic review of the UN human rights council.” *Journal of Language an Politics*.

Kneuer, Marianne, Thomas Demmelhuber, Raphael Peresson, and Tobias Zumbrägel. 2019. “Playing the regional card: why and how authoritarian gravity centres exploit regional organisations.” *Third World Quarterly* 40:451–470.

Lake, David A., Lisa L. Martin, and Thomas Risse. 2021. “Challenges to the liberal order: Reflections on international organization.” *International Organization* 75(2):225–257.

Lebovic, James H. and Erik Voeten. 2006. “The politics of shame: The condemnation of country human rights practices in the UNCHR.” *International Studies Quarterly* 50:861– 888.

Levitsky, Steven and Lucan A. Way. 2010. *Competitive Authoritarianism: Hybrid Regimes After the Cold War*. Cambridge: Cambridge University Press.

Libman, Alexander and Anastassia V. Obydenkova. 2018. “Regional international organizations as a strategy of autocracy: The Eurasian Economic Union and Russian foreign policy.” *International Affairs* 94:1037–1058.

Linder, Fridolin, Bruce Desmarais, Matthew Burgess, and Eugenia Giraudy. 2020. “Text as policy: Measuring policy similarity through bill text reuse.” *Policy Studies Journal* 48(2):546–574.

Lipps, Jana and Marc S. Jacob. 2025. “Undermining liberal international organizations from within: Evidence from the parliamentary assembly of the council of Europe.” *The Review of International Organizations* 20:357—383.

Little, Andrew T. and Anne Meng. 2024. “Measuring democratic backsliding.” *PS: Political Science & Politics* 57(2):149–161.

Lu, Lucie. 2024. “Be My Friendly Reviewers: How China Shapes its UN Human Rights Reviews.” *IGCC Working Paper*. URL: *https://ucigcc.org/publication/be-my-friendly-reviewers-how-china-shapes-its-unhuman-rights-reviews/*

Luhrmann, Anna and Staffan I. Lindberg. 2019. “A third wave of autocratization is here: what is new about it?” *Democratization* 26:1095–1113.

Maerz, Seraphine F., Amanda Edgell, Matthew C. Wilson, Sebastian Hellmeier, and Staffan I. Lindberg. 2021. “A framework for understanding regime transformation: introducing the ERT dataset.” *V-Dem Working Paper* 113.

Manning, Christopher D. 2008. *Introduction to information retrieval*. Cambridge University Press.

Meyerrose, Anna M. 2020. “The Unintended Consequences of Democracy Promotion: International Organizations and Democratic Backsliding.” *Comparative Political Studies* 53.

Meyerrose, Anna M. 2024. “Building Strong Executives and Weak Institutions: How European Integration Contributes to Democratic Backsliding.” *The Review of International Organizations* pp. 307–343.

Meyerrose, Anna M. 2026. *Eroding Democracy from the Outside In: International Policy Delegation and Democratic Backsliding*. Oxford University Press.

Meyerrose, Anna M. and Irfan Nooruddin. 2025. “Trojan Horses in Liberal International Organizations?: How Democratic Backsliders Undermine the UNHRC.” 20:125–156. *The Review of International Organizations*.

Morrison, Kelly, Daniela Donno, Burcu Savun, and Perisa Davutoglu. 2025. “Competing judgments: Multiple election observers and post-election contention.” *The Review of International Organizations* 20:293–321.

Morse, Julia C. and Bridget Coggins. 2024. “Your silence speaks volumes: Weak states and strategic absence in the UN General Assembly.” *The Review of International Organizations* pp. 1–30.

Mudde, Cas and Cristóbal Rovira Kaltwasser. 2017. *Populism: A Very Short Introduction*. Oxford University Press.

Nord, Marina, Martin Lundstedt, David Altman, Fabio Angiolillo, Cecilia Borella, Tiago Fernandes, Lisa Gastaldi, Ana Good God, Natalia Natsika, and Staffan I. Lindberg. 2024. “Democracy Winning and Losing at the Ballot.” *Democracy Report*.

Pauselli, Gino, Francisco Urdínez, and Federico Merke. 2023. “Shaping the liberal international order from the inside: A natural experiment on China’s influence in the UN human rights council.” *Research & Politics* 10(3).

Pevehouse, Jon. 2005. *Democracy from above: regional organizations and democratization*. Cambridge: Oxford University Press.

Prasad, Shubha Kamala and Irfan Nooruddin. 2024. “States living in glasshouses...: Why fighting domestic insurgency changes how countries vote in the UN human rights council.” *Conflict Management and Peace Science* 41(5):556–573.

Seligman, Steven. 2011. “Politics and principle at the UN human rights commission and council (1992-2008).” *Israel Affairs* 17:520–541.

Spirling, Arthur. 2011. “US treaty-making with American Indians.” *American Journal of Political Science* 56(1):84–97.

Steinert, Christoph V. and David Weyrauch. 2024. “Belt and road initiative membership and voting patterns in the United Nations General Assembly.” *Research & Politics* 11(1):20531680241233784.

Terman, Rochelle and Erik Voeten. 2018. “The relational politics of shame: Evidence from the universal periodic review.” *The Review of International Organizations* 13:1–23.

Terman, Rochelle and Joshua Byun. 2022. “Punishment and Politicization in the International Human Rights Regime.” *American Political Science Review* 116:385–402.

Voeten, Erik. 2000. “Clashes in the Assembly.” *International Organization* 54(2):185–215.

Voeten, Erik, Anton Strezhnev, and Michael Bailey. 2009. “United Nations General Assembly Voting Data.”

von Borzyskowski, Inken and Felicity Vabulas. 2025. *Exit from International Organizations:*
*Costly Negotiation for Institutional Change*. Cambridge University Press.

Voss, M. Joel. 2019. “The Use (or Misuse) of Amendments to Contest Human Rights Norms at the UN Human Rights Council.” *Human Rights Review* 20:397–422.

Vreeland, James Raymond. 2019. “Corrupting International Organizations.” *Annual Review of Political Science* 22:205–222.

Vreeland, James Raymond and Axel Dreher. 2014. *The political economy of the United Nations Security Council: Money and influence*. Cambridge University Press.

Waldner, David and Ellen Lust. 2018. “Unwelcome Change: Coming to Terms with Democratic Backsliding.” *Annual Review of Political Science* 21(5):1–21.

Watanabe, Kohei and Yuan Zhou. 2022. “Theory-driven analysis of large corpora: Semisupervised topic classification of the UN speeches.” *Social Science Computer Review* 40(2):346– 366.

Weiss, Jessica Chen and Jeremy L Wallace. 2021. “Domestic politics, China’s rise, and the future of the liberal international order.” *International Organization* 75(2):635–664.

Winzen, Thomas. 2025. “How backsliding governments keep the European Union hospitable for autocracy: Evidence from intergovernmental negotiations.” *The Review of International Organizations* 20:385–414.

World Bank. 2024. “World Development Indicators 2024.”

Ziblatt, Daniel and Steven Levitsky. 2018. *How Democracies Die*. New York: Crown.

# The Backsliding Bloc? Selective Contestation and Alliance Realignment in the Liberal International Order

**Online Appendices**

## 1. Identifying Episodes of Democratic Backsliding

Episodes of democratic backsliding are listed below, with asterisks indicating backsliders that were voting members of the UNHRC at least once between 2006–23.[39] We focus exclusively on states that begin their autocratization episode as at least minimally democratic states, i.e., ones with a (pre-autocratization) electoral democracy score of 0.5 or higher (Haggard and Kaufman, 2021). Second, while the ERT dataset uses the electoral democracy index to capture the relevant regressions, we instead rely on the liberal democracy index. Democratic backsliding occurs when elected officials erode or undermine often *liberal* democratic institutions, such as the constitution, rule of law, civil and minority rights, judicial and media independence, and separation of power within governments. The liberal democracy index importantly includes these other liberal democratic institutions, while the electoral democracy index focuses primarily on election and participation-based indicators. Therefore, we adopt the V-Dem approach, but instead use the liberal democracy index for the subset of states that began as democracies, to identify backsliders.

- Argentina, 2002-2010*
- Armenia, 2020-2023*
- Benin, 2018-2021
- Bolivia, 2005-2020*
- Botswana, 2015-2023*
- Brazil, 2015-2021*
- Bulgaria, 2001-2018
- Burkina Faso, 2018-2023*
- Croatia, 2013-2017*
- Cyprus, 2017-2023
- Czechia, 2010-2021*
- Ecuador, 2007-2013*
- El Salvador, 2018-2023
- Ghana, 2013-2023*
- Greece, 2013-2023
- Guatemala, 2018-2023
- Guyana, 2019-2023
- Honduras, 2006-2018
- Hungary, 2006-2023*
- India, 2009-2023*
- Indonesia, 2009-2023*
- Ivory Coast, 2017-2023*
- Lesotho, 2015-2017
- Maldives, 2012-2015*
- Mali, 2017-2023
- Mauritius, 2013-2023
- Mexico, 2019-2023*
- Moldova, 2013-2017*
- Mongolia, 2008-2023*
- Nicaragua, 2006-2018*
- Niger, 2016-2023
- North Macedonia, 2007-2013
- Palestine/West Bank, 2006-2008
- Peru, 2016-2023*
- Philippines, 2016-2023*
- Poland, 2015-2023*
- Romania, 2017-2018 and 2021-2023*
- Senegal, 2017-2023*
- Serbia, 2010-2023
- Slovenia, 2012-2021*
- South Korea, 2008-2013 and 2019-2023*
- Sri Lanka, 2019-2020
- Thailand, 2005-2006
- Tunisia, 2015-2022*
- Türkiye, 2005-2017
- Venezuela, 1994-2006*

[39] Our statistical approach codes the United States as a backslider from 2016–18. Our results remain the same if we code the United States as a backslider, and exclude it from the HRF group, in the relevant years. But, for intellectual consistency, since our focus in this paper is on the extent to which backslider states are shifting away from the original architects of the LIO—including, arguably most prominently, the US—we exclude it from our list of backslider states.

## 2. Marginal Effects of Liberal Democracy on Vote Choice in the UNHRC

The effect of backsliding on a state's propensity to vote with the autocratic bloc in the UNHRC may depend on the state's level of liberal democracy at the time of the vote. To explore this possibility, Figure A1 plots the marginal effects of being a backsliding state (or not) on vote choice for different levels of liberal democracy, using the results from Model 4 in Table 1. As states' levels of liberal democracy increase, they are less likely to vote with autocracies; this is true for both backslider and nonbackslider democracies. However, at all levels of liberal democracy, backslider states are more likely to vote with autocrats than nonbackslider democracies.

**Figure A1.** Holding all other co-variates at their mean, we find that states are less likely to vote with the autocratic bloc as their liberal democracy score increases. This is equally the case for backslider and nonbackslider democracies. This plot is created using the results from Model 4 from Table 1.

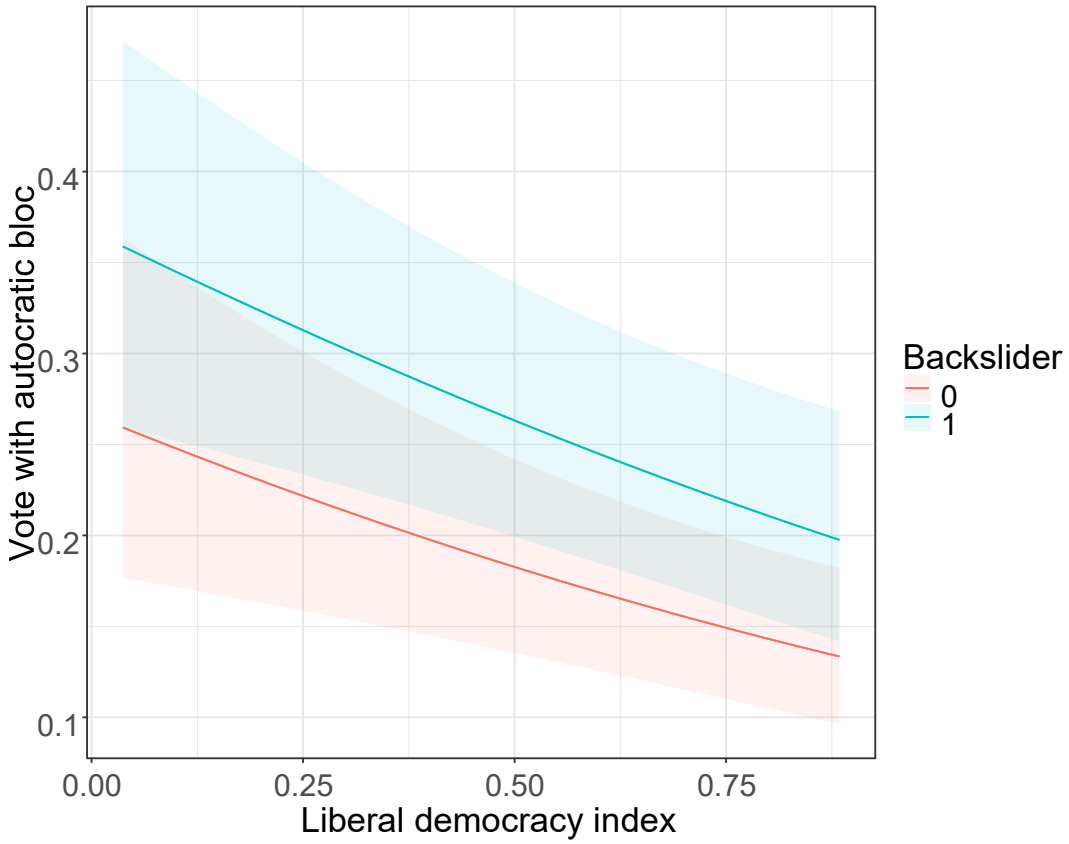


**Table A1.** Voting with the Autocratic Bloc in the UNHRC, 2006–2023, 0.2 Backsliding Threshold

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Vote with autocratic bloc (1 = yes, 0 = no) | | | |
| | All | General | Targeted | Contentious targeted |
| | (1) | (2) | (3) | (4) |
| Backslider state | 0.28*** | -0.13 | 0.70*** | 0.96*** |
| | (0.08) | (0.13) | (0.13) | (0.16) |
| Consolidated autocracy | 0.53*** | 0.45*** | 0.68*** | 0.61*** |
| | (0.08) | (0.12) | (0.11) | (0.13) |
| Voting lib dem score | -0.37* | -0.69** | 0.02 | -0.64* |
| | (0.16) | (0.24) | (0.24) | (0.29) |
| Voting GDP pc (log) | -0.07*** | -0.13*** | -0.01 | -0.14*** |
| | (0.02) | (0.03) | (0.03) | (0.03) |
| Voting UNGA U.S. agree | -9.70*** | -11.53*** | -8.94*** | -14.18*** |
| | (0.19) | (0.28) | (0.29) | (0.48) |
| Percent demo sponsors | -0.01*** | -0.003*** | -0.001 | -0.002 |
| | (0.0005) | (0.001) | (0.001) | (0.001) |
| Target = Israel | | | 1.40*** | 1.83*** |
| | | | (0.23) | (0.28) |
| Target = Syria | | | -0.36*** | -0.51*** |
| | | | (0.09) | (0.12) |
| Target lib dem score | | | 0.75* | -0.32 |
| | | | (0.38) | (0.47) |
| Target GDP pc (log) | | | -0.06 | 0.09 |
| | | | (0.04) | (0.05) |
| Target UNGA U.S. agree | | | -0.04 | 1.45** |
| | | | (0.37) | (0.48) |
| Voting-target UNGA affinity | | | 0.10 | 2.51*** |
| | | | (0.22) | (0.40) |
| Voting-target same region | | | -0.24* | -0.20 |
| | | | (0.12) | (0.14) |
| Year FE | Yes | Yes | Yes | Yes |
| Clustered SE (Country) | Yes | Yes | Yes | Yes |
| Logit model | Yes | Yes | Yes | Yes |
| Observations | 21,292 | 11,244 | 9,256 | 7,304 |
| Log Likelihood | -8,935.17 | -4,002.34 | -4,135.76 | -2,637.97 |
| Akaike Inf. Crit. | 17,918.34 | 8,052.68 | 8,333.51 | 5,337.94 |
| *Note:* | | | | *p<0.05; **p<0.01; ***p<0.001 |

**Table A2.** Voting with the Autocratic Bloc in the UNHRC, 2006–2023, 0.42 Democracy Threshold

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Vote with autocratic bloc (1 = yes, 0 = no) | | | |
| | All | General | Targeted | Contentious targeted |
| | (1) | (2) | (3) | (4) |
| Backslider state | 0.10* | -0.10 | 0.29*** | 0.49*** |
| | (0.05) | (0.08) | (0.07) | (0.09) |
| Consolidated autocracy | 0.58*** | 0.58*** | 0.71*** | 0.72*** |
| | (0.08) | (0.12) | (0.11) | (0.14) |
| Voting lib dem score | -0.50*** | -0.64** | -0.20 | -0.65* |
| | (0.15) | (0.22) | (0.22) | (0.27) |
| Voting GDP pc (log) | -0.11*** | -0.17*** | -0.05 | -0.18*** |
| | (0.02) | (0.03) | (0.03) | (0.03) |
| Voting UNGA US agree | -9.26*** | -11.21*** | -8.36*** | -13.52*** |
| | (0.19) | (0.28) | (0.28) | (0.46) |
| Percent demo sponsors | -0.01*** | -0.004*** | -0.005*** | -0.004* |
| | (0.0005) | (0.001) | (0.001) | (0.001) |
| Target = Israel | | | 1.16*** | 1.84*** |
| | | | (0.23) | (0.28) |
| Target = Syria | | | -0.45*** | -0.64*** |
| | | | (0.09) | (0.12) |
| Target lib dem score | | | 0.88* | -0.17 |
| | | | (0.38) | (0.47) |
| Target GDP pc (log) | | | -0.06 | 0.07 |
| | | | (0.04) | (0.05) |
| Target UNGA US agree | | | -0.24 | 1.22** |
| | | | (0.37) | (0.47) |
| Voting-target UNGA affinity | | | -0.19 | 2.31*** |
| | | | (0.22) | (0.39) |
| Voting-target same region | | | -0.25* | -0.17 |
| | | | (0.12) | (0.14) |
| Year FE | Yes | Yes | Yes | Yes |
| Clustered SE (Country) | Yes | Yes | Yes | Yes |
| Logit model | Yes | Yes | Yes | Yes |
| Observations | 21,709 | 11,487 | 9,423 | 7,485 |
| Log Likelihood | -9,059.94 | -4,056.00 | -4,197.37 | -2,697.69 |
| Akaike Inf. Crit. | 18,167.89 | 8,160.01 | 8,456.74 | 5,457.37 |

*Note:* *p<0.05; **p<0.01; ***p<0.001

**Table A3.** Voting with the Autocratic Bloc in the UNHRC, 2006–2012

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Vote with autocratic bloc (1 = yes, 0 = no) | | | |
| | All | General | Targeted | Contentious targeted |
| | (1) | (2) | (3) | (4) |
| Backslider state | 0.16 | 0.09 | 0.13 | 0.61* |
| | (0.13) | (0.20) | (0.20) | (0.31) |
| Consolidated autocracy | 1.07*** | 1.20*** | 1.00*** | 1.41*** |
| | (0.17) | (0.26) | (0.26) | (0.34) |
| Voting lib dem score | 0.51 | 0.08 | 1.15* | 0.99 |
| | (0.33) | (0.50) | (0.53) | (0.71) |
| Voting GDP pc (log) | -0.08* | -0.15* | 0.02 | 0.07 |
| | (0.04) | (0.07) | (0.07) | (0.09) |
| Voting UNGA US agree | -13.31*** | -16.76*** | -10.91*** | -21.07*** |
| | (0.46) | (0.72) | (0.70) | (1.45) |
| Percent demo sponsors | -0.01*** | -0.01** | 0.01*** | 0.01** |
| | (0.001) | (0.002) | (0.003) | (0.01) |
| Target = Israel | | | 2.89 | 7.38** |
| | | | (1.51) | (2.40) |
| Target = Syria | | | -0.49 | -3.55*** |
| | | | (0.41) | (0.79) |
| Target lib dem score | | | -0.72 | -2.82 |
| | | | (1.82) | (2.38) |
| Target GDP pc (log) | | | -1.33*** | -0.99* |
| | | | (0.32) | (0.46) |
| Target UNGA US agree | | | 1.57 | -2.46 |
| | | | (1.82) | (2.55) |
| Voting-target UNGA affinity | | | -5.17*** | -1.15 |
| | | | (0.55) | (1.29) |
| Voting-target same region | | | 0.77* | 0.62 |
| | | | (0.36) | (0.51) |
| Year FE | Yes | Yes | Yes | Yes |
| Clustered SE (Country) | Yes | Yes | Yes | Yes |
| Logit model | Yes | Yes | Yes | Yes |
| Observations | 5,852 | 3,360 | 2,312 | 1,678 |
| Log Likelihood | -2,093.80 | -934.81 | -859.58 | -450.04 |
| Akaike Inf. Crit. | 4,213.61 | 1,895.62 | 1,759.16 | 940.08 |
| *Note:* | | | | *p<0.05; **p<0.01; ***p<0.001 |

**Table A4.** Voting with the Autocratic Bloc in the UNHRC, 2013–2023

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Vote with autocratic bloc (1 = yes, 0 = no) | | | |
| | All | General | Targeted | Contentious targeted |
| | (1) | (2) | (3) | (4) |
| Backslider state | 0.11 | -0.08 | 0.32*** | 0.43*** |
| | (0.06) | (0.09) | (0.09) | (0.11) |
| Consolidated autocracy | 0.24** | 0.04 | 0.45*** | 0.26 |
| | (0.09) | (0.13) | (0.13) | (0.15) |
| Voting lib dem score | -0.94*** | -1.16*** | -0.78** | -1.55*** |
| | (0.18) | (0.26) | (0.26) | (0.32) |
| Voting GDP pc (log) | -0.06** | -0.11*** | -0.004 | -0.17*** |
| | (0.02) | (0.03) | (0.03) | (0.04) |
| Voting UNGA US agree | -8.52*** | -10.13*** | -7.64*** | -11.87*** |
| | (0.20) | (0.30) | (0.31) | (0.49) |
| Percent demo sponsors | -0.01*** | -0.002** | -0.004*** | -0.003* |
| | (0.001) | (0.001) | (0.001) | (0.002) |
| Target = Israel | | | 1.67*** | 1.79*** |
| | | | (0.24) | (0.29) |
| Target = Syria | | | -0.58*** | -0.32** |
| | | | (0.10) | (0.12) |
| Target lib dem score | | | 1.34*** | -0.38 |
| | | | (0.40) | (0.50) |
| Target GDP pc (log) | | | -0.06 | 0.13* |
| | | | (0.04) | (0.05) |
| Target UNGA US agree | | | -0.73 | 1.97*** |
| | | | (0.39) | (0.51) |
| Voting-target UNGA affinity | | | 1.14*** | 4.15*** |
| | | | (0.25) | (0.44) |
| Voting-target same region | | | -0.43*** | -0.30* |
| | | | (0.13) | (0.15) |
| Year FE | Yes | Yes | Yes | Yes |
| Clustered SE (Country) | Yes | Yes | Yes | Yes |
| Logit model | Yes | Yes | Yes | Yes |
| Observations | 15,857 | 8,127 | 7,111 | 5,761 |
| Log Likelihood | -6,904.78 | -3,067.07 | -3,205.04 | -2,139.96 |
| Akaike Inf. Crit. | 13,843.56 | 6,168.13 | 6,458.08 | 4,327.93 |
| *Note:* | | | | *p<0.05; **p<0.01; ***p<0.001 |

**Table A5.** Voting with the Autocratic Bloc in the UNGA

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | Vote with autocratic bloc (1 = yes, 0 = no) | | | | | |
| | All | Contentious | Important | Important contentious (IC) | IC 1990–2012 | IC 2013–2022 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Backslider state | -0.11*** | -0.11*** | -0.05 | 0.12* | -0.06 | 0.21** |
| | (0.02) | (0.02) | (0.03) | (0.06) | (0.09) | (0.08) |
| Consolidated autocracy | 0.07*** | 0.17*** | 0.19*** | 0.33*** | 0.33*** | 0.28** |
| | (0.02) | (0.02) | (0.04) | (0.05) | (0.07) | (0.09) |
| Voting lib dem score | -0.06 | -0.04 | -0.14 | -0.65*** | -1.02*** | -0.68*** |
| | (0.03) | (0.05) | (0.08) | (0.12) | (0.16) | (0.20) |
| Voting GDP pc (log) | -0.07*** | -0.07*** | -0.06*** | -0.06*** | -0.06*** | -0.10*** |
| | (0.004) | (0.01) | (0.01) | (0.01) | (0.02) | (0.02) |
| Voting UNGA US agree | -7.21*** | -13.83*** | -5.18*** | -15.40*** | -16.17*** | -14.43*** |
| | (0.04) | (0.06) | (0.09) | (0.20) | (0.27) | (0.30) |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Clustered SE (Country) | Yes | Yes | Yes | Yes | Yes | Yes |
| Logit model | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 673,691 | 371,541 | 76,581 | 41,383 | 23,679 | 14,286 |
| Log Likelihood | -248,566.50 | -117,232.60 | -38,752.40 | -15,406.08 | -8,720.81 | -5,041.01 |
| Akaike Inf. Crit. | 497,268.90 | 234,601.10 | 77,594.80 | 30,900.15 | 17,497.62 | 10,112.02 |
| *Note:* | | | | | | *p<0.05; **p<0.01; ***p<0.001 |

**Table A6.** Cosine Similarity of UPR Reports, 2008–2022, 0.2 Backsliding Threshold.

| Dependent variable | *Cosine similarity* | | | | |
|---|---|---|---|---|---|
| | All | 2008-2012 | 2013-2022 | HRF | Autocracies |
| Model: | (1) | (2) | (3) | (4) | (5) |
| *Variables* | | | | | |
| DB-DB | 3.37*** | 4.98*** | 3.03*** | 1.96 | 4.70*** |
| | (0.621) | (0.782) | (0.695) | (1.81) | (0.850) |
| DB-HRF | -0.326 | 0.548 | -0.501 | 0.065 | -0.076 |
| | (0.438) | (0.572) | (0.457) | (1.16) | (0.633) |
| DB-Auto | 3.10*** | 4.69*** | 2.76*** | 0.790 | 4.34*** |
| | (0.601) | (0.858) | (0.676) | (1.76) | (0.815) |
| Auto-Auto | 3.48*** | 5.87*** | 2.96*** | 0.148 | 4.66*** |
| | (0.606) | (1.10) | (0.681) | (1.80) | (0.819) |
| HRF-Auto | -1.44*** | -0.654 | -1.66*** | -1.56** | -2.11*** |
| | (0.242) | (0.525) | (0.243) | (0.613) | (0.332) |
| Demo-HRF | 1.17*** | 1.68*** | 0.918*** | -0.131 | 1.50*** |
| | (0.255) | (0.481) | (0.253) | (0.630) | (0.350) |
| Demo-Auto | 2.30*** | 3.20*** | 1.98*** | 0.340 | 3.07*** |
| | (0.602) | (1.03) | (0.676) | (1.79) | (0.815) |
| Demo-DB | 3.45*** | 4.22*** | 3.17*** | 2.46 | 4.74*** |
| | (0.601) | (0.828) | (0.675) | (1.76) | (0.816) |
| Demo-Demo | 3.83*** | 4.60*** | 3.45*** | 2.07 | 5.13*** |
| | (0.603) | (0.995) | (0.676) | (1.80) | (0.817) |
| Both rec same region | 1.42*** | 1.19*** | 1.47*** | 1.62*** | 1.53*** |
| | (0.080) | (0.109) | (0.086) | (0.205) | (0.093) |
| US-agree diff between recs | -0.382** | -1.23*** | -0.870*** | 0.110 | -0.299 |
| | (0.183) | (0.406) | (0.219) | (0.995) | (0.231) |
| SuR HR score | -0.019*** | 0.002 | -0.015*** | -2.10 | -0.043*** |
| | (0.002) | (0.024) | (0.002) | (17,832.0) | (0.003) |
| SuR UNHRC member | -0.085*** | -0.996*** | -0.160*** | | -0.232*** |
| | (0.022) | (0.215) | (0.027) | | (0.031) |
| Either rec under review | -0.128*** | -0.080** | -0.157*** | -0.067 | -0.047** |
| | (0.017) | (0.041) | (0.019) | (0.093) | (0.022) |
| *Fixed-effects* | | | | | |
| Year | Yes | Yes | Yes | Yes | Yes |
| SuR | Yes | Yes | Yes | Yes | Yes |
| Rec 1 | Yes | Yes | Yes | Yes | Yes |
| Rec 2 | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 1,008,665 | 169,213 | 839,452 | 26,783 | 492,386 |
| $R_2$ | 0.06 | 0.09 | 0.07 | 0.09 | 0.09 |
| Mean of DV | 5.94 | 5.86 | 5.95 | 6.19 | 5.72 |
| SD of DV | 6.93 | 7.00 | 6.91 | 6.82 | 6.74 |

*Note: Clustered standard-errors by dyad in parentheses.* *p<0.05;**p<0.01;***p<0.001.
*DB = democratic backslider; HRF = human rights founder; Demo = democracy; Auto = autocracy.*

**Table A7.** Cosine Similarity of UPR Reports, 2008–2022, 0.42 Democracy Threshold.

| Dependent variable | *Cosine similarity* | | | | |
|---|---|---|---|---|---|
| | All | 2008-2012 | 2013-2022 | HRF | Autocracies |
| Model: | (1) | (2) | (3) | (4) | (5) |
| *Variables* | | | | | |
| DB-DB | 3.16*** | 2.43*** | 3.18*** | 1.54 | 4.48*** |
| | (0.544) | (0.897) | (0.628) | (1.55) | (0.738) |
| DB-HRF | -0.219 | -0.282 | -0.366 | -0.203 | -0.290 |
| | (0.344) | (0.555) | (0.353) | (0.696) | (0.472) |
| DB-Auto | 2.48*** | 2.26** | 2.46*** | 0.511 | 3.55*** |
| | (0.538) | (0.905) | (0.623) | (1.55) | (0.732) |
| Auto-Auto | 3.62*** | 4.49*** | 3.41*** | 0.345 | 4.92*** |
| | (0.542) | (0.963) | (0.628) | (1.57) | (0.736) |
| HRF-Auto | -1.72*** | -1.62*** | -1.79*** | -1.51*** | -2.43*** |
| | (0.212) | (0.462) | (0.217) | (0.431) | (0.288) |
| Demo-HRF | 0.851*** | 0.577 | 0.859*** | -0.662 | 1.17*** |
| | (0.214) | (0.418) | (0.226) | (0.423) | (0.302) |
| Demo-Auto | 1.84*** | 1.67* | 1.82*** | -0.067 | 2.79*** |
| | (0.537) | (0.931) | (0.623) | (1.56) | (0.730) |
| Demo-DB | 3.03*** | 2.37*** | 3.06*** | 1.55 | 4.36*** |
| | (0.538) | (0.894) | (0.623) | (1.55) | (0.731) |
| Demo-Demo | 3.23*** | 2.55*** | 3.30*** | 1.27 | 4.73*** |
| | (0.537) | (0.919) | (0.622) | (1.56) | (0.729) |
| Both rec same region | 1.43*** | 1.38*** | 1.45*** | 1.64*** | 1.58*** |
| | (0.078) | (0.111) | (0.083) | (0.203) | (0.093) |
| US-agree diff between recs | -0.458** | -1.38*** | -0.955*** | -0.202 | -0.397* |
| | (0.183) | (0.406) | (0.219) | (0.995) | (0.236) |
| SuR HR score | -0.019*** | 0.001 | -0.015*** | 1.17 | -0.044*** |
| | (0.002) | (0.024) | (0.002) | (17,853.5) | (0.003) |
| SuR UNHRC member | -0.084*** | -1.01*** | -0.156*** | | -0.229*** |
| | (0.022) | (0.214) | (0.027) | | (0.031) |
| Either rec under review | -0.125*** | -0.099** | -0.156*** | -0.035 | -0.047** |
| | (0.017) | (0.041) | (0.019) | (0.093) | (0.023) |
| *Fixed-effects* | | | | | |
| Year | Yes | Yes | Yes | Yes | Yes |
| SuR | Yes | Yes | Yes | Yes | Yes |
| Rec 1 | Yes | Yes | Yes | Yes | Yes |
| Rec 2 | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 1,008,665 | 169,213 | 839,452 | 26,783 | 478,462 |
| $R_2$ | 0.06 | 0.08 | 0.07 | 0.09 | 0.08 |
| Mean of DV | 5.94 | 5.86 | 5.95 | 6.19 | 5.70 |
| SD of DV | 6.93 | 7.00 | 6.91 | 6.82 | 6.75 |

*Note: Clustered standard-errors by dyad in parentheses.* *p<0.05;**p<0.01;***p<0.001.
*DB = democratic backslider; HRF = human rights founder; Demo = democracy; Auto = autocracy*

**Table A8.** Jaccard Similarity of UPR Reports, 2008–2022.

| Dependent variable | *Jaccard similarity* | | | | |
|---|---|---|---|---|---|
| | All | 2008-2012 | 2013-2022 | HRF | Autocracies |
| Model: | (1) | (2) | (3) | (4) | (5) |
| *Variables* | | | | | |
| DB-DB | 2.43*** | 1.62* | 2.41*** | 1.18 | 3.41*** |
| | (0.507) | (0.717) | (0.611) | (1.16) | (0.686) |
| DB-HRF | -0.093 | -0.215 | -0.240 | -0.636 | -0.107 |
| | (0.325) | (0.468) | (0.336) | (0.519) | (0.443) |
| DB-Auto | 1.62** | 1.89** | 1.51* | -0.298 | 2.50*** |
| | (0.501) | (0.724) | (0.606) | (1.16) | (0.680) |
| Auto-Auto | 2.17*** | 3.80*** | 1.83** | -1.31 | 3.25*** |
| | (0.506) | (0.796) | (0.611) | (1.20) | (0.685) |
| HRF-Auto | -1.72*** | -1.09** | -1.89*** | -1.49*** | -2.34*** |
| | (0.200) | (0.380) | (0.208) | (0.333) | (0.277) |
| Demo-HRF | 0.973*** | 0.918** | 0.884*** | 0.097 | 1.34*** |
| | (0.201) | (0.329) | (0.214) | (0.308) | (0.281) |
| Demo-Auto | 1.02* | 1.36 | 0.853 | -0.500 | 1.71* |
| | (0.500) | (0.743) | (0.605) | (1.16) | (0.678) |
| Demo-DB | 2.39*** | 1.87** | 2.34*** | 1.36 | 3.42*** |
| | (0.501) | (0.708) | (0.606) | (1.15) | (0.680) |
| Demo-Demo | 2.74*** | 2.50*** | 2.62*** | 1.50 | 3.94*** |
| | (0.499) | (0.725) | (0.605) | (1.17) | (0.677) |
| Both rec same region | 1.15*** | 0.927*** | 1.20*** | 1.13*** | 1.27*** |
| | (0.062) | (0.086) | (0.068) | (0.147) | (0.072) |
| US-agree diff between recs | -0.315* | -0.847** | -0.875*** | 1.51* | -0.375 |
| | (0.155) | (0.324) | (0.190) | (0.751) | (0.193) |
| SuR HR score | -0.004** | 0.015 | 0.003 | -1.03 | -0.015*** |
| | (0.001) | (0.019) | (0.002) | (14,015.2) | (0.002) |
| SuR UNHRC member | -0.081*** | -0.944*** | -0.137*** | | -0.166*** |
| | (0.017) | (0.155) | (0.021) | | (0.024) |
| Either rec under review | -0.147*** | -0.176*** | -0.148*** | -0.348*** | -0.077*** |
| | (0.015) | (0.032) | (0.016) | (0.073) | (0.019) |
| *Fixed-effects* | | | | | |
| Year | Yes | Yes | Yes | Yes | Yes |
| SuR | Yes | Yes | Yes | Yes | Yes |
| Rec 1 | Yes | Yes | Yes | Yes | Yes |
| Rec 2 | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 1,008,665 | 169,213 | 839,452 | 26,783 | 490,067 |
| $R_2$ | 0.08 | 0.11 | 0.08 | 0.14 | 0.09 |
| Mean of DV | 7.11 | 6.56 | 7.20 | 7.05 | 7.00 |
| SD of DV | 5.68 | 5.44 | 5.72 | 5.57 | 5.58 |

*Note: Clustered standard-errors by dyad in parentheses.* *p<0.05; **p<0.01; ***p<0.001.
*DB = democratic backslider; HRF = human rights founder; Demo = democracy; Auto = autocracy.*