Coordinated Contestation: The Shifting Allegiances of Backsliding States in the UN Human Rights Institutions

Anna M. Meyerrose*and Irfan Nooruddin†

Abstract

In recent years, consolidated autocracies —led by China and Russia— have been actively challenging the liberal international order (LIO) and the advanced western democracies at its helm. While noteworthy, this autocratic push-back would perhaps be of less concern for the LIO if advanced democracies could reliably mobilize the broader coalition of democratic states they constructed in the initial years of the post-Cold War era. However, over the last decade, the West has also been forced to confront another international shift: the consequences of an ongoing global democratic recession in which previously democratic states are advocating increasingly illiberal norms and preferences in ways that echo their consolidated autocratic counterparts. Building on previous work showing that backsliding states contest Western liberal values, in this paper, we explore the consequences of this democratic backsliding for the West's broader international democratic coalition, asking: have backsliding states shifted their allegiances toward autocracies within western liberal international organizations (IOs). If so, is there evidence that this behavior is the result of coordination, rather than independent decisions made by self-interested states? Combining data from the UN Human Rights Council (UNHRC) and the Universal Periodic Review (UPR), we argue and show that backsliding states are more likely to vote with consolidated autocracies, especially when this bloc opposes the position taken by advanced western democracies. We also estimate the semantic similarity of reports written by pairs of states in the UPR to show that reviews written by backsliding states, developing democracies, and autocratic states have converged over time, while simultaneously becoming significantly different from reports written by pairs of advanced democracies. Taken together, this evidence suggests that backsliding states, in ways comparable to consolidated autocracies, are coordinating their behavior at the international level in efforts to undermine international human rights norms. Western democratic states are consequently increasingly isolated in these fora, with concerning implications for the future of the LIO.

*Assistant Professor, School of Politics and Global Studies, Arizona State University.

†Hamad bin Khalifa Al-Thani Professor of Indian Politics, School of Foreign Service and Department of Government, Georgetown University.

1 Introduction

Advanced democracies are increasingly isolated within liberal international organizations (IOs). After the end of the Cold War, when democracy was ascendant, advanced democracies, led by the United States and its allies, built a powerful coalition of democratic states that sought to institutionalize Western liberal values at the international level. In recent years, however, consolidated autocracies —this time led by China and Russia— have become more active in their efforts to challenge the liberal international order (LIO) established by the West. While more active challenges by autocracies are noteworthy unto themselves, the threat they pose would be less existential for Western hegemony if advanced democracies could reliably mobilize their broader coalition of democratic states to act as a counterweight against growing autocratic push-back. We argue here that this is no longer the case: over the last decade, the West has had to confront not only the rise of powerful autocracies, but also the consequences of an ongoing global democratic recession.

Democratic backsliding involves the erosion of liberal democratic institutions by elected officials. By definition, therefore, democratic backsliding is a process that begins *in democracies*, and so its consequences for global governance are asymmetric. Stated differently, a key international effect of democratic backsliding is the erosion of the broad democratic coalition that the West relied on in the post-Cold War era to create and maintain the LIO since the backsliding members of that coalition are no longer liberal democracies or committed to democratic values. Focusing on dynamics within the United Nations' (UN) human rights institutions, this paper asks two related questions: Has democratic backsliding hurt the efforts of advanced democracies to maintain and advance their liberal agenda internationally? And do these backsliding states continue to support issues of importance to the West, or does their increasingly illiberal regime lead them to become more sympathetic to those consolidated autocracies that are championing an alternative to the LIO *status quo*?

The UN human rights institutions have been foundational to the LIO since the years shortly following the end of World War 2. Furthermore, recent research shows that back-

sliding states display distinctive behavioral patterns within these institutions: they use their membership to challenge advanced democracies while also advocating a substantively different set of human rights values — namely, social and economic, rather than civil and political, rights (Meyerrose and Nooruddin, 2023). These patterns suggest the interests of backsliding states are shifting away from the West and increasingly towards those typically espoused by consolidated autocracies. A critical question, however, is whether these behavioral changes accumulate to constitute a true shift away from positions taken by advanced Western democracies, or if instead they are simply fragmented and uncoordinated actions.

Our findings suggest that these behavioral changes are indeed indicative of a true shift away from Western democracies by these backsliding states. First, using an original dataset on voting behavior and resolution sponsorship in the UNHRC between 2006 and 2023, we show that two distinct blocs have formed within this institution: an autocratic one and one made up of advanced democracies. In line with our argument, we find that backsliding states have become significantly more likely to vote with the autocratic bloc over time; this is particularly the case when the autocratic bloc votes in opposition to the advanced democracies. We find further evidence that backsliding states are coordinating with autocracies, and against the leaders of the LIO, with text data from UPR reports. Using cosine similarity scores, we measure how semantically similar any two reports written about the same state during the same session are. Regression analyses show that reports written by pairs of backsliding states are significantly more similar than reports written by two advanced democracies. We also find that backsliding states' reports diverge from those written by advanced democracies when reviewing autocracies; this suggests that backsliding states are not only coordinating with one another, but are also diverging from the West in their evaluations of human rights practices in autocratic states with whom their interests are increasingly aligned. Taken together, these findings suggest that advanced democracies are increasingly isolated at the international level as formerly democratically committed developing states are backsliding, espousing increasingly illiberal values, and, as result, more actively allying

with consolidated autocracies that seek to fundamentally alter the international order.

2 The LIO and the Authoritarian Turn in International Organizations

The liberal international order (LIO) is grounded in ideals first championed by Woodrow Wilson and other Anglo-American liberals. Its origins can be traced to the years after World War 1, but expanded significantly after World War 2 with the founding of prominent multilateral institutions, with the United Nations (UN) system, the Bretton Woods Institutions, and the General Agreement on Tariffs and Trade at its core (Börzel and Zürn, 2021). These pillars of the LIO sought to institutionalize a new era of global governance that protected and furthered Western (especially the United States') interests by promoting international economic exchange, and, to a lesser extent, by encouraging international deliberation and policy coordination through the UN Security Council, General Assembly, the UN human rights institutions, and a proliferation of more technical bodies under the UN umbrella. Building on these foundational institutions, after the Cold War ended, additional Western-led IOs proliferated, with politically oriented ones (such as the UN) in particular becoming inextricably linked to the construction of a liberal global order centered on democracy, elections, rule of law, and civil and political human rights (Barnett and Finnemore, 2021). Throughout the twentieth and into the twenty-first century, this liberal international order was championed by the United States (US) with the help of its (mostly) Western European allies (Ikenberry, 2009).

Today, however, there is growing concern that the US-backed LIO —and, by extension, the liberal international organizations that undergird it— is under real threat for the first time since 1930 (Ikenberry, 2018; Mearsheimer, 2019; Eilstrup-Sangiovanni and Hofmann, 2020; Kornprobst and Paul, 2021). 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). But as the US' relative power and influence has waned (Cooley and Nexon, 2020; Hyde, 2020) in the face of an increasingly powerful and assertive China and other autocratic powers, the post-Cold War LIO has been more forcefully contested (Weiss and Wallace, 2021; Boyle, 2023; Cottiero et al., 2024). This challenge against the US-backed order has come, on the one hand, from the emergence of IOs composed primarily if not entirely of autocratic countries (as a shorthand, we refer to these as 'autocratic IOs') (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, 2021), reshape international legal standards (Ginsburg, 2020*b*), and justify their rule to both international and domestic audiences (Debre and Morgenbesser, 2018; Bush, Cottiero and Prather, 2024; Hafner-Burton, Pevehouse and Schneider, 2024; Morrison et al., 2024). Importantly, these illiberal regimes' influence is not limited to autocratic IOs. Rather, as democratic backsliding has become increasingly common (Waldner and Lust, 2018; Ziblatt and Levitsky, 2018; Meyerrose, 2020; Haggard and Kaufman, 2021), smaller illiberal states are joining powerful autocrats to unravel long-established consensus within western-backed IOs (Kelemen, 2020, 2024). Mounting evidence suggests these states are also more assertively using their memberships in these organizations to advance their own illiberal interests (Baturo, 2023; Winzen, 2023; Meyerrose and Nooruddin, 2023; Lipps and Jacob, 2024).

3 Illiberal Regimes in the UN Human Rights Institutions

While not without its limitations, the UN Human Rights Council (UNHRC) (and its predecessor, the UNCHR) is the epitome of the ambition of the LIO currently in crisis. The UNHRC is but one of a generation of UN bodies that form a critical part of the LIO firmament, but by espousing a commitment to global enforcement of human rights standards championed —if not always practiced— by the West, it arguably was the most intrusive

institution of the LIO. Admittedly, the UNCHR lacked real teeth, though its condemnations were not without some material consequence (Lebovic and Voeten, 2006). But the real threat these institutions posed to human rights violators, especially autocratic states outside the orbit of protective Western patronage, was through naming and shaming (Finnemore and Sikkink, 1998; Hafner-Burton, 2005; Kelley and Simmons, 2015; Terman and Voeten, 2018). States targeted by the UNHRC/UNCHR perhaps worried less about material international consequences than 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. Undermining the normative legitimacy and procedural efficacy of the UNHRC was in the collective interests of any states that feared someday being in its cross-hairs.

The most cohesive and consistent bloc of supporters of the UNHRC's mandate has been EU members and the United States (Burmester and Jankowski, 2014; Hug and Lukács, 2014; Burmester and Jankowski, 2018). Extensive evidence shows that the US and other advanced industrial democracies have historically been able to influence voting in multilateral institutions, especially the UN (Dreher, Nunnenkamp and Thiele, 2008; Vreeland and Dreher, 2014; Vreeland, 2019). Consistent with this work, and as we show in Figure 2 below, advanced democracies in the UNHRC vote with each other more than 90% of the time.

Against these Western democracies is arrayed a set of consolidated autocratic regimes, elected to the UNHRC from their world regions. Recent research has shown 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 like-minded states to challenge and constrain the UN human rights system (Inboden, 2021; Dukalskis, 2023). These consolidated autocratic states 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 (this is similar to Hug and Lukács (2014)).

Between these two camps sit countries we label "developing democracies." These are principally Global South states that range from stable democracies to those with more fragile and emerging democratic institutions.[1] Historically, the ability of the LIO's champions to exert influence through the UNHRC rests on their internal cohesion and their ability to attract support from Global South democracies, especially those more closely aligned with Western preferences (Voeten, 2000; Lebovic and Voeten, 2006).

During the last decade, this "coalition of the willing" has grown more rickety and began to splinter. We argue that a primary cause is widespread democratic backsliding —a process that occurs when democratically elected officials weaken or erode institutional checks on government power (Bermeo, 2016), including the constitution, rule of law, civil and minority rights, the independence of the judiciary and the media, and separation of power within governments (Haggard and Kaufman, 2021; Meyerrose, 2025). Democratic backsliding, we argue, has two primary consequences for the UNHRC: (a) LIO principles have lost legitimacy and power in backsliding states as populist leaders reject their universal pretensions by appealing to principles of state sovereignty and non-interference in domestic politics, and (b) these same leaders seek to protect themselves against intrusive international scrutiny that might embolden domestic opposition and threaten their grip on power.[2] China's growing assertiveness internationally at the same time as democratic backsliding has become more common is a perfect storm for the LIO.

Our argument is consistent with recent evidence from Meyerrose and Nooruddin (2023), who show that backsliding states in the UNHRC are more likely to vote against or abstain from targeted resolutions, and are more likely to stress economic and cultural rights rather than the political and civil liberties that are central to the LIO. A related argument is made by Prasad and Nooruddin (2024), who document similar voting patterns among states

[1] More on how we operationalize these categories in Section 5 below.

[2] This argument complements concerns about China's growing efforts to solidify its own voting bloc in the UN and other multilateral institutions (Inboden, 2021; Dukalskis, 2023).

dealing with domestic insurgencies. Their argument is that such states often use extra-judicial tactics to quash insurgents and therefore they seek to defang the UNHRC before it turns the spotlight on them. Earlier research similarly shows that polarization in the UNHRC grows when votes are "controversial," that is, when they are introduced by countries with poor human rights records (Hug and Lukács, 2014). These efforts by such states are also evidence of strategic action to divide any pro-LIO voting coalition and thereby to undermine the normative power of UNHRC deliberations.

We advance this prior scholarship by arguing that democratic backsliding internationally poses a grave —even existential— threat to the UN human rights institutions. A persistent question in this literature is whether voting patterns in the UNHRC reflect individual state preferences or coordinated blocs. While observationally equivalent on the surface, the theoretical and policy implications of the answer to this question matter. We argue that coordinated bloc voting is increasingly prevalent in the UNHRC, a phenomenon we label "coordinated contestation." Here, the bloc of advanced democracies —comprised primarily of the US, the EU, and their closest allies— is increasingly isolated in the UN human rights institutions against a bloc of consolidated autocracies who use their membership in the Council 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 LIO is the global phenomenon of democratic backsliding, which has splintered the always fragile pro-LIO coalition. Instead, like an ice floe detaching itself and floating away, this set of backsliding states has moved away from the Western pole and has 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: *Backsliding states increasingly vote with the autocratic bloc, and against the bloc comprised of advanced Western democracies, in the UNHRC.*

Our framework suggests two other hypotheses in the context of the UNHRC. First, we

recognize that not all votes in the UNHRC are created equal. A great many resolutions are general in nature and/or not especially controversial and so are passed by broad consensus. To uncover underlying voting blocs requires identifying more controversial votes that necessitate the explicit choosing of sides. This is the intuition behind the use of so-called "party unity" votes analyzed by scholars of the U.S. Congress (see Desposato (2005) for a discussion). We translate that intuition to our context by identifying resolutions on which a majority of consolidated autocratic states voted against a super-majority (75%) of Western democratic states. These resolutions have the virtue of unambiguously pitting the two ideological camps against each other, thereby forcing other UNHRC members to clarify their positions. Further, by limiting our attention to those resolutions backed by a super-majority of mature democracies, 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 backsliding states are moving to 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. Failure to uncover such movement for this set of resolutions would require us to revisit the utility of our argument.

Hypothesis 2: *Backsliding states will vote with the autocratic bloc on contentious UNHRC resolutions that pit consolidated autocracies against advanced democracies.*

The era of backsliding is widely dated as reaching its acme by 2013. We argue that the dynamics of coordinated contestation expected by our framework should be particularly visible from 2013 onward, a time period in which backsliding became largely complete and consolidated in many erstwhile democracies, and, when Western champions of democratic values were distracted by their own political crises at home. While the earlier decades of the UNHRC (and its predecessor) might have been broadly characterized as experiencing competition between democratic and autocratic blocs with developing democracies lying between these two poles, more recent experience suggests we need to consider backsliding states as a potentially distinct "camp" unto themselves, separate from the developing democracies of

which they were once a part and yet also different from the consolidated autocracies towards which they have drifted. We therefore expect that the re-alignment of backsliding states with autocratic states in the UNHRC to be more pronounced after 2013.

Hypothesis 3: *Backsliding states' shift toward autocracies in the UNHRC will be particularly relevant after 2013, which marks the height of the global democratic recession.*

This alignment of backsliding states with the autocratic bloc should be evident not only in their voting behavior, but also in the rhetoric they use. Meyerrose and Nooruddin (2023) analyze text data from the Universal Periodic Review (UPR) reports submitted by states to show that backsliding states are more likely to raise alternative conceptions of human rights that emphasize economic, social, and cultural rights, rather than the political and civil liberties prioritized by the LIO's champions. We build on that insight to pursue a different tack: since the UPR process involves (a series of) pairs of states serving as reviewers for a given state, we assess the degree of similarity in the language used by reviewers as a function of their shared regime status.[3] We expect like pairs of states to be most similar in their language, but autocratic pairs should be distinguishable from democratic pairs if in fact the language used in reviews reflects underlying values. Mixed regime-pairs should be more heterogeneous: for instance, a democratic-autocratic pair should be less similar than, say, a democratic pair. But a backslider-autocrat pair should, if our argument is correct, be more similar than a democratic pair. Such patterns of language similarity would be evidence in favor of greater collusion by backsliding states with autocratic partners in the UPR, and coupled with the voting analysis would indicate that backsliding states might be exhibiting patterns of coordinated behavior in that forum.

Hypothesis 4: *The language backsliding states use in their UPR reports will show evidence of coordination.*

[3]Others, including Kim (2023) and Lu (2024) have similarly measured language similarity in the context of the UPR.

Finally, similar to the UNHRC, we expect these dynamics to be particularly salient after 2013, and test this intuition in the UPR context as well.

Hypothesis 5: *Coordination —as measured by semantic similarity of reports— between backsliding states in the context of the UPR should increase over time, and particularly after 2013.*

In the following sections, we describe our data collection strategy and the research design we employ to analyze these data.

4 The UN Human Rights Institutions

Today's UN human rights system consists of two distinct institutions: the UNHRC and the UPR. The UN Human Rights Council (UNHRC) was founded in 2006 after its predecessor —the UNCHR— 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 for or against the resolution; alternatively, members can abstain or fail to participate entirely. The UNHRC also votes on resolutions in support of 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 passed by mere consensus (or not).

We scrape information on these recorded votes directly from the UNHRC online library for all resolutions from 2006 through 2023.[4] The resulting dataset contains information on how each member of the UNHRC voted on any given targeted resolution,[5] the state that is

[4] https://searchlibrary.ohchr.org/search?ln=en&cc=Voting. Last accessed 30-July-2024.

[5] The possible outcomes are "yes," "no," "abstain," or failure to participate altogether.

the subject of the resolution, if any, as well as details about which state(s) sponsored the resolution. Therefore, our dataset contains observations at the resolution-UNHRC member state level of analysis 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 and 2023.[6]

In 2008 the UN introduced the Universal Periodic Review (UPR) as an additional mechanism for monitoring the human rights practices of its member states. Under the UPR, *all* UN member states, rather than simply those identified by members of the UNHRC, are subject to periodic reviews of their domestic human rights practices. Furthermore, all states, not just members elected to the UNHRC, have the opportunity to comment on the human rights situation in their fellow member states, and to recommend changes and/or improvements (McMahon and Ascherio, 2012).

All UN member states are reviewed every four and a half years under the UPR process; these review sessions take place three times per year in Geneva. In these sessions, the state under review's human rights record is "peer reviewed" through an interactive dialogue between the state under review and all other UN member states. This process begins with the state under review presenting a self-assessment of its domestic human rights practices. All other states then have the opportunity to respond and evaluate the state. These exchanges are recorded in an outcome report (Cox, 2010; Terman and Byun, 2022). Data on the content of these outcome reports are publicly available online.[7] We use these to create a dataset that consists of 90,938 individual recommendations made for all UN member states via the UPR mechanism from 2008 through 2020. This dataset also includes information on the recommending state.

[6]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).

[7]https://upr-info-database.uwazi.io/en/. Last accessed 9-March-2023.

5 Voting Blocs in the UNHRC

To test Hypotheses 1, 2, and 3, we explore the dynamics of voting blocs in the UNHRC over time. Beyond ensuring representation across geographic regions, there are no limits on who can be a member of the UNHRC. As such, both democracies and autocracies can be UNHRC members. Furthermore, over time, as democratic backsliding has progressed and become more common, the number of backsliding states represented on the Council has also increased. With whom do these backsliding states vote on (recorded) UNHRC resolutions? Have they emerged as a distinct bloc?

To answer these questions, we first identify the universe of backsliding states that have been represented in the UNHRC since its inception in 2006. To do this, we build on the approach used by the Varieties of Democracy (V-Dem) team. V-Dem's Episodes of Regime Transformation (ERT) dataset (Maerz et al., 2021) identifies episodes of autocratization, defined as moves away from democracy —or, toward autocracy— in any type of state. Our analytical focus —democratic backsliding— constitutes one type of autocratization: backsliding is a process of autocratization that *begins in democracies*. Therefore, we adopt Maerz et al. (2021)'s approach to identifying periods of backsliding, with three important differences. First, we focus exclusively on states that begin their autocratization episode as at least minimally democratic states, *i.e.*, ones with an 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 concentrate instead on the liberal democracy index. Democratic backsliding occurs when elected officials erode or undermine often liberal institutions (Bermeo, 2016), such as the constitution, rule of law, civil and minority rights, the independence of the judiciary and the media, and separation of power within governments (Haggard and Kaufman, 2021; Meyerrose, 2024). The liberal democracy index importantly includes these other institutions that are often targeted in cases of backsliding, while the electoral democracy index focuses primarily on election and participation-based indicators of democracy. Finally, the V-Dem dataset codes an autocratization episode as ending when

a state's democracy level stabilizes, or, ceases to register substantial negative changes. However, from a theoretical perspective, we argue that states that that have backslid have a unique set of preferences and interests that distinguish them from other democracies (or semi-autocracies) that have not backslid. As such, we argue stabilization alone is insufficient to categorize them as "non-backsliding." Rather, we argue that unless or until a state returns to its previous level of (liberal) democracy, it should still be considered a case of backsliding, and we code it as such.[8]

Accordingly, in our dataset a backsliding episode 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 over the entire backsliding period. The episode continues if, according to the liberal democracy index, there is:

- An annual (negative) change in 1 of every 5 consecutive years;
- There is no reverse (positive) annual change of 0.03 or more;
- And there is no cumulative (positive) reversal of 0.1 over a 5-year period.

Once this episode ends, a state continues to be considered a backslider unless and until it returns to its pre-episode liberal democracy level. Based on these criteria, we identify all episodes of democratic backsliding, with asterisks indicating backsliding states that were members of the UNHRC at least once between 2006 and 2023:[9] We use these episodes to identify all backsliding states represented in the UNHRC each year since 2006.[10] As

[8]Specifically, in our approach, a state continues to be coded as a backsliding state after its regressions have stopped unless or until that state returns to its pre-backsliding levels of liberal democracy. For example, if a state begins its backsliding episode with a liberal democracy score of 0.7, and the substantial regressions end in 2015, but its liberal democracy score does not return to 0.7 until 2020, then we code it as "backsliding" until 2020. This approach adds several cases (Belarus, Ukraine, and Venezuela) that otherwise would have been coded as autocracies, in the case of Belarus and Venezuela, or fluctuating between a democracy and autocracy, in the case of Ukraine, to our set of backsliding states. We note that all results reported throughout the paper are substantively similar if we instead use the V-Dem approach of only coding country-years within an autocratization episode, per the V-Dem coding, as cases of backsliding.

[9]This approach also codes the United States from 2016 through 2023 as a case of democratic backsliding. However, since our focus in this paper is on the extent to which backsliding states coordinate their behavior in opposition to the leaders of the liberal international order —including, perhaps most prominently, the United States— we exclude the US from our list of backsliding states.

[10]Here and in our analysis of UPR reports below, we code a state as backsliding only for the time period

Figure 1 shows, the percent of UNHRC member states that are backsliding states has steadily increased over time.

- Argentina, 2002-2023*
- Armenia, 1994-2018 & 2020-2023*
- Belarus, 1994-2023
- Benin, 2018-2023*
- Bolivia, 2005-2023*
- Botswana, 2015-2023*
- Brazil, 2015-2023*
- Bulgaria, 2001-2023*
- Burkina Faso, 2018-2023*
- Croatia, 2013-2023*
- Cyprus, 2017-2023
- Czech Republic, 2010-2023*
- Ecuador, 2007-2018*
- El Salvador, 2018-2023
- Ghana, 2013-2023*
- Greece, 2013-2023
- Guatemala, 2018-2023
- Guyana, 2019-2023
- Honduras, 2006-2022*
- Hungary, 2006-2023*
- India, 2009-2023*
- Indonesia, 2009-2023*
- Ivory Coast, 2017-2023*
- Lesotho, 2015-2019
- Madagascar, 1995-2023*
- Maldives, 2012-2019*
- Mali, 2017-2023
- Mauritius, 2013-2023
- Mexico, 2019-2023*
- Moldova, 2000-2010 & 2013-2021*
- Mongolia, 2008-2023*
- Nicaragua, 2006-2023*
- Niger, 2016-2023
- North Macedonia, 2007-2023*
- Palestine, 2006-2023
- Peru, 2016-2023*
- Philippines, 2016-2023*
- Poland, 2015-2023*
- Romania, 2017-2019 & 2021-2023*
- Senegal, 2017-2023*
- Serbia, 2010-2023
- Slovenia, 2012-2023*
- South Korea, 2008-2017 & 2019-2023*
- Sri Lanka, 1977-2016 & 2019-2022*
- Thailand, 2005-2011*
- Tunisia, 2015-2023*
- Turkey, 2005-2023
- Ukraine, 1995-2006*
- Venezuela, 1994-2023*

Theoretically, we are interested in whether backsliding states are shifting their allegiances to autocracies and away from advanced democracies. To begin, we examine the levels of voting cohesion among four distinct and mutually exclusive groups of states within the UNHRC: consolidated autocracies, backsliding states, developing democracies[11], and advanced democ-

corresponding to its backsliding episode. However, we are unsure if this is the best approach, and wonder if we should instead code a case as backsliding even after the episode ends, unless or until that state returns to its pre-backsliding level of (liberal) democracy. For example, if a state begins its backsliding episode with a liberal democracy score of 0.7, and the episode ends in 2015, but its score stays below 0.7 in subsequent years, should those additional years also be coded as cases of backsliding? We would value any feedback on this point for future iterations of the paper.

[11]This group includes all democracies that are neither backsliding nor one of the 19 advanced democracies

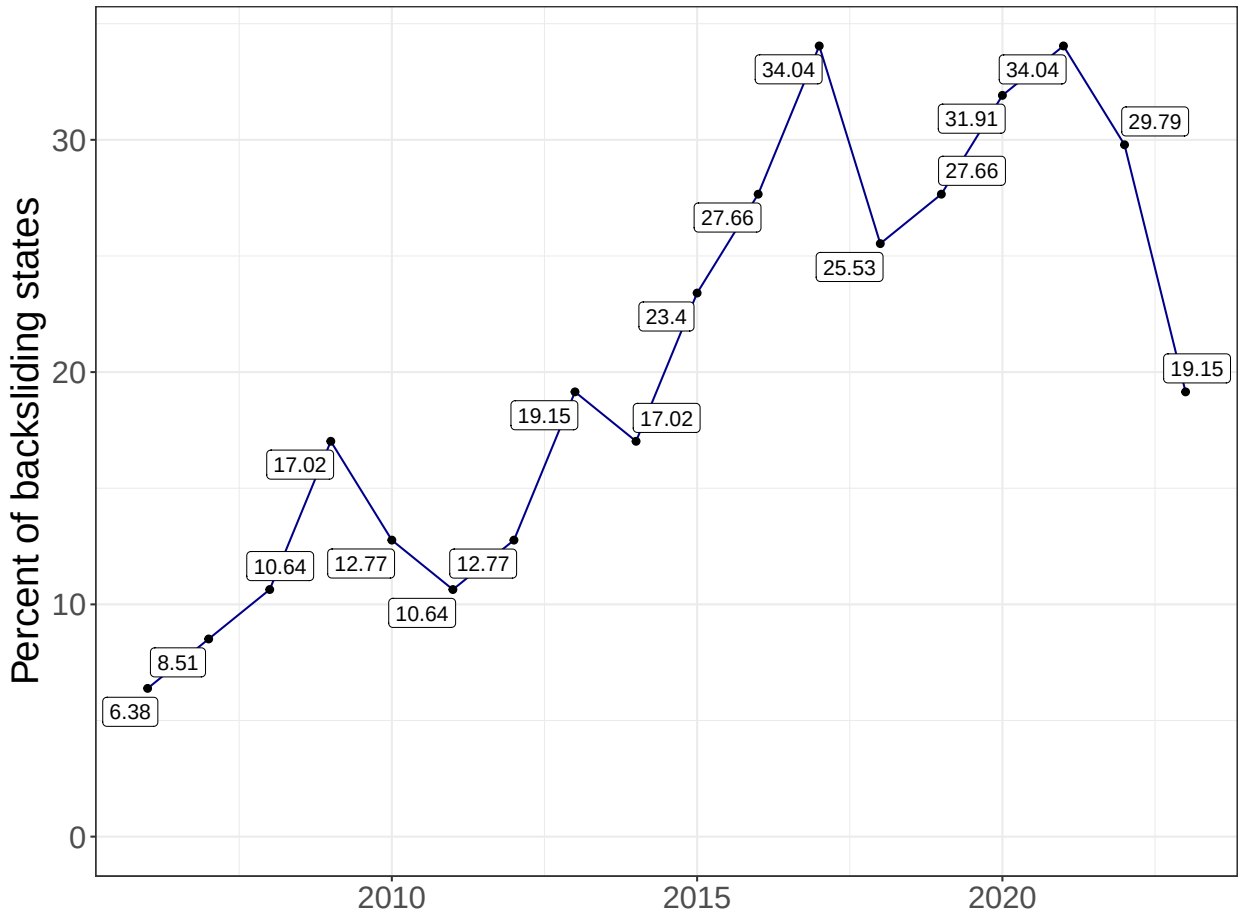


Figure 1: The percent of UNHRC member states that are backsliding has increased over time.

racies.[12] To measure voting cohesion, we follow Burmester and Jankowski (2018) (see also Desposato (2005)). 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.[13] Then, for each individual state in that group, we code whether they voted with or against the average decision of their relevant regime-type group. We then calculate the mean number of times that all members of a group voted with their group each year to create Figure 2.

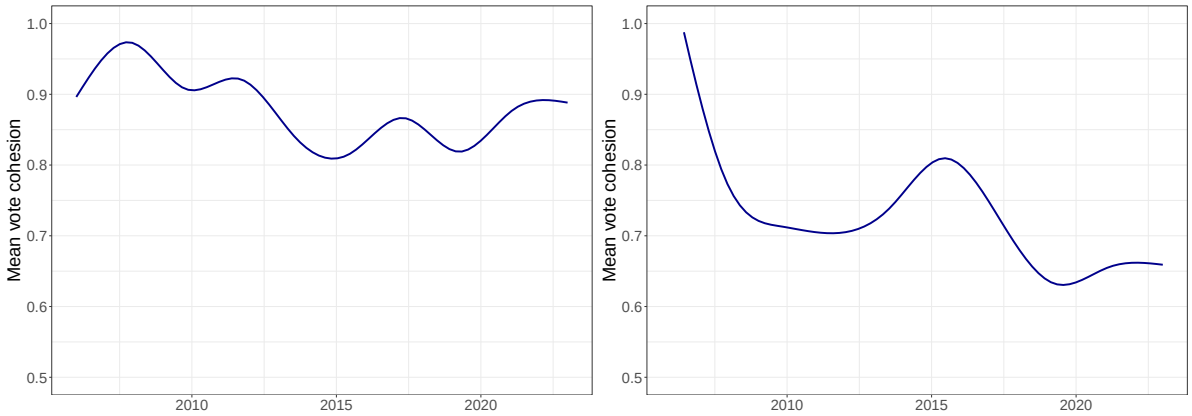
As Figure 2 shows, advanced democracies —the main architects and supporters of the LIO— consistently vote together at high rates. But, to only a slightly lesser degree, so do consolidated autocracies. As such, they represent two poles within the UNHRC. Backsliding states and developing democracies, on the other hand, show far less cohesion in their voting decisions. Indeed, although backsliding states have gained greater representation in the UNHRC over time, as illustrated in Figure 1, they have actually become *less* cohesive during this same period. This is likely because backsliding states, like developing democracies, are a significantly more heterogeneous group than either consolidated autocracies or advanced democracies.

The variation described in Figure 2 raises an important question: are these backsliding states more likely to gravitate toward advanced democracies, or are these increasingly illiberal regimes instead moving closer to the consolidated autocracy bloc? The implications for the

listed below. In other words, this group largely includes third-wave democracies that have not succumbed to backsliding.

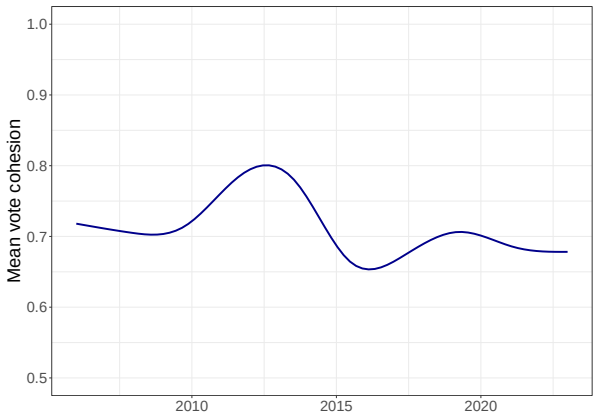
[12] Flores and Nooruddin (2016) identify non-consolidated or advanced democracies as states "for whom a democratic system was not a certainty in 1946 or in the year of its birth as a sovereign country, whichever came second." This excludes —or in other words, designates as advanced, consolidated democracies— the following 19 countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States.

[13] 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 analysis below. See Meyerrose and Nooruddin (2023) and Prasad and Nooruddin (2024) for a similar treatment of abstentions. 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, here and below we exclude observations where a UNHRC member did not participate in a vote; these absences constitute only 0.006% of the observations in our dataset, and results are the same if we include these observations in the "opposition" group of votes.

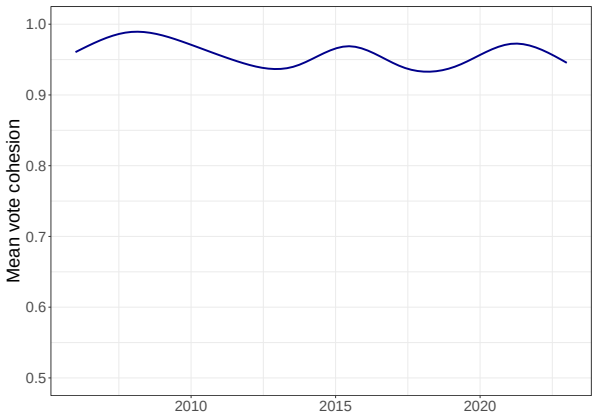


(a) Autocracies

(b) Backsliders



(c) Developing democracies



(d) Advanced democracies

Figure 2: Voting Cohesion in the UNHRC

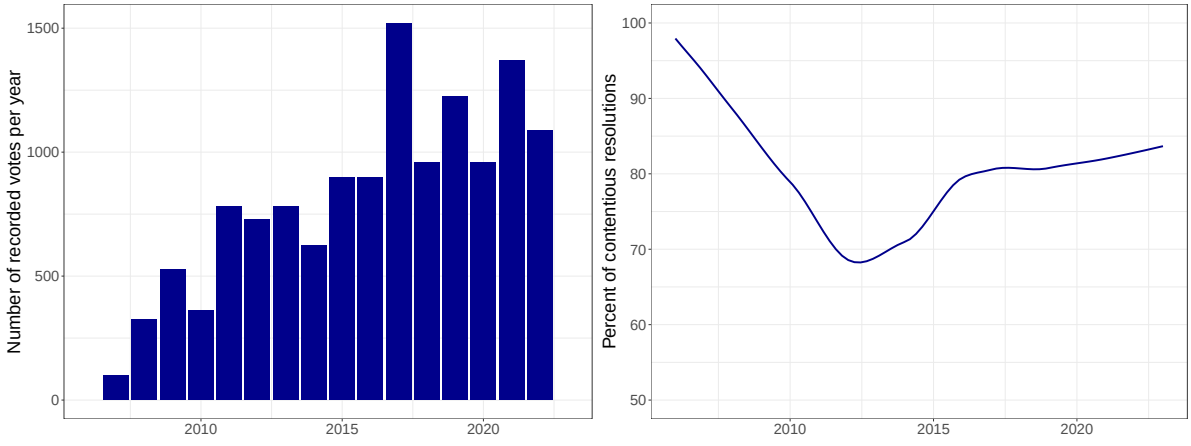
LIO's longer-term viability cannot be overstated since the first possibility maintains the pro-LIO coalition led by advanced democracies while the second splinters the normative consensus around the LIO's legitimacy from within a centerpiece of its ambition: the UN human rights regime. Descriptively, the data suggest that backsliding states align themselves more closely with autocracies than they do with advanced democracies, especially when it comes to contentious votes.[14] As Table 1 shows, backsliding states are more likely to vote with, rather than against, the autocratic majority bloc on all recorded votes. On average, backsliding states vote with the autocratic bloc more than half (roughly 65%) of the time. In the case of non-contentious votes, backsliding states vote with the autocratic bloc 86% of the time, while for contentious votes they agree with autocrats 57% of the time. This is a nuanced finding, indicating that backsliding states are closer to autocracies in their voting preferences, but that, on contentious votes, at least some backsliding states still vote with the advanced democracies. A possible explanation, which we test below, is that this variation is a function of the backsliding state's liberal democracy score since some backsliding states still remain quite democratic, albeit less so than they once were.

A second descriptive observation is also worth noting explicitly. Mirroring the trends in Figure 1, Panel B of Figure 3 shows that as the percent of UNHRC states that are backsliding has increased, so too has the percent of contentious votes on UNHRC resolutions; this increase in contentious votes has coincided with a rise in the number of recorded votes on resolutions in the UNHRC, as shown in Panel A of Figure 3. This is important because it makes patently clear that the UNHRC is increasingly a site of contestation: where once votes passed by consensus or acclimation, a much greater number of votes in recent years are subject to recorded votes that split between ideological poles. This is further evidence that the LIO is under strain, and bolsters our contention that democratic backsliding is worsening the stress on these global governance institutions.

[14]We define contentious votes as ones where 75% or more of advanced democracies voted in one direction, while the majority of consolidated autocracies voted in the opposite direction on the same resolution. For example, if 78% of advanced democracies vote in favor of a resolution, and the majority from consolidated autocracies is "no" or "abstain," then it is a contentious resolution.

	Non-contentious	Contentious
Votes with autocracy bloc		
No	152	1765
Yes	918	2385

Table 1: On average, backsliding states vote with the autocratic bloc on the majority of recorded UNHRC resolutions.



(a) Recorded votes

(b) Contentious votes

Figure 3: As the number of recorded UNHRC has increased over time, so too has the percent of contentious UNHRC resolutions.

We use regression analysis to 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. Following the approach described above, we identify 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 backsliding state, a consolidated autocracy, or an advanced democracy, with developing democracies as the omitted reference category.[15]

We control for a range of factors that should impact how any given state votes on a resolution. First, we control for the voting state's democracy score (Coppedge et al., 2021), with the expectation that higher levels of democracy will make states more likely to vote against the autocratic bloc. Doing so also helps control for some of the heterogeneity within our regime categories since not all backsliding states, for instance, start or end at the same point on the liberal democracy index. Second, voting patterns against Israel in the UNHRC are distinct (Seligman, 2011). Indeed, as noted above, Israel is the most frequent target of UNHRC resolutions in our data. 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) during this time period. Next, we control for the human rights scores of both the voting and target states (Herre and Roser, 2016) and for whether the voting state and the target state are from the same geographic region. Lastly, we calculate the percent of sponsors

[15]For transparency, we summarize the coding rules for each of these categories here: (1) Consolidated autocracies are states that were below the 0.5 threshold of VDem's liberal democracy score; (2) Advanced democracies are states that were consolidated as such in 1945 or the year of their birth; (3) Backsliding states are states that begin an autocratization period with a score above 0.5 on the liberal democracy index but have a cumulative decline of at least 0.1 over the period and meet the other criteria spelled out in section 5 above; and (4) Developing democracies are countries that score above 0.5 on liberal democracy but do not qualify as advanced democracies and have not experienced backsliding. These four groups of states are mutually exclusive.

of each resolution that are democracies (whether developing and advanced) to capture the logic that resolutions sponsored mainly by democracies should generate greater autocratic and backsliding state opposition.

To test our overarching prediction that backsliding states will be more likely to vote with the autocratic bloc, and against advanced democracies (Hypothesis 1), we estimate three models. Since our dependent variable is a binary variable, we estimate logistic regressions with year fixed effects.[16] 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. Our baseline model includes observations for all resolutions for which votes were recorded in the UNHRC between 2006 and 2023. These can be those resolutions targeted at a specific state, or they can be more general resolutions concerning broader human rights issues.[17] Once we begin controlling for target state-level characteristics, however, the subset of observations necessarily is limited to those resolutions targeting a specific country. Therefore, our second model, which includes the full set of controls listed above, focuses exclusively on how states voted relative to the autocratic bloc on *targeted* resolutions. Finally, in our third model we focus exclusively on *contentious targeted* resolutions to test Hypothesis 2. Contentious targeted resolutions are those in which a super-majority (75% or greater) of advanced democracies voted one way, while the autocratic bloc voted in opposition. The results of these models are reported in Table 2.

We find compelling evidence that backsliding states are increasingly aligning with the consolidated autocratic voting bloc of the UNHRC. In all three models in Table 2, compared to developing democracies, backsliding states are significantly more likely to vote with the autocratic bloc on all resolutions, on targeted resolutions, and on contentious resolutions. As expected, consolidated autocracies are similarly particularly likely to vote with their own

[16]Because the number of years that a state serves in the UNHRC can vary significantly from one country to the next, we do not include voting state fixed effects.

[17]For example, the details of one such general resolution from 2018 are the "promotion of a democratic and equitable international order." (Resolution A/HRC/RES/39/4 was passed on 27 September 2018 with 28 votes in favor, 14 against, and 5 abstentions).

Table 2: Voting with the Autocratic Bloc in the UNHRC, 2006–2023

	Dependent variable:		
	Vote with autocracy bloc (1 = yes, 0 = no)		
	All	Targeted	Contentious targeted
	(1)	(2)	(3)
Backsliding state	0.29***	0.19**	0.38***
	(0.04)	(0.08)	(0.08)
Mature democracy	−1.94***	−0.67***	−2.34***
	(0.05)	(0.10)	(0.10)
Consolidated autocracy	1.72***	0.53***	0.60***
	(0.04)	(0.13)	(0.13)
Voting state demo score		−1.37***	−2.24***
		(0.25)	(0.25)
Target state = Israel		1.24***	0.74***
		(0.10)	(0.10)
Target state = Syria		−0.29***	−0.60***
		(0.09)	(0.09)
Voting state HR score		−0.39***	−0.76***
		(0.03)	(0.03)
Perc. demo sponsors		−0.002**	−0.005***
		(0.001)	(0.001)
Target state HR score		−0.08	0.28***
		(0.05)	(0.05)
Voting and target in same region		−0.46***	−0.71***
		(0.12)	(0.12)
Year fixed effects	Yes	Yes	Yes
Clustered standard errors	Yes	Yes	Yes
Model type	Logit	Logit	Logit
Observations	22,556	6,996	4,722
Log Likelihood	−11,760.05	−3,496.73	−1,945.97
Akaike Inf. Crit.	23,562.09	7,041.45	3,939.94
Note:			*p<0.1; **p<0.05; ***p<0.01

group, while advanced democracies are significantly less likely to vote in line with autocrats. The results for the control variables also largely align with expectations. First, states with higher democracy and human rights scores are less likely to vote with the autocrats. Inter-

estingly, we also find that resolutions targeting Israel make states significantly more likely to vote with autocrats, whereas those targeting Syria have the opposite effect. This may be indicative of a dynamic of states opposing the US by aligning with the autocratic (and therefore anti-US) position on resolutions concerning Israel. Furthermore, model 3 suggests that, on contentious votes, states are more likely to vote with the autocratic bloc as the target state's human rights score increases.[18]

As Table 3 shows, we also find that the coefficients for all main independent variables across all three models in Table 2 are significantly different from one another, with one exception. There is no significant difference between backsliding states and autocracies in model 3. This suggests that, in the case of particularly contentious votes, backsliding states and autocratic states are equally likely to vote with the autocratic bloc and, therefore, against the position advocated by advanced democracies. This provides further evidence that mature democracies are increasingly isolated within the UNHRC.

	All	Targeted	Contentious targeted
Backsliding state–mature democracy	✓	✓	✓
Backsliding state-consolidated autocracy	✓	✓	✗
Mature democracy-consolidated autocracy	✓	✓	✓

Table 3: 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.

We also predicted in Hypothesis 3 that these dynamics would be more pronounced after 2013, which marks an inflection point when backsliding became particularly widespread and extensive. To test this prediction, we re-estimate the same models from Table 2, disaggregating the observations into pre- and post-2013 samples.[19] The results for the main independent variables are reported in Figure 4.[20] As predicted, we find that backsliding

[18]This may be evidence that autocratic and autocracy-leaning states are using their voice in the UNHRC to target more democratic states that are allies of leaders of the liberal international order.

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

[20]These models include all control variables reported in Table 2, where relevant, but are not included here to simplify the visualization. Full results are reported in Appendix Tables 6 and 7.

states' shifts toward autocratic states is limited to recorded votes from 2013 onward. However, from 2013 onward, backsliding states were significantly more likely to vote with the autocratic bloc on both targeted and contentious targeted resolutions. This is an important finding, confirming that concerns about growing autocratic influence in international organizations over the past few years is not unwarranted (Meyerrose and Nooruddin, 2023; Cottiero et al., 2024), but also providing new evidence that the growing ability of autocratic states to challenge the LIO is because of the global phenomenon of democratic backsliding that has considerably weakened advanced democracies' ability to marshal 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.

In short, data from the UNHRC make clear that there are now two distinct voting blocs within this institution: the advanced democracies and a bloc composed of consolidated autocracies, with backsliding states and developing democracies lying somewhere in the middle. On their own, backsliding states do not constitute a distinct bloc, and their voting patterns have actually become less cohesive over time. This suggests that the support of backsliding states is up for grabs, so to speak. Our findings show that these backsliding states —whose governments are moving away from democracy and are becoming increasingly illiberal— are more likely to support the positions taken by consolidated autocracies, even when this brings them in direct conflict with the preferences of Western democracies. These dynamics have been particularly pronounced since 2013, after which backsliding has become a pervasive global phenomenon.

6 Coordination in the UPR

Not all backsliding 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.

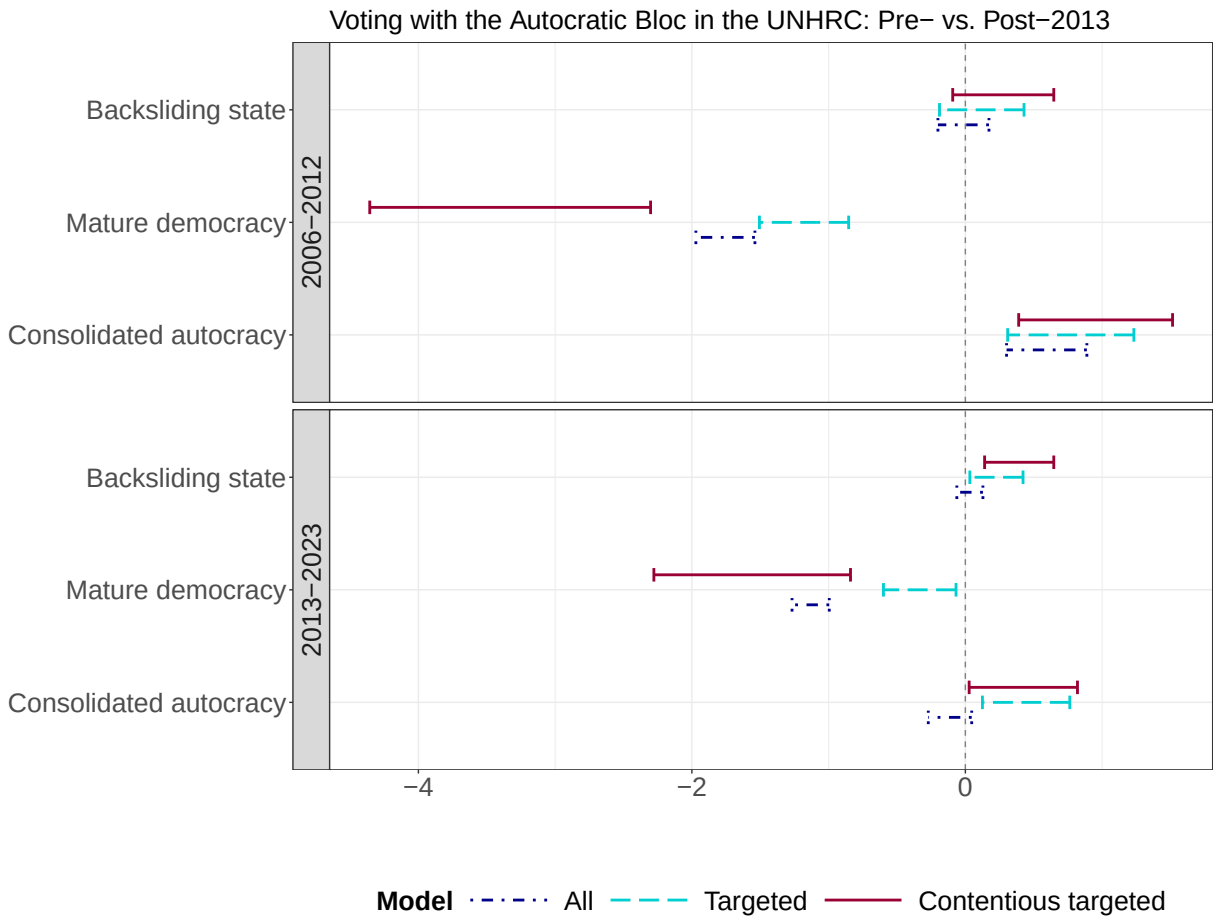


Figure 4: Prior to 2013, there is no evidence that backsliding states were shifting towards autocracies. However, following 2013, backsliding states were significantly more likely to vote the autocratic bloc on both targeted and contentious targeted resolutions. The plot depicts 95 percent confidence intervals.

Therefore, to explore further the extent to which backsliding states align with autocracies and against advanced democracies, and also to test for evidence of coordination among backsliding states, we turn next to data from the Universal Periodic Review (UPR).

Drawing on evidence from the UPR, Terman and Búzás (2021) show there are a distinct set of communities within the international human rights regime, each with their own identifiable normative positions. Backsliding states specifically have been using the UPR mechanism both to challenge advanced democracies by identifying more issues when evaluating the human rights practices of those states, and also by emphasizing social and economic, rather than civil and political, human rights values in the same reports (Meyerrose and Nooruddin, 2023). To what extent is this the result of coordinated behavior among backsliding states, rather than simply one-off, self-interested decisions? We explore that question here by analyzing the free-form text content of UPR reports.

Specifically, we use the text from the UPR reports to measure the semantic similarity of any two given reports written in the same session about the same state (the state under review, or, SuR). Recent work from political science has leveraged text-as-data methods to measure state preferences in the UN by studying 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 (Meyerrose and Nooruddin, 2023; Kim, 2024; Dai and Lu, 2024; Lu, 2024). The recommendations that states write in their UPR reports are relatively heterogeneous in terms of 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 other states.

Our dataset includes all potential dyads of state reviews written about the same SuR during the same UPR session.[21] To measure the similarity between any given two of these reports, we use a common measure of semantic similarity: cosine similarity.[22] The co-

[21] We use reports from all UPR sessions from 2008 through 2020.

[22] For other applications of cosine similarity in political science, see: Diodati, Marino and Carlotti (2018), Hager and Hilbig (2020), Kim (2024), and Lu (2024).

sine 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.

To estimate the cosine similarity for all dyads in our data, we first pre-process the text data contained in all UPR reports, using the `quanteda` package in R to remove numbers, punctuation, symbols, and stop-words (e.g., "the," or "and"). Next, we convert the text data to a matrix that records the number of times any given word appears in each report. Finally, again using the `quanteda` package, we use this matrix to estimate the cosine similarity for all pairs of documents in our dataset. Theoretically, cosine scores can range from 0, which indicates complete disagreement, to 1, which indicates complete agreement. In our dataset, the values range from 0.002 to 1, with a mean of 0.129 and a standard deviation of 0.087. We re-scale these scores to range from 0 to 100 for ease of interpretation. These cosine similarity scores on a 0 to 100 scale are the dependent variable in our analysis below.

Our theory predicts that backsliding states will shift away from the West, and increasingly toward autocracies, as they transition from developing to backsliding democracies. We explore this descriptively first by focusing on four prominent cases of backsliding: Brazil, India, Peru, and the Philippines, with Sweden, a stable advanced democracy, included as a comparison case. For each year from 2008 through 2020, we calculate these states' average cosine similarity scores first with the US, and then with China, with the expectation that over time, as these states backslide, their UPR reports will be more similar to those written by China, and less similar to those written by the US. We capture these trends in Figure 5, which plots each state's average difference in cosine similarity with China and the US (average China similarity minus average US similarity). By this measure, positive values would indicate a state has moved toward China and away from the US. As Figure 5 shows, backsliding has indeed resulting in these states becoming more similar to China with respect to the content of their UPR reports, on average, as they move away from the US.

We also more formally analyze the cosine similarity of backsliding states' UPR reports.

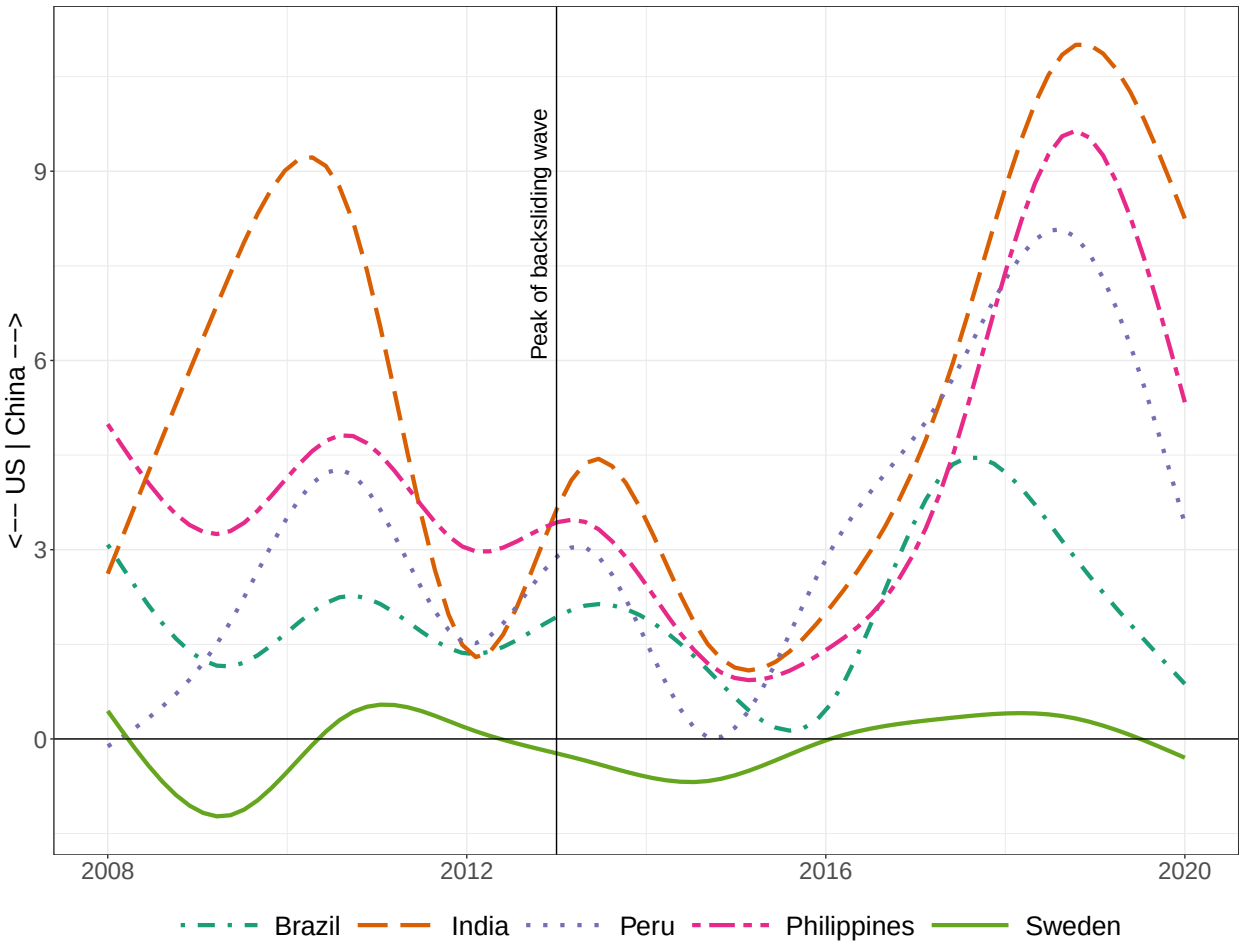


Figure 5: Over time, as backsliding has progressed, the content of backsliding states' UPR reports has become more similar to those written by China, and less similar to reports written by the US.

Similar to the results reported in Table 2 and Figure 4, in order to create the main independent variables in our models, we again divide the reviewing states into four mutually exclusive groups: advanced democracies, consolidated autocracies, backsliding states, and developing democracies. Using these four categories, our main independent variables are each possible pairing between these states,[23] with the advanced democracy-advanced democracy dyad as our omitted reference category. We choose to use the advanced democracy 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. Furthermore, using these states as the comparison group allows us to directly explore the extent to which these mature democracies are an isolated group in the UPR.

We also include a set of controls that might impact the extent to which two states' reports of the same state under review (SuR) are similar. 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 US in the UN General Assembly (Bailey, Strezhnev and Voeten, 2017); 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 and an indicator for whether the state under review was simultaneously a member of the UNHRC.

We use these data to test Hypotheses 4 and 5, estimating five OLS models with fixed effects for year, both recommending states, 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 4, and we bold coefficients that are of particular relevance to our argument (discussed below). In the first model, we include observations for all UPR reports written between 2008 and 2023. In the second and third models, we focus on different time periods: 2008–2012 and 2013–2020, respectively. Since the UPR was created in 2008, it is likely that, in the

[23]Namely, backsliding-backsliding; backsliding-advanced democracy; backsliding-autocracy; autocracy-autocracy; advanced democracy-autocracy; developing democracy-developing democracy; developing democracy-backsliding; developing democracy-autocracy; and developing-democracy-advanced democracy.

early years, states were still learning how to write their UPR reviews. Furthermore, this time period largely pre-dates the current wave of democratic backsliding, which expanded significantly starting around 2013. For both of these reasons, we expect to see meaningful differences between these two time periods. 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 19 advanced democracies we identify, and in model 5 we only include observations for reviews written about consolidated autocracies.

Several findings in particular stand out in Table 4. First, in all five models we find that two reports written by two backsliding states (DB-DB); a backsliding state and an autocracy (DB-Auto); two autocracies (Auto-Auto); two developing democracies (Dev Demo-Dev Demo); and a developing democracy and a backsliding state (Dev Demo-DB) are all significantly more similar than two reports written by a pair of advanced democracies (our baseline, omitted dyad). This may be suggestive of coordination —or at the very least shared interests— across all non-advanced democracies.

We also see several interesting patterns across models. First, we find that when the state under review is an advanced democracy (model 4), a backsliding state-advanced democracy pair is actually more similar than two advanced democracies; however, when the state under review is an autocracy, the sign flips and is no longer significant. This suggests that when an advanced democracy is under review, backsliding states are reluctant to challenge the prevailing western human rights norms, but, when an autocracy is under review, they are more reticent to align with the West.

Next, Table 4 shows that reports written by an advanced democracy and an autocracy (Adv Demo-Auto) are significantly less similar than reports written by a pair of advanced democracies, except for in models 2 and 3. On the one hand, the results for this pair in models 1, 4, and 5 validate conceptually our measure of text similarity; we of course would expect advanced democracies and autocracies to write less similar reports. But the results

Table 4: Cosine Similarity of UPR Reports, 2008–2020

Dependent variable	*Cosine similarity (0-100)*				
	All	2008-2012	2013-2023	Adv. democracies	Autocracies
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
DB-DB	1.63***	1.67***	1.69***	0.765***	1.99***
	(0.158)	(0.262)	(0.155)	(0.159)	(0.203)
DB-Adv Demo	0.047*	-0.119**	0.006	0.184**	-0.048
	(0.026)	(0.059)	(0.028)	(0.082)	(0.041)
DB-Auto	1.55***	1.87***	2.48***	0.691***	1.49***
	(0.172)	(0.298)	(0.171)	(0.172)	(0.221)
Auto-Auto	2.22***	3.24***	3.94***	0.805***	2.09***
	(0.237)	(0.456)	(0.235)	(0.243)	(0.299)
Adv Demo-Auto	-1.01***	-0.745***	-0.123	-0.046	-2.01***
	(0.115)	(0.225)	(0.117)	(0.119)	(0.146)
Dev Demo-Dev Demo	2.19***	2.08***	2.05***	0.467**	2.59***
	(0.181)	(0.351)	(0.183)	(0.190)	(0.229)
Dev Demo-DB	1.86***	1.80***	1.80***	0.666***	2.21***
	(0.157)	(0.285)	(0.156)	(0.157)	(0.202)
Dev Demo-Auto	1.11***	1.27***	1.96***	0.353*	0.776***
	(0.186)	(0.372)	(0.185)	(0.193)	(0.236)
Dev Demo-Adv Demo	0.570***	0.497***	0.457***	-0.037	0.683***
	(0.088)	(0.170)	(0.089)	(0.092)	(0.113)
Both rec same region	0.852***	0.543***	0.937***	0.690***	1.04***
	(0.089)	(0.104)	(0.100)	(0.091)	(0.094)
US-agree diff between recs	0.008	-1.07***	0.033	1.19***	-0.188
	(0.078)	(0.174)	(0.102)	(0.229)	(0.133)
SuR HR score	-0.151***	-0.869***	-0.060***	-0.356***	-0.493***
	(0.014)	(0.113)	(0.019)	(0.085)	(0.023)
SuR UNHRC member	-0.035***	-0.089*	-0.223***	0.131**	-0.250***
	(0.012)	(0.053)	(0.016)	(0.062)	(0.019)
Either rec under review	-0.101***	-0.187***	-0.144***	-0.216***	-0.053**
	(0.021)	(0.039)	(0.024)	(0.029)	(0.027)
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	28,374,994	5,761,437	22,613,557	3,771,939	8,504,754
R^2	0.05357	0.07640	0.06041	0.07933	0.05771
Mean of DV	12.95	12.78	12.99	12.61	13.39
SD of DV	8.75	8.53	8.79	8.84	9.04

Clustered standard-errors by dyad in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: DB = democratic backslide; Adv demo = advanced democracy; Dev demo = developing democracy; auto = autocracy.

from model 3 suggest that, in the early years of the UPR, agreement across reports in general was relatively high, and that learning has occurred over time as the UPR has developed. Perhaps more interestingly, the results in model 4 may signal that autocracies are less likely to push back against the west directly (when reviewer advanced democracies).

Finally, while there is no significant relationship between reports written by a developing democracy and an advanced democracy when the state under review is an advanced democracy (model 4), these reports are significantly more similar when a pair of these states is reviewing an autocracy (model 5) This suggests that, unlike backsliding states, the more stable, developing democracies will still align with advanced democracies in their evaluation of autocracies and perhaps communicate human rights values and preferences that more closely align with those that are foundational to the West and the LIO.

One important question is: how large at the estimated effect sizes? At the bottom of Table 4, we report the mean and standard deviation for the dependent variable for each relevant subset of the data. Among the significant coefficients, the effect sizes range from a change of 0.01 (backslide-advanced democracy, model 1) to 0.45 (autocracy-autocracy, model 3) standard deviations.[24] We also report in Table 5 whether the coefficients for each distinct pair of our (main) independent variables are significantly different from one another.[25]

In our view, of particular note in Table 5 are the pairs of coefficients that are *not* statistically significantly different from one another.[26] In particular, pairs of backsliding states are not significantly different than pairs of consolidated autocracies, providing further suggestive evidence that backsliding states are coordinating their behavior in the UPR in ways comparable to fully consolidated autocracies.

[24]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)).

[25]We report these pairwise comparisons for model 5 from Table 4.

[26]We highlight these pairs in bold, and add red text for those pairs we find to be of particular theoretical interest in Table 5.

	DB-Adv	DB-Auto	Auto-Auto	Adv-Auto	Dev-Dev	Dev-DB	Dev-Auto	Dev-Adv
DB-DB	✓	✓	✗	✓	✓	✗	✗	✓
DB-Adv		✓	✓	✓	✓	✓	✓	✓
DB-Auto			✓	✓	✓	✓	✓	✓
Auto-Auto				✓	✓	✗	✓	✓
Adv-Auto					✓	✓	✓	✓
Dev-Dev						✓	✓	✓
Dev-DB							✓	✓
Dev-Auto								✗

Table 5: Pairwise differences between coefficient estimates for UPR model 5 (Autocracies) from Table 4. "✓" indicates pairs of coefficients that are statistically different from one another ($p < 0.05$), while "✗" indicates no statistically significant difference.

Note: DB = democratic backslide; Adv demo = advanced democracy; Dev demo = developing democracy; auto = autocracy.

7 Conclusion

Rising challenges from China, Russia, and other powerful autocratic states, coupled with political weaknesses among its founders and champions, has called into question the future of the liberal international order. At the heart of this existential crisis for the LIO is the question of whether the advanced Western democracies that created it are able to maintain and continue to expand their coalition of like-minded 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 the United States, the United Kingdom, and their partners were able to maintain the support of those states that are neither mature democracies nor fully autocratic regimes.

We argue that the ongoing global democratic recession threatens the West's efforts to maintain the LIO. Not only are advanced democracies increasingly facing contestation from a bloc of consolidated autocracies, but they are also losing members of their liberal coalition as former developing democracies backslide, become increasingly illiberal and, as a result, drift closer toward established autocracies.

We explore these dynamics in the context of the UN human rights institutions, providing new evidence that backsliding states are more likely to vote with autocratic states on all types

of resolutions (general, targeted, and contentious ones) in the UNHRC, especially after 2013. Drawing on the text of reports from the UPR, we also find that reports written by pairs of backsliding states; autocratic states; a backsliding state and an autocratic state; two developing democracies; and a developing democracy and a backsliding state are all significantly more similar than two reports written by a pair of advanced democracies. We also show that backsliding states become less aligned with advanced autocracies when the state under scrutiny is a consolidated autocracy. Taken together, this evidence suggests that advanced democracies are becoming increasingly isolated in the UN human rights institutions.

References

Bailey, Michael A., Anton Strezhnev and Erik Voeten. 2017. "Estimating dynamic state preferences from United Nations voting data." *Journal of Conflict Resolution* 61:430–456.

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

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

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* pp. 1–31.

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, Nazifa Alizada, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Allen Hicken, Garry Hindle, Nina Ilchenko, Joshua Krusell, Anna L uhrmann, Seraphine F. Maerz, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundström, Ei tan Tzelgov, Yi ting Wang, Tore Wig, Steven Wilson and Daniel Ziblatt. 2021. "V-Dem Country-Year/Country-Date Dataset v11.".

Cottiero, Christina, Emilie M Hafner-Burton, Stephan Haggard, Lauren Prather and Christina J Schneider. 2024. "Illiberal regimes and international organizations." *The Review of International Organizations* pp. 1–29.

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. 2021. "Clubs of autocrats: Regional organizations and authoritarian survival." *Review of International Organizations* .

Debre, Maria J and Lee Morgenbesser. 2018. Out of the shadows: autocratic regimes, election observation and legitimation. In *Justifying Dictatorship*. Routledge pp. 78–97.

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.

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.

Eilstrup-Sangiovanni, Mette and Stephanie C Hofmann. 2020. "Of the contemporary global order, crisis, and change." *Journal of European Public Policy* 27(7):1077–1089.

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

Flores, Thomas and Irfan Nooruddin. 2016. *Elections in Hard Times: Building Stronger Democracies in the 21st Century*. Cambridge: 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 CW Pevehouse and Christina J Schneider. 2024. "Good governance in autocratic international organizations." *The Review of International Organizations* pp. 1–34.

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

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.

Herre, Bastian and Max Roser. 2016. "Human Rights." *Our World in Data* . https://ourworldindata.org/human-rights.

Hug, Simon and Richard Lukács. 2014. "Preferences or blocs? Voting in the United Nations Human Rights Council." *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. 2009. "Liberal Internationalism 3.0: America and the Dilemmas of Liberal World Order." *Perspectives on Politics* 7:71–87.

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.

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 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. 2023. "A network analysis of naming and shaming in the universal periodic review." *International Interactions* 0(0):1–28.

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 and 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.

Kornprobst, Markus and Thazha V Paul. 2021. "Globalization, deglobalization and the liberal international order." *International Affairs* 97(5):1305–1316.

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.

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.

Lipps, Jana and Marc S Jacob. 2024. "Undermining liberal international organizations from within: Evidence from the parliamentary assembly of the council of Europe." *The Review of International Organizations* pp. 1–27.

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-un-human-rights-reviews/*

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.".

McMahon, Edward and Marta Ascherio. 2012. "A Step Ahead in Promoting Human Rights? The Universal Periodic Review of the UN Human Rights Council." *Global Governance* 18:231–248.

Mearsheimer, John J. 2019. "Bound to fail: The rise and fall of the liberal international order." *International security* 43(4):7–50.

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." *Review of International Organizations* pp. 307–343.

Meyerrose, Anna M. 2025. *Eroding Democracy from the Outside In: International Policy Delegation and Democratic Backsliding*. Oxford University Press.

Meyerrose, Anna M. and Irfan Nooruddin. 2023. "Trojan Horses in Liberal International Organizations?: How Democratic Backsliders Undermine the UNHRC.". Review of International Organizations.

Morrison, Kelly, Daniela Donno, Burcu Savun and Perisa Davutoglu. 2024. "Competing judgments: Multiple election observers and post-election contention." *The Review of International Organizations* pp. 1–29.

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.

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." *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.

Terman, Rochelle and Zoltán I Búzás. 2021. "A house divided: Norm fragmentation in the international human rights regime." *International Studies Quarterly* 65(2):488–499.

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

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. 2023. "How backsliding governments keep the European Union hospitable for autocracy: Evidence from intergovernmental negotiations." *The Review of International Organizations* pp. 1–30.

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

Coordinated Contestation: The Shifting Allegiances of Backsliding States in the UN Human Rights Institutions

Online Appendices

8 UNHRC Regression Table: 2006–2012

Table 6: Voting with the Autocratic Bloc in the UNHRC, 2006–2012

	Dependent variable:		
	Vote with autocracy bloc (1 = yes, 0 = no)		
	All	Targeted	Contentious targeted
	(1)	(2)	(3)
Backsliding state	−0.01	0.12	0.28
	(0.10)	(0.16)	(0.19)
Mature democracy	−1.75***	−1.18***	−3.33***
	(0.11)	(0.17)	(0.52)
Consolidated autocracy	0.59***	0.77***	0.95***
	(0.15)	(0.24)	(0.29)
Voting state demo score	−3.27***	−1.16**	−1.68***
	(0.28)	(0.45)	(0.55)
Target state = Israel		1.72***	0.80***
		(0.16)	(0.23)
Target state = Syria		1.95***	−0.38
		(0.26)	(0.48)
Voting state HR score		−0.30***	−0.61***
		(0.05)	(0.07)
Perc. demo sponsors		0.003	−0.01**
		(0.002)	(0.003)
Target state HR score		0.34***	0.45***
		(0.12)	(0.15)
Voting and target in same region		0.51*	0.50
		(0.28)	(0.33)
Year fixed effects	Yes	Yes	Yes
Clustered standard errors	Yes	Yes	Yes
Model type	Logit	Logit	Logit
Observations	6,206	2,545	1,756
Log Likelihood	−2,813.56	−1,172.69	−686.84
Akaike Inf. Crit.	5,649.12	2,379.38	1,407.68

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

9 UNHRC Regression Table: 2013–2023

Table 7: Voting with the Autocratic Bloc in the UNHRC, 2013–2023

	Dependent variable:		
	Vote with autocracy bloc (1 = yes, 0 = no)		
	All	Targeted	Contentious targeted
	(1)	(2)	(3)
Backsliding state	0.03	0.23**	0.39***
	(0.05)	(0.10)	(0.13)
Mature democracy	−1.13***	−0.33**	−1.56***
	(0.07)	(0.14)	(0.37)
Consolidated autocracy	−0.11	0.44***	0.42**
	(0.08)	(0.16)	(0.20)
Voting state demo score	−4.05***	−1.48***	−2.53***
	(0.15)	(0.30)	(0.39)
Target state = Israel		1.12***	0.47**
		(0.14)	(0.24)
Target state = Syria		−0.64***	−0.67***
		(0.10)	(0.13)
Voting state HR score		−0.45***	−0.84***
		(0.03)	(0.05)
Perc. demo sponsors		−0.005***	−0.004**
		(0.001)	(0.002)
Target state HR score		−0.10*	0.23***
		(0.06)	(0.08)
Voting and target in same region		−0.83***	−1.22***
		(0.15)	(0.19)
Year fixed effects	Yes	Yes	Yes
Clustered standard errors	Yes	Yes	Yes
Model type	Logit	Logit	Logit
Observations	16,350	4,451	2,966
Log Likelihood	−8,460.21	−2,244.70	−1,233.27
Akaike Inf. Crit.	16,950.43	4,523.40	2,500.53

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