



Trojan horses in liberal international organizations? How democratic backsliders undermine the UNHRC

Anna M. Meyerrose¹ · Irfan Nooruddin²

Accepted: 14 September 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Liberal democracy is facing renewed challenges from a growing group of states undergoing democratic backsliding. While entrenched autocrats have long resented and contested the established liberal order, we know far less about how newer backsliding states behave on the international stage. We argue these states, who joined prominent western liberal institutions prior to their backsliding, will use their established membership in these organizations both to protect themselves from future scrutiny regarding adherence to liberal democratic values and to oppose the prevailing western liberal norms that increasingly conflict with their evolving interests. Using voting data from the UN Human Rights Council (UNHRC) from 2006–2021, we show that backsliding states are more likely to vote against targeted resolutions that name and shame specific countries. We supplement this analysis with detailed data from the Universal Periodic Review (UPR) and combine regression analysis and a structural topic model (STM) to show that backsliding states are more critical in their UPR reports when evaluating advanced western democracies, and more likely to emphasize issues that align with their own interests while de-emphasizing ones that might threaten government power and control over citizens.

Keywords United Nations · Illiberalism · Democratic backsliding · International organizations · Liberal international order · Human rights · UNHRC · UPR

Responsible editor: Axel Dreher

✉ Anna M. Meyerrose
Anna.Meyerrose@asu.edu

✉ Irfan Nooruddin
in62@georgetown.edu

¹ Assistant Professor, School of Politics and Global Studies, Arizona State University, Tempe, AZ 85281, USA

² Hamad bin Khalifa Al-Thani Professor of Indian Politics, School of Foreign Service, Georgetown University, Washington, DC NW, 20057, USA

JEL Classification: F5

Liberal democracy faces its most severe global challenge since the fall of the Soviet Union: in 2021, the number of liberal democracies around the world declined to levels not seen since 1989 and such regimes currently represent only 13% of the world's population (Alizada et al., 2022). Liberal democracy's decline has accelerated precipitously over the last decade in large part due to democratic backsliding, an historically unprecedented phenomenon that occurs when democratically elected officials intentionally weaken or erode institutional checks on government power (Bermeo, 2016; Waldner and Lust, 2018; Haggard and Kaufman, 2021). While research has begun to explore the consequences of backsliding for domestic democratic institutions and seeks to provide both domestic- (Haggard and Kaufman, 2016; Mudde Rovira Kaltwasser, 2017; Ziblatt and Levitsky, 2018) and international-level (Meyerrose, 2020, 2023; Kelemen, 2020) explanations for its occurrence, less attention has been paid to the implications of backsliding for international outcomes.

There exists extensive work on the international interests and behavior of *entrenched* autocrats, with an emphasis on how these autocrats use both new and existing multilateral international organizations (IOs) to advance their agendas, protect themselves from international intervention and scrutiny, and more broadly challenge established western liberal values (Finnemore and Sikkink, 1998). Powerful autocratic states such as Russia and China have long contested western liberal values, particularly in the area of human rights (Binder and Payton, 2022), from within established western IOs (Lebovic and Voeten, 2006; Johnston, 2007; Donno, 2010; Smith, 2016; Carraro, 2017). More recently, these powerful autocrats have even begun to create their own regional IOs (Cottiero and Haggard, 2023) that promote the spread of authoritarian rules and regimes in their neighborhoods (Libman and Obydenkova, 2018; Kneuer et al., 2019) and provide an ideological counter-narrative to the current international environment (Ginsburg, 2020; Debre, 2021; Hafner-Burton et al., 2023). These and related developments have led pundits and politicians to warn the liberal international order (LIO) is under threat (Ikenberry, 2018; Börzel and Zürn, 2021; Way, 2022).

Backsliding states are fundamentally altering the composition of long-established western liberal IOs, yet we know little about how they behave within these organizations. In the post-Cold War era, western-dominated multilateral organizations increasingly sought to promote and protect western liberal values and at the same time widened their membership to bring emerging and fragile democracies into the fold (Pevehouse, 2005; Donno, 2013; Genna and Hiroi, 2014; Poast and Urpelainen, 2018). However, this tactic might now be backfiring. Today, many of the backsliding states of greatest concern are established members of western IOs, with troubling implications for these organizations' efforts to protect and promote values central to the liberal international order. We explore these dynamics by focusing on one core aspect of the LIO, the international human rights regime, asking: do backsliding states exhibit distinct behavior and promote different types of issues in the United Nations (UN) human rights institutions? Answering this question sheds light on the ways and extent to which these backsliding states leverage their membership in the UN to protect and

advance their evolving and increasingly illiberal domestic interests, with implications for other established western IOs.

Democratic backsliding is unique when compared to historic cases of democratic collapse. Today, moves away from democracy occur much more gradually than in the past (Bermeo, 2016; Waldner and Lust, 2018), and contemporary cases of backsliding suggest the outcome of ongoing erosions is an illiberal democratic or semi-autocratic regime, rather than a full-fledged autocracy (Luhmann and Lindberg, 2019). Backsliding states, while increasingly illiberal, heed incentives to appear at least minimally democratic internationally (Hyde, 2011) and also to maintain the patina of democracy to assuage their domestic constituencies. Therefore, while we expect backsliding states to display illiberal interests that are increasingly similar to those of fully entrenched autocrats, the methods backsliding states employ to protect these interests are likely distinct from those of entrenched autocrats who are less concerned with maintaining a minimally democratic veneer.

Specifically, we argue backsliding states will use their established membership in the UN human rights institutions both to protect themselves from future scrutiny regarding their adherence to western human rights values, and also more directly to oppose the prevailing norms that dominate the international system and threaten the domestic legitimacy of these newly emerging regimes. As these states move away from democracy, they can expect to face greater scrutiny for their pseudo-democratic and illiberal domestic practices. In anticipation, these regimes have incentives to oppose and undermine international efforts to enforce western liberal values. At the same time, these countries are part of a growing global wave of backsliding states, all of which have illiberal interests advocated by powerful would-be autocrats. This emboldens them to challenge the West on the universality of liberal values, and also to promote an alternative set of norms that better align with their evolving regime. In these ways, we argue the very international fora created, subsidized, and promulgated by the West to promote liberal hegemony, such as the UN human rights institutions, might now be used by backsliding states to stymie and undo those efforts.

We test our argument using data from the UN Human Rights Council (UNHRC), founded in 2006, and the Universal Periodic Review (UPR), which was introduced in 2008 as an additional tool to promote human rights. The UPR is arguably one of the most elaborate multilateral institutions devoted to promoting and protecting human rights (Terman and Byun, 2022) and to naming-and-shaming states that violate these rights (Kim, 2023).¹ To our knowledge, no fully updated dataset of UNHRC resolutions exists; therefore, we create a comprehensive dataset that includes voting behavior for all members of the UNHRC from 2006 through 2021, and collect detailed information on UPR reports and their content.

Using these data, we show backsliding states are more likely to vote against resolutions that target specific states, consistent with our argument that backsliders use

¹ Although the UNHRC's predecessor—the UN Commission on Human Rights (UNCHR)—was founded in 1946, we focus on the time period from 2006 to the present since democratic backsliding is an historically recent phenomenon that began in the early- to mid-2010s (Haggard and Kaufman, 2021).

their voting powers in the UNHRC to reinforce the notion of sovereign immunity and non-interference. The collective impact of such behavior is to undermine the legitimacy of these resolutions that are being passed, which in turn guards illiberal states against future criticism in similar fora. We supplement these findings with analyses of the UPR reports. We show that backsliding states use the UPR mechanism to target consolidated western democracies for their own alleged domestic human rights abuses. We also use the UPR data to estimate a structural topic model (STM), finding that backsliding states' reports are more likely to emphasize topics related to protecting vulnerable groups and human development, while de-emphasizing ones that might pose a challenge to government power and control over citizens, such as civil and political liberties. Taken together, these findings suggest backsliding states are using international human rights fora—a critical component of the liberal international order—simultaneously to protect their own interests, to promote a new set of international human rights practices, and increasingly to scrutinize and counter efforts by western democracies to enforce international liberal norms and values.

1 Western multilateral fora and the liberal international order

The liberal international order (LIO) has a long history that can be traced to the years after World War I, but it expanded significantly in the post-Cold War era. Beginning in the early 1990s, western liberal democracies increasingly worked through international organizations (IOs) with the overarching goal of creating “an international ‘space’ for liberal democracy” (Ikenberry 2018, 8). They sought to promote liberal values such as free and fair elections, enhanced rule of law, and civil and political human rights (Finnemore and Sikkink, 1998; Barnett and Finnemore, 2004; Hyde, 2011; Hafner-Burton, 2012; Flores and Nooruddin, 2016). Throughout the 1990s and early 2000s, these international institutions gained significantly more authority and became decisively more liberal, and as democracy promotion became a core foreign policy objective for western powers (Collier and Levitsky, 1997; Carothers, 2002; Levitsky and Way, 2010; Johnstone and Snyder, 2016), membership in these IOs was increasingly conditioned on regime type (Hyde, 2011). It was in this context that countries in east Asia, eastern Europe, and Latin America transitioned to democracy and sought out and secured membership in these organizations (Pevehouse, 2005; Donno, 2010; Poast and Urpelainen, 2018). The belief in the power of the liberal international order was so great that the West even anticipated that tougher cases for democracy, such as China and Russia, could eventually liberalize politically as they became more integrated in these liberal international institutions (Walker, 2016); therefore, their membership and participation in these organizations was also encouraged and expanded.

The United Nations institutions, in particular, were pivotal in these efforts. The UN was founded with the goal of saving “succeeding generations from the scourge of war..., to reaffirm faith in fundamental human rights..., to establish conditions under which justice...can be maintained..., [and] to promote social progress and better

standards of life in larger freedom” (United Nations Charter, 1945). This emphasis on the importance of protecting and spreading human rights led to the creation of the UN Commission on Human Rights (UNCHR) in 1946 and the Universal Declaration of Human Rights in 1948. Together these served as the foundation for a global framework for promoting and monitoring states’ human rights practices. The UNCHR was widely viewed as a platform for democratic states to advance liberal norms via socialization (Finnemore and Sikkink, 1998) and an institution that could be used to formalize commitments to norms of liberal behavior (Voss, 2019). Scholars found that new democracies joined these human rights institutions and treaties at higher rates than both dictatorships and consolidated democracies, using the sovereignty costs associated with membership to commit credibly to upholding international law and to becoming consolidated democracies (Moravcsik, 2000; Landman, 2005; Hafner-Burton, 2012).

The UNCHR faced significant criticism throughout its lifespan, with many arguing it was too heavily exposed to political influence and that its rules for choosing members left open the possibility that states with poor human rights records could become members and de-legitimize its work (Edwards et al., 2008). In response to such concerns, the UNCHR was dissolved and replaced by the UN Human Rights Council (UNHRC) in 2006. The UNHRC reports directly to the UN General Assembly (UNGA), and therefore has higher status than its predecessor. While few amendments were made to membership criteria (Cox, 2010; Hug and Lukács, 2014), in other ways the UNHRC implemented significant changes: the Council now meets more often (Edwards et al., 2008) and, in 2008, it initiated a new tool—the Universal Periodic Review (UPR)—to assess human rights situations in *all* UN member states.

Over time, western liberal democracies have augmented these formal human rights institutions with other tools to promote human rights (Koliev, 2020; Kim, 2023). The UNCHR’s activities expanded significantly following the end of the Cold War, becoming particularly far-reaching in the mid-1990s, and focusing increasingly on the worst human rights violators (Lebovic and Voeten, 2006). New peer review tools, such as the UPR, appear to be effective tools for naming and shaming (Carraro et al., 2019), though the particular human rights norms that are enforced in this context are contingent on the nature of the relationship between the reviewing and target states (Terman and Byun, 2022). Western democracies have also found less direct ways of punishing states that violate liberal norms, for example by tying human rights rules directly to market access via preferential trade agreements (Hafner-Burton, 2005). Scholars disagree, however, about the utility of these tactics. While there is evidence that naming and shaming is an effective tool used by IOs and NGOs to punish human rights violations (Risse and Sikkink, 1999; Lebovic and Voeten, 2006; Woo and Murdie, 2017; Terman and Voeten, 2018), other research suggests human rights treaties and naming-and-shaming are unable to influence human rights adherence in target states meaningfully (Tsutsui and Wotipka, 2004; Hafner-Burton and Tsutsui, 2005; Neumayer, 2005; Vreeland, 2008) and in some cases might even fuel additional human rights abuses (Hafner-Burton, 2008).

2 The global democratic recession and autocratic international cooperation

Despite extensive efforts to promote democracy, human rights, and other liberal values through IOs in the post-Cold War environment, the last decade and a half has witnessed a sharp rise in cases of democratic backsliding and rampant illiberalism in new democracies whose transitions to democracy were once heavily influenced and supported by the international community (Meyerrose, 2020, 2023), and in particular by the West.

Luhrmann and Lindberg (2019) identify three waves of “autocratization,” each of which followed their corresponding waves of democracy (Huntington, 1991). Autocratization indicates moves away from democracy (or, toward autocracy) that can occur in any type of regime. When it occurs in democracies, autocratization denotes both abrupt regime changes, such as those that result from a classic coup d’état, as well as more gradual and incremental erosions that result in either an illiberal, diminished democracy or involve an eventual transition to autocracy. The current third wave of autocratization is distinguishable from previous waves because today’s moves away from democracy occur more gradually (Bermeo, 2016; Waldner and Lust, 2018). Such regressive moves—commonly termed ‘backsliding’—are increasingly evident in an unprecedented number of erstwhile democracies (Luhrmann and Lindberg, 2019). Given this unique characteristic of the third wave of autocratization, we focus on cases of democratic backsliding.

Democratic backsliding 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 powers within governments (Haggard and Kaufman, 2021). Democratic backsliding is an historically recent and unprecedented phenomenon that has impacted all types of democracies, including those new ones that joined the UN and other liberal IOs in the post-Cold War era. Contemporary cases suggest the outcome of ongoing erosions is a form of illiberal democracy or semi-autocratic regime, rather than full-fledged authoritarianism (Luhrmann and Lindberg, 2019). As democratic backsliding has become increasingly common, illiberal democracies and semi-autocracies have proliferated, expanding the number of non-liberal democratic states both in the world and also within the very western liberal IOs charged with promoting and maintaining the liberal international order.

Concurrent with this global democratic recession, consolidated autocracies and backsliding states have taken active measures at the international level to evade western pressures to adhere to liberal values. To date, most research on autocratic cooperation has focused in particular on autocratic regional IOs, or on IOs dominated by autocratic members (Cottiero and Haggard, 2023). Powerful autocratic states such as Russia, China, Venezuela, and Saudi Arabia have created regional organizations including the Shanghai Cooperation Organization, the Bolivarian Alliance for the Peoples of Our America, and the Gulf Cooperation Council as alternatives to established liberal IOs. By providing member states with economic and military support, increasing regime legitimacy, and challenging international pressures to democratize (Weyland, 2017), these regional organizations serve as one tool for powerful autocratic states to promote

the spread of authoritarian rules and regimes in their neighborhoods (Libman and Obydenkova, 2018; Kneuer et al., 2019) and to legitimize an ideological counter-narrative to liberal hegemony (Ginsburg, 2020; Debre, 2021). These regional organizations not only provide an alternative to the prevailing liberal democratic values, but can also be used to help autocratic leaders maintain power at home, in part by appeasing international actors committed to promoting democracy, good governance, and liberal values (von Soest, 2015; Hafner-Burton et al., 2023). By passing regional Human Rights Charters or sending accommodating election observers to member states, such regional organizations allow autocratic leaders to signal to their domestic audiences that (western liberal) international accusations of flawed elections or abuses of human rights are unfounded (Debre and Morgenbesser, 2018; Debre, 2021; Donno et al., 2023; Bush et al., 2023).

While these autocratic regional organizations will undoubtedly continue to gain influence in the coming decades, western-based IOs still dominate the multilateral landscape. Yet, even there, evidence is growing that powerful, established autocracies use their membership in these IOs in ways that subvert from within institutional efforts to promote liberal democratic values. Binder and Payton (2022) analyze voting behavior in the UN General Assembly to reveal that rising powers such as Brazil, Russia, India, China, and South Africa have formed a bloc of dissatisfied states. Focusing on the UNCHR, Hug and Lukács (2014) similarly find authoritarian states pursued membership on the Commission to dilute international human rights norms and to advance illiberal ones; these states also worked to shift the focus from civil and political rights to more purely economic issues (a finding we will corroborate below). Repressive states also used their membership in the UNCHR to insulate themselves from investigation and to deflect attention and blame toward other states (Edwards et al., 2008; Cox, 2010; Seligman, 2011; Hug and Lukács, 2014; Vreeland, 2019). More recent data from the current UNHRC also suggest states from the Global South, including Russia, use amendments to UNHRC resolutions to protect their own human rights preferences and to challenge the existing international order (Voss, 2019). Looking beyond established autocracies, Meyerrose (2018) shows that new democracies, which are over-represented in contemporary cases of backsliding, exhibit distinct voting patterns in western liberal IOs such as the European Parliament.

The post-Cold War push to expand the LIO resulted in then-emerging and fragile democracies also becoming members of western liberal IOs. Many of these states have more recently begun backsliding toward autocracy (Luhmann and Lindberg, 2019). Today, these backsliding states are joining fully autocratic states as established members of the very IOs that historically were at the forefront of promoting and supporting the LIO. While initially this membership expansion was viewed favorably by proponents of liberal democracy, broadening the reach and potential influence of western liberal IOs via expanded or universal membership may be incompatible with deepening the ability of these organizations to accomplish their goals of promoting and supporting western liberal values.² We develop our argument in the following section.

² Pahre (1995) illustrates this theoretical tension in the context of the European Union's enlargement.

3 Undermining the LIO from within: Backsliding states in western liberal IOs

The fall of the Soviet Union left the United States as the undisputed hegemon of the international system in the 1990s. In the preceding decades the US and its partners had spearheaded a proliferation of international organizations, covering every topic imaginable, from the purely political to the very technical. The decisions and rulings rendered by this dense web of IOs generated a robust body of international law, the enforcement of which, while uneven, was backed by uncontested US power. A particular focus of this new liberal international order was the promotion of democracy, which in turn necessitated a more intrusive examination of member state's domestic politics.

The power of the LIO and the absence of an alternative champion left developing countries with scant choice but to sign on to the corresponding international organizations, and also to accept the greater scrutiny of their domestic policy records. As Susan Hyde has argued, this spawned a generation of "pseudo-democrats" who hoped that being seen to hold free and fair elections could unlock material benefits, while fearing that being caught explicitly violating this new norm would invite punishment (Hyde, 2011). Elsewhere, as documented in the literature reviewed above, membership in the LIO meant subjecting oneself to new pressures to improve human rights records, to protect labor and the environment from corporate predation, and, even, to accept new international accounting standards to facilitate trade and investment.

For the developing world, being admitted to the liberal club was important symbolically for the domestic legitimacy and international rewards it conferred but also risky because of the challenges to unfettered domestic sovereignty. But with democracy ascendant and unrivaled, these countries had little choice but to go along with the wave. However, this status quo changed faster than anyone had anticipated, and a mere two decades after the "end of history" (Fukuyama, 1989) was celebrated, the world was on the cusp of a serious democratic recession. The United States was distracted by its wars in Iraq and Afghanistan, and the moral suasion power it exercised was severely undermined by the loss of legitimacy caused by those wars, and particularly the false pretenses under which the Iraq war was waged (Sky, 2023). The staggering growth of China's economy provided alternative, non-western sources of aid and investment with fewer policy conditions attached for many developing countries, but in doing so also legitimated a different model of governance that privileged stability and performance over multi-party elections. Democratic backsliding grew more common, as the elections that had become commonplace in the 1990s no longer yielded any gains in democracy but, if anything, often harmed the broader democratic project (Flores and Nooruddin, 2016; Meyerrose et al., 2019; Bartels et al., 2023). A key point to remember is that backsliding did not necessarily mean new leadership; rather it was the incumbent governments that engineered the undermining of electoral integrity, indulged in electoral violence, and repressed their oppositions to ensure their continued hold on power (Flores and Nooruddin, 2023). These democrats-turned-autocrats, having cemented their rule at home, now had to limit and deflect international criticism of their actions.

Membership in western liberal IOs provides democratic backsliders an ideal forum in which to undermine the LIO's pressure on them. Because they are already members, having been admitted to these IOs either at independence in the case of the universal IOs, or on the backs of their democratizing credentials for the more exclusive ones, they can exercise their voting privileges in these organizations to advance their increasingly illiberal domestic interests. The UNHRC is a perfect example. Membership in the Council is not predicated on a country's own democratic or human rights records, but rather the UNHRC's membership is elected by regional peers. Over time, this has meant that the global democratic recession also resulted in a backsliding of democracy *within* the Council. This has created the opportunity for backsliding states both to use their membership on the Council to oppose resolutions that punish states for violating western human rights values, and to emphasize different types of human rights issues that better align with their domestic interests.

We expect backsliding states to exhibit distinct behaviors in the UN human rights institutions. First, backsliding countries, like their fully autocratic counterparts (Dukalskis, 2023), should be less likely to vote in favor of UNHRC resolutions targeting other states for alleged human rights violations.³ These backsliding countries have the most to gain from undermining the ability of IOs to scrutinize the domestic politics of nations, and have the most to gain from reinforcing the notion of sovereign immunity within the international system. Voting against these resolution is one tool these states can use to push back against established western efforts to promote and protect liberal human rights values. This logic suggests the following testable hypothesis:

Hypothesis 1 *Backsliding states are less likely to vote in favor of UNHRC resolutions targeting other states for alleged human rights violations.*

Voting on specific resolutions allows states to register visibly their support for or opposition to the principles of the LIO. Lakatos (2022) argues that states interested in weakening the international human rights monitoring system are increasingly using the UNHRC to escalate tensions between the Global North and the Global South. However, their ability to do so in a setting governed by majoritarian voting rules is limited. The creation of the UPR mechanism inadvertently provided an additional means for backsliding democracies and autocracies to challenge prevailing liberal norms and western dominance. Because the UPR allows countries to comment on other countries' human rights situations in a bilateral setting, it serves not only as a way for liberal democracies to promote human rights, but also as a tool for backsliding states to challenge directly their would-be critics. One way to do this is to place these western democracies under increased scrutiny for their own alleged domestic human rights shortcomings.⁴ This helps deflect attention from backsliding and non-democratic states and also undermines the credibility of states at the forefront of

³ For a related project that is consistent with our approach, see Lipps and Jacob (2023).

⁴ A fascinating case-study of this approach is provided by Baturro (2023) who shows how Russia has used the language of anti-Nazism to dull criticism of its domestic policies.

the promotion of western-backed human rights. This suggests the following testable hypothesis:

Hypothesis 2 *Backsliding states will be more critical in UPR reports targeting Western liberal democracies.*

The UPR also reveals the issues countries deem most worthy of discussion by the international community. Autocratic-leaning states have long argued that western states' emphasis on individual political and civil liberties is at best partial and not reflective of their country's needs. Rather they contend that greater focus should be paid to economic and social rights (Inboden, 2021). While both sets of rights have been formally enshrined under international human rights law, we see in the UPR mechanism an opportunity to understand whether backsliding states, in ways similar to China and other long-established autocracies, are more likely to raise a different set of issues than consolidated democracies. In particular, we test the following hypothesis:

Hypothesis 3 *Backsliding states are less likely to emphasize democratic rights and civil liberties in their UPR recommendations.*

Taken as a set, if confirmed by data, these hypotheses have implications for how the global trend in democratic backsliding may be undermining existing IOs such as the UNHRC in promulgating the core ideas of the LIO. If such states are less likely to vote in favor of UNHRC resolutions condemning human rights violations in particular countries, thereby generating fewer unanimous condemnations, and if such states are more likely to de-emphasize political and civil liberties in favor of other articulations of rights in their use of the UPR mechanism, then the overall efficacy of the UNHRC is undermined since the organization can no longer claim to speak with one voice. Over time, the toll of such internal division within the UNHRC would be to undermine its very reason for existence. In the next section, we analyze country voting data in the UNHRC and data from UPR recommendations to assess if these concerns are warranted, or hyperbolic.

4 Data and results

We test our argument using data from human rights organizations within the United Nations. Given its universal membership, the UN is particularly vulnerable to backsliding states interfering with its efforts to support the liberal values of its founding Charter. We focus on two tools for human rights promotion and protection in particular: the UN Human Rights Council (UNHRC) and the Universal Periodic Review (UPR).

4.1 Voting in the UNHRC

The UNHRC was founded in 2006 after its predecessor the UNCHR was dissolved (Hug and Lukács, 2014). The present-day UNHRC consists of 47 member states that

are elected by the UN General Assembly for three-year terms. However, there are no requirements for membership; as such, even backsliding states—and states with abysmal human rights records—have the opportunity to be elected to the UNHRC (Edwards et al., 2008; Hug and Lukács, 2014). As Fig. 1 illustrates, while the average democracy score of the UNHRC and its predecessor has always been relatively low, this average has declined even more recently, reflecting the ongoing global democratic recession.

Building on work that finds regime type influences how states vote on targeted UNHRC resolutions, we test if backsliding states exhibit distinct voting patterns. Existing research suggests states use international human rights institutions to advance norms that characterize their domestic politics (Edwards et al., 2008), and that democracies are more likely to vote in favor of targeted UNHRC resolutions than their non-democratic counterparts (Seligman, 2011). Therefore, we expect backsliding states will seek to protect state sovereignty, and also to dilute efforts at human rights promotion, by voting against targeted UNHRC resolutions.

We explore how backsliding states behave in the UNHRC with a dataset that includes all UNHRC resolutions that were a) targeted at a specific state and b) decided using a recorded roll-call vote among all members of the UNHRC between 2006 and 2021. We focus exclusively on this time period since it limits the sample to resolutions voted on in the current UNHRC, rather than its predecessor organization. From a theoretical perspective, this time period is also of particular relevance since democratic backsliding is a recent phenomenon that began in the early-to-mid-2000s.

To our knowledge, no fully updated dataset of votes on targeted UNHRC resolutions currently exists. Therefore, we scrape this information directly from the UNHRC

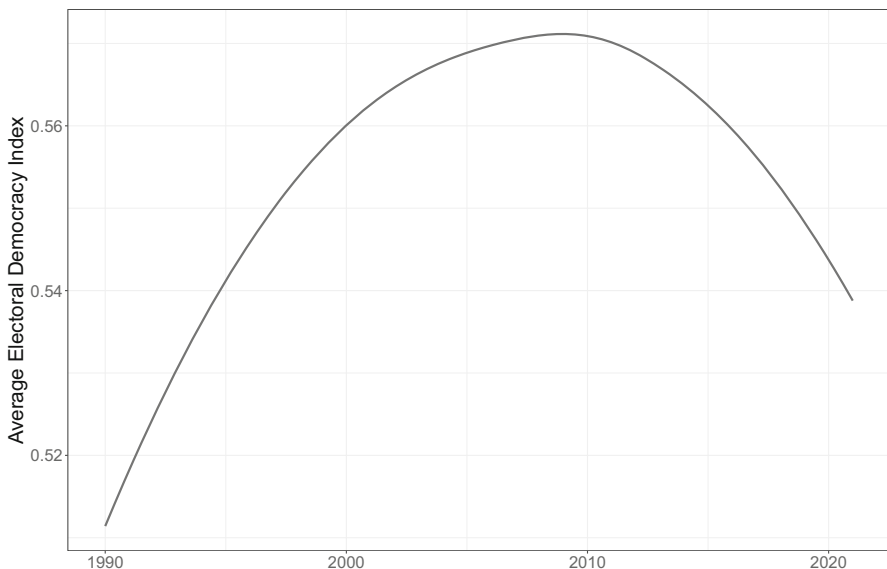


Fig. 1 Average electoral democracy index scores of UNHRC members by year. Note that the electoral democracy index is continuous between 0 and 1. Data source: Varieties of Democracy dataset (Coppedge et al., 2021)

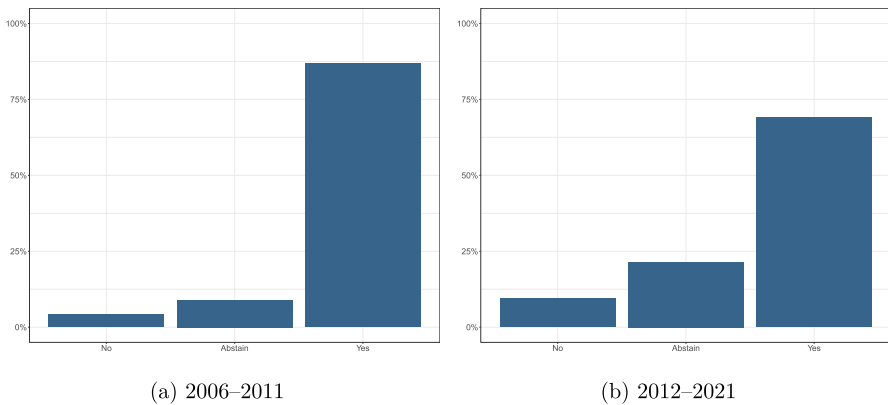


Fig. 2 Distribution of Votes in UNHRC, 2006–2021

online library for all resolutions from 2006 through 2021.⁵ The resulting dataset contains information on how each member of the UNHRC voted on any given targeted resolution⁶ and the state that is the subject of the resolution. Therefore, our dataset contains observations at the resolution-UNHRC member state level of analysis and consists of 13,721 observations covering 120 unique resolutions for which votes were recorded and adopted by the UNHRC targeting 18 states between 2006 and 2021.⁷

We predict that backsliding states will be less likely to vote in favor of targeted UNHRC resolutions than their more democratic counterparts. Figure 2 shows that between 2006 and 2021 “yes” votes are by far the most common type of vote. However, when we compare votes in the period before backsliding became more common (2006–2011) to the more recent backsliding era of 2012–2021, we see a marked increase in the proportion of “no” and “abstain” votes in the UNHRC. Figure 3 illustrates this starkly and plots the percent of “yes” votes in the UNHRC per year, making clear the precipitous decline in consensus on resolutions that begins around the same time when backsliding became increasingly common both globally and also among UNHRC members, as illustrated in Fig. 1 above.

The dependent variable in our analysis is a UNHRC member state’s vote choice on any given targeted resolution. Since UNHRC membership rotates, not all states vote on targeted resolutions, and there is variation among those who do with respect to the number of resolutions on which they vote. Figure 4 shows the number of targeted resolutions each UNHRC member state voted on between 2006 and 2021.⁸ During this time period, 131 states cast a vote on at least one targeted resolution.

⁵ <https://searchlibrary.ohchr.org/search?ln=en&cc=Voting>. Last accessed 9-March-2023.

⁶ The possible votes are “yes,” “no,” or “abstain”.

⁷ The states targeted in these resolutions, and the number of times they were targeted, are: Belarus (10), Burundi (7), Democratic People’s Republic of Korea (1), Democratic Republic of the Congo (1), Eritrea (3), Ethiopia (1), Georgia (4), Iran (7), Israel (29), Myanmar (6), Nicaragua (3), Philippines (1), South Sudan (1), Sri Lanka (2), Syria (30), Ukraine (6), Venezuela (5), and Yemen (3).

⁸ Note that the unit of observation for Fig. 4 is UNHRC member state-resolution-vote.

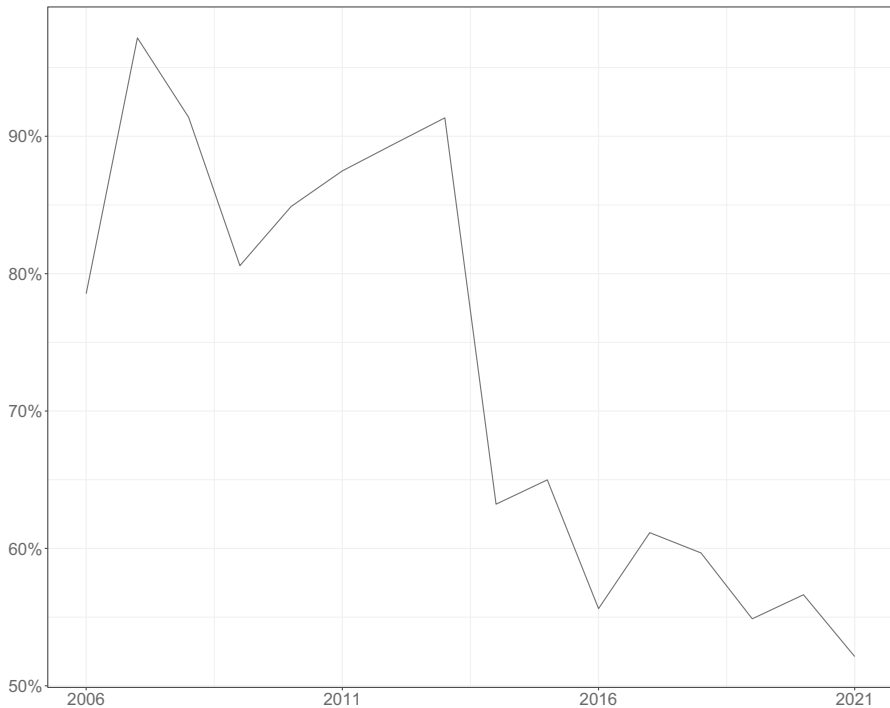


Fig. 3 Percent of yes votes in the UNHRC by year, 2006–2021

Our main independent variable is a binary indicator for whether or not the voting state has shown evidence of democratic backsliding. To create this variable, we use the electoral democracy index (EDI) from the Varieties of Democracy (V-Dem) dataset (Coppedge et al., 2021). The EDI captures the extent to which a state meets Dahl (1971)'s definition of polyarchy, defined by the presence of universal suffrage, free and fair elections, varying sources of information, and freedom of expression and association. The EDI is a continuous variable taking values between 0 and 1, with higher values indicating more democratic states; this indicator has been used in a number of recent studies to capture backsliding (Jee et al., 2022). As noted above, we define backsliding as a process that uniquely occurs (or at least begins) in democratic states; therefore, as a starting point we identify all countries that were democracies in 1995; any country with an EDI score greater than 0.5⁹ in 1995 is coded as a democracy in that year.¹⁰ Among these democracies, we identify cases that have experienced

⁹ The 0.5 threshold on the EDI is frequently used in recent work on backsliding to distinguish democracies from autocracies (e.g., Haggard and Kaufman 2021).

¹⁰ We choose 1995 as our starting point for several reasons. First, 1995 roughly corresponds with the end of the third wave of democracy (Huntington, 1991; Alizada et al., 2022). Furthermore, 1995 also marks the beginning of what scholars have termed the third wave of autocratization, which followed the third wave of democracy and continues into the present. One of the defining features of this third wave of autocratization is that, unlike previous waves, democracies are over-represented in the universe of countries moving toward autocracy (Luhmann and Lindberg, 2019); in other words, democratic backsliding is a core characteristic of the third wave of autocratization.

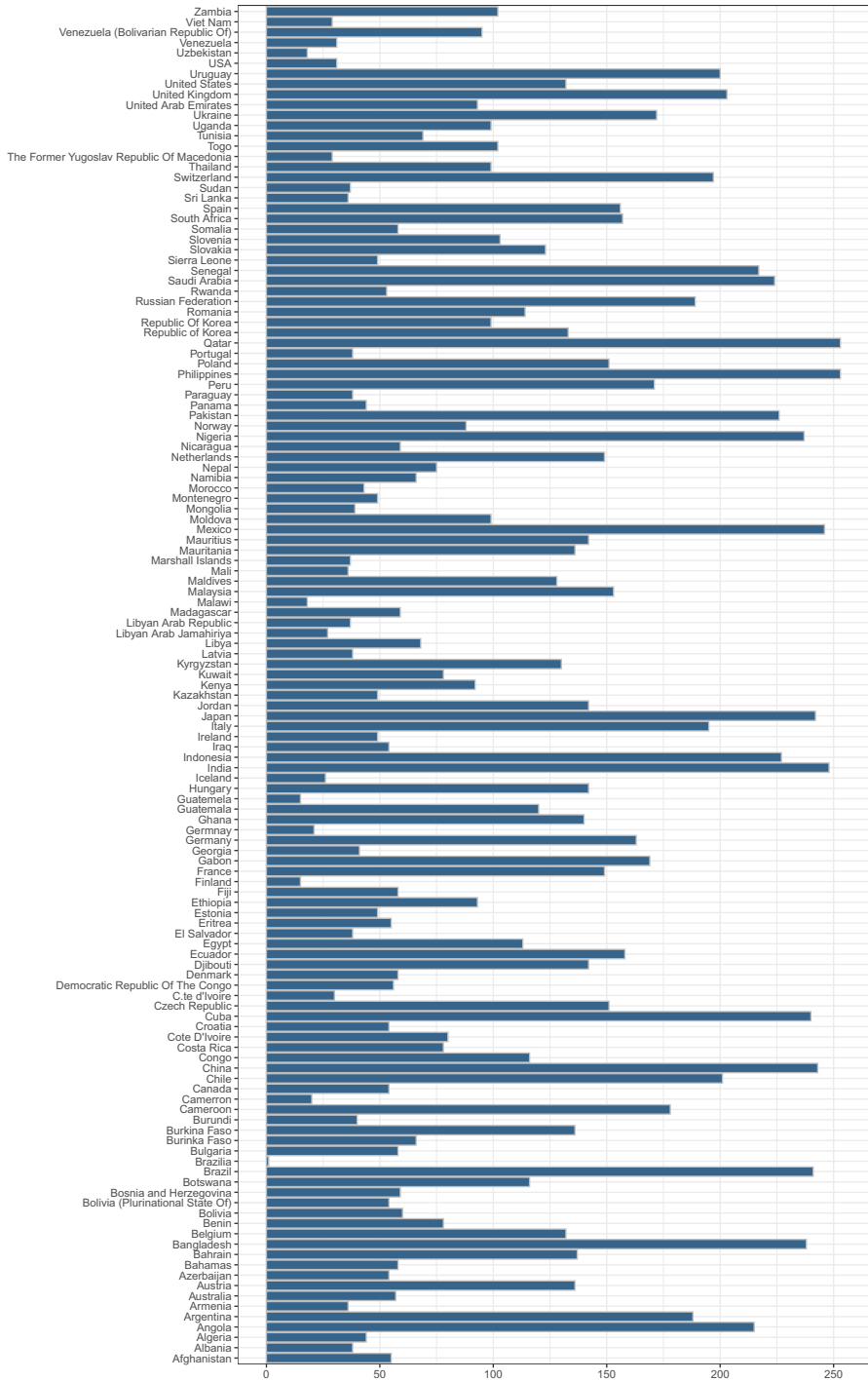


Fig. 4 Total number of targeted votes in the UNHRC per country, 2006–2021

democratic decline since 1995. Since small declines along the EDI may simply be due to measurement error or other idiosyncratic factors, we follow Luhrmann and Lindberg (2019) and set a threshold to identify substantial changes in a state's level of democracy. Specifically, we operationalize backsliding as any country-year observation for which a UNHRC member state's EDI has declined by 0.1 (i.e., 10%) or more since 1995. Our main independent variable is coded as 1 for all country-year observations where the voting state's EDI was above 0.5 in 1995 and has decreased by 0.1 or more since 1995; otherwise, it takes a value of 0.¹¹

We also include a relevant set of control variables for both the voting and target states. On the voting state side, we control for its current electoral democracy score, in addition to the backsliding indicator. While we expect backsliding to make a state more likely to vote against targeted UNHRC resolutions, backsliding can result in an illiberal democratic, semi-autocratic, or fully autocratic regime; this endpoint is likely also relevant for predicting how a state will vote. We also control for the proportion of UN General Assembly (UNGA) votes on which the voting state agreed with the US in a given year; the UNGA agreement score between the voting state and the target state (s-scores) (Bailey et al., 2017); and for whether the voting state is located in the same (UN) geographical region as the target state.

Our model includes relevant characteristics of the state targeted by a UNHRC resolution as well. First, we include an indicator for whether the resolution is targeted at Israel, as research suggests voting patterns against Israel in the UNHRC are distinct (Seligman, 2011). Similarly, since Syria was the subject of nearly as many UNHRC resolutions as Israel during this time period, we also include an indicator for resolutions targeting Syria. In addition, we capture the target state's current electoral democracy score, its agreement score with US votes in the UNGA, and its human rights score (Herre and Roser, 2016).¹² Summary statistics for all variables included in our analysis are reported in Appendix 2.

To test the extent to which backsliding influences voting behavior on targeted UNHRC resolutions, we estimate a multinomial logistic regression model with year fixed effects. As noted above, the voting outcome can take one of three values: "yes," "abstain," or "no." We designate "yes" votes as the baseline category for the dependent

¹¹ The following UNHRC member states meet these criteria for backsliding in one or more of the years between 2006 and 2021: Bangladesh, Bolivia, Brazil, Ecuador, Fiji, Hungary, India, Madagascar, Nicaragua, Philippines, Poland, Sri Lanka, Ukraine, and Venezuela. While our rationale for using 1995 as a baseline to identify these cases of backsliding states in the UNHRC is related to the end of the third wave of democracy and the start of the subsequent wave of autocratization, there is no single year that clearly marks the end of the former and beginning of the latter. We explore the effect of using 1995 as our baseline year in the Appendix, available on the *Review of International Organizations'* webpage. Specifically, in Appendix Figure A1 we show the democratic trajectory of the states we identify as cases of backsliding among UNHRC members from 1985 until 2021, finding that overall these states' levels of democracy increased in the 10 years prior to 1995. Furthermore, with a few exceptions, these states experienced steady, monotonic declines in their democracy scores sometime after 1995. In Appendix Table A1 we also show what our universe of cases of backsliding members of the UNHRC would be if we set the baseline years as 2000 and 2006, the year in which the UNHRC was created. Here, we find that most of the cases of backsliding we identify with the 1995 baseline are also coded as backsliding with these alternative baseline years. This gives us further confidence in our choice of 1995 as the baseline.

¹² While many studies use data on human rights violations from the CIRI human rights data project, those data only go through 2011. We use the Our World in Data dataset, which provides coverage through 2019.

Table 1 Democratic backsliding and vote choice in the UNHRC, 2006–2021

	<i>Dependent variable: Vote choice</i>	
	Abstain	No
Backslide since 1995	−0.08 (0.12)	0.76*** (0.13)
Voting EDI score	−0.36*** (0.14)	−2.67*** (0.21)
Israel	0.93*** (0.22)	1.79*** (0.36)
Syria	0.51*** (0.10)	0.07 (0.14)
Target EDI score	2.00*** (0.33)	2.00*** (0.50)
Voting and target in same region	−0.92*** (0.13)	−0.63*** (0.16)
Voting UNGA US agree	−0.43 (0.27)	2.51*** (0.43)
Target UNGA US agree	−0.99*** (0.36)	−2.55*** (0.61)
Target-voting UNGA affinity	3.11*** (0.23)	4.25*** (0.35)
Target state HR score	0.27*** (0.05)	0.03 (0.07)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Multinomial logit	Multinomial logit
Observations	11,809	11,809
K	3	3
Akaike Inf. Crit.	12,932.72	12,932.72

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

variable in our model and report the results in Table 1. Although we estimate a single model, we report the log-odds for abstentions in the first column, and the same for “no” votes in the second. Since most states vote on multiple resolutions, the observations in our data are not independent. Therefore, we also cluster standard errors by voting state.¹³

Overall, we find evidence that backsliding states are more likely to vote against targeted UNHRC resolutions. Specifically, we find that, for a democracy that has backslid since 1995, its log-odds of voting “no” relative to “yes” increase by 0.76. To

¹³ Since several of our control variables contain missing data, only 11,809 of the total 13,721 recorded votes on targeted UNHRC resolutions we discussed above are represented in the model reported in Table 1.

facilitate interpretation, we also calculate the relative risk of a backsliding state voting “no” rather than “yes” by exponentiating the reported coefficient: a backsliding state is over 2 times more likely to vote “no” as opposed to “yes” on a targeted UNHRC resolution.¹⁴

As we would expect, the results of this model also suggest that less democratic states, which typically exhibit weaker commitments to protecting human rights, are more likely to both vote “no” and “abstain,” rather than “yes,” on targeted UNHRC resolutions. UNHRC members during this time period are also more likely to vote against or abstain from votes targeting Israel, while resolutions concerning Syria were more likely to draw abstentions, but not “no” votes. We also find that, as the democracy level of the target state increases, UNHRC members are more likely to abstain or vote against a resolution, suggesting a possible bias that favors more democratic states when evaluating human rights practices. Interestingly, when the voting and target state are in the same geographic region, abstentions and “no” votes become less likely, suggesting states are more willing to punish their neighbors for alleged domestic human rights abuses. However, UNHRC members are more likely to abstain and vote against resolutions against states with which they share greater UNGA voting affinity. Furthermore, states more closely aligned with the US in the UNGA are less likely to abstain or vote “no,” while resolutions targeting states more closely aligned with the US are more likely to garner abstentions or opposition, suggesting US alignment provides some level of protection from sanction. Finally, states are more likely to abstain on resolutions against states with better overall human rights scores.

Since the magnitude of backsliding may influence how a state votes on targeted UNHRC resolutions, we also re-estimate the model from Table 1 using a continuous, rather than binary, variable to capture backsliding. When backsliding is operationalized as the change in a state’s EDI since 1995, we again find that backsliding states are more likely to vote against and less likely to abstain relative to the baseline “yes.”¹⁵

Finally, it is possible that there is a regional and income bias in our sample of backsliding states in the UNHRC. Asian-Pacific and Latin American countries, along with states that are classified as low or low-middle income countries, are over-represented in this group; therefore, the behavior of backsliding states may merely proxy for how low and middle income states from these regions would vote on UNHRC resolutions, regardless of their backsliding status. To ensure this is not the case, we re-estimate the models from Table 1 controlling for the GDP per capita of the voting state. In addition, we include fixed effects for the geographic region of the voting state. The results, reported in Appendix 5, confirm our findings: controlling for the region and income of the voting country, backsliding states remain significantly more likely to vote “no,” as opposed to “yes,” on targeted UNHRC resolutions.

¹⁴ We define backsliding as a process unique to democracies in that the starting point for any case of backsliding is within a democratic state. However, the literature recently has focused more broadly on autocratization, which is defined as any move away from democracy in both democracies *and* autocracies (Luhmann and Lindberg, 2019). When we operationalize backsliding as a 10% or greater shift away from democracy in any state, regardless of starting point, we find similar results. See Appendix 3.

¹⁵ The opposite signs on voting ‘no’ and ‘abstaining’ is mechanical since states can only vote once. As Fig. 2 above illustrates, what is happening since 2012 is that states are moving away from voting “yes.” Backsliding states in particular are now more likely to vote “no.” See also Appendix 4.

To summarize, Table 1 provides support for our expectations regarding the behavior of backsliding states as outlined in Hypothesis 1: backsliding democracies are more likely to vote against targeted UNHRC resolutions. This effect is robust to alternate definitions and thresholds of democratic backsliding. We next test the extent to which these states use the UN's human rights mechanisms to challenge the western liberal democracies who work to promote and sustain the liberal international order more directly.

4.2 Targeting advanced democracies in the UPR

In 2008 the UN introduced the Universal Periodic Review (UPR) process. Under the UPR, *all* UN member states, rather than simply ones identified by members of the UNHRC, are subject to periodic reviews of their domestic human rights practices. Furthermore, unlike most functions in the UNHRC, which are subject to regional affiliations and North-South conflicts, the UPR process emphasizes bilateral state-to-state relations (McMahon and Ascherio, 2012). We take advantage of this feature of the UPR to observe more directly the individual behavior of backsliding states.

All UN member states are reviewed every four and a half years; 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. In response, all UN members have the option to provide feedback to the state under review in the form of zero, one, or more specific recommendations about how to remedy any identified human rights issues. The state under review then publicly states whether they choose to "accept" or "take note" of these recommendations; the latter effectively indicates they reject the suggestion, though this term is not explicitly used. These proceedings are all recorded in an outcome report, and the state under review is expected to address the issues it "accepted" before its next review four and a half years later (Cox, 2010; Terman and Byun, 2022).

Data on the content of these outcome reports are publicly available online.¹⁶ We use these to create a dataset that consists of 32,598 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. Similar to the UNHRC (Fig. 1), we find the average democracy score for states writing UPR recommendations has declined substantially over time: while in 2008 the average electoral democracy score for recommending states in the UPR was 0.71, by 2020 that number had fallen by more than 10% to 0.59.¹⁷ Reflecting this trend, as Fig. 5 illustrates, the percent of UPR reviewers that meet our criteria for cases of backsliding has increased steadily over time from 3.7% in 2008 to 9.8% in 2020.

We use these UPR data to investigate how backsliding states use UN human rights tools differently than their more stable, non-backsliding counterparts. Since all states are subject to UPR reviews, we begin by examining the extent to which these back-

¹⁶ <https://upr-info-database.uwazi.io/en/>. Last accessed 9-March-2023.

¹⁷ See Figure A2 in Appendix 6.

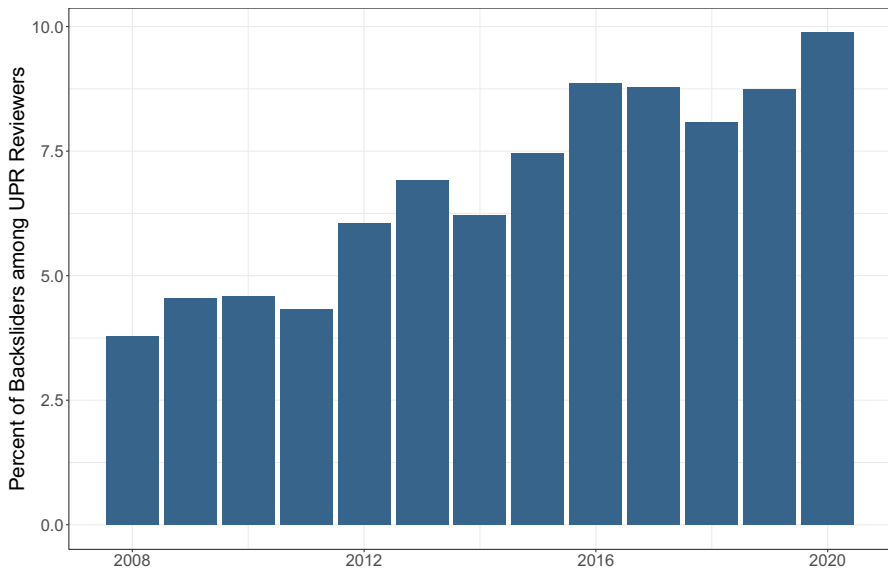


Fig. 5 The percent of UPR reviewers that qualify as cases of backsliding has steadily increased over time

sliding states use their UN membership to confront western liberal states, namely, by testing if they use the recommendations they make through the UPR's interactive dialogue portion of the review process to place more scrutiny and pressure on consolidated liberal democracies regarding their own human rights practices.

We focus on a subset of our UPR data. Flores and Nooruddin (2016, 85) define developing 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. Since we are interested in how backsliding and non-liberal democratic states use the UPR tool to pressure and scrutinize consolidated democracies, we restrict our data to include only observations where the state under review is one of these 19 established democracies and the recommending state is *not* one of these 19 countries.

To test Hypothesis 2, our main dependent variable is a count of the number of issues a reviewing state identifies in the recommendation it makes to a state under review. There is a finite (though substantially long) list of issues that can be highlighted in these reviews.¹⁸ The number of issues in any given review in our subset of the data ranges from 0 to 19, with a mean of 3.49 issues identified per report. The number of issues identified in a report is a plausible proxy for how critical that report is; more critical reports will identify more issues in the state under review.

¹⁸ There are over 60 types of issues in our dataset, and they include topics such as: “death penalty,” “elections,” “extrajudicial executions,” “freedom of the press,” “labour rights,” “minority rights,” “trafficking,” and “women’s rights.”

As before, our main independent variable is a binary variable indicating whether the reviewing state was a democracy in 1995, and, if so, whether its level of democracy has deteriorated by 0.1 or more since 1995 for each year in our sample.¹⁹ Furthermore, we use the same control variables as those reported in Table 1, with several exceptions. First, we no longer control for whether the state under review is Israel or Syria, since only reviews of the 19 democracies listed above are included in our models. We also control for whether or not the recommending state is also under review by the UPR in the same cycle to account for any incentives this may create.²⁰

Since the count of issues identified in the UPR reports is under-dispersed,²¹ we employ a quasi-Poisson model as our primary specification. Furthermore, to account for potential selection bias, we follow Terman and Byun (2022) and estimate an additional model that includes only observations where the recommending state was concurrently a member of the UNHRC. Since only oral comments are recorded in UPR reports, states with more representatives in Geneva—including UNHRC member states that are already present and expected to participate in UPR sessions—may be over-represented among UPR recommending states. However, there is no reason to expect the types of recommendations made by UNHRC members to be substantively distinct from those of other states, all else equal. Each of these models include year fixed effects. Since any UN member states can provide recommendations for more than one state through the interactive dialogue process, the observations in our data are not independent. Therefore, the models also include standard errors clustered by recommending state. The results are reported in Table 2. Through these models, we show that backsliding states identify a greater number of issues in reports targeting consolidated western democracies.

Our UPR results are robust to alternative modeling specifications. First, models that use the continuous measure of backsliding discussed in Section 4.1 return similar results.²² Next, we consider the possibility that power differentials between the state writing a UPR recommendation and the state under review might influence the number of issues the former identifies in the latter. To test this explicitly, we re-estimate the models from Table 2, this time controlling for the relative power of both the recommending state and the state under review, measured using each state's CINC (Composite Index of National Capability) score.²³ As Appendix Table A8 shows, our main results are robust to controlling for relative power. More relevant to our theory, these models also suggest that states with lower CINC scores identify significantly more issues in their UPR reports on advanced democracies. This suggests that the

¹⁹ It is important to note that since all UN member states, rather than just members of the UNHRC, can write UPR reviews, the subset of cases of backsliding in these models is larger than in the UNHRC models, which only includes states that were members of the Council between 2006 and 2021. The following states meet our operationalization of backsliding and wrote at least one UPR review of an advanced democracy between 2008 and 2020: Bangladesh, Belarus, Benin, Bolivia, Brazil, Ecuador, Fiji, Honduras, Hungary, India, Mali, Mauritius, Nicaragua, Philippines, Poland, Sri Lanka, Turkey, Ukraine, Venezuela, and Zambia.

²⁰ Summary statistics for these data are reported in Appendix 7.

²¹ The dispersion parameters for models 1 and 2 in Table 2 are 0.61 and 0.54, respectively.

²² See Appendix 8.

²³ See Appendix 9.

Table 2 UPR issues identified against advanced democracies, 2008–2020

	<i>Dependent variable: Count of issues identified in country report</i>	
	All reports	Reports by UNHRC members
Backslide since 1995	0.16*** (0.02)	0.25*** (0.03)
Recommend EDI score	0.13*** (0.03)	0.42*** (0.04)
State under review EDI score	-0.37 (0.31)	0.30 (0.48)
Same region	0.24*** (0.02)	0.35*** (0.03)
Recommend UNGA US agree	-0.05 (0.06)	0.77*** (0.16)
State under review UNGA US agree	-0.11 (0.11)	-1.08*** (0.22)
Recommend-under review UNGA affinity	-0.90*** (0.07)	-1.91*** (0.20)
State under review HR score	0.03*** (0.01)	0.03*** (0.01)
Recommending state under review	-0.06*** (0.01)	-0.15*** (0.02)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Quasi-Poisson	Quasi-Poisson
Observations	6,301	2,654

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

UPR mechanism is being used by relatively weak states as a tool to challenge powerful, established democracies. Finally, as discussed in Section 4.1, we again consider the possibility of income and regional bias in our sample of backsliding states and re-estimate the models from Table 2 with controls for the GDP and geographic region of the recommending state and find similar results.²⁴

To summarize, we find evidence that backsliding states identify a greater number of issues in reports targeting consolidated western democracies. These results suggest that backsliding states, in addition to being more likely to vote against targeted resolutions in the UNHRC, also use the UPR mechanism deliberately to challenge the established liberal international order by placing western democracies under greater scrutiny for their human rights practices.

²⁴ See Appendix 10.

4.3 The content of UPR recommendations

We have shown thus far that backsliding states exhibit distinct behaviors in the UN human rights institutions: they are more likely to vote “no” on targeted UNHRC resolutions and they identify a greater number of human rights issues in their peer reviews of advanced western liberal democracies. Having established these distinct patterns of behavior, as an additional test of our theory we examine the extent to which the substance of backsliding states’ input into these human rights mechanisms reflects their changing and often increasingly illiberal domestic interests that directly contrast with the liberal norms underlying the existing international order.

Broadly, the rights outlined by the Universal Declaration of Human Rights can be divided into two categories. First generation human rights, as detailed in the International Covenant on Civil and Political Rights (ICCPR), emphasize individual civil and political rights—including the right to life, liberty, the freedom of movement, freedom from torture and other violations of physical integrity, and equal standing before the law—that historically have been advocated and supported by western states and are closely aligned with liberal democratic values. As such, in many ways these first-generation human rights reflect the values underlying the current international liberal order.

The International Covenant on Economic, Social, and Cultural Rights (ICESCR), on the other hand, focuses on second generation collective economic, social, and cultural rights of equality, such as the right to equal pay for equal work and universal access to education. While western states and, by extension, many consolidated liberal democracies, have long-emphasized first generation civil and political rights, members of the Global South and rising powers tend to focus instead on these second-generation collective social and economic issues. There is evidence that China and other states in the Global South have increasingly used their membership on the UNHRC as a means of advancing this alternative human rights agenda. These states are more likely to introduce UNHRC resolutions focused on issues such as the right to economic development, the protection of family, and the rights of minorities—including women, children, migrants, and racial minorities (Voss, 2020; Dukalskis, 2023).

Though equally legitimate and important in theory, in practice different human rights norms and topics vary in how sensitive and threatening they are to a state under review for its human rights practices. Norms that directly challenge the power or legitimacy of the target regime are more threatening and sensitive; these often fall under the umbrella of first generation human rights. For example, governments condemned for violating the physical integrity rights of their citizens face potential international sanctions, such as loss of foreign aid, as well as challenges to their domestic legitimacy and support (Lebovic and Voeten, 2009; Krain, 2012). Similarly, civil and political liberties are typically viewed as particularly sensitive human rights since promoting them could pose a challenge to incumbents by opening space for political opposition (Linz and Stepan, 1996; Haggard and Kaufman, 2016).

It is less costly to enforce second generation rights since these often do not pose a direct challenge to the target government. For example, socioeconomic rights that emphasize human development in the form of improved health care, poverty reduction, and educational access, pose little threat to regime stability and, in fact, most

regimes have an interest in promoting these second generation human rights (Teets, 2014). Similarly, women’s and children’s rights —such as protections against domestic violence and human trafficking— are typically viewed as “safe” norms since they do not concern abuses committed by the state (Terman and Byun, 2022). We expect that backsliding states, which are by definition in a transitional phase and seeking to consolidate their power, will seek to avoid challenges to their rule in part by de-emphasizing threatening norms in their reviews of other states, focusing instead on less sensitive human rights topics.²⁵

To test this expectation, we return to our UPR dataset. In addition to the number and type of issues identified and actions recommended in these reports, our dataset includes information on the recommendations a reviewing state makes to the state under review. Unlike the data on issues, these recommendations are free-form text responses. These allow us to examine the content and topics that backsliding states choose to emphasize in their reviews of other states’ human rights practices.²⁶

We draw on the recommendations contained in UPR reports to estimate a structural topic model (STM) with 6 topics.²⁷ The `stm` package generates common words associated with each topic identified by the model. Using these common words in combination with sample excerpts from the UPR review documents, we manually label each of the 6 topics. These topics and their corresponding key words are shown in Fig. 6.²⁸

In our STM model, we explore the types of issues backsliding states emphasize in their UPR reports. When estimating the model, we include the same set of independent variables that was included in Table 2. The results of this model are presented in Figs. 7 and 8.²⁹

Figure 7 plots the change in topic prevalence as our binary indicator for democratic backsliding changes from 0, which indicates a stable regime, to 1, which indicates a case of backsliding. The results in Fig. 7 provide additional support for our theory: backsliding states are significantly more likely to emphasize human development and the protection of minorities and vulnerable groups than their non-backsliding counterparts. They are also significantly less likely to advocate establishing national human rights institutions. However, the confidence intervals for these topics are relatively large, and the results for the other three topics identified in the model are not statistically significant. This is likely due to the fact that the absolute number of backsliding states represented among the universe of UPR recommending states is small by construction. For a different perspective, therefore, we explore the change in topic prevalence as a recommending state’s level of democracy declines along the

²⁵ See Winzen (2023) for a related argument in the EU context. Prasad and Nooruddin (2023) advance a similar argument in the context of states fighting domestic insurgency.

²⁶ We note that while in Section 4.2 we focus exclusively on UPR reports where the state under review is one of 19 established democracies, in the analyses below we include all available UPR reports, regardless of the state under review, to explore the types of human rights issues backsliding states emphasize.

²⁷ We use the `stm` package in R to tune the number of topics and estimate parameters. We select the number of topics (6) that maximizes semantic coherence and exclusivity (see Appendix Figure A3). We pre-process all text by eliminating common English stopwords, numbers, and punctuation. We also lowercase all text.

²⁸ See Appendix 12 for representative excerpts from UPR reports in each topic.

²⁹ Confidence intervals in both figures are at the 95% level.

Topic	Representative words
1. Human development	Education, poverty, rural, quality, empower, health, strengthen, improve
2. Ratify international human rights conventions	Convention, international, protocol, ratify, inhumane, death, degrade
3. Women and LGBTQ+ rights	Violence, women, sexual, fetus, marriage, transgender, lesbian, harm
4. Civil and political liberties	Ensure, freedom, right, law, independent, journalist, accountable, association, assembly, express
5. Establish national human rights institutions	Nation, human, right, institution, implement, mandate
6. Protect minorities and vulnerable groups	Discrimination, minor, roma, racism, xenophobia, intolerance, hate, ensure, asylum

Fig. 6 This figure shows the labels we assign to each topic in our model and the most common words appearing in each STM topic

electoral democracy index, rather than requiring that they meet our higher threshold of democratic backsliding; these results are more precisely estimated and are reported in Fig. 8.

As Fig. 8 shows, as a reviewing state's level of democracy declines, it becomes significantly more likely to focus on non-threatening human rights topics. While the positive and large coefficient on human development is particularly noteworthy, we also find that these states are more likely to reference issues related to protecting minorities and vulnerable groups and efforts to establish national human rights institutions. We think this emphasis on national, rather than international human rights institutions, reflects concerns related to sovereignty and growing aversion to international influence, both of which have been characteristics of contemporary backsliding regimes and their often populist leaders (Haggard and Kaufman, 2021).

Figure 8 also shows that these states are less likely to emphasize issues that pose a threat to government control and stability. As states' levels of democracy decline, they are significantly less likely to use the UPR mechanism to advocate for civil and political liberties often associated with western liberalism, such as freedom of expression, religion, and association, as well as rights for political opposition groups that could serve as a challenge to the ruling regime. These states are also significantly less likely to support ratifying international human rights conventions, which again may reflect issues related to sovereignty, and are less likely to advocate the protection

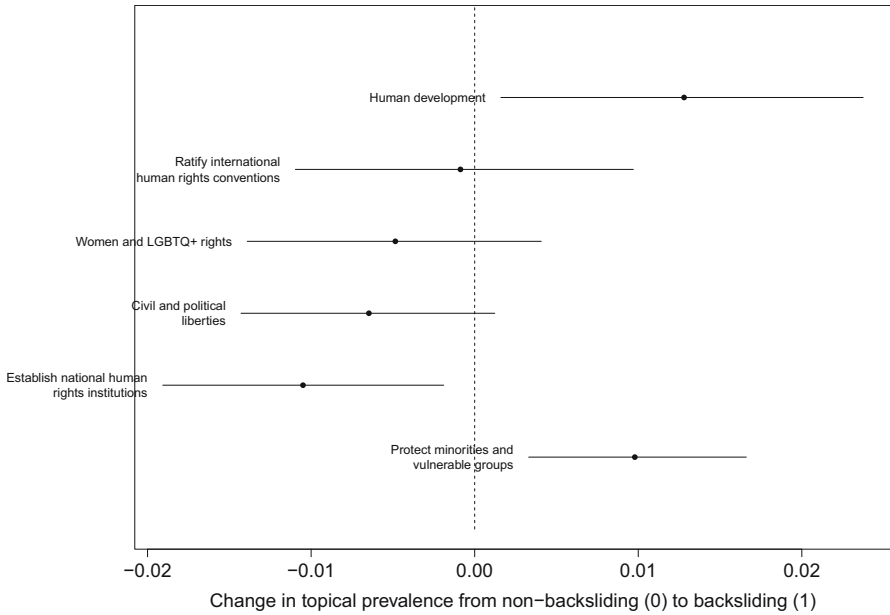


Fig. 7 Structural Topic Model Results (Binary Indicator of Backsliding). Substantively, the effect size corresponds to the change in the proportion of the text relevant to a given topic when our indicator for backsliding changes from 0 (not backsliding) to 1 (backsliding)

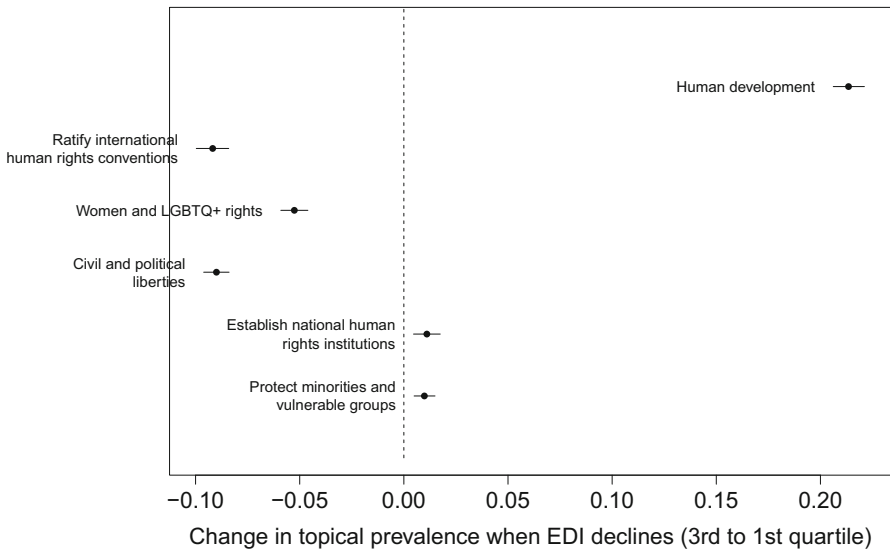


Fig. 8 Structural Topic Model Results (Continuous Regime Measure). Substantively, the effect size corresponds to the change in the proportion of the text relevant to a given topic when a recommending state's electoral democracy index declines from the 3rd quartile (0.87) to the 1st quartile (0.33)

of gender-based and LGBTQ+ rights. The latter is likely related to the fact that many contemporary cases of backsliding occur under the leadership of far-right, populist politicians who advocate returns to more traditional cultural and family values (Mudde Rovira Kaltwasser, 2017). Taken together, these results provide further suggestive evidence that backsliding states are using international human rights fora to promote their own interests and to counter efforts by democracies to spread liberal democratic norms and values.

5 Conclusion

The democratic upsurge that followed the collapse of the Soviet Union was accompanied by a broadening of membership of many western-backed IOs and their constitutive institutions, including the UNHRC, as these new and aspiring democracies were brought into the fold. Our findings suggest that the now backsliding states among these recent democracies are distinct from their long-time non-democratic counterparts, and thus present a historically unique challenge for liberal IOs. In the Cold War era, the countries that were non-democracies had never before been functioning democracies, and many enjoyed the patronage of either the United States or the Soviet Union that granted them immunity from international criticism. Today's backsliders, on the other hand, are defensive of their sovereignty precisely because they were once democracies and are conscious therefore of their vulnerability to criticism both domestically and internationally for their democratic regression. We argue—and believe our data show—that this set of states therefore behaves differently than both their stable democratic and non-democratic counterparts.

We present evidence that backsliding states' representation in the United Nations' human rights organizations has increased over time, and find that they exhibit distinct behavior. First, we show that backsliding states are more likely to vote against targeted UNHRC resolutions. We also draw on data from the Universal Periodic Review to demonstrate how these states are more actively using the UN human rights institutions to challenge consolidated liberal democracies by identifying a larger number of issues in their UPR reports on these countries. Furthermore, the types of recommendations they make in their UPR reports illustrate how these states use existing international institutions to promote policies and norms more in line with their domestic environments and interests. Specifically, we show backsliding states are more likely to emphasize issues such as human development, while de-emphasizing topics that might threaten their power and control over citizens, such as civil and political liberties, in their UPR reports.

Taken together, our findings hold several implications for the functioning and efficacy of the UN's human rights institutions, and suggest the possibility that similar dynamics could be at play in other western liberal IOs whose members are similarly backsliding. First, our findings regarding voting behavior in the UNHRC suggest recent global backsliding has impacted the efficacy of this institution, most concretely by decreasing the unanimity of UNHRC resolutions (see Fig. 3 above). Passing resolutions with less than unanimous support risks diminishing the amount of pres-

sure target states feel to implement reforms. At a more fundamental level it also undermines the legitimacy of the UNHRC and the values it was created to uphold. Our findings illustrate the rise of illiberalism in IOs and its consequence for international cooperation.

The ways in which we show that these backsliding states strategically leverage recommendations in their UPR reports are perhaps even more troubling. The evidence suggests these backsliding states are not only using their voting rights to decrease the efficacy and legitimacy of existing institutions, but are also more actively working to undermine prevailing fundamental conceptions of human rights that are centered on civil and political liberties, and replacing them with an alternative set of second-generation economic rights—such as human development—that better align with their domestic, illiberal interests. The importance of economic rights is not debatable, but that their apparent champions are the backsliding states raises the question of the sincerity of this position. Our inclination is to see the preponderance of evidence presented in this paper as suggesting that such espousal is more cynical than sincere, and that we should reject the false binary between political and economic rights. To the extent that the efforts of backsliding states succeed, the past three decades of progress made in expanding core first generation human rights, such as individual freedoms, could potentially be undone.

We identify several areas for additional research. First, future studies should explore the extent to which these dynamics we identify within the UN human rights institutions are at play in other liberal IOs. To what extent, for example, have backsliding states been able to leverage their membership privileges in historically democratic regional organizations to undermine political and liberal aspects of these institutions from within? How, if at all, are these IOs in turn working to counteract actions being taken by their now backsliding member states? Furthermore, additional research is needed to evaluate whether backsliding states' actions within existing liberal IOs are having the effects we argue they desire. This requires a better understanding of who these backsliding states view as their intended audiences when they employ these tactics, as well as the effect these efforts have on these audiences. From whom do these backsliding states seek legitimacy? Do these backsliding states face any costs for their subversive behavior?

Our evidence that backsliding states use the UN human rights institutions to promote a different set of human rights also points to the underlying tensions between liberalism and democracy. While liberalism as embodied by the post-Cold War liberal international order promotes and supports a distinct set of values and perspectives, an idealized democracy provides a venue that allows for contestation over competing ideas; in this way, the liberal international order and democracy may not match but instead merely “rhyme.”³⁰ While our focus in this paper takes the existing order as given and shows evidence of the ways in which backsliding states seek to challenge and alter it, this of course raises interesting normative questions regarding the Eurocentric bias underpinning the current world order that merit greater consideration in future research on international consequences and even conceptualizations of democratic backsliding.

³⁰ We owe this turn of phrase to an anonymous reviewer.

The implications of our findings should concern supporters of the LIO agenda on human rights. The absence of any credible global leadership on these issues, coupled with the growing presence of autocracies and problematic democracies within the very IOs charged with protecting rights, threatens to spell the end of an era in which human rights were a meaningful topic of discussion internationally. Of particular concern is the inability of western states to tackle democratic backsliding within their own borders and in their strategic partners globally. At least at this moment, as global backsliding trends show little sign of reversal, it seems likely that these states will continue to use their established positions in these IOs to shape the international order to better suit their illiberal goals.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11558-023-09511-6>.

Acknowledgements We thank Christina Cottiero, Emilie Hafner-Burton, Stephan Haggard, Lauren Prather, Christina Schneider, and all participants at two workshops hosted by the UC Institute for Global Conflict and Cooperation at the University of California, San Diego for invaluable detailed comments and suggestions, as well as the support and encouragement to develop this project. We also thank participants at the ISA 2023, PEIO 2023, and APSA 2023 conferences and three anonymous reviewers for their helpful comments on previous drafts.

Author Contributions The order of authors reflects the significance of the authors' contributions
 a. research design and conceptualization: A.M. (55%) I.N. (45%)
 b. statistical analysis: A.M. (60%), I.N. (40%)
 c. writing: A.M. (75%), I.N. (25%)

Data Availability Statement The datasets generated and analyzed by the authors for the current study are available in the Dataverse repository,

Declarations

Conflict / Competing Interests The authors have no relevant financial or non-financial interests to disclose. The authors have no competing interests to declare that are relevant to the content of this article. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

References

- Alizada, N., Boese, V.A., Lundstedt, M., Morrison, K., Natsika, N., Sato, Y., Tai, H., Lindberg, S.I., Gastaldi, L., Grahn, S., Hindle, G., Ilchenko, N., Pernes, J., & von Römer, J. (2022). *Democracy Report 2022: Autocratization Changing Nature?* V-Dem Institute: University of Gothenburg.
- Bailey, M. A., Strezhnev, A., & Voeten, E. (2017). Estimating dynamic state preferences from United Nations voting data. *Journal of Conflict Resolution*, *61*, 430–456.
- Barnett, M., & Finnemore, M. (2004). *The Power of Liberal International Organizations*. In *Power in Global Governance*. Cambridge University Press.
- Bartels, L. M., Daxecker, U. E., Hyde, S. D., Lindberg, S. I., & Nooruddin, I. (2023). The Forum: Global Challenges to Democracy? Perspectives on Democratic Backsliding. *International Studies Review*, *25*(2), viad019.

- Baturo, A. (2023). Autocrats and the Moral High Ground: Russia and the Politics of Anti-Nazism in International Organizations. Working paper.
- Bermeo, N. (2016). On Democratic Backsliding. *Journal of Democracy*, 27(1), 5–19.
- Binder, M., & Payton, A. L. (2022). With frenemies like these: Tising power voting behavior in the UN general assembly. *British Journal of Political Science*, 52(1), 381–398.
- Börzel, T. A., & Zürn, M. (2021). Contestations of the liberal international order: From liberal multilateralism to postnational liberalism. *International Organization*, 75(2), 282–305.
- Bush, S., Cottiero, C., & Prather, L. (2023). Zombies Ahead: Explaining the Rise of Low-Quality Election Monitoring. Working paper.
- Carothers, T. (2002). The End of the Transition Paradigm. *Journal of Democracy*, 13, 5–21.
- Carraro, V. (2017). The United Nations Treaty Bodies and Universal Periodic Review: Advancing Human Rights by Preventing Politicization? *Human Rights Quarterly*, 39(4), 943–970.
- Carraro, V., Conzelmann, T., & Jongen, H. (2019). Fears of peers? Explaining peer and public shaming in global governance. *Cooperation and Conflict*, 54, 335–355.
- Collier, D., & Levitsky, S. (1997). Democracy with Adjectives: Conceptual Innovation in Comparative Research. *World Politics*, 49(3), 430–451.
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Alizada, N., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Hicken, A., Hindle, G., Ilchenko, N., Krusell, J., L uhrmann, A., Maerz, S. F., Marquardt, K. L., McMann, K., Mechkova, V., Medzihorsky, J., Paxton, P., Pemstein, D., Pernes, J., von Römer, J., Seim, B., Sigman, R., Skaaning, S.-E., Staton, J., Sundström, A., Tzelgov, E., Wang, Y., Wig, T., Wilson, S., & Ziblatt, D. (2021). V-Dem Country-Year/Country-Date Dataset v11. Varieties of Democracy (V-Dem) Project.
- Cottiero, C., & Haggard, S. (2023). Stabilizing Authoritarian Rule: The Role of International Organizations. *International Studies Quarterly*, 67(2), sqad031.
- Cox, E. (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.
- Dahl, R. (1971). *Polyarchy: participation and opposition*. Yale University Press.
- Debre, M. J. (2021). Clubs of autocrats: Regional organizations and authoritarian survival. *Review of International Organizations*, 17(3), 485–511. <https://doi.org/10.1007/s11558-021-09428-y>
- Debre, M. J., & Morgenbesser, L. (2018). Out of the shadows: Autocratic regimes, election observation and legitimation. In *Justifying Dictatorship*, pp. 78–97. Routledge.
- Donno, D. (2010). Who Is Punished? Regional Intergovernmental Organizations and the Enforcement of Democratic Norms. *International Organization*, 64(04), 593–625.
- Donno, D. (2013). *Defending Democratic Norms: International Actors and the Politics of Electoral Misconduct*. Oxford: Oxford University Press.
- Donno, D., Morrison, K., Savun, B., & Davutoglu, P. (2023). Competing Verdicts: Multiple Election Monitors and Post-Election Contention. Working paper.
- Dukalskis, A. (2023). Afox in the henhouse: China, normative change, and the UN Human Rights Council. *Journal of Human Rights*, 22(3), 334–350. <https://www.tandfonline.com/doi/full/10.1080/14754835.2023.2193971>.
- Edwards, M. S., Scott, K. M., Allen, S. H., & Irvin, K. (2008). Sins of commission? Understanding membership patterns on the United Nations human rights commission. *Political Research Quarterly*, 61, 390–402.
- Finnemore, M., & Sikkink, K. (1998). International Norm Dynamics and Political Change. *International Organization*, 52(4), 887–917.
- Flores, T. E., & Nooruddin, I. (2023). Why incumbents perpetrate election violence during civil war. *Conflict Management and Peace Science*, 40(5), 533–553.
- Flores, T., & Nooruddin, I. (2016). *Elections in Hard Times: Building Stronger Democracies in the 21st Century*. Cambridge University Press.
- Fukuyama, F. (1989). The End of History? *The National Interest*.
- Genna, G. M., & Hiroi, T. (2014). *Regional Integration and Democratic Conditionality: How Democracy Clauses Help Democratic Consolidation and Deepening*. Routledge.
- Ginsburg, T. (2020). Authoritarian international law? *American Journal of International Law*, 114, 221–260.
- Hafner-Burton, E., Pevehouse, J., & Schneider, C. (2023). Enlightened Autocrats? Good Governance in Autocratic International Organizations. Working paper.
- Hafner-Burton, E. M. (2005). Trading Human Rights: How Preferential Trade Agreements Influence Government Repression. *International Organization*, 59, 593–629.

- Hafner-Burton, E. M. (2008). Sticks and Stones: Naming and Shaming the Human Rights Enforcement Problem. *International Organization*, 62, 689–716.
- Hafner-Burton, E. M. (2012). International regimes for human rights. *Annual Review of Political Science*, 15, 265–286.
- Hafner-Burton, E. M., & Tsutsui, K. (2005). Human Rights in a Globalizing World: The Paradox of Empty Promises. *American Journal of Sociology*, 110(5), 1373–1411.
- Haggard, S., & Kaufman, R. (2021). *Backsliding: Democratic Regress in the Contemporary World*. Cambridge University Press.
- Haggard, S., & Kaufman, R. R. (2016). Democratization During the Third Wave. *Annual Review of Political Science*, 19(8), 1–20.
- Herre, B., & Roser, M. (2016). Human Rights. *Our World in Data*. <https://ourworldindata.org/human-rights>. Accessed 9 March 2023.
- Hug, S., & Lukács, R. (2014). Preferences or blocs? Voting in the United Nations Human Rights Council. *Review of International Organizations*, 9, 83–106.
- Huntington, S. P. (1991). *The Third Wave: Democratization in the Late Twentieth Century*. University of Oklahoma Press.
- Hyde, S. D. (2011). *The Pseudo-Democrat's Dilemma: Why Election Observation Became An International Norm*. Cornell University Press.
- Ikenberry, G. J. (2018). The end of liberal international order? *International Affairs*, 94, 7–23.
- Inboden, R. Siu. (2021). *China and the International Human Rights Regime*. Cambridge University Press.
- Jee, H., Lueders, H., & Myrick, R. (2022). Towards a unified approach to research on democratic backsliding. *Democratization*, 29, 754–767.
- Johnston, A. I. (2007). *Social States: China in International Institutions, 1980–2000*. Princeton University Press.
- Johnstone, I., & Snyder, M. (2016). Democracy promotion. In *The Oxford Handbook of International Organizations*. Oxford University Press.
- Kelemen, R. D. (2020). The European Union's authoritarian equilibrium. *Journal of European Public Policy*, 27, 481–499.
- Kim, Y. (2023). A network analysis of naming and shaming in the universal periodic review. *International Interactions*, 49(3), 287–314.
- Kneuer, M., Demmelhuber, T., Peresson, R., & Zumbrägel, T. (2019). Playing the regional card: why and how authoritarian gravity centres exploit regional organisations. *Third World Quarterly*, 40, 451–470.
- Koliev, F. (2020). Shaming and democracy: Explaining inter-state shaming in international organizations. *International Political Science Review*, 41, 538–553.
- Krain, M. (2012). J'accuse! Does Naming and Shaming Perpetrators Reduce the Severity of Genocides or Politicides? *International Studies Quarterly*, 56, 574–589.
- Lakatos, I. (2022). The United Nations Human Rights Council: A Platform for the Clash of Civilizations or a Forum of Cross-Regional Cooperation? In *Comparative Human Rights Diplomacy*. Springer International Publishing pp. 335–363.
- Landman, T. (2005). The Political Science of Human Rights. *British Journal of Political Science*, 35(3), 549–572.
- Lebovic, J. H., & Voeten, E. (2006). The politics of shame: The condemnation of country human rights practices in the UNCHR. *International Studies Quarterly*, 50, 861–888.
- Lebovic, J. H., & Voeten, E. (2009). The cost of shame: International organizations and foreign aid in the punishing of human rights violators. *Journal of Peace Research*, 46, 79–97.
- Levitsky, S., & Way, L. A. (2010). *Competitive Authoritarianism: Hybrid Regimes After the Cold War*. Cambridge University Press.
- Libman, A., & Obydenkova, A. V. (2018). Regional international organizations as a strategy of autocracy: The Eurasian Economic Union and Russian foreign policy. *International Affairs*, 94, 1037–1058.
- Linz, J., & Stepan, A. (1996). *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and Post-Communist Europe*. Johns Hopkins University Press.
- Lipps, J., & Jacob, M. S. (2023). Undermining Liberal International Organizations from Within: Evidence from the Parliamentary Assembly of the Council of Europe. Working paper.
- Luhmann, A., & Lindberg, S. I. (2019). A third wave of autocratization is here: what is new about it? *Democratization*, 26, 1095–1113.
- McMahon, E., & Ascherio, M. (2012). A Step Ahead in Promoting Human Rights? The Universal Periodic Review of the UN Human Rights Council. *Global Governance*, 18, 231–248.

- Meyerrose, A. M. (2018). It is all about value: How domestic party brands influence voting patterns in the European Parliament. *Governance*, 31(4), 625–642.
- Meyerrose, A. M. (2020). The Unintended Consequences of Democracy Promotion: International Organizations and Democratic Backsliding. *Comparative Political Studies*, 53(10–11), 1547–1581.
- Meyerrose, A. M. (2023). Building Strong Executives and Weak Institutions: How European Integration Contributes to Democratic Backsliding. *Review of International Organizations*.
- Meyerrose, A. M., Flores, T. E., & Nooruddin, I. (2019). From Elections to Democracy in Hard Times. In *Oxford Research Encyclopedia of Politics*. Oxford University Press.
- Moravcsik, A. (2000). The Origins of Human Rights Regimes: Democratic Delegation in Postwar Europe. *International Organization*, 54(2), 217–252.
- Mudde, C. & Kaltwasser, C.R. (2017). *Populism: A very short introduction*. Oxford University Press.
- United Nations Charter (1945). United Nations, Office of Public Information.
- Neumayer, E. (2005). Do International Human Rights Treaties Improve Respect for Human Rights? *Journal of Conflict Resolution*, 49(6), 925–953.
- Pahre, R. D. (1995). Wider and deeper: The links between expansion and integration in the European Union (pp. 111–136). In *Towards a new Europe: Stops and starts in regional integration*. Praeger.
- Pevehouse, J. (2005). *Democracy from above: Regional organizations and democratization*. Cambridge University Press.
- Poast, P., & Urpelainen, J. (2018). *Organizing democracy: How international organizations assist new democracies*. University of Chicago Press.
- Prasad, S. K., & Nooruddin, I. (2023). States living in glasshouses....: Why fighting domestic insurgency changes how countries vote in the UN human rights council. *Conflict Management and Peace Science*, 0(0). <https://doi.org/10.1177/07388942231198489>
- Risse, T., & Sikkink, K. (1999). The Power of Human Rights: International Norms and Domestic Change. In Risse, T., Sikkink, K., & Ropp, S. C. (Eds.), *The Socialization of International Human Rights Norms into Domestic Practices*. Cambridge University Press.
- Seligman, S. (2011). Politics and principle at the UN human rights commission and council (1992–2008). *Israel Affairs*, 17, 520–541.
- Sky, E. (2023). The Iraq Invasion at Twenty: The Iraq War and Democratic Backsliding. *Journal of Democracy*, 34(2), 135–149.
- Smith, A. (2016). Leader Turnover, Institutions, and Voting at the UN General Assembly. *Journal of Conflict Resolution*, 60(1), 143–163.
- Teets, J. C. (2014). *Civil Society under Authoritarianism: The China Model*. Cambridge University Press.
- Terman, R., & Voeten, E. (2018). The relational politics of shame: Evidence from the universal periodic review. *Review of International Organizations*, 13, 1–23.
- Terman, R., & Byun, J. (2022). Punishment and Politicization in the International Human Rights Regime. *American Political Science Review*, 116(2), 385–402.
- Tsutsui, K., & Wotipka, C. M. (2004). Global Civil Society and the International Human Rights Movement: Citizen Participation in Human Rights International Nongovernmental Organizations. *Social Forces*, 83(2), 587–620.
- von Soest, C. (2015). Democracy prevention: The international collaboration of authoritarian regimes. *European Journal of Political Research*, 54, 623–638.
- Voss, M. J. (2019). The Use (or Misuse) of Amendments to Contest Human Rights Norms at the UN Human Rights Council. *Human Rights Review*, 20, 397–422.
- Voss, M. Joel. (2020). The global south and norm advocacy at the United Nations Human Rights Council. *International Journal on Minority and Group Rights*, 27, 728–764.
- Vreeland, J. R. (2008). Political Institutions and Human Rights: Why Dictatorships Enter into the United Nations Convention Against Torture. *International Organization*, 62(1), 65–101.
- Vreeland, J. R. (2019). Corrupting International Organizations. *Annual Review of Political Science*, 22, 205–222.
- Waldner, D., & Lust, E. (2018). Unwelcome Change: Coming to Terms with Democratic Backsliding. *Annual Review of Political Science*, 21(5), 1–21.
- Walker, C. (2016). The Authoritarian Threat: The Hijacking of Soft Power. *Journal of Democracy*, 27, 49–63.
- Way, L. A. (2022). The Rebirth of the Liberal World Order? *Journal of Democracy*, 33, 5–17.
- Weyland, K. (2017). Autocratic diffusion and cooperation: The impact of interests vs. ideology. *Democratization*, 24, 1235–1252.

- Winzen, T. (2023). Walking the tightrope: How backsliding governments keep the EU hospitable for autocracy. Working paper.
- Woo, B., & Murdie, A. (2017). International organizations and naming and shaming: Does the International Monetary Fund care about the human rights reputation of its client? *Political Studies*, 65(4), 767–785.
- Ziblatt, D., & Levitsky, S. (2018). *How Democracies Die*. Crown.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.