

The Authoritarian Trojan Horse Threatening Liberal International Organizations

Anna M. Meyerrose* and Irfan Nooruddin†

Abstract

International organizations (IOs) dominated by autocracies are becoming increasingly common. Do autocracies and backsliding states use their membership in established western IOs to undermine liberal values from within? How? And what impact do these backsliding and autocratic member states have on the liberalizing goals and overall functioning of these western IOs? In this working paper, we argue that backsliding and autocratic states' continuing membership in historically liberal and western-dominated IOs provides these states with the opportunity to undermine and challenge the liberal international order while also re-orienting it to emphasize values that better align with their domestic 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. 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.

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

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

1. Introduction

Democratic backsliding, which occurs when elected officials weaken or erode liberal democratic institutions, has been on the rise over the past decade. Backsliding states are partnering with more established autocracies to create and join their own international and regional organizations to offset the power of western democracies and the international organizations (IOs) they created to promote liberal democratic values around the world (von Soest 2015; Libman and Obydenkova 2018; Kneuer et al. 2019; Cottiero and Haggard 2021; Debre 2021). Indeed, these and related developments have led pundits and politicians to warn that the liberal international order (LIO) is under threat (Ikenberry 2018; Börzel and Zürn 2021; Way 2022). While IOs created by entrenched autocratic powers such as Russia and China will likely continue to gain influence in the coming decades, western-centered organizations still dominate the multilateral landscape, yet we know comparably less about how backsliding states behave within these established western IOs.

In the post-Cold War era, western-based international institutions increasingly sought to widen their membership to bring emerging and fragile democracies into the fold (Pevehouse 2005; Donno 2013; Genna and Hiroi 2014; Poast and Urpelainen 2018). This tactic might now be backfiring as many of the backsliding and autocratizing states of greatest concern are established members of these very western liberal organizations. The global democratic reversal, which began in the mid-2010s, just twenty-odd years after liberal democracy triumphantly proclaimed the “end of history,” raises serious questions about the expansionist policies pursued by IOs in the 1990s and early 2000s. With democracy in rapid retreat globally, research must explore how entrenched and emerging autocrats cooperate in an international system long dominated by liberal democracies. What impact do these backsliding member states have on the liberalizing goals and overall functioning of these western IOs? As the share of non- and pseudo-democratic members of these organizations increases, does the behavior of autocratic and especially backsliding states within these organizations begin to change? And if democratic backsliding cannot be staunched, what is the future of the liberal democratic IOs when their core members no longer qualify as liberal or democratic?

There is extensive work showing how advanced democracies—most prominently, the United States—have used their dominance over the Bretton Woods institutions and the United Nations (UN) to further their own strategic goals.¹ At the same time, there is growing evidence that autocratic members are learning to use these same institutions to block international scrutiny of their domestic practices and undermine the LIO, particularly in the area of human rights (Ginsburg 2020; Binder and Payton 2022).

¹ Vreeland (2019) offers a comprehensive review.

As states backslide, and therefore as the number of autocratic members of these IOs grows, we should expect autocrats to work together to use these institutions more actively to further their shared interests to dull the power of the United States and other advanced democracies. Specifically, we argue that backsliding and autocratic states' continuing membership in historically liberal and western dominated IOs provides these states with the opportunity to undermine and challenge the liberal international order while also re-orienting the emphasis of these IOs to values that better align with their evolving and often illiberal domestic interests. In that sense, the very international fora created, subsidized, and promulgated by the West to promote liberal hegemony might now be used by backsliding and autocratic states to stymie and undo those efforts.

This paper evaluates this argument by exploring the ways in which autocratic and in particular backsliding states use their influence and voting rights within one core institution of the liberal international order—the UN—to undermine its efforts to promote and support democracy and human rights. Following the end of World War II and particularly after the Cold War, the UN played a pivotal role in promoting and supporting liberal democratic values; human rights were heavily emphasized. Therefore, we explore these dynamics in the context of the UN human rights mechanisms, focusing on 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.²

2. 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 expanded significantly after World War II with the founding of prominent multilateral institutions, with the United Nations (UN) System, the Bretton Woods Institutions, and the General Agreement on Tariffs and Trade at the core (Ikenberry 2009; Börzel and Zürn 2021). The founders of these and subsequent western international organizations (IOs) viewed them in part as tools to further expand and entrench the LIO. While the Bretton Woods institutions promoted open trade and embedded liberalism, the UN was crucial in managing the transition from empires to sovereign states and founding the international human rights regime. After the end of

² Although the UNHRC's predecessor, the UN Commission on Human Rights (UNCHR) was founded in 1946, we choose to focus on the time period from 2006 to the present since democratic backsliding is a historically recent phenomenon that began in the early to mid 2010s.

the Cold War, western-based IOs proliferated, with politically oriented ones (such as the UN) in particular becoming inextricably linked to the construction of a liberal global order centered on democracy, elections, rule of law, and human rights (Barnett and Finnemore 2021).

Indeed, the LIO expanded significantly in the post-Cold War era, a period when western liberal democracies increasingly worked through IOs with the overarching goal of creating “an international ‘space’ for liberal democracy, reconciling the dilemmas of sovereignty and interdependence, seeking protections and preserving rights within and between states” (Ikenberry 2018, 8). Throughout the 1990s and early 2000s, these international institutions gained significantly more authority and became decisively more liberal with a strong emphasis on human rights, rule of law, democracy promotion, the free movement of people, global cooperation, liberal values, and material progress via free markets (Pevehouse 2005; Barnett and Finnemore 2021; Börzel and Zürn 2021). Indeed, 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 and the holding of internationally certified free and fair elections (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 western liberal IOs.

These IOs have been credited with facilitating transitions to democracy, and even democratic consolidation, among these emerging democratic states in the 1990s and 2000s (Pevehouse 2005; Donno 2010; Poast and Urpelainen 2018).³ The belief in the power of the liberal international order was so great that the West even believed that tougher cases for democracy, such as China and Russia, could eventually liberalize politically as they become more integrated into these liberal international institutions (Walker 2016); therefore, their membership and participation in these organizations was also encouraged and expanded.

Going beyond election monitoring, these IOs have also actively worked to promote liberal values such as rule of law and human rights; the UN agencies 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” (*Charter of the United Nations* 1945; Vreeland 2019). This emphasis on the importance of protecting and spreading human

³ More recent evidence, however, suggests these IOs, and globalization more broadly, can unintentionally contribute to backsliding in new democracies (Meyerrose 2020, 2021).

rights led to the creation of the UN Commission on Human Rights (UNCHR) in 1946. 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). Indeed, scholars find that new democracies join these human rights institutions and treaties at higher rates than both dictatorships and consolidated democracies, using the sovereignty costs associated with membership to credibly commit to upholding international law and to becoming consolidated democracies (Moravcsik 2000; Landman 2005; Hafner-Burton 2012).

Nevertheless, the UNCHR faced significant criticism throughout its lifespan, with many arguing it was too heavily exposed to political influence, and its method of choosing members left open the possibility (too often verified by reality) that states with poor human rights records could become members and inhibit its work (Edwards et al. 2008). Considering these issues, the UNCHR was dissolved and replaced by the UN Human Rights Council (UNHRC) in 2006. In its current form, the Council consists of 47 member states elected for 3-year terms. The UNHRC reports directly to the UNGA, and therefore has higher status than its predecessor. Nevertheless, few amendments were made to membership criteria, which has led some to argue there has been little improvement (Cox 2010; Hug and Lukács 2014). However, in other ways, the UNHRC has implemented significant changes when compared to its predecessor. For example, the UNHRC meets more often (Edwards et al. 2008) and has a new tool at its disposal—the Universal Periodic Review (UPR)—to assess human rights situations in *all* UN member states and emphasizes bilateral state-to-state relations.

Over time, western liberal democracies have used and expanded these formal human rights institutions with other tools to both promote human rights and, when necessary, to name and shame states who violate them (Koliev 2020). Indeed, the UNCHR's activities increased significantly following the end of the Cold War, becoming particularly far-reaching in the mid-1990s, and its activities increasingly targeted and punished the worst human rights violators (Lebovic and Voeten 2006). Scholars also find peer review groups, such as the UPR, can be effective tools for naming and shaming (Carraro, Conzelmann, and Jongen 2019), though the types of 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 means 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, regarding 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 (Lebovic and Voeten 2006; Risse and Sikkink 1999; Woo and Murdie 2017; Terman and Voeten 2018), there is also research suggesting human rights treaties and naming and shaming are unable to meaningfully

influence human rights adherence in target states (Tsutsui and Wotipka 2004; Hafner-Burton and Tsutsui 2005; Neumayer 2005; Vreeland 2008) and in some cases can even fuel additional human rights abuses (Hafner-Burton 2008).

3. The Global Democratic Recession and Autocratic International Cooperation

Despite these extensive efforts to promote democracy, human rights, and other liberal values via IOs in the post-Cold War environment, these efforts have more recently been accompanied by a sharp rise in cases of democratic backsliding and rampant illiberalism in new democracies whose transitions to democracy were heavily influenced and supported by the international community (Meyerson 2020), and in particular by the West. These trends reflect a wider phenomenon that Luhrmann and Lindberg (2019) term the “third wave of autocratization,” characterized by moves away from democracy in both democracies and autocracies beginning around 1993. Regime changes in this current wave tend to be more gradual than in previous waves of autocratization and have occurred in an unprecedented number of democracies when compared to the first and second waves of autocratization.⁴ Nevertheless, they are increasingly common: 2020 was the first year since 2001 when there were more autocracies than democracies around the world (Lührmann Marquardt and Mechkova 2020). Contemporary cases suggest the outcome of ongoing erosions, at least in the medium term, is a sort of illiberal democracy or semi-autocratic regime, rather than full-fledged authoritarianism (Luhrmann and Lindberg 2019). Thus, 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 within the very western liberal IOs charged with promoting and maintaining the liberal international order.

⁴ These three waves of autocratization all occurred after their respective waves of democratization. In other words, the current third wave of autocratization corresponds to and follows the widely cited third wave of democratization that began with Portugal’s transition to democracy in 1974 and expanded in the early 1990s primarily in post-communist eastern European countries (Huntington, 1991).

Concurrent with this global democratic recession, autocracies and backsliding states have begun to take active measures at the international level to evade western pressures to adhere to liberal values, and even to directly challenge the established liberal order (Weyland 2017). To date, most research on autocratic cooperation has focused on autocratic regional IOs, or IOs dominated by autocratic members (Cottiero and Haggard 2021). Powerful autocratic states such as Russia, China, Venezuela, and Saudi Arabia have created regional organizations such as the Shanghai Cooperation Organization, the Bolivarian Alliance for the Peoples of Our America, and the Gulf Cooperation Council as alternatives to the established western 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 provide an ideological alternative to the current international environment (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 (von Soest 2015), in part by appeasing international actors committed to promoting democracy, good governance, and liberal values. By passing regional Human Rights Charters or sending accommodating election observers to member states (Debre and Morgenbesser 2017), these 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 2021).

While these autocratic regional organizations will undoubtedly continue to gain influence in the coming decades, western-based IOs still dominate the multilateral landscape. Focusing instead on these established IOs, there is also evidence that powerful, established autocracies have and continue to use their membership in these institutions to subvert from within their efforts to promote liberal democratic values. As these states' economic power has grown, they have become increasingly dissatisfied with the international status quo and are therefore working to challenge western dominance more actively. Binder and Payton (2022) show that rising powers such as Brazil, Russia, India, China, and South Africa have begun to form a bloc of dissatisfied states, as exemplified by their voting behavior in the UN General Assembly. Focusing on the UNCHR, Hug and Lukács (2014) similarly find authoritarian states pursued membership on the Commission to dilute international human rights norms and instead advance illiberal ones; they also worked to shift the focus from civil and political rights to more purely economic ones. Repressive states also used their membership in the UNCHR to insulate themselves from investigation and 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 suggests states from the Global South, including Russia, use amendments to UNHRC resolutions to

protect their own human rights preferences and perhaps also challenge the existing international order (Voss 2019). Looking beyond established autocracies, there is also evidence that new democracies, which are overrepresented in contemporary cases of backsliding, exhibit distinct voting patterns in western liberal IOs, such as the European Parliament (Meyerrose 2018).

The post-Cold War push to expand the LIO resulted in then-emerging and fragile democracies—many of which, more recently, have begun backsliding toward autocracy (Luhmann and Lindberg 2019)—also becoming members of western liberal IOs. Today, these backsliding and autocratizing states of greatest concern are joining fully autocratic states as established members of the very IOs that, historically, have been 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, or the ability of these organizations to effectively achieve their goals of supporting and further promoting western liberal values.⁵ We explore these issues and develop our argument in the following section.

4. 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 U.S. 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 guaranteed by U.S. power. A particular focus of this new liberal international order was the promotion of democracy, which in turn involved a more intrusive examination of member states' 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 to accept the greater scrutiny of their domestic policy records. As Susan Hyde has argued, this spawned a generation of 'pseudo-democrats' who understood that being seen to hold free and fair elections could unlock material benefits, while violating this new norm might 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, protect

⁵ Pahre (1995) illustrates this tension in the context of the European Union.

labor and the environment from corporate predation, and, even, 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 declared, 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 justified. The inexorable growth of China’s economy provided both alternative sources of aid and investment for many developing countries chafing under Western rules, but 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, Flores and Nooruddin 2019). 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. These democrats-turned-autocrats, having secured their rule at home, now had to limit international criticism of their actions.

Membership in IOs provided these democratic backsliders an ideal forum in which to undermine the LIO’s pressure on them. Because they were 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 could exercise their voting privileges in these organizations to limit the efficacy of the IOs. The UNHRC is a perfect example. Here was an organization whose sole purpose is to use the power of the United Nations system to shine a light on countries that are violating human rights and to issue condemnations that have the force of the collective Human Rights Council. But because the UNHRC’s membership is elected by regional peers,

membership in the Council is not predicated on a country's own democratic or human rights records. Over time, this meant that the global democratic recession also resulted in a backsliding of democracy within the Council's membership.

We argue that non-democratic and backsliding countries should be less likely to vote in favor of UNHRC resolutions targeting other states for alleged human rights violations. These 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. If correct, this has an important consequence for the UNHRC more generally because it means that increasingly its resolutions, even if passed, will do so with narrower majorities than before due to a larger number of 'nay' votes. This should have the effect of undermining the moral force of UNHRC resolutions and, as such, their influence. This logic suggests the following testable hypothesis:

Hypothesis 1: *Backsliding states are less likely to vote in favor of UNHRC resolutions.*

Voting on specific resolutions is a visible way in which states register their support for or opposition to the principles of the LIO. Indeed, recent research suggests 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 South (Lakatos 2022). However, their ability to do so in a setting governed by majoritarian voting rules is constrained. The introduction of the UPR mechanism provides an additional means for non-democracies 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 and autocracy-leaning states to directly challenge these dominant western democracies. One way to do this is to place these western states under increased scrutiny for their domestic human rights practices to deflect attention from nondemocracies and also undermine the credibility of states at the forefront of the promotion of 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 allows countries to reveal the issues they think are 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 argue that greater focus should be paid to economic and social rights. We do not disagree that all rights are worthy of

elevation by the international community but see in the UPR mechanism an opportunity to understand whether autocratic and backsliding states are more likely to engage a different set of issues than their consolidated democracy counterparts. In particular, we seek to 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 shed new light on how the global trend in democratic backsliding undermines 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 directly confront the West in their UPR reports while also de-emphasizing political and civil liberties and stressing other framings 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 must be to undermine its very reason for existence. In the next section, we analyze country voting data in the UNHRC and a text dataset from UPR recommendations to assess if these concerns are warranted, or hyperbolic.

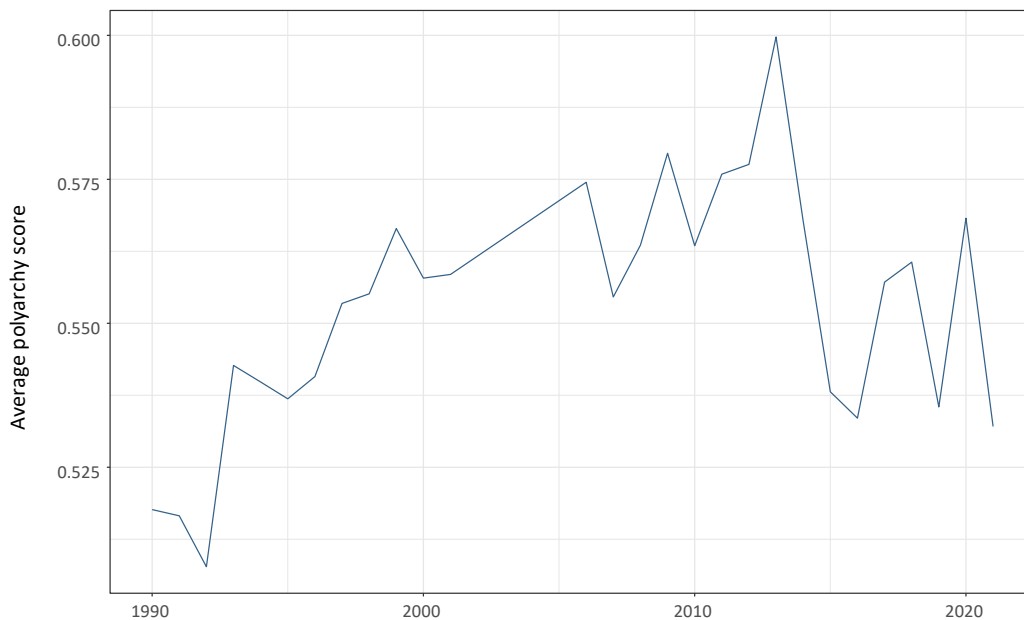
5. Research Design

We predict backsliding and autocratic states will use their membership in existing western liberal IOs to undermine the pressures on themselves and on states like them to adhere to liberal democratic values. We explore these dynamics in the context of the United Nations (UN). The UN was founded with the goal of promoting peace amongst nations in large part by supporting and spreading the observance of universal freedoms and human rights around the world. This emphasis on fundamental human rights is at the core of the UN's efforts to promote and sustain the LIO. Nevertheless, 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. Therefore, we test our argument using data from human rights organizations within the UN. 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).

5.1 Voting in the UNHRC

We test our theory first in the context of the UNHRC. The UNHRC was founded in 2006 after its predecessor, the UNCHR, was dissolved amidst growing criticism of its politicization, dubbed the “shame of the UN” (Hug and Lukács 2014). The present-day UNHRC consists of 47 member states that are elected by the UNGA 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 Figure 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.

Figure 1. Average polyarchy scores of UNHRC members by year. Note that the polyarchy index takes values between 0 and 1. Data source: Varieties of Democracy dataset (Coppedge et al., 2021).



Building on work that finds regime type influences how states vote on targeted UNHRC resolutions (Edwards et al. 2008; Seligman 2011), we test if these new and still emerging 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 should expect that backsliding states will have incentives not only

to protect state sovereignty, but 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 mid-2010s.

To our knowledge, no fully updated dataset of votes on targeted UNHRC resolutions currently exists. Therefore, we scrape this information directly from the UNHRC online library for all resolutions from 2006 through 2021⁶ and combine this with existing data on targeted resolutions in the UNCHR prior to 2006. The resulting dataset contains information on how each member of the UNHRC voted on any given targeted resolution⁷ 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, for the time period of interest, consists of 13,749 observations covering 189 unique resolutions for which votes were recorded and adopted by the UNHRC targeting 18 states between 2006 and 2021.⁸

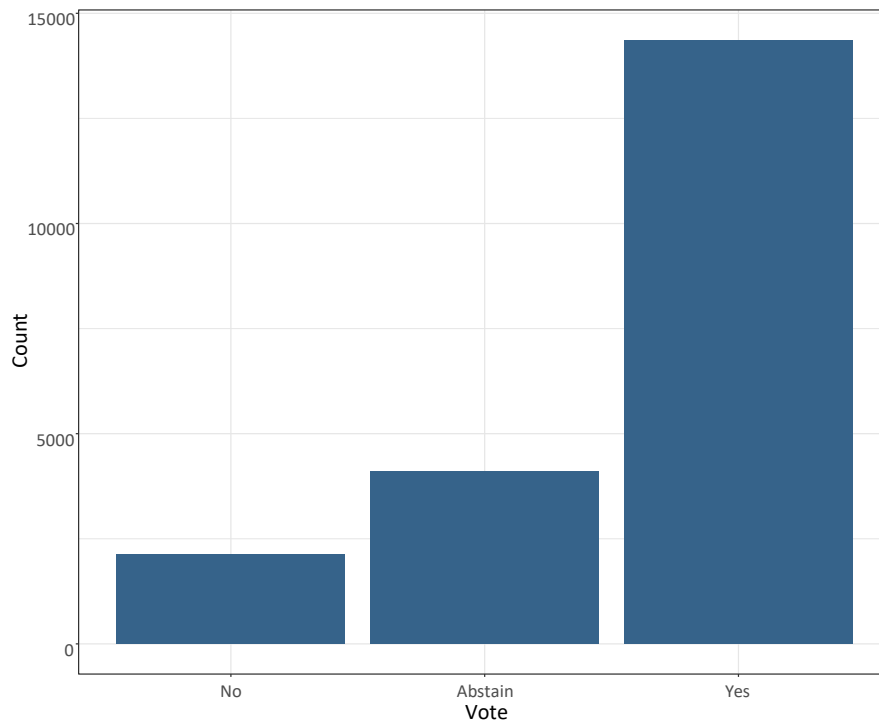
We predict that backsliding states, along with fully autocratic ones, will be less likely to vote in favor of targeted UNHRC resolutions than their more democratic counterparts. Figure 2 shows that for the entire time period for which we have roll call data (1973–2021), “yes” votes are by far the most common type of vote. However, Figure 3 shows the proportion of “yes” votes in the UNHRC per year begins to decline around the same time when backsliding became increasingly common both globally and also amongst UNHRC members, as illustrated in Figure 1.

⁶ <https://searchlibrary.ohchr.org/search?ln=en&cc=Voting>.

⁷ The possible votes are “yes,” “no,” or “abstain”

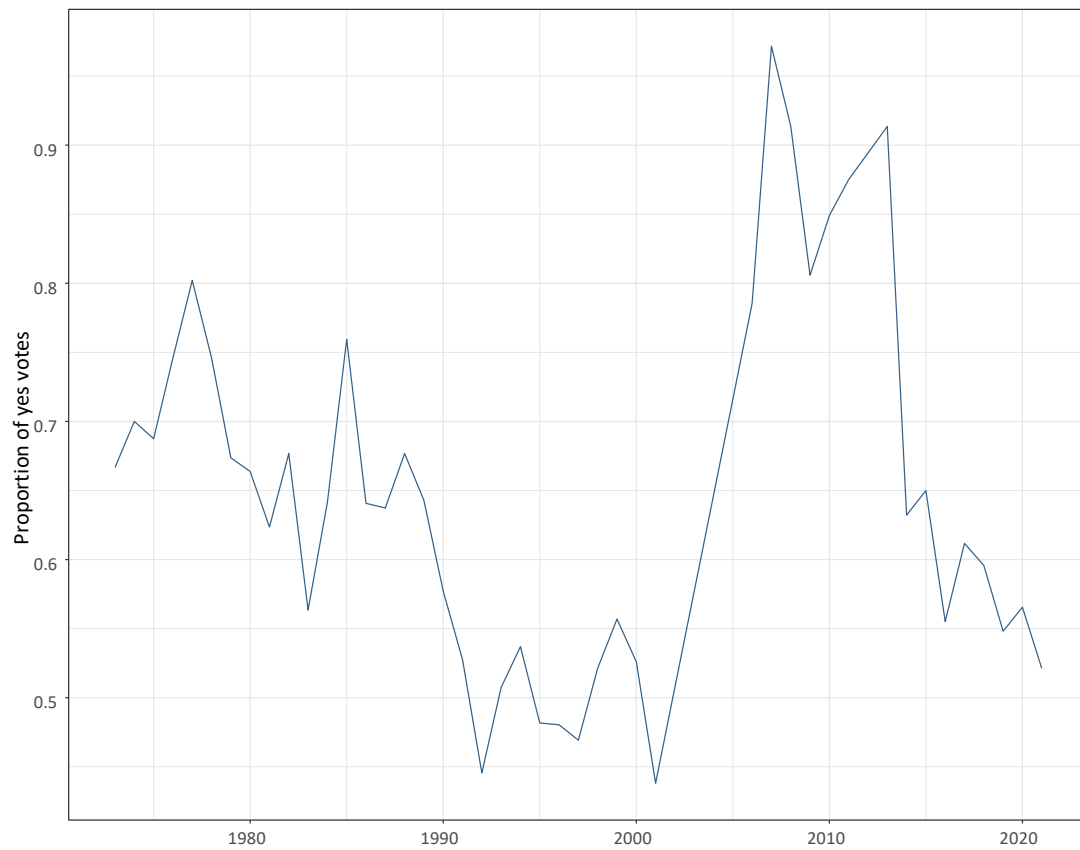
⁸ The states targeted in these resolutions, and the number of times they were targeted, are: Belarus (10), Burundi (7), Democratic People’s Republic of Korea (1), Democratic Republic of the Congo (1), Eritrea (3), Ethiopia (1), Georgia (4), Iran (7), Israel (29), Myanmar (6), Nicaragua (3), Philippines (1), South Sudan (1), Sri Lanka (2), Syria (30), Ukraine (6), Venezuela (5), and Yemen (3).

Figure 2. Distribution of Votes in UNHRC, 1973–2021



The main 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 amongst 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. During this time period, 131 states cast a vote on at least one targeted resolution.

Figure 3. Proportion of yes votes in the UNHRC by year, 1973–2021

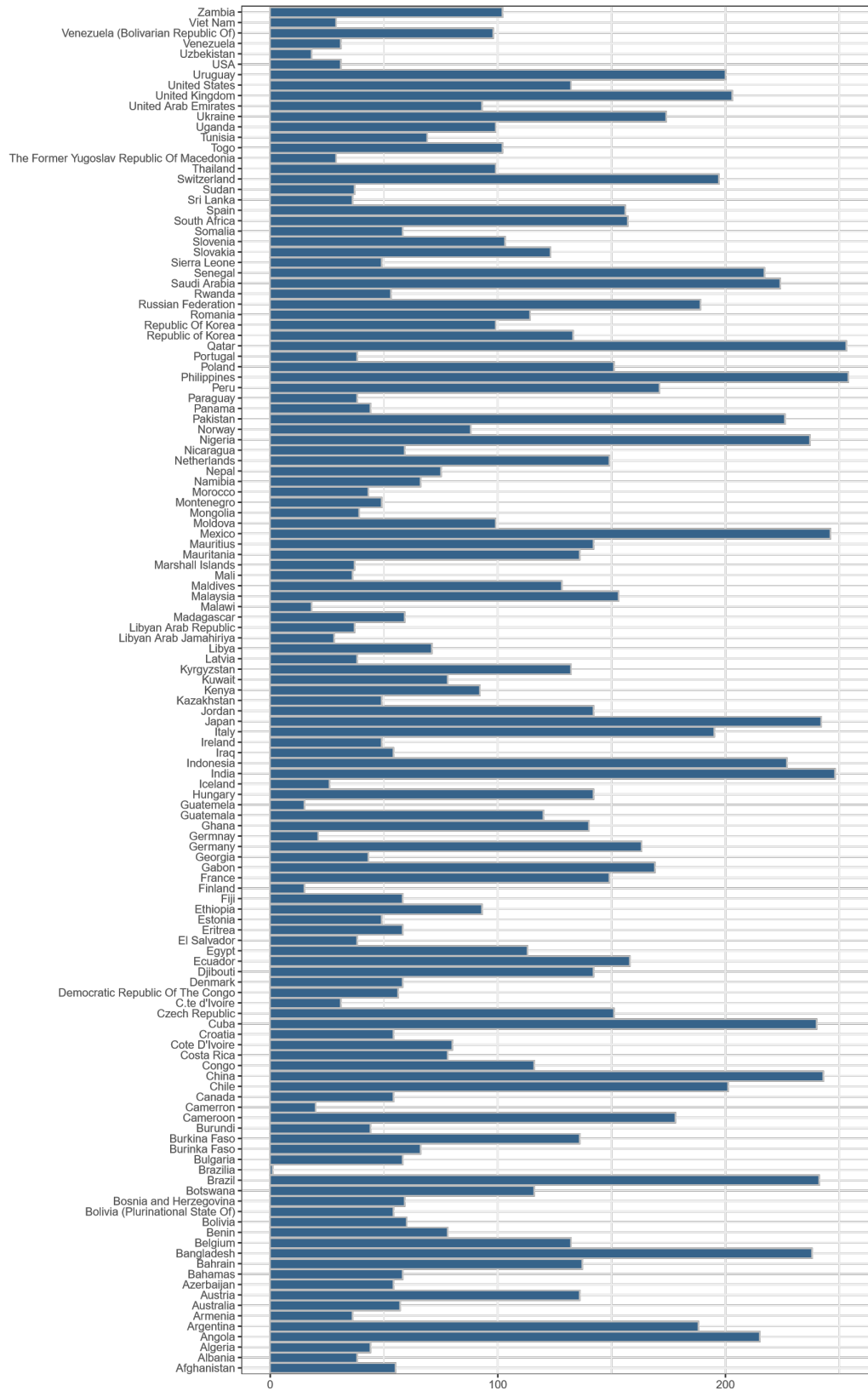


Our main independent variable is a binary indicator for whether the voting state has undergone democratic backsliding since 2006, the year when the UNHRC was founded. To create this variable, we use data from the Varieties of Democracy (V-Dem) dataset (Coppedge et al. 2021). Specifically, we use the electoral democracy index (EDI), which 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. 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 (or, autocratization) as any country-year observation for which a state's EDI has declined by 0.1 or more since 2006. Our main variable is coded as 1 if the voting state's EDI has decreased by 0.1 or more since 2006; otherwise, it takes a value of 0.

We also include a relevant set of control variables for both the voting and target states. On the voting state side, we control for the current electoral democracy score, in addition to the backsliding indicator. While we expect backsliding to make a state more likely to vote against UNHRC resolutions, backsliding can result in an illiberal democratic, semi-autocratic, or fully autocratic regime; this endpoint is likely also relevant in 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 (Bailey, Strezhnev and Voeten 2017), as well as for whether or not the voting state is located in the same (UN) geographical region as the target state. Finally, since states with better human rights records are more likely to shame other countries for human rights violations in the post-Cold War era (Lebovic and Voeten 2006), we also control for the voting state's human rights score from the Our World in Data dataset (Herre and Roser 2016).⁹

⁹ While many studies use data on human rights violations from the CIRI human rights data project, those data only go through 2011. The Our World in Data dataset, on the other hand, extends through 2019.

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



We also control for relevant characteristics of the state targeted by a UNHRC resolution. 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). We also control for the target state's current electoral democracy score, its agreement score with US votes in the UNGA, and its human rights score. Summary statistics for all variables included in our analysis are reported in Table 1.

Table 1. Descriptive Statistics: Voting in the UNHRC, 2006–2021

<u>Statistic</u>	<u>N</u>	<u>Mean</u>	<u>St. Dev.</u>	<u>Min</u>	<u>Pctl(25)</u>	<u>Pctl(75)</u>	<u>Max</u>
Vote choice	13,721	2.68	0.60	1	3	3	3
Backslide since 2006	13,062	0.08	0.27	0.00	0.00	0.00	1.00
Current dem score	13,626	0.57	0.27	0.02	0.33	0.83	0.92
Target state = Israel	13,721	0.27	0.44	0	0	1	1
Target dem score	13,721	0.38	0.25	0.07	0.17	0.70	0.76
Voting and target in same region	13,626	0.07	0.25	0.00	0.00	0.00	1.00
Voting UNGA US agree	12,716	0.38	0.15	0.00	0.25	0.50	0.81
Target UNGA US agree	12,879	0.44	0.29	0.04	0.20	0.81	0.96
Voting state HR score	11,993	0.52	1.43	-2.00	-0.64	1.56	5.16
Target state HR score	11,993	-1.27	0.69	-2.30	-1.76	-0.99	1.04

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 variable in our model and report the results in Table 2. 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.

Overall, we find evidence that backsliding and autocratic states are more likely to vote against targeted UNHRC resolutions. Specifically, we find that, for a state that has backslid since 2006, its log-odds of voting “no” relative to “yes” increases by 0.69. To 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 roughly 2 times more likely to vote “no” as opposed to “yes” on a targeted UNHRC resolution. The negative coefficient for the current democracy score variable in the “no” column of Table 2 also suggests that as a state's level of democracy decreases, it is more likely to vote “no,” rather than “yes,” on targeted UNHRC resolutions.

Table 2. Democratic backsliding and vote choice in the UNHRC, 2006–2021

	<i>Dependent variable:</i>	
	Abstain (1)	No (2)
Backslide since 2006	−0.22** (0.11)	0.67*** (0.12)
Current dem score	−0.03 (0.15)	−2.58*** (0.22)
Target state = Israel	−0.03 (0.20)	0.13 (0.32)
Target dem score	1.82*** (0.32)	1.85*** (0.50)
Voting and target in same region	−0.81*** (0.14)	−0.58*** (0.16)
Voting UNGA US agree	−0.46 (0.30)	0.57 (0.48)
Target UNGA US agree	−1.64*** (0.35)	−2.82*** (0.60)
Voting state HR score	−0.20*** (0.03)	−0.16*** (0.04)
Target state HR score	0.28*** (0.05)	0.10 (0.07)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Multinomial logit	Multinomial logit
Observations	11,245	11,245
K	3	3

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

Interestingly, we also find that backsliding states are less likely to abstain on UNHRC resolutions (with a log-odds of -0.22 and corresponding relative risk ratio of 0.79). To better understand why this may be the case, we further subset our data to re-estimate this model. As noted above, one artifact of our operationalization of backsliding is that democratic, semi-autocratic, or fully autocratic regimes can all be coded as cases of backsliding. However, since our theory is focused particularly on states that fall somewhere between fully democratic and fully autocratic, rather than on countries such as North Korea that were closed dictatorships prior to the recent global democratic recession but may have declined even further since 2006, we re-estimate our main model, this time excluding all observations for which the voting country was a closed autocracy. Following Lührmann, Tannenberg, and Lindberg (2018), we use the V-Dem

dataset to identify closed autocracies as ones with a) an electoral democracy index score below 0.5 and b) a lack of multi-party elections at the national level.¹⁰ The results of this model are reported in Table 3.

Table 3. Democratic backsliding and vote choice in the UNHRC, 2006–2021, Exclude Closed Autocracies

	<i>Dependent variable:</i>	
	Abstain (1)	No (2)
Backslide since 2006	−0.05 (0.16)	2.26*** (0.25)
Current EDI score	−2.10*** (0.41)	−0.68 (0.91)
Target state = Israel	0.41 (0.27)	−0.38 (0.52)
Target EDI score	1.48*** (0.43)	−0.19 (0.97)
Same region	−0.38 (0.24)	0.55 (0.41)
Voting UNGA US agree	−0.37 (0.36)	1.98*** (0.68)
Target UNGA US agree	−1.33*** (0.48)	1.68* (1.00)
Voting state HR score	−0.06 (0.04)	0.09 (0.08)
Target state HR score	0.25*** (0.06)	−0.39** (0.16)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Multinomial logit	Multinomial logit
Observations	6,603	6,603
K	3	3

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

¹⁰ We use the multiparty elections variable (v2elmulpar) from V-Dem, and code a case as lacking multiparty elections if this variable equals 0 (no meaningful competition) or 1 (multiple candidates but no-party or single-party).

When we exclude these fully autocratic states, the relationship between backsliding states and abstentions is no longer statistically significant, suggesting that these closed dictatorships that have regressed even further since 2006 are driving the abstention results reported in Table 2. More relevant to our theory, however, we also find that the magnitude of the coefficient for “no” votes for the backsliding states of interest—ones that result in illiberal democratic or semi-autocratic regimes—grows substantially once we excluded these closed autocracies. The log-odds of these backsliding states voting “no” relative to “yes” increases by 2.26; in other words, we find that backsliding states are roughly 9.6 times more likely to vote “no” as opposed to “yes” on targeted UNHRC resolutions.

We also find in Table 2 that states are *less* likely to abstain or vote against (or, in other words, are more likely to vote in favor of) resolutions targeting states in their same geographical region. However, when we remove the closed autocracies from the data, this relationship is no longer significant, again suggesting that full dictatorships, but not other types of countries, are more likely to vote in favor of resolutions targeting their neighbors.

Since the magnitude of backsliding may influence how a state votes on targeted UNHRC resolutions, we also re-estimate the model from Table 2 using a continuous, rather than binary, variable to capture backsliding. When backsliding is operationalized as the change in a state’s EDI since 2006, we again find that backsliding states are less likely to abstain and more likely to vote no, relative to the baseline yes.¹¹

To summarize, Tables 2 and 3 provide initial support for our expectations regarding the behavior of backsliding states as outlined in Hypothesis 1: backsliding and autocratic states are more likely to vote against targeted UNHRC resolutions. Having shown that these backsliding states use their voting power to block and undermine attempts to promote human rights in a subset of states identified by the Council, we next test the extent to which these states use the UN’s human rights mechanisms to more directly challenge the western liberal democracies who work to promote and sustain the liberal international order.

¹¹ See Appendix Table A1.

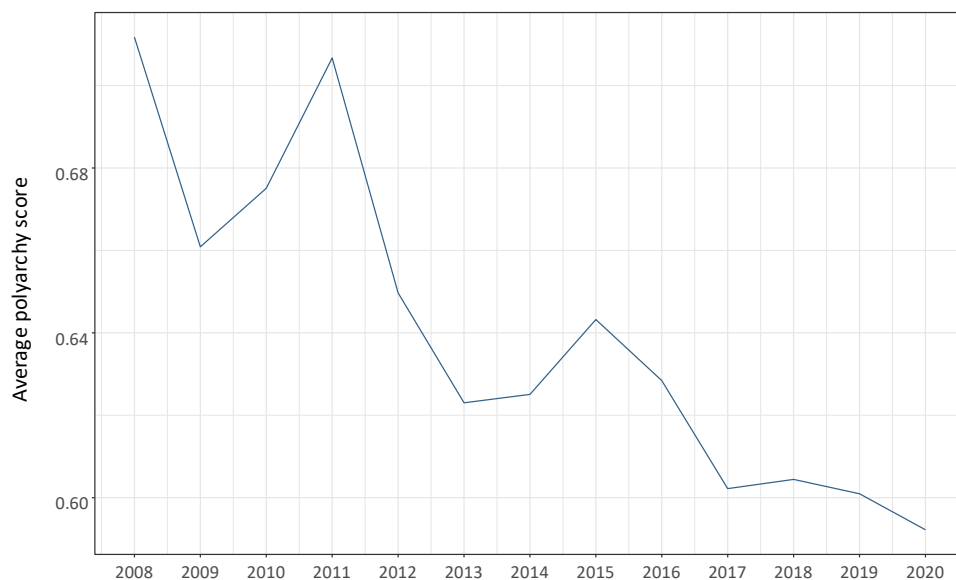
5.2 Targeting Advanced Democracies in the UPR

In its efforts to reform the UNCHR in 2006, in addition to creating the current UNHRC, the UN also introduced the Universal Periodic Review (UPR) in 2008 as another tool to promote human rights. Under the UPR, *all* UN member states, rather than simply ones identified by members of the UNHRC, are subjected to periodic reviews of their domestic human rights practices. All states are reviewed every four and a half years, and these review sessions take place three times per year in Geneva. In these sessions, three randomly assigned “troika” states lead the discussion of the state under review. Reviews are based on information provided by the state under review, independent human rights commissions and experts, and national human rights or non-governmental groups. The reviews are conducted by the 47 sitting members of the UNHRC, but any state can be involved in the discussion. In the reviews, recommending states identify issues in the state under review and make suggestions for how to address them. The state under review can either accept or reject these recommendations (Cox 2010). 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, thereby giving states more leeway to make decisions and recommendations outside of pressures from their regional groups (McMahon and Ascherio 2012). We take advantage of this feature of the UPR to more directly observe the individual behavior of backsliding states.

Data on the content of these reviews are publicly available online.¹² We use these to create a dataset that consists of 32,598 individual recommendations made for all UN member states between 2008 and 2021. This dataset also includes information on the reviewing state. Similar to the UNHRC (Figure 1), we find the average democracy score for states writing UPR recommendations has also declined in recent years, as illustrated in Figure 5.

¹² <https://upr-info-database.uwazi.io/en/>.

Figure 5. Average polyarchy score of UPR recommending states by year. Data source: Varieties of Democracy dataset (Coppedge et al., 2021).



We use these UPR data to further explore how backsliding and authoritarian states use UN human rights tools differently than their democratic counterparts. Since all states are subject to UPR reviews, we begin by examining the extent to which these backsliding states use their UN membership to more directly confront western liberal states, namely, by testing if they use these UPR reports to place more scrutiny and pressure on consolidated western liberal democracies regarding their own human rights practices.

To test this, 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 western democracies, we restrict our data to include only observations where the state under review is one of these 19 advanced democracies and the recommending state is *not* one of these 19 countries. For our first test of Hypothesis 2, our main dependent variable is a count of the number of issues a recommending state identifies in their UPR report. There is a finite (though

substantially long) list of issues that can be highlighted in these reports,¹³ and we find the number of issues in any given report in our subset of the data ranges from 0 to 19, with a mean of 3.49 issues identified per report. We argue that the number of issues identified in a report serves as a proxy for how critical that report is; more critical reports will identify more issues in the state under review.

As before, our main independent variable is a binary variable indicating whether the reviewing state's level of democracy has deteriorated by 0.1 or more since 2006.¹⁴ Furthermore, we use the same control variables as those reported in Table 2, with several exceptions. First, we no longer control for whether the state under review is Israel, since only reviews of the 19 advanced 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 that same year to account for any incentives this may create. Summary statistics for these data are reported in Table 4.

Since the count of issues identified in the UPR reports is under-dispersed,¹⁵ 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 amongst UPR recommending states. However, there is no reason to expect the types of recommendations made by UNHRC members will be substantively distinct from those of other states, all else equal. Each of these models include fixed effects for year. Since states can submit multiple reviews, the observations in our data are not independent. Therefore, the models also include standard errors clustered by recommending state. The results can be found in Table 5.¹⁶ We find evidence that backsliding states identify a greater number of issues in reports targeting advanced western democracies. Similar to Table 3, these results are robust to models that exclude observations for UPR reports written by closed autocracies.¹⁷

¹³ 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."

¹⁴ Although the UPR began in 2008, we continue to use 2006 as the benchmark for the sake of consistency.

¹⁵ When we subset the data to the relevant universe of cases, the dispersion parameter is 0.63.

¹⁶ Alternative models that use the continuous measure of backsliding discussed in Section 5.1 return similar results. See Appendix.

¹⁷ See Appendix.

Table 4. Descriptive Statistics: UPR Reports of Advanced Democracies

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Number of issues per review	7,777	3.49	1.99	0	2	4	19
Action recommended in review (interval)	7,745	3.98	1.01	1.00	4.00	5.00	5.00
Action recommended in review (binary)	7,745	0.76	0.42	0.00	1.00	1.00	1.00
Backslide since 2006	7,640	0.12	0.32	0.00	0.00	0.00	1.00
Current EDI score	7,645	0.52	0.25	0.02	0.28	0.74	0.92
State under review EDI score	7,777	0.88	0.03	0.81	0.87	0.90	0.92
Same region	7,777	0.10	0.30	0	0	0	1
Recommend UNGA US agree	7,710	0.38	0.16	0.00	0.25	0.48	0.94
State under review UNGA US agree	6,956	0.61	0.09	0.44	0.53	0.69	0.85
Recommend HR score	6,856	0.25	1.35	-2.30	-0.99	1.18	4.61
State under review HR score	6,931	2.86	1.12	0.31	2.33	3.31	5.33
Recommending state under review	7,777	0.19	0.39	0	0	0	1

In addition to identifying issues in the state under review, the reviewing state also makes recommendations regarding the level of action that should be taken to rectify the human rights issues identified in their country report. The recommended action can fall into one of five increasingly onerous categories: 1) Minimal Action, 2) Continuing Action, 3) Considering Action, 4) General Action, and 5) Specific Action. As an additional test of Hypothesis 2 and the extent to which backsliding and autocratic states use the UPR mechanism to target advanced democracies, we estimate additional models where the outcome is the action recommended in a report written about any one of those 19 consolidated democracies. Specifically, for the dependent variable we create a binary variable based on the five levels of possible action recommendations; this variable is coded as 1 if the recommendation is either “General Action” or “Specific Action,” the two most demanding types of actions that can be recommended, and 0 otherwise. While the other three types of action do not require the state under review to take additional or new measures, the “General” and “Specific” actions suggest much more active efforts are required on the part of the state under review. Our main independent variable is again the indicator for whether a state has backslid since 2006.

Table 5. UPR Issues Identified against Advanced Democracies, 2008–2020

<i>Dependent variable: Count of issues identified in country report</i>		
	All Reports (1)	Reports by UNHRC Members (2)
Backslide since 2006	0.10*** (0.02)	0.22*** (0.03)
Current polyarchy score	0.11*** (0.03)	0.34*** (0.04)
State under review polyarchy score	−0.46 (0.31)	−0.005 (0.50)
Same region	0.22*** (0.02)	0.32*** (0.03)
Recommend UNGA US agree	−0.48*** (0.05)	−0.61*** (0.09)
State under review UNGA US agree	0.61*** (0.10)	0.43*** (0.15)
Recommend HR score	−0.03*** (0.01)	−0.03*** (0.01)
State under review HR score	0.02*** (0.01)	0.02** (0.01)
Recommending state under review	−0.07*** (0.01)	−0.16*** (0.02)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Quasi-Poisson	Quasi-Poisson
Observations	6,259	2,654
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01		

We estimate two logistic regression models, where, as before, the first includes all reports targeting advanced democracies, and the second includes only those written by standing UNHRC member states. These models incorporate the same controls variables as before, and again include year fixed effects and clustered standard errors. The results are reported in Table 6.

Table 6. UPR Actions Recommended for Advanced Democracies, 2008–2020

<i>Dependent variable: 1 for General or Specific Action, otherwise 0</i>		
	All Reports (1)	Reports by UNHRC Members (2)
Backslide since 2006	0.42*** (0.06)	0.39*** (0.09)
Current polyarchy score	−0.94*** (0.17)	−0.85*** (0.25)
State under review polyarchy score	2.41 (1.67)	−1.12 (2.66)
Same region	0.31*** (0.12)	0.55** (0.21)
Recommend UNGA US agree	0.47 (0.32)	0.74 (0.54)
State under review UNGA US agree	1.24** (0.55)	1.12 (0.88)
Recommend HR score	0.04 (0.03)	−0.02 (0.04)
State under review HR score	0.03 (0.04)	0.03 (0.06)
Recommending state under review	−0.02 (0.08)	0.01 (0.12)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Logit	Logit
Observations	6,254	2,654

Note:

*p<0.1; **p<0.05; ***p<0.01

We find that backsliding states are more likely to recommend either “General” or “Specific” action in their reports on advanced democracies, as are states with lower current democracy scores. As reported in the Appendix, these results are robust to models that exclude closed autocracies from the recommending states, ones that use a continuous measure of backsliding as the main independent variable, and ones where the outcome variable is coded as 1 only if “Specific Action”—the strongest recommendation—is suggested in the UPR review. These results, combined with those reported in Table 5, suggest that backsliding and autocratic states use the UPR mechanism deliberately to challenge the established liberal international order by

placing western democracies under greater scrutiny for their human rights practices by identifying more issues and recommending more extensive actions to address these issues.

5.3. The Content of UPR Recommendations

In previous sections we showed that backsliding and autocratic 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 and recommend more extensive actions in their UPR reports of advanced western liberal democracies. Having established these distinct patterns of behavior, we now examine the extent to which the substance of backsliding states’ input into these human rights mechanisms reflects their (often illiberal) domestic interests that directly contrast with the norms underlying existing international liberal values.

To explore this, we again use the UPR dataset. In addition to the number and type of issues identified and actions recommended in these reports, our dataset also 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 how the content and topics that backsliding states choose to emphasize in their UPR reports.

We draw on the recommendations contained in UPR reports to estimate a structural topic model (STM). To do so, we use the `stm` package in R to tune the number of topics and estimate parameters. Before estimating these models, we pre-process all text by eliminating common English stopwords, numbers, and punctuation. We then estimate models for 8, 10, 12, and 15 topics and compare the output. Our diagnostic tests suggest that the model with 10 topics maximizes coherence and exclusivity, as illustrated in Figure 6; as such, we focus on the 10-topic model. 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 10 topics. These topics and their corresponding common words are shown in Figure 7.

Figure 6. Semantic Coherence and Exclusivity Topic Model Comparisons

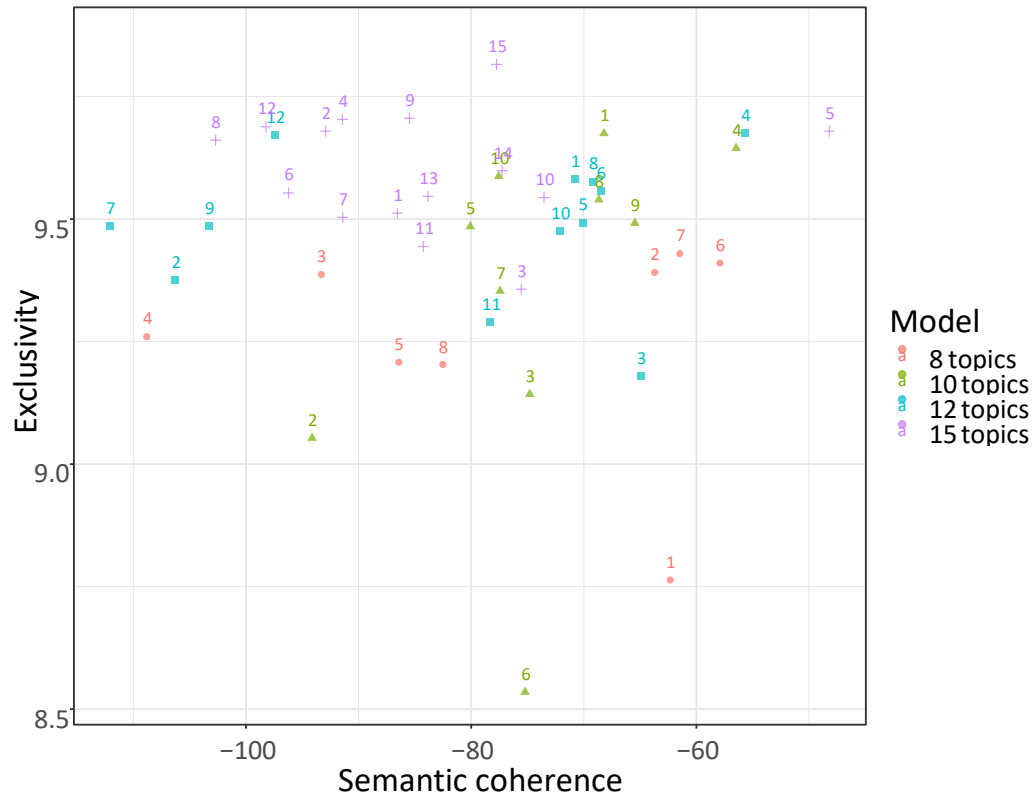
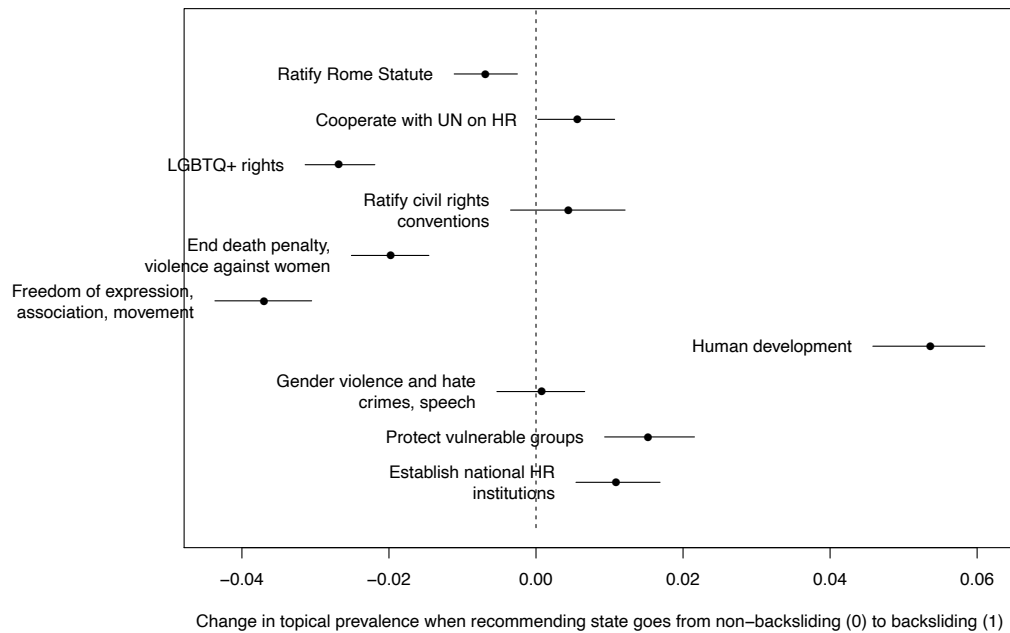


Figure 7. Semantic Coherence and Exclusivity Topic Model Comparisons

Topic	Representative words
1. Ratify Rome Statute of ICC	Procedure, human, right, court, statue, invite, Rome, ICC
2. Cooperative with UN on human rights	Human, nation, report, right, intern, implement, treaty
3. LGBTQ+ rights	Sexual, gender, discrimination, transgender, abortion, bisexual, lesbian, gay
4. Ratify international civil rights conventions	Intern, protect, punishment, families convention, cruel, protocol, disappearances, convention
5. End death penalty, violence against women	Death, penalty, women, abolish, marriage, divorce
6. Freedom of expression, association movement	Freedom, express, ensure, journalist, assemble, dissent, right
7. Human development	Education, development, access, social, health, poverty, rural
8. Gender violence and hate crimes, speech	Violence, combat, xenophobia, trafficking, women, hate, racism
9. Protect vulnerable group	Protect, promote, strengthen, conditions, women, disabilities, undocumented
10. Establish national human rights institutions	Establish, human, right, institution, nation, indigenous, stateless

In our STM model, we look specifically at the extent to which a shift from 0 (no backsliding) to 1 (backsliding) on our binary indicator of backsliding discussed above impacts the types of issues that states emphasize in their UPR reports. As before, we include year fixed effects when estimating the impact of backsliding on topic emphasis. Figure 8 plots the results.

Figure 8. Change in topic prevalence as democracy level of recommending state decreases



Overall, we find that backsliding states are more likely to focus on topics related to protecting vulnerable groups, establishing national human rights institutions, and cooperating with various UN agencies on human rights issues. The heavy emphasis these states place on human development in particular stands out. At the same time, these backsliding states are significantly less likely to mention more sensitive issues surrounding political and civil liberties that directly challenge state sovereignty and government control over its citizens, such as ratifying the International Criminal Court (ICC)'s Rome Statute, allowing freedom of expression, association, and movement, and eliminating the death penalty. We also find these backsliding states are less likely to advocate the protection of LGBTQ+ rights; this 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. Taken together, these results suggest these backsliding states are using international human rights fora to promote their own interests and to counter efforts by democracies to further spread liberal democratic norms and values.

6. Conclusion

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

Taken together, these findings have several implications for the functioning and efficacy of the UN's efforts at promoting human rights and suggest the possibility that similar dynamics could be at play in other western, liberal IOs whose established 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, as illustrated in Figure 3. Resolutions with less unanimous support not only risk diminishing the amount of pressure target states feel to implement reforms, but at a more fundamental level also serve to further undermine the legitimacy of the UNHRC and the values it was created to uphold.

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 fundamentally undermine existing conceptions of human rights that are centered on civil and political liberties, and replacing them with an alternative set of softer, economic rights—such as human development—that better align with their domestic, illiberal interests. To the extent that these efforts succeed, the past three decades of progress made in expanding individual freedoms could potentially be undone, with a series of corresponding normative implications.

The evidence we present also suggests 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 western liberal IOs. To what

extent, for example, have states such as Hungary and Poland been able to leverage their membership privileges to undermine political and liberal aspects of the European Union and related European institutions from within? How, if at all, are these IOs 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 desired effects. This requires both 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?

The implications of our findings presented here should concern any 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, should 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, it appears that human rights IOs are simply theater, but where the once-upon-a-time understudies are increasingly seizing center stage and changing the script while they are at it.

References

- Bailey, Michael A., Anton Strezhnev, and Erik Voeten. 2017. "Estimating dynamic state preferences from United Nations voting data." *Journal of Conflict Resolution* 61:430–456.
- Barnett, Michael and Martha Finnemore. 2021. *The Power of Liberal International Organizations*. Johns Hopkins University Press.
- Binder, Martin and Autumn Lockwood Payton. 2022. "With Frenemies Like These: Rising Power Voting Behavior in the UN General Assembly." *British Journal of Political Science* 52(1):381–398.
- Börzel, Tanja A. and Michael Zürn. 2021. Contestations of the Liberal International Order: From Liberal Multilateralism to Postnational Liberalism. Vol. 75 Cambridge University Press: 282–305.
- Carothers, Thomas. 2002. "The End of the Transition Paradigm." *Journal of Democracy* 13:5–21.

Carraro, Valentina, Thomas Conzelmann, and Hortense Jongen. 2019. "Fears of peers? Explaining peer and public shaming in global governance." *Cooperation and Conflict* 54:335–355.

Charter of the United Nations. 1945.

Collier, David and Steven Levitsky. 1997. "Democracy with Adjectives: Conceptual Innovation in Comparative Research." *World Politics* 49:430–451.

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

Cottiero, Christina and Stephan Haggard. 2021. "The Rise of Authoritarian Regional International Organizations." *Working Paper, UC Institute on Global Conflict and Cooperation*.

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

Dahl, Robert. 1971. *Polyarchy: participation and opposition*. Yale University Press.

Debre, Maria J. 2021. "The dark side of regionalism: how regional organizations help authoritarian regimes to boost survival." *Democratization* 28:394–413.

Debre, Maria J. and Lee Morgenbesser. 2017. "Out of the shadows: autocratic regimes, election observation and legitimation." *Contemporary Politics* 23:328–347.

Donno, Daniela. 2010. "Who Is Punished? Regional Intergovernmental Organizations and the Enforcement of Democratic Norms." *International Organization* 64:593–625.

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

Edwards, Martin S., Kevin M. Scott, Susan Hannah Allen, and Kate Irvin. 2008. "Sins of commission? Understanding membership patterns on the United Nations human rights commission." *Political Research Quarterly* 61:390–402.

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

Flores, Thomas Edward and Irfan Nooruddin. 2016. *Elections in hard times: Building stronger democracies in the 21st century*. Cambridge University Press.

Fukuyama, Francis. 1989. "The End of History?" *The National Interest* (16):3–18.

Genna, Gaspare M. and Taeko Hiroi. 2014. *Regional Integration and Democratic Conditionality: How Democracy Clauses Help Democratic Consolidation and Deepening*. Florence, GB: Routledge.

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

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

Hafner-Burton, Emilie M. 2008. "Sticks and Stones: Naming and Shaming the Human Rights Enforcement Problem." *International Organization* 62:689–716.

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

Hafner-Burton, Emilie M. and Kiyoteru Tsutsui. 2005. "Human Rights in a Globalizing World: The Paradox of Empty Promises." *American Journal of Sociology* 110(5): 1373–1411.

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

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

Huntington, Samuel P. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. Norman: University of Oklahoma Press.

- Hyde, Susan D. 2011. *The Pseudo-Democrat's Dilemma: Why Election Observation Became An International Norm*. Cornell University Press.
- Ikenberry, G. John. 2009. "Liberal Internationalism 3.0: America and the Dilemmas of Liberal World Order." *Perspectives on Politics* 7:71–87.
- Ikenberry, G. John. 2018. "The end of liberal international order?" *International Affairs* 94:7–23.
- Johnstone, Ian and Michael Snyder. 2016. Democracy promotion. In *The Oxford handbook of international organizations*.
- Kneuer, Marianne, Thomas Demmelhuber, Raphael Peresson, and Tobias Zumbrägel. 2019. "Playing the regional card: why and how authoritarian gravity centres exploit regional organisations." *Third World Quarterly* 40:451–470.
- Koliev, Faradj. 2020. "Shaming and democracy: Explaining inter-state shaming in international organizations." *International Political Science Review* 41:538–553.
- Lakatos, István. 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, Todd. 2005. "The Political Science of Human Rights." *British Journal of Political Science* 35(3):549–572.
- Lebovic, James H. and Erik Voeten. 2006. "The politics of shame: The condemnation of country human rights practices in the UNCHR." *International Studies Quarterly* 50:861–888.
- Levitsky, Steven and Lucan A. Way. 2010. *Competitive Authoritarianism: Hybrid Regimes After the Cold War*. Cambridge University Press.
- Libman, Alexander and Anastassia V. Obydenkova. 2018. "Regional international organizations as a strategy of autocracy: The Eurasian Economic Union and Russian foreign policy." *International Affairs* 94:1037–1058.
- Lührmann, Anna, Kyle L. Marquardt, and Valeriya Mechkova. 2020. "Constraining Governments: New Indices of Vertical, Horizontal, and Diagonal Accountability." *American Political Science Review* pp. 811–820.

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

Lührmann, Anna, Marcus Tannenberg, and Staffan I. Lindberg. 2018. "Regimes of the world (RoW): Opening new avenues for the comparative study of political regimes." *Politics and Governance* 6:60–77.

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

Meyerrose, Anna M. 2018. "It is all about value: How domestic party brands influence voting patterns in the European Parliament." *Governance* 31:625–642.

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

Meyerrose, Anna M. 2021. In *International Sources of Democratic Backsliding*. Routledge.

Meyerrose, Anna M., Thomas Edward Flores and Irfan Nooruddin. 2019. In *From Elections to Democracy in Hard Times*. Oxford University Press.

Moravcsik, Andrew. 2000. "The Origins of Human Rights Regimes: Democratic Delegation in Postwar Europe." *International Organization* 54(2):217–252.

Neumayer, Eric. 2005. "Do International Human Rights Treaties Improve Respect for Human Rights?" *Journal of Conflict Resolution* 49(6):925–953.
<https://doi.org/10.1177/0022002705281667>

Pahre, Robert David. 1995. Wider and deeper: The links between expansion and integration in the European Union. In *Towards a new Europe: Stops and starts in regional integration*. Praeger pp. 111–136.

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

Poast, Paul and Johannes Urpelainen. 2018. *Organizing democracy: How international organizations assist new democracies*. University of Chicago Press.

- Risse, Thomas and Kathryn Sikkink. 1999. *The Socialization of International Human rights Norms into Domestic Practices: Introduction*. Cambridge University Press.
- Seligman, Steven. 2011. "Politics and principle at the UN human rights commission and council (1992-2008)." *Israel Affairs* 17:520–541.
- Terman, Rochelle and Erik Voeten. 2018. "The relational politics of shame: Evidence from the universal periodic review." *Review of International Organizations* 13:1–23.
- Terman, Rochelle and Joshua Byun. 2022. "Punishment and Politicization in the International Human Rights Regime." *American Political Science Review* 116:385–402.
- Tsutsui, Kiyoteru and Christine Min Wotipka. 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, Christian. 2015. "Democracy prevention: The international collaboration of authoritarian regimes." *European Journal of Political Research* 54:623–638.
- Voss, M. Joel. 2019. "The Use (or Misuse) of Amendments to Contest Human Rights Norms at the UN Human Rights Council." *Human Rights Review* 20:397–422.
- Vreeland, James Raymond. 2008. "Political Institutions and Human Rights: Why Dictatorships Enter into the United Nations Convention Against Torture." *International Organization* 62(1):65–101.
- Vreeland, James Raymond. 2019. "Corrupting International Organizations." *Annual Review of Political Science* 22:205–222.
- Walker, Christopher. 2016. "The Authoritarian Threat: The Hijacking of 'Soft Power.'" *Journal of Democracy* 27:49–63.
- Way, Lucan Ahmad. 2022. "The Rebirth of the Liberal World Order?" *Journal of Democracy* 33:5–17. <https://muse.jhu.edu/article/852741>
- Weyland, Kurt. 2017. "Autocratic diffusion and cooperation: the impact of interests vs. ideology." *Democratization* 24:1235–1252.
- Woo, Byungwon and Amanda Murdie. 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.

Trojan Horse Autocracies, International Organizations, and the Liberal International Order

Online Appendices

1. UNHRC Voting: Continuous Measure of Backsliding

While the main models in the manuscript operationalize backsliding as a binary variable, we re-estimated the same models using a continuous measure of backsliding in Table A1. In this model the main independent variable captures the change in a state's level of electoral democracy since 2006, with negative changes indicating a state has backslid. Since the main variable is now a change variable, we also control for a state's current democracy score to isolate the impacts of the change specifically on vote choice. Our results are robust to this alternative operationalization: states that have backslid since 2006 are less likely to abstain, relative to voting yes, but are more likely to vote no, rather than yes.

Table A1. Democratic backsliding and vote choice in the UNHRC, 2006–2021,
Continuous Measure of Democratic Backsliding

	<i>Dependent variable:</i>	
	Abstain (1)	No (2)
Demo score change since 2006	1.43*** (0.27)	−2.52*** (0.46)
Current dem score	−0.11 (0.15)	−2.43*** (0.22)
Target state = Israel	−0.03 (0.20)	0.13 (0.32)
Target dem score	1.83*** (0.32)	1.83*** (0.50)
Same region	−0.81*** (0.13)	−0.58*** (0.16)
Voting UNGA US agree	−0.21 (0.30)	0.11 (0.49)
Target UNGA US agree	−1.64*** (0.35)	−2.81*** (0.60)
Voting state HR score	−0.20*** (0.03)	−0.16*** (0.04)
Target state HR score	0.28*** (0.05)	0.10 (0.07)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Multinomial logit	Multinomial logit
Observations	11,245	11,245
K	3	3
Akaike Inf. Crit.	12,283.81	12,283.81

Note:

*p<0.1; **p<0.05; ***p<0.01

2. UPR Issues: Exclude Closed Autocracies

We re-estimate the models from Table 5, this time excluding observations for UPR reports written by closed autocracies. As mentioned in the manuscript, following Lührmann, Tannenberg and Lindberg (2018), we use the V-Dem dataset to identify closed autocracies as ones with a) an electoral democracy index score below 0.5 and b) a lack of multi-party elections at the national level. We use the multiparty elections variable (*v2elmulpar*) from V-Dem, and code a case as lacking multiparty elections if this variable equals 0 (no meaningful competition) or 1 (multiple candidates but no-party or single-party). The results of these models are reported in Table A2.

Table A2. UPR Issues Identified against Advanced Democracies, 2008–2020, Exclude Closed Autocracies

<i>Dependent variable: Count of issues identified in country report</i>		
	All Reports (1)	Reports by UNHRC Members (2)
Backslide since 2006	0.20*** (0.03)	0.25*** (0.03)
Current polyarchy score	0.39*** (0.08)	0.16 (0.12)
State under review polyarchy score	−1.57*** (0.42)	−0.82 (0.63)
Same region	0.28*** (0.03)	0.41*** (0.04)
Recommend UNGA US agree	−0.63*** (0.06)	−0.80*** (0.10)
State under review UNGA US agree	0.33*** (0.12)	0.30* (0.18)
Recommend HR score	0.01* (0.01)	0.02 (0.01)
State under review HR score	0.01 (0.01)	0.03*** (0.01)
Recommending state under review	−0.09*** (0.02)	−0.17*** (0.03)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Quasi-Poisson	Quasi-Poisson
Observations	3,289	1,630

Note:

*p<0.1; **p<0.05; ***p<0.01

3. UPR Issues: Continuous Measure of Backsliding

As in Table A1, in the models in Table A5 the main independent variable captures the change in a state's level of electoral democracy since 2006, with negative changes indicating a backsliding state.

Table A3. Issues identified in UPR reports on OECD states, 2008–2020, Continuous Measure of Democratic Backsliding

<i>Dependent variable: Count of issues identified in country report</i>		
	All Reports (1)	Reports by UNHRC Members (2)
Demo score change since 2006	−0.65*** (0.05)	−0.63*** (0.09)
Current demo score	0.15*** (0.03)	0.34*** (0.04)
State under review demo score	−0.47 (0.31)	−0.003 (0.50)
Same region	0.21*** (0.02)	0.31*** (0.03)
Recommend UNGA US agree	−0.57*** (0.05)	−0.70*** (0.10)
State under review UNGA US agree	0.59*** (0.10)	0.42*** (0.15)
Recommend HR score	−0.03*** (0.005)	−0.03*** (0.01)
State under review HR score	0.02*** (0.01)	0.02** (0.01)
Recommending state under review	−0.07*** (0.01)	−0.15*** (0.02)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Quasi-Poisson	Quasi-Poisson
Observations	6,259	2,654

Note:

*p<0.1; **p<0.05; ***p<0.01

4. UPR Actions: Exclude Closed Autocracies

Table A4. UPR Actions Recommended for Advanced Democracies, 2008–2020, Exclude Closed Autocracies

<i>Dependent variable: 1 for General or Specific Action, otherwise 0</i>		
	All Reports (1)	Reports by UNHRC Members (2)
Backslide since 2006	0.41*** (0.08)	0.59*** (0.12)
Current EDI score	−0.29 (0.42)	−0.49 (0.63)
State under review EDI score	−1.87 (2.24)	−3.41 (3.36)
Same region	0.49*** (0.15)	0.89*** (0.26)
Recommend UNGA US agree	0.22 (0.36)	0.60 (0.61)
State under review UNGA US agree	−0.21 (0.72)	−0.06 (1.08)
Recommend HR score	0.17*** (0.04)	0.07 (0.06)
State under review HR score	0.02 (0.05)	0.01 (0.07)
Recommending state under review	0.11 (0.10)	−0.02 (0.15)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Logit	Logit
Observations	3,287	1,630

Note:

*p<0.1; **p<0.05; ***p<0.01

5. UPR Actions: Continuous Measure of Backsliding

Table A5. Issues identified in UPR reports on OECD states, 2008–2020, Continuous Measure of Democratic Backsliding

<i>Dependent variable: Count of issues identified in country report</i>		
	All Reports (1)	Reports by UNHRC Members (2)
Demo score change since 2006	−0.65*** (0.05)	−0.63*** (0.09)
Current demo score	0.15*** (0.03)	0.34*** (0.04)
State under review demo score	−0.47 (0.31)	−0.003 (0.50)
Same region	0.21*** (0.02)	0.31*** (0.03)
Recommend UNGA US agree	−0.57*** (0.05)	−0.70*** (0.10)
State under review UNGA US agree	0.59*** (0.10)	0.42*** (0.15)
Recommend HR score	−0.03*** (0.005)	−0.03*** (0.01)
State under review HR score	0.02*** (0.01)	0.02** (0.01)
Recommending state under review	−0.07*** (0.01)	−0.15*** (0.02)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Quasi-Poisson	Quasi-Poisson
Observations	6,259	2,654

Note:

*p<0.1; **p<0.05; ***p<0.01

6. UPR Actions: Specific Action Only

In Table 6, we code reports that recommend either “General Action” or “Specific Action” as 1, and all others as 0. As an additional test of the extent to which backsliding states use these reports to target advanced liberal democracies, we create another dependent variable that is coded as 1 only if a report recommends “Specific Action,” the most demanding potential recommendation contained in a report. As Table A6 shows, our results are robust to focusing exclusively on reports recommending Specific Action.

Table A6. UPR Actions Recommended for Advanced Democracies, 2008–2020, Specific Action Only

	<i>Dependent variable: 1 for Specific Action, otherwise 0</i>	
	All Reports (1)	Reports by UNHRC Members (2)
Backslide since 2006	0.24*** (0.06)	0.31*** (0.09)
Current EDI score	−0.02 (0.15)	−0.20 (0.22)
State under review EDI score	−5.82*** (1.44)	−3.95* (2.27)
Same region	0.29*** (0.10)	0.67*** (0.16)
Recommend UNGA US agree	0.02 (0.26)	0.09 (0.46)
State under review UNGA US agree	−0.61 (0.48)	−0.36 (0.77)
Recommend HR score	0.12*** (0.03)	0.11** (0.04)
State under review HR score	−0.04 (0.04)	−0.15*** (0.05)
Recommending state under review	−0.08 (0.07)	−0.01 (0.11)
Year fixed effects	Yes	Yes
Clustered standard errors	Yes	Yes
Model type	Logit	Logit
Observations	6,254	2,654

Note:

*p<0.1; **p<0.05; ***p<0.01