Back from the Brink: Countering Illiberalism in Liberal Democracies

Building Tolerance for Backsliding by Trash-Talking Democracy: Theory and Evidence From Mexico

Lautaro Cella[1], Ipek Çınar[1], Susan Stokes[1], and Andres Uribe[2] ⓘ

Abstract
Leaders who seek to build public toleration for democratic backsliding have a little-noticed strategy at their disposal: degrading their democracies in the eyes of their citizens. If voters can be induced to believe that their democracy is already broken, then nothing of value is lost when leaders attack the courts, vilify the press, or undermine confidence in elections. We call this strategy *trash-talking democracy*, and study it in the context of contemporary Mexico. We use text-as-data methods to show that President Andrés Manuel López Obrador spent more time trash-talking his democracy than he did deepening partisan polarization. With a survey experiment we show that exposure to López Obrador's trash-talking of the courts elicits anti-democratic attitudes among Mexicans — both among his supporters and among supporters of the opposition. Strategies to resist backsliding should include not just efforts at de-polarization but also at restoring confidence in democratic institutions.

[1]University of Chicago, Chicago, IL, USA
[2]University of Wisconsin-Madison, Madison, WI, USA

Corresponding Author:
Andres Uribe, University of Wisconsin-Madison, 110 North Hall, Madison, WI 53706, USA.
Email: andres.uribe@wisc.edu

Data Availability Statement included at the end of the article.

Keywords

democratization and regime change, Latin American politics, quality of democracy

Backsliding leaders face a dilemma. They seek to weaken their country's democratic institutions without sparking a backlash from voters. Their challenge is to undermine horizontal accountability (e.g., from courts, the civil bureaucracy, or the legislature) without triggering vertical accountability from voters.[1] How do they deal with this dilemma?

Researchers offer several answers to this question. Some suggest that a prior decline in support for democracy paves the way for backsliders (see, e.g., Foa & Mounk, 2016). Other researchers find little evidence of such a decline (see Bartels, 2023; Voeten, 2016). Lupu et al. (2023) report drops in support for democracy in some Latin American countries. But they come after the onset of democratic erosion in several Latin American countries. Backsliding leaders can also sidestep electoral accountability. They can manipulate electoral districts (as occurred in Hungary), attempt to limit ballot access for opposition voters (as in the U.S.), or attempt to nullify the results of elections when they lose (as in the U.S., Brazil, and Venezuela). Public support for backsliding leaders is stronger when the leader displays competence, for instance in managing the economy (Frederiksen, 2022).

Another prominent explanation links erosion with partisan polarization. Even voters who favor democracy may be willing to trade away institutional integrity in pursuit of other prized goals. Because these trade-offs are more appealing when politics is polarized on partisan grounds, erosion is easier to carry out in polarized societies. Svolik (2019, p. 23) explains that "countries' acute society-wide political conflicts raise the stakes in elections and, in turn, the price their supporters have to pay for putting democratic principles above partisan interests." (See also McCoy, Rahman, & Somer, 2018; Graham & Svolik, 2020; Simonovits et al., 2022.) Somer et al. (2021) observe that backsliding leaders do not merely benefit from prior polarization, they actively try to further polarize their polities.

But partisan polarization has some downsides. When politicians mobilize their supporters by vilifying their opponents, they risk provoking an equal and opposite reaction on the other side. Thus when Donald Trump decries Democrats as "an angry mob, bent on destroying anything or anyone in their path," he risks turning off independents and mobilizing Democrats in ways that could nullify the gains he makes with his own base.[2]

In this paper we explore a second, less noticed strategy, and one that reduces the risk of backlash. Backsliders can *trash-talk democracy* — they can denigrate democratic institutions, depicting them as corrupt, ineffective,

elitist, or defective in other ways. We define democratic trash-talk (used interchangeably in this paper with *democracy-denigrating speech*) as rhetoric that diminishes the public's perceptions of the quality, fairness, or effectiveness of democracy or of its institutions as they operate in the speaker's country.

After developing the concept of trash-talking democracy in the next section, in the third one we study it in one eroding democracy, Mexico.[3] Under President Andrés Manuel López Obrador (2018–2024), Mexico's democratic institutions frayed. López Obrador did not attempt to loosen term limits or otherwise remain in power past his constitutional term in office. In this way and others that we discuss later in this paper, though his "attempts to weaken independent centers of power" and his "propensity to demonize opponents" did not escalate into a full authoritarian reversal, still he "eroded the conditions for pluralistic politics" (Sánchez-Talanquer, 2020).

López Obrador's presidency offers an excellent opportunity to systematically study the rhetoric of a backsliding leader. He gave lengthy morning press conferences nearly every weekday when he was in Mexico City. He thus created an enormous trove of rhetoric, which we analyze. We employ text-as-data analysis to identify the distinct registers of party-polarizing and trash-talking statements. We show that in López Obrador's communications, trash-talking outweighs polarization.

In the fourth section we report on a survey experiment, which we fielded in Mexico, that probes the effectiveness of López Obrador's rhetorical denigration of Mexican democracy. Do claims that horizontal institutions are corrupt and self-serving sully their image in voters' eyes? Did López Obrador's rhetorical attacks only shape the perceptions of his supporters, or did they also influence non-partisans and opposition-supporting voters? And does trash-talk avoid spurring a backlash among opposition supporters and non-partisans?[4]

What Does it Mean to Trash-Talk Democracy?

When backsliders trash-talk democracy in the sense we develop here, they do not question democracy as a *system of government*. Rather, they claim that particular institutions in their country are in shambles. And the particular institutions they criticize are the ones they are trying to subvert. Trash-talk can involve a range of criticisms: that the institution in question is led by corrupt individuals, that these leaders are elitist and unresponsive to voters, that they cost too much to run, or several others. More important than the precise grounds of the complaint is the message that the country would be better off tearing these institutions down and starting over, preferably with bodies that are more subservient to the executive — the president or prime minister.

When a leader deploys party-polarizing rhetoric, he or she implicitly says to supporters, *You may not like my attacks on our democracy, but they are the price you pay for keeping the hated other side out of power.* When a leader trash-talks democratic institutions, he or she implicitly says, *You shouldn't worry about my attacks on our democracy, our institutions are already in shambles.* To the extent that leaders succeed in vilifying the other party, they contribute to the polarization of the electorate. To the extent that they succeed in tarnishing their democracy in the public's eyes, they may contribute to public cynicism about institutions. Cynicism among opposition supporters and independents is preferable to counter-mobilization against the back-sliders, and against their project of erosion.

To be clear, presidents and prime ministers sometimes denigrate institutions *and* blame opposition parties for the sorry state that these institutions are in. We call this a mixed strategy, one that combines pure polarizing rhetoric, which emphasizing the perfidy and dangers of the opposition, and pure trash-talk, which emphasizes the corruption and self-serving nature of institutions and their leaders and staffs but does not link them to opposition parties.

A recognition of trash-talk as a weapon in backsliders' arsenal opens up fresh perspectives on strategies to counter democratic erosion (Capoccia, 2023). Resistance to illiberals in power should include actively countering their rhetorical onslaught against democratic institutions. And indeed, actors who resist backsliding do try to convince the public that democratic institutions are worth defending. The Mexican experience offers an example of the success of this strategy, when mass protests erupted 2023 in defense of the country's electoral administration body, as we explain at the end. The president's response was more trash-talk — of the National Electoral Institute (*Instituto Nacional Electoral*, INE) and also of the protesters. In this instance, citizens' protests and the courts helped block the backslider's effort to undermine a key democratic institution. But the episode suggests that claims and counter-claims about the value of institutions will be an ongoing dynamic in eroding democracies.

Regarding partisan polarization, how it helps backsliding leaders can be explained at the aggregate level of electorates: the larger the partisan divide among voters, the greater their tolerance for their leader's erosion of democracy. It can also be explained at the level of the individual voter: the more catastrophic she views the eventuality of the other side gaining power to be, the more tolerant she is of her leader's undermining of democracy.

But notice that there are two ways in which a person can become more inured to the erosion of democracy. A politician can increase the voter's tolerance of democratic erosion *either* by boosting her level of partisan polarization *or* by downgrading the value that she sees in her democracy. In aggregate, holding levels of polarization constant, the more degraded the public's perceptions of its democracy, the greater its tolerance for backsliding.

The effects of partisan polarization and of denigrating democracy are unlikely to be the same. The politician who uses rhetoric to drive up her society's level of polarization may succeed in mobilizing her supporters to get to the polls. But she may also mobilize her opponent's supporters to go to the polls and vote for him; and a polarizing discourse might push some non-partisans toward the opposition's side, as well. If her messages degrading institutions are effective and avoid mobilizing the opposition, trash-talk is more beneficial to the backslider. Because trash-talk avoids negative state-ments about the opposition party, it seems less likely to cause a negative backlash among them.[5] But whether this is true is an empirical question, and one we explore later in this paper.

Perhaps because they have these tradeoffs in mind, backsliders do make statements that seem designed to drive down support for democracy. As mentioned earlier, trash-talk is rhetoric that diminishes the public's percep-tions of the quality, fairness, or effectiveness of a country's democratic in-stitutions. In the current context, we focus on these rhetorical strategies as deployed by backsliding leaders. Democratic trash-talk is distinct from other kinds of criticisms of democracy in that the former (1) presents a biased picture of the shortcomings of the country's democratic institutions, over-stating some and ignoring others; and (2) proposes as a solution the ag-grandizement of the powers of the executive – the president or prime-minister.

To trash-talk democracy is not to question its status as a superior, if flawed, system of government (the Churchill proposition that democracy is the worst system of government except all the others). Hungary's Viktor Orbán stands out as a contemporary backsliding leader who comes close to questioning the value of democracy as a system. He claims that today's dynamic countries are all au-tocracies or eroding democracies and decries the bullying posture of the dem-ocratic West. More typically, backsliding leaders embrace democracy in general while disparaging its condition in their own countries. Their message: *Democracy may be a desirable system of government, but our own democratic institutions are hollow, controlled by actors who are self-serving and ill-intentioned. To restore true democracy, I must be allowed to take a wrecking ball to our institutions.* Spreading a culture of cynicism and *institutional nihilism* (Rau & Stokes, 2024) with regard to institutions like the courts, the press, and the civil administration is at the heart of the backsliding leader's strategy.

When politicians use rhetoric to *polarize* their electorates, they deride opposition parties – their policies, their leaders, their likely effect on the country should they prevail in elections. When they use rhetoric to *denigrate their democracies*, they deride institutions without tying the institutions' presumed failings to the influence of opposition parties.

Consider the harsh words about democracy deployed by Hugo Chávez in 1998, when he first campaigned for the presidency of Venezuela. Chávez invited his listeners to envisage Venezuela under a new constitution. An

entirely revamped institutional structure was "the only peaceful and democratic path …a leap away from a moribund democracy, a cauterized system, where the Republic came to an end …to a new democracy, a true democracy…" (July 30, 1998).

Chávez's words shed light an important aspect of democratic trash-talk. Though the criticisms of democracy may be cast in general terms, they frequently excuse specific actions that will aggrandize the power of the executive. Venezuela's 1999 constitution, drafted after Chávez's victory in the 1998 election, transferred many powers from the legislature to the executive, and eliminated the legislature's power to impeach the president.[6] Indeed, the constituent assembly chose "to disband all the existing state institutions that the Chavistas did not control" (Corrales, 2005, p. 107).

Other backsliding leaders similarly denigrate particular institutions that they are preparing to weaken. In the sampling that follows, leaders attack, in turn, their bureaucracy, their courts, and their national legislature. Note that none of these diatribes mentions a political party:

- In the 2018 presidential campaign, López Obrador of Mexico asserted that "corruption roams about with total impunity in *the public bureaucracy.* Never before have we suffered as much corruption as we do today, never in the history of Mexico since the colonial era have we suffered so much corruption" (April, 1, 2018).
- Israeli Prime Minister Benjamin Netanyahu said of his government's proposed overhaul of the *judiciary*, "The claim that this reform is the end of democracy is baseless. The truth is that the balance between the branches in the governmental system has been violated over the last two decades, and even more so in recent years."[7]
- Evo Morales, then recently elected president of Bolivia, reflected back on his days in the Bolivian Congress: "Always before there was talk of democracy, there was a fight for democracy, there was talk of a pact for democracy, a pact for governance. In 1997, when I came to this *parliament*, I did not personally see any pact for democracy or governance, but rather the pacts of corruption, a pact on how to get money from where and how …" (January 22, 2006).

Trash-talk often intensifies at moments when the president is trying to undermine a specific horizontal institution, such as the courts in Israel or the national election-administration body in Mexico. Again, the message: *these institutions' corruption, overreach, and incompetence justifies my meddling with them.*

When backsliding leaders try to exacerbate partisan polarization, they sound quite different. They vilify the opposition and warn voters about the disasters they will face should the other party come to power. Jair Bolsonaro,

running for president of Brazil in January 2018, deployed such speech. He asserted that the Workers Party (*Partido dos Trabahadores*, PT), which governed the country from 2003 to 2016, "plunged Brazil into the most absolute corruption, something never seen anywhere [else] in the world" (January 10, 2018). In the run-up to the 2018 U.S. mid-term elections, Donald Trump described the opposing party as wanting to "impose socialism on our incredible nation, make it Venezuela." They want to "throw open your borders to deadly and vicious gangs." Democrats "have truly turned into an angry mob, bent on destroying anything or anyone in their path" (October 18, 2018).

Polarizing and democracy-denigrating speech are conceptually distinguishable. A close reading of backsliding leaders' speeches suggests that they in fact can be assigned into separate categories, as explained in the next section. Leaders sometimes rail against institutions and denigrate democracy (trash talk), and sometimes rail against opposing parties and exacerbate partisanship in the electorate (polarization). Sometimes they do both: they decry the sorry state of the country's democracy and its institutions and place the blame squarely on the opposing party. As an example, Donald Trump persistently depicts the U.S. *judiciary* as infiltrated by *Democratic* judges, as in August 2023 when he accused the judge overseeing his federal trial in Washington D.C. of being "highly partisan."[8]

Our expectation, then, is that an aspiring backsliding leader's speech will contain three distinct types of statements: ones aimed at denigrating democracy (trash talk), ones aimed at exacerbating partisanship in the electorate (polarization), and ones that mix elements of the two (mixed). Figure 1 is a Venn diagram depicting the categories of pure trash-talk (left), pure polarization (right), and the mix of the two (center).

But perhaps the conceptual distinction we are drawing has little bearing in the real world. Two patterns would cast doubt on our distinction. Most rhetoric would fall into the right portion of Figure 1; polarizing speech would dominate. Or it would fall into the middle section; mixed messages, combining disparagement of opposition parties and disparagement of institutions, would prevail. To evaluate the explanatory power of our propositions, we turn to evidence from Mexico.

Trash Talk and Partisan Polarization in Mexico

Backsliding Under Andrés Manuel López Obrador

Mexico's transition to democracy, after more than 70 years of single-party rule, culminated in a major opposition party winning the presidential election in 2000. The long-ruling party was the Institutional Revolutionary Party (*Partido Revolucionario Institucional*, PRI).[9] Vicente Fox, from the center-right wing National Action Party (*Partido Acción Nacional*, PAN) won the
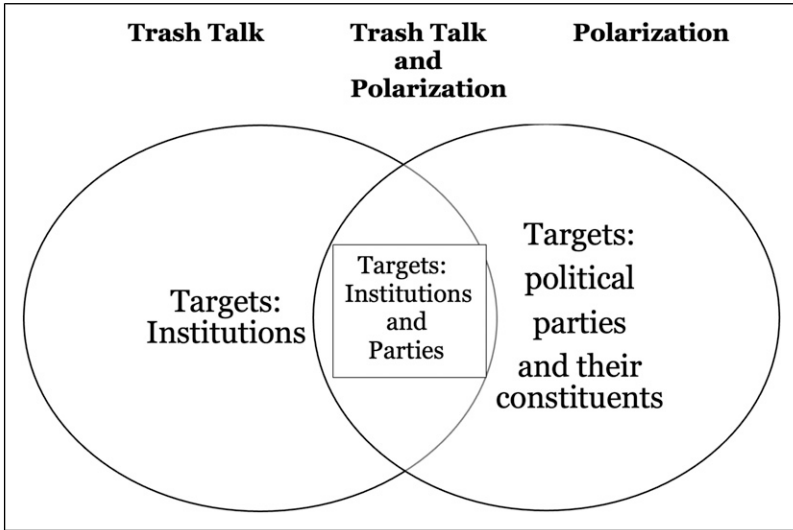
Figure 1. Venn diagram of trash talking and partisan polarization.

election and incumbent PRI President Ernesto Zedillo peacefully stepped down. After two PAN administrations, the PRI returned to power in 2012, holding the presidency until 2018. From 2000 to 2018, in presidential elections, as well as congressional and state elections, the PRI and the PAN competed along with a party of the left — for most of this period the Democratic Revolution Party (*Partido de la Revolución Democrática*, PRD).

López Obrador began his career as a PRI politician but in 1989 switched to the PRD. He served as the elected mayor of Mexico City from 2000–2005. As the PRD's candidate in the 2006 presidential elections, he was narrowly defeated by the candidate of the PAN. López Obrador claimed fraud and demanded a recount. The Federal Electoral Court demurred. It declared that evidence of fraud was insufficient to warrant a recount of all 41 million votes. In the 2012 presidential election López Obrador was defeated decisively, this time by the candidate of the PRI. López Obrador again made accusations of fraud. Soon after, he left the PRD and established a new leftist party, the National Regeneration Movement (*Movimiento de Regeneración Nacional*, MORENA). In his third presidential run in 2018, López Obrador was elected president with 53% of the vote. The runner-up was Ricardo Anaya, the PAN candidate.

Voter support for MORENA in 2018 was associated with affective polarization between MORENA, on one side, and both the PAN and the PRI, on the other. López Obrador encouraged the perception that the PRI and the PAN represented the same political alternative (Castro Cornejo, 2023a). Many

voters saw the PAN and PRI as responsible for disappointing economic performance, rising crime and violence, and persistent corruption (Greene & Sánchez-Talanquer, 2018).

During his presidency, López Obrador emphasized problems of social injustice, inequality, and corruption in Mexico. He also attempted to concentrate power in the executive and undercut the independence of horizontal institutions and semi-autonomous agencies. As two leading scholars of Mexican politics explained mid-way through the president's term,

> López Obrador used his mandate to centralize power in the executive and challenge core democratic norms and institutions, raising concerns about democratic backsliding. Similar to other populist figures around the world, [López Obrador] has capitalized on widespread citizen discontent to weaken checks and balances and to lock in competitive advantages for MORENA. The constitutional safeguards built into Mexican democracy to prevent the arbitrary use of power are now under stress (Sánchez-Talanquer & Greene, 2021).

Regarding the courts, Ríos-Figueroa notes that López Obrador's government "forced the resignation of one Justice and rushed the appointment of two Justices whose partisan affiliation and personal connections to one of the administration's main contractors were highly questionable." At the president's behest, the MORENA-dominated Congress advanced initiatives that threatened judicial independence "by increasing the number of justices and packing the court. All this, plus the daily doses [at morning news conferences] of the President's hostility towards judges and the judiciary, send a chilling message" (Ríos-Figueroa, 2023).

The López Obrador government undertook a similar effort to undermine the federal election-administration body, the National Electoral Institute (*Instituto Nacional Electoral*, INE) (Gamboa, 2023; Sánchez-Talanquer & Greene, 2021). The INE's predecessor, the Instituto Federal Electoral (IFE), had been crucial to Mexico's democratic transition (Eisenstadt, 2003; Magaloni, 2006).[10] In 2022, López Obrador proposed dismantling the INE. This plan was met with massive citizen protests in defense of the agency. The president then pivoted to a "Plan B." This reform, approved by Congress, aimed to reduce the INE's autonomy, budget, and personnel. In 2023, the Supreme Court struck down the reform; tensions between López Obrador and the judiciary escalated. Throughout this episode, the president launched bitter, often highly personal attacks at the INE leadership.

As a departing salvo in his conflict with the institutions that had impeded his plans, at the end of his term López Obrador's party in Congress approved a massive change whereby all federal judges would be elected. As Aguiar Aguilar et al. (2025) suggest, López Obrador advertised these changes as improving Mexican democracy by beating back elitism and corruption.

López Obrador's rhetoric has been characterized as polarizing and populist (Aguilar-Rivera, 2022; Mainwaring & Pérez-Liñán, 2023; Sánchez-Talanquer & Greene, 2021). In his daily press conferences, he galvanized "the people" against the "neoliberal elite" of the PRI and the PAN. His verbal attacks on institutions went beyond the courts and the INE to include the central bank, universities, the press, and civil society organizations. The president's critics accused him of eroding the independence of public entities which were by design meant to be autonomous of the executive branch. He responded by questioning the value of horizontal accountability and arguing that all parts of government should be accountable to "the people." Playing on the concept of *autonomy*, he criticized the courts, the INE, the Bank, and others for being "autonomous of the people."[11] And López Obrador did not mince words. Mexican analysts note that his favorite epithets included *hypocrite*, *racist*, *social climber*, *know-it-all* and *thief (ratero)*.[12]

Analyzing López Obrador's Morning News Conferences

To systematically examine López Obrador's discourse, we compiled a collection of his morning news conferences or *mañaneras*. These news conferences commenced at 7 am nearly every weekday, and lasted an average of one-and-a-half hours. López Obrador answered questions from journalists, opined about the news of the day, aired grievances, and discussed politics and policy. He was typically accompanied by one or more officials, who also took part. Though López Obrador occasionally gave speeches or sat for interviews, the *mañaneras* were the main way he communicated with the public. The *mañaneras* tended to set the national news cycle each day: they were broadcast on public television and streamed on YouTube, and government accounts and media organizations disseminated clips and wrote news stories based on them, which made the rounds on social media. The office of the Presidency claimed that roughly 10 million people watched the *mañanera* every morning (Martínez, 2021).

Helpfully for our purposes, López Obrador's office also transcribed each *mañanera* and uploaded the transcription to its website.[13] We scraped the transcripts corresponding to every *mañanera* that took place from the beginning of his term until October 6, 2023, comprising 1155 press conferences in total. This corpus represents a remarkable resource for studying López Obrador's political messaging. It is very large: transcription breaks the 1155 mañaneras into just under 300,000 sentences spoken by López Obrador.[14] It also offers valuable temporal coverage and granularity, allowing us to examine how López Obrador's rhetoric changed over time and in response to political developments.

We use this corpus to explore whether López Obrador's messages sort themselves into ones that denigrate democracy and institutions (trash talk),

ones that dwell on the presumed perfidy of opposing parties such as the PRI and PAN (polarizing speech), and ones that combine the two. After cleaning the corpus and subsetting it to only sentences uttered by López Obrador himself, we begin by identifying which sentences mention democratic institutions and which discuss partisan opponents.

We rely primarily on simple keyword searches. To identify trash-talking statements, we search for an exhaustive list of government agencies and their acronyms, supplemented with general terms like "judge" and "court." To identify polarizing statements, we use a list of party names, acronyms, and names of previous heads of state from opposition parties, complemented by generic terms like "conservative." Keywords used are listed in the Appendix.

A keyword search alone, however, is likely to miss a set of potentially polarizing statements: references to individual politicians belonging to opposition parties. To record these references, we use a large language model (LLM). Trained on vast collections of human speech, including political rhetoric, Wikipedia pages, and massive troves of news stories, the current generation of large language models are well suited to zero-shot text classification tasks like this one (Törnberg, 2024). We use the GPT-4o model to identify references to opposition politicians.[15] This combined approach yields 37,617 sentences referencing political institution or political opponents, representing roughly 13% of López Obrador's full set of utterances.

To be sure, not all of these paragraphs are rhetorical attacks on the institutions or opponents they mention. To distinguish criticisms from value-neutral or positive references, or from passing mentions, we turn again to the large language model. For each sentence that mentions government institutions or political opponents, we query GPT-4o as to whether that sentence represents a criticism of one or both of the targets.[16] In the Appendix, we validate the performance of this classifier and present the results of a simulation-based sensitivity analysis showing that classification error is unlikely to affect our main findings.

As discussed in the previous section, we argue that politicians can employ trash-talking, polarization, or a mix of the two strategies (see Figure 1). To allow for all three of these possibilities, we labeled a sentence as *trash-talk* when it contains a criticism of a democratic institution without referencing partisan opponents, as *polarizing* when it condemns partisan opponents without alluding to institutions, and as *mixed* when it merged both elements.

As an example, pure instances of López Obrador trash-talking Mexican democracy often centered on accusations of corruption. "The judicial branch," he argued on May 7, 2021, "has been completely spoiled, corrupted, dominated by corruption." He accused federal judges of being "accomplices of corruption." He frequently aimed his fire at the National Electoral Institute (INE). On June 8, 2023, for example, he accused the INE of "justify[ing] and legitimiz[ing] electoral fraud," and of enjoying "privileges" and "extremely high salaries."[17]

At other times, López Obrador's rhetorical attacks were in a pure *polarization* frame, attacking opposition politicians and parties as evil, corrupt, and fraudulent. López Obrador accused the PRI and PAN governments that preceded him of "looting the country, seiz[ing] control of the government, kidnapp[ing] it, and turn[ing] it into a committee serving a few" (November 5, 2020). He applied this Manichean lens to his current opponents, too, claiming that "there is a conservative bloc that is against the transformation of the country, and I have been saying this all along, that it is a very corrupt conservative bloc" (August 4, 2023).

Sometimes the president combined democratic trash-talk with partisan polarization, depicting the target institutions as beholden to opposing partisan forces; these are the *mixed* statements. The INE, he claimed, was not just corrupt but also captured by conservative political parties: leaders of the INE "have become the supreme conservative power, they decide who is a candidate and who is not" (March 26, 2021). So, too, were the courts. "During the neoliberal period," the courts "protect[ed] foreign investors," and "conservative judges" made decisions "to defend groups of vested interests" (May 26, 2021).

Figure 2 shows the frequency of trash-talking statements, polarizing ones, and combinations of the two, from December 2018 through September 2023. Statements are aggregated to the monthly level. Dotted vertical indicators denote election dates, with local elections held every June and a national legislative election in June 2021. The final months of our sample coincided
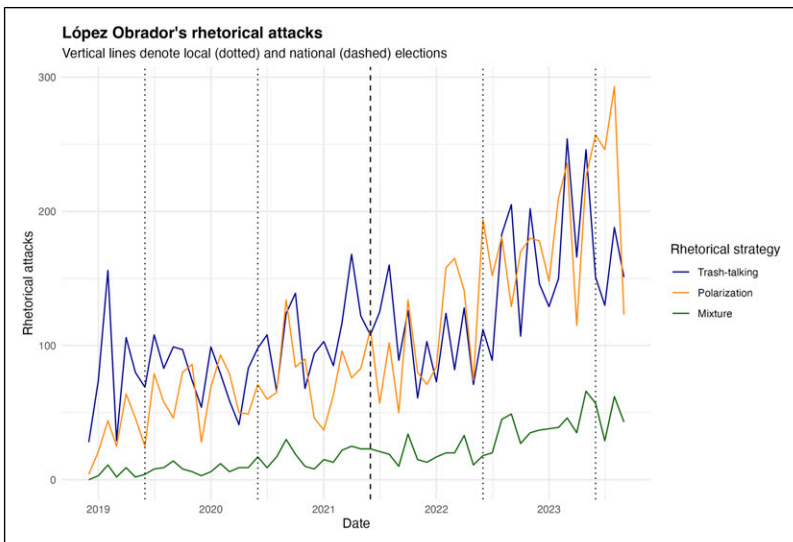

Figure 2. Trash talking, polarizing, and mixed statements by López Obrador.

with early campaigning in anticipation of the 2024 presidential elections that ultimately selected López Obrador's successor.

The plot underscores two main conclusions. First, trash-talking and partisan polarization appear to be distinct rhetorical tactics. To be sure, as in the examples mentioned earlier, López Obrador sometimes decried institutions as captured by his partisan opponents. But this mixed rhetorical strategy is comparatively rare. Of our combined sample of criticisms, only 9% incorporate elements of both trash-talking and polarization; the rest constitute pure examples of one tactic or the other.

The second conclusion suggested by Figure 2 is that trash-talking is common: most of López Obrador's hostile statements targeted institutions without relating their presumed defects to political parties. In 33 of 58 full months in our sample, López Obrador employed trash-talking more than partisan polarization. The second largest category were barbs with strong partisan elements, followed by statements that combined elements of the two strategies. The large number of democracy-denigrating messages, and the relatively small number of mixed ones, are confirmations that trash-talk is not just a theoretical possibility but the dominant rhetorical strategy López Obrador employed.

It is beyond our scope here to develop a full theory of backsliders' choices of rhetorical strategies. But note that the denigration of democratic institutions and attacks on opposing parties have somewhat distinct objectives. Both are in a sense electoral strategies, aimed at forestalling a negative reaction in voters to the leader's attacks on democratic institutions. But the denigration of democracy has the additional aim of blunting adverse popular responses to those attacks in real time. Mexico offers a good example: López Obrador unleashed much vitriol against the national electoral institute at same time that the Congress was voting on his proposals to de-fund and weaken it. In contrast, partisan polarization is more closely linked to elections. In Figure 2, López Obrador generally beat up on institutions at a higher rate until the run-up to the 2024 presidential elections, when, amid an increasingly vitriolic turn, his polarizing statements began to outstrip trash-talking.

To further place López Obrador's rhetorical attacks on institutions in the context of his broader political strategy, note that his harsh words do not extend to all of Mexico's public institutions. Notably exempt are the armed forces; like backsliding leaders elsewhere, he assiduously cultivated them.

When López Obrador denigrated non-military public institutions, he implied that the bad behavior of individuals is symptomatic of rot in the institution in which they serve. But with the military, he turned the argument around: bad behavior of individuals did not reflect on the institution in which they operated. They were the proverbial bad applies in a bushel of good ones. Hence, responding to evidence of the army's complicity in a case of mass disappearances of college students, in July 2023 López Obrador insisted that

"one thing is the institution and another the officials and public servants, and the bad behavior of one functionary does not stain an institution" (morning press conference, July 27, 2023).[18]

For more systematic evidence about López Obrador's institutional leanings, we repeated the procedures described above, this time directing the large language model to discern negative statements by the president about a set of military institutions.[19] It located 2347 sentences in which the president mentioned one or more of these entities. Of these mentions, it identified only 4% as criticisms. This in sharp contrast to statements about non-military public institutions – the courts, the INE, the bank, and so forth — 30% of which were critical.

The results in this section raise a number of additional questions. Under what conditions are leaders likely to talk trash, and under which to pursue polarization? What are the advantages of each, and the trade-offs they entail? Full answers to these questions will require additional research. In the conclusions we offer some possible answers.

Does Trash Talking Democracy Change Voters' Views of It?

Up to this point, we have argued that backsliding leaders try to keep the public on their side by using polarizing and democracy-denigrating rhetoric. And we have demonstrated that, in the case of a recent Mexican president, it makes sense to consider these two strategies as distinct and, frequently, undertaken one at a time. It remains to explore whether backsliding leaders' trash talk gets through to voters. If the leader's aim is to spread a mood of cynicism and thus boost their tolerance of democratic erosion, does the strategy work? And which kinds of voters are influenced by trash-talk — just the president's supporters, or opposition voters and non-partisans as well?

To assess the impact of trash-talking rhetoric, in September 2023 we fielded a survey experiment in Mexico, with 3001 participants. They encountered two distinct statements by López Obrador: one with trash-talking rhetoric targeting the judiciary, the other a statement about the judiciary that was neutral in tone. Both texts drew on the president's actual statements at morning news conferences.[20]

Respondents in the **control** group were presented with the following text:

> At a recent morning conference, President Andrés Manuel López Obrador said: "Good morning. It is important to bear in mind the responsibilities of the Judiciary. The Constitution of Mexico formally establishes that there are three Powers: the Executive represented by the president; the Legislative personified by deputies and senators, as well as the Judicial branch which is made up of judges, magistrates and ministers.

The Judiciary is responsible for ensuring that justice is delivered; they are the ones who must always be on the lookout. The Judiciary, the Supreme Court, the Judicial Council, address this matter. Their job and basic function is to assure justice. There are public ministries and judges, and magistrates and ministers, who are the ones who decide who is guilty and who is not guilty. The various judicial bodies, with their specific roles, contribute to the functioning of the judiciary power."[21]

Respondents in the **treatment** group were presented with the following text:

At a recent morning conference, President Andrés Manuel López Obrador said: "Good morning. It is important to bear in mind the responsibilities of the Judiciary. The Constitution of Mexico formally establishes that there are three Powers: the Executive represented by the president; the Legislative personified by deputies and senators, as well as the Judicial branch which is made up of judges, magistrates and ministers.

However, the sad reality is that the judiciary today is riddled with inefficiency and corruption. It is taken over by white-collar crime and organized crime. Judges and magistrates are often influenced by money and grant protection to criminals. They are not people characterized by honesty. The judiciary is rotten. We have one of the world's priciest judicial systems and one of the most inefficient. We're wasting citizens' taxes on a broken system. The judicial power needs reform."[22]

Respondents were then asked a series of outcome questions. We crafted the questions to probe for anti-democratic attitudes and levels of institutional nihilism. They were asked to indicate whether they strongly agree, agree, neither agree nor disagree, disagree, or strongly disagree with each of the following statements:

- The President should be able to fire judges who oppose him.
- The President should obey court rulings even when they go against his government.[23]
- The President should have more power over institutions such as the Judiciary, the INE (National Electoral Institute), and the Banxico [Bank of Mexico].
- We cannot fix the problems in our political institutions, so we need to tear them down and start over.

We incorporate the four outcome measures into an index of tolerance for anti-democratic actions. We reversed the response scale for the "president should obey the courts" question, so in all cases, higher scores indicate more anti-democratic attitudes. Our expectation is that exposure to democracy-denigrating statements

increases tolerance for anti-democratic actions (higher levels of support for firing judges, disobeying judicial decisions and for reducing the autonomy of institutions), and encourages institutional nihilism (support for tearing down institutions and starting over).

Previous research suggests that most Mexicans perceive actions such as the purge of disloyal civil servants in the judiciary as anti-democratic. An experiment conducted in Mexico in 2020 found that most respondents viewed a hypothetical incumbent's purging of independent-minded career prosecutors in the justice department as at odds with democracy (Albertus & Grossman, 2021).[24] This finding aligns with previous research indicating that a liberal understanding of democracy is prevalent, though not universal, among the population in Latin America (Canache, 2012; Svolik, 2020).

Earlier we posited that the rhetorical denigration of democracy might allow backsliders to avoid mobilizing opposition or non-partisan voters. Of interest, then, are possibly distinct reactions among partisan sub-groups. Before the treatments, we probed respondents' partisanship, asking them which party they feel closest to.[25] Those who answered "None" were coded as non-partisans (27% of the sample). Those closest to MORENA represented 47% of the sample. Those closest to an opposition party (PAN, PRI, PRD) were 18%.[26]

The treatment language comes from López Obrador's actual statements, and the treatments make this clear. This, then, is a hard test; we might also expect opposition-party supporters to brush aside the diatribes of a president whom they do not admire, or even to react against them.

Experimental Results

The key results from our survey are: (1) exposure to trash-talking rhetoric increases anti-democratic attitudes among our Mexican samples; and (2) it does so across the board — among the president's supporters, those who oppose his party, and nonpartisan voters.[27]

Figure 3 displays the effect of the trash-talking treatment on people's average scores on the anti-democratic index and on each anti-democratic item. All effects are compared to the responses of people in the control group. The dots represent point estimates, the lines 95-percent confidence intervals (with robust standard errors).[28] Exposure to the treatment boosts scores on the anti-democratic index by 6.2%.[29]

The pattern is consistent across all anti-democratic attitudes. López Obrador's attacks on the judiciary increase levels of support for firing judges, for disobeying judicial decisions, and for reducing the autonomy of institutions; it also increases institutional nihilism. For example, exposure to López Obrador's rhetoric increases agreement that the president should be able to fire judges who oppose him by 7.4%. It encourages institutional nihilism — the
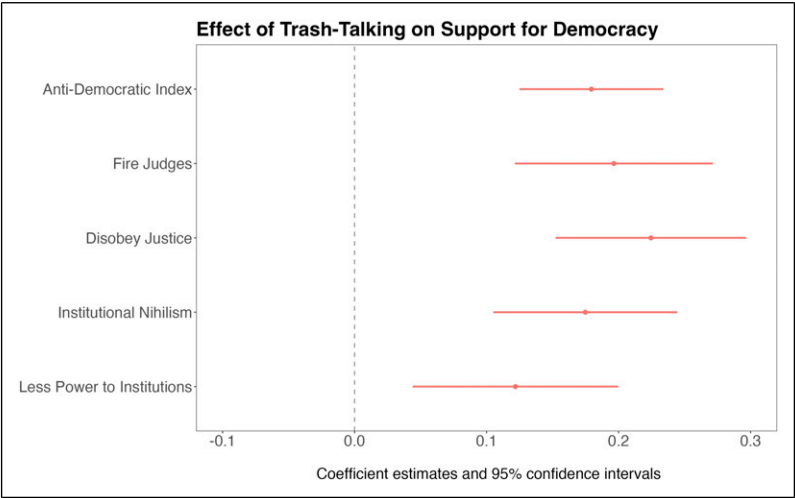
Figure 3. Experiment results.

view that political institutions should be "torn down" and Mexico should start over — by 5.4%.

In line with our proposition that a trash-talking strategy can avoid a backlash against the backsliding leader among opposition voters, *all* partisan types in our treatment group became more tolerant of democratic erosion. When López Obrador attacks the judiciary, the reactions of MORENA supporters, nonpartisans, and supporters of opposition parties run in parallel.

Beyond the overall similarities, are there any differences in the responses of partisan groups? We estimate treatment effects separately, with the results depicted in Figure 4. It shows the average treatment effects of the treatment on the pooled sample — on MORENA supporters (red), on nonpartisans (green), and on supporters of opposition parties (blue). Again, the comparison group in each case is people with the same partisan identities who were assigned to the control group.[30] The treatment effects on MORENA supporters appear slightly larger than those for nonpartisans. Yet these differences are not statistically significant.

The uniformity of reactions across partisan groups might be surprising, given the wide scholarship demonstrating the power of partisanship to color Mexicans' political perceptions (Castro Cornejo, 2023b; Samuels & Zucco, 2018). Some scholars do find that these voters are sometimes receptive to messages from the opposition and sometimes hold their co-partisans accountable (Aruguete et al., 2021; Klašnja et al., 2021). Our results underscore the benefits to backsliding leaders of avoiding a partisan lens and of focusing sharply on the presumed shortcomings of Mexico's democratic institutions.
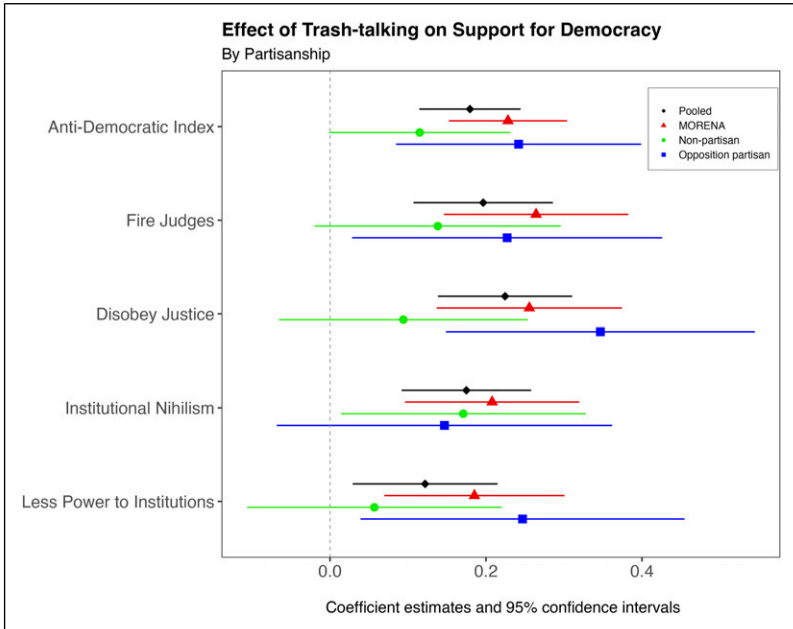
Figure 4. Experiment results by partisanship.

We have seen that partisans of different stripes react in similar ways to their president's diatribe against the courts. But it is also true that these partisans came into our study with different attitudes toward democracy — differences that can be seen by comparing anti-democracy scores in our control group. A substantial subset of respondents in our sample harbor anti-democratic attitudes. In the control group, anti-democratic attitudes are highest for MORENA partisans, significantly lower for nonpartisans and opposition partisans. Figure 5 plots the mean values on the anti-democratic index for the control and treatment groups, broken down by partisanship.[31] The tails of the arrows represent the control group means; the arrowheads represent the treatment means. The solid arrowheads indicate statistically significant results, the transparent one (for non-partisans) indicate that this effect falls short of significance at the 95-percent level.

Since partisanship is not randomly assigned, we can be less confident that differences in partisanship are what lie behind differences in baseline views of democracy. But if indeed MORENA supporters are more willing to sacrifice democratic practices, even without treatment, this might indicate that they have already been "treated," repeatedly, in real life. They are likely to have heard López Obrador's diatribes more regularly and to have internalized them
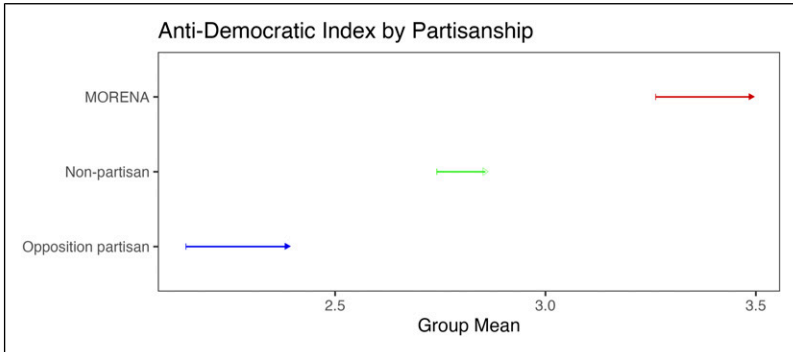
Figure 5. Control and treatment group means. *Note*: Vertical lines are the control group mean, and arrowheads the treatment group mean. Arrowheads are either filled or hollow to indicate statistical significance at the 95% level.

more fully. More research needs to be done to reach firm conclusions regarding partisanship and tolerance for anti-democratic actions in Mexico.

Trash-talk, then, builds tolerance for democratic erosion. But what lies behind this reaction? One interpretation is that the leader provides information (whether accurate or not) about the nature of an institution. In this case, he informs voters that the judiciary is very corrupt.[32] This new information — that corruption is, presumably, widespread — induces people to update their beliefs and, by extension, to distrust the broader array of democratic institutions.

Responses to a post-treatment question about the level of corruption in the Mexican judiciary allow us to probe this informational mechanism. The question was designed as a manipulation check — were survey subjects paying attention to López Obrador's diatribe? In fact we found stark differences across partisan groups in response to the manipulation check. MORENA supporters — and only MORENA supporters — were persuaded that the courts were, indeed, highly corrupt (see Figure 6). But the same cannot be said about non-partisans or supporters of opposition parties. The figure shows that the treatment had no significant effect on the perceptions of levels of corruption in the Mexican judiciary among opposition party supporters or among non-partisans.

Recall, however, that the treatment *was* effective in making these non-MORENA voters more anti-democratic. But if not because they viewed the judiciary as more corrupt, what explains this reaction? A distinct mechanism that could be at work involves people's emotions. Negative messages about one's environment tend to elicit emotions like anger. Perhaps harsh pictures of the Mexican judiciary cause the recipients of those messages to feel angry, an emotion that might turn them off democracy more broadly. We gathered self-

reported emotional responses, post-treatment, asking people to what extent they feel angry after being exposed to the trash-talking treatment. Figure 7 shows the average treatment effects of trash-talking on feelings of anger. We find that *all* partisan types in our treatment group became angrier. (For a similar finding, see Castro Cornejo et al., 2020).

In the Appendix, we report mediation analyses, which allow us to evaluate whether the president's diatribe induced an anti-democratic stance because it triggered negative emotions, like anger. The answer is that it did. Among MORENA respondents, we find that the treatment effect is channeled through increased corruption perceptions and anger. Among nonpartisans, the treatment effect is not channeled through corruption perceptions but is channeled through anger. The same is true for opposition partisans, though here the mediation effect is in the expected direction but not statistically significant.
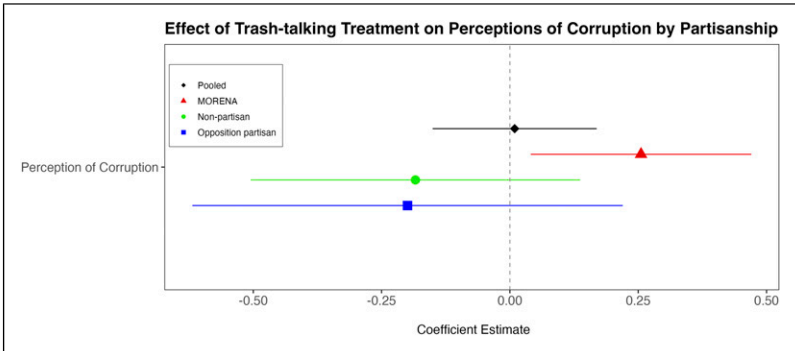

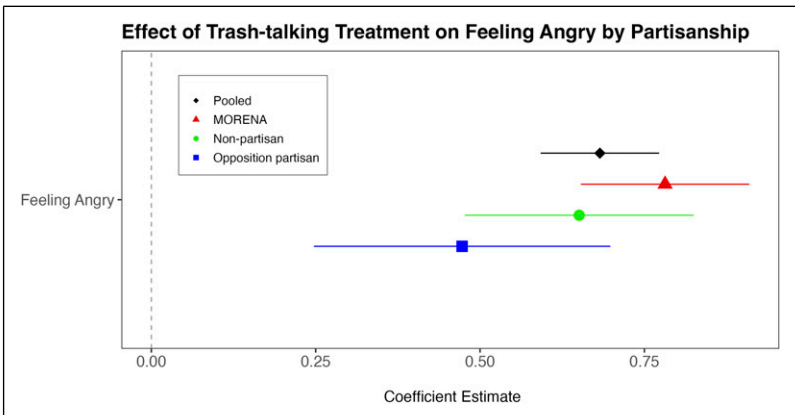**Figure 6.** Treatment on corruption perceptions within the judicial Power.

Figure 7. Treatment on feeling angry.

Conclusions and Discussion

Why do voters not always turn against backsliding leaders who threaten democratic institutions? The question is especially germane given that most citizens of backsliding countries report being in favor of democracy. The usual answer that scholars offer focuses on partisan polarization. When partisans in the electorate despise the other party, they will tolerate undemocratic actions by their own side as the price to be paid for keeping the detested others out of power.

But backsliding leaders can weaken horizontal accountability with a different communicative strategy, one that softens support for democracy with less risk of a backlash among independent and opposition voters. They can talk trash about their democracies. Judges, bureaucrats, journalists, election administrators: all are depicted as corrupt or ill-intentioned. The implicit message the leader sends is *don't worry about my attacks on our democracy, it's not a real, high-quality, functioning democracy at all.*

Our paper offers evidence that trash-talk is real and can have nefarious effects on voters. It also raises several questions for future research. One is, Under what conditions are backsliders likely to denigrate their countries' political institutions? Under which will they be more likely to double down on partisan polarization? One feature that may matter is the relative sensitivity of voters to both strategies. Will polarizing messages rev up their own base more than it will alienate and hence mobilize opponents and non-partisans? If a central point of trash-talk is to spread a demobilizing cynicism among voters, are the backsliders' own partisans in danger of being demobilized as well?

Another factor we would expect to drive backsliders' choice of rhetorical strategies is the prior level of polarization of the electorate. Partisan polarization is both an exogenous factor that politicians confront, and an endogenous result of their efforts. Our expectation is that polarized electorates invite polarizing speech. This dynamic leads to the kind of downward spiral of "pernicious polarization" that scholars of democratic erosion warn about (see Somer et al., 2021). But would-be backsliders who find themselves presiding over an electorate that is not especially polarized would risk driving away non-partisans, independents, and weak supporters of other parties, should they build a narrative of a nefarious, rather than merely wrong-headed, opposition party.

Negative attacks, like partisan polarization and trash-talk, are not the only strategies available to politicians, backsliders or conventional ones. Sub-national research lends support to the claim that polarization encourages more polarizing speech. Çınar (2022) shows, in the Turkish setting, that a would-be autocrat will engage in polarizing speech in districts in which their followers are preponderant, playing a game of mobilization rather than of persuasion. But in more mixed and opposition-dominant districts the backslider does not

just trash-talk Turkish democracy but focuses on positive messages, touting his record of accomplishments.[33]

In sum, the role of electoral vulnerability, anti-institutional efforts, and level of prior partisan polarization are likely to be factors that influence backsliders' choice of rhetorical strategy. More research, in a broader set of countries and regions, will be needed to assess these propositions.

Our research raises questions about effective resistance to backsliding. Up to now, efforts to inoculate voters against backsliding leaders have focused on de-polarization. For instance, commentators and researchers offer many recommendations about how to "fix" American politics by depolarizing it.[34] These include emphasizing disagreement within political parties, avoiding dehumanizing language, and expressing empathy for people in the other "tribe." These measures are well-suited to reducing partisan polarization but would seem to have little direct effect on shoring up people's sense of democratic institutions as worth saving.

There are exceptions, in which efforts to depolarize and to restore confidence in democratic institutions overlap. Both require the debunking of misinformation. Polarization involves, in part, exaggerating the differences between supporters of competing parties, or the number of dimensions on which they differ. Trash talk involves, in part, exaggerating the rottenness of institutions, as well as selectivity in which institutions are called out. Corruption is certainly a pressing issue in Mexico; but is the country "more corrupt than at any time in its history," going back to the colonial period, as López Obrador insisted? His denunciations of corruption were also selective, never mentioning cases of government contracts, or official party misallocation of funds, both of which have been documented by Mexican transparency groups (Monasterio & Maite Laris, 2020). Are Polish judges really doddering alcoholics, as allies of President Duda claimed? Are leaders in the West the bullies and hypocrites of Viktor Orbán's diatribes?

More research will be required to establish the most effective strategies for debunking this kind of misinformation. What is unlikely to work is whitewashing a country's politics, ignoring the very real problems that all democracies face. Nor should counter-strategies to democratic denigration be purely rhetorical. When the public sees democratic institutions working in their interests, they will be more resistant to backsliders' trash talk.

Is it quixotic to hope that the public may rally to the defense of democratic institutions? In Mexico, as we have seen, resistance to attacks on federal election administration brought masses of demonstrators to the streets of cities throughout the country, and may have hardened the spine of the court that eventually rejected the government's proposed reforms. In Israel, a leader who was flirting with the backsliders' playbook by trying to weaken an independent judiciary faced massive protests by civil society organizations, opposition parties, and individuals. These actors sustained protests over eight

months in 2023, in defense of the judiciary and against executive aggrandizement.

The lesson from these experiences is that the public can believe that democratic institutions are worth saving from an encroaching executive. More than that, they can be actively rallied in support. What is likely to be required to counter trash-talk about democracy are effective institutions that people view as worth saving.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iD

Andres Uribe https://orcid.org/0000-0002-2026-0462

Data Availability Statement

Replication materials and code can be found at Uribe (2025).

Supplemental Material

Supplemental material for this article is available online.

Notes

1. On democratic backsliding as a decline in both dimensions of accountability, see Laebens (2022).
2. "Remarks at a 'Make America Great Again' Rally in Missoula, Montana." https://www.presidency.ucsb.edu/documents/remarks-make-america-great-again-rally-missoula-montana.
3. See also Ríos-Figueiroa (2023); Sánchez-Talanquer and Greene (2021).
4. Replication materials and code can be found at Uribe (2025).

5. There are likely downsides of trash-talk, as well, such as spurring cynicism that spreads to the backsliders' own followers, and may therefore discourage them from going to the polls.

6. Like López Obrador, Chávez preferred at least an appearance of vertical over horizontal accountability, making broad use of popular referendums, including in the creation of a constituent assembly and approval of the constitution that it drafted.

7. Israel is not considered an eroding democracy, according to standard criteria (see Laebens, 2022; Rau & Stokes, 2024). But the ruling coalition's recent judicial reforms in Israel are a classic instance of the leader's playbook.

8. https://abcnews.go.com/Politics/trump-lashes-judge-warned-inflammatory-remarks-campaign-trail/story?id=102256338.

9. The Partido Nacional Revolucionario (PNR) was founded in 1929. In 1938, it was reconstituted as Partido de la Revolución Mexicana (PRM). In 1946, the party was refounded and acquired its current name, PRI.

10. In 2014, the INE supplanted the IFE.

11. For instance, at a March 6, 2023 morning news conference, the president said of the central bank, the Bank of Mexico, "They are autonomous also, autonomous of the people, not of financiers."

12. https://www.infobae.com/america/mexico/2022/11/17/cuales-son-los-insultos-que-mas-veces-ha-dicho-amlo-en-la-mananera-durante-estos-cuatro-anos/.

13. https://lopezobrador.org.mx/.

14. We discard sentences with 10 or fewer characters, which are usually thought fragments or instances of cross-talk with interlocutors.

15. GPT-4o was at the time of writing the latest iteration of the general-purpose language model that underpins OpenAI's chatGPT application. The model version we use was trained on data through the end of 2023, after our sample period concludes.

16. We used the following prompt: "Does this statement by Mexican President Andrés Manuel López Obrador criticize government institutions or political opponents? Answer only yes or no".

17. Other references to government institutions were not coded as criticisms. Some praised institutions: "We must protect the Army because it is a fundamental institution of the Mexican state" (April 10, 2023). Others were neutral policy announcements: López Obrador announced that he had received "a resolution from the Supreme Court of Justice, they gave us a deadline to decide on the legalization of marijuana" (January 27, 2020).

18. "*Pero una cosa son las instituciones y otra cosa son los funcionarios o los servidores públicos, y no por el hecho del mal comportamiento de un funcionario se va a manchar una institución*".

19. These included the Secretaría de la Defensa Nacional, Secretaría de Marina, Ejército Mexicano, and Fuerza Aérea Mexicana.

20. The survey was administered in September 2023, using an online panel of respondents recruited by the firm NetQuest. Respondents were randomly assigned to either the control group (neutral statements) or the treatment group (democracy-denigrating statements), using an equal probability-based simple random selection. Random assignment ensured that the treatment and control groups were, on average, similar in all relevant respects. A balance table comparing treatment and control groups can be found in the Appendix. Differences in responses can probably be attributed to the treatments and not to other factors that might influence people's views of democracy. We used two pre-treatment attention checks and discard inattentive respondents. The full survey instrument, in Spanish and English, can be found in the Appendix to this paper.

21. This experimental statement includes statements from López Obrador's *mañaneras* on 1/30/2020, 7/15/2021, 7/21/2021, and 7/8/2022.

22. This experimental treatment features statements sourced from López Obrador's *mañaneras* on dates 8/30/2022, 3/2/2023, 4/13/2023, and 5/25/2023.

23. We reverse the response scale in analyses that use this outcome question. We present it this way to discourage respondents from simply giving identical answers, without reflection.

24. The same holds regardless of whether the incumbent transgressed formal laws or only norms. The experimental conditions survey respondents were exposed to did not mention the word "democracy." The experiment was also conducted in Argentina, Brazil, and the US, with similar results. In Mexico, individuals who voted for AMLO were more likely to judge the action as consistent with democracy than those who did not vote for AMLO.

25. "Of the following political parties, which of them do you most identify or sympathize with?"

26. Supporters of Partido del Trabajo (PT), Partido Verde Ecologista de México (PVEM), Movimiento Ciudadano, and other parties were placed in a blanket category (8%). In some elections, PT, PVEM, and Movimiento Ciudadano made alliances with MORENA or endorsed MORENA candidates.

27. We did not find significant results for an outcome question about having a strong leader, see Appendix for details.

28. All outcomes vary from 1 to 5, with higher values indicating more anti-democratic positions.

29. The mean anti-democratic index is 2.89 for the control and 3.07 for the treatment. So, exposure to the treatment boosts the ant-democratic index by 6.2%. The average treatment effect is 0.18.

30. In the Appendix, we include estimates for non-MORENA supporters, grouping together nonpartisans and supporters of opposition parties.

31. The mean values for other outcomes for the control and treatment groups can be found in the Appendix.

32. Subjects assigned to the trash-talking treatment read a short debrief at the end of the survey. "In this survey, you read statements about the Judicial Power in

Mexico. Please read the following clarification. Regarding expenses in the judicial system, other countries in the world and in the Americas spend more than Mexico, both in absolute terms and per capita. For example, Argentina, Brazil, Canada, and the United States spend more. On the other hand, most cases of corruption in Mexico do not occur in the judicial branch."

33. This logic is not exclusive to backsliding leaders, like Recep Tayyip Erdoğan in Turkey. What is distinctive is these leaders' particular interest in deepening partisan polarization in order to reduce popular opposition to erosion.

34. https://carnegieendowment.org/2020/07/23/7-ideas-to-reduce-political-polarization.-and-save-america-from-itself-pub-82365.

References

Aguiar Aguilar, A. A., Cornejo, R. C., & Monsiváis-Carrillo, A. (2025). Is Mexico at the gates of authoritarianism? *Journal of Democracy*, *36*(1), 50–64. https://doi.org/10.1353/jod.2025.a947883

Aguilar-Rivera, J. A. (2022). Dinámicas de La Autocratización: México 2021. *Revista de Ciencia Politica*, *42*(2), 355–382. https://doi.org/10.4067/s0718-090x2022005000109

Albertus, M., & Grossman, G. (2021). The Americas: When do voters support power grabs? *Journal of Democracy*, *32*(2), 116–131. https://doi.org/10.1353/jod.2021.0023

Aruguete, N., Calvo, E., Cantú, F., Ley, S., Scartascini, C., & Ventura, T. (2021). Partisan cues and perceived risks: The effect of partisan social media frames during the COVID-19 crisis in Mexico. *Journal of Elections, Public Opinion, and Parties*, *31*(sup1), 82–95. https://doi.org/10.1080/17457289.2021.1924740

Bartels, L. M. (2023). *Democracy erodes from the top: Leaders, citizens, and the challenge of populism in Europe*. Princeton University Press.

Canache, D. (2012). Citizens' conceptualizations of democracy: Structural complexity, substantive content, and political significance. *Comparative Political Studies*, *45*(9), 1132–1158. https://doi.org/10.1177/0010414011434009

Capoccia, G. (2023). What countries are at risk? Inequality and democratic erosion. Working Paper.

Castro Cornejo, R. (2023b). Same scandal, different interpretations: Politics of corruption, anger, and partisan bias in Mexico. *Journal of Elections, Public Opinion, and Parties*, *33*(3), 497–518. https://doi.org/10.1080/17457289.2022.2120487

Castro, C., Rodrigo, S. L., & Beltrán, U. (2020). Anger, partisanship, and the activation of populist attitudes in Mexico. *Politica y Gobierno*, *27*(2), 1–37.

Castro Cornejo, R. (2023a). The AMLO voter: Affective polarization and the rise of the left in Mexico. *Journal of Politics in Latin America*, *15*(1), 96–112. https://doi.org/10.1177/1866802x221147067

Çınar, I. (2022). Creeping authoritarianism beyond 'mere words': Rhetorical strategems of democratic eroders. Unpublished manuscript. University of Chicago.

Corrales, J. (2005). In search of a theory of polarization: Lessons from Venezuela, 1999-2005. *European Review of Latin American and Caribbean Studies/Revista Europea de Estudios Latinoamericanos y del Caribe*, 79, 105–118. https://doi.org/10.18352/erlacs.9666

Eisenstadt, T. A. (2003). *Courting democracy in Mexico: Party strategies and electoral institutions*. Cambridge University Press.

Foa, R. S., & Mounk, Y. (2016). The danger of deconsolidation: The democratic disconnect. *Journal of Democracy*, 27(3), 5–17. https://doi.org/10.1353/jod.2016.0049

Frederiksen, K. V. S. (2022). Does competence make citizens tolerate undemocratic behavior? *American Political Science Review*, 116(3), 1147–1153. https://doi.org/10.1017/s0003055422000119

Gamboa, L. (2023). How oppositions fight back. *Journal of Democracy*, 34(3), 90–104. https://doi.org/10.1353/jod.2023.a900435

Graham, M. H., & Svolik, M. W. (2020). Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States. *American Political Science Review*, 114(2), 392–409. https://doi.org/10.1017/s0003055420000052

Greene, K., & Sánchez-Talanquer, M. (2018). Latin America's shifting politics: Mexico's party system under stress. *Journal of Democracy*, 29(4), 31–42.

Klašnja, M., Lupu, N., & Tucker, J. A. (2021). When do voters sanction corrupt politicians? *Journal of Experimental Political Science*, 8(2), 161–171. https://doi.org/10.1017/xps.2020.13

Laebens, M. G. (2022). Beyond democratic backsliding: Incumbent takeover attempts, 1989-2019. Working Paper, Oxford University.

Lupu, N., Rodriguez, M., Wilson, C., & Zechmeister, E. (2023). Pulse of democracy. LAPOP Technical Report.

Magaloni, B. (2006). *Voting for autocracy: Hegemonic party survival and its demise in Mexico*. Cambridge Studies in Comparative Politics Cambridge University Press.

Mainwaring, S., & Pérez-Liñán, A. (2023). Why Latin America's democracies are stuck. *Journal of Democracy*, 34(1), 156–170. https://doi.org/10.1353/jod.2023.0010

Martínez, L. (2021). The Mexican president's COVID diagnosis has paused his exhausting daily press conferences. Slate.

McCoy, J., Rahman, T., & Somer, M. (2018). Polarization and the global crisis of democracy: Common patterns, dynamics, and pernicious consequences for democratic polities. *American Behavioral Scientist*, 62(1), 16–42. https://doi.org/10.1177/0002764218759576

Monasterio, L. O., & Maite Laris, G. U. (2020). Corrupción y confianza en las instituciones. https://contralacorrupcion.mx/corrupcion-y-confianza-en-las-instituciones/

Rau, E., & Stokes, S. (2024). *Institutional nihilism: How inequality puts democracy in danger. Working Paper.* Chicago Center on Democracy.

Rau, E., & Stokes, S. (2024). Income inequality and the erosion of democracy in the twenty-first century. *Proceedings of the National Academy of Sciences*, 122(1), e242254312.

Ríos-Figueiroa, J. (2023). Democratic backsliding and the supreme court in Mexico. Verfassungsblog: On Matters Constitutional.

Samuels, D. J., & Zucco, C. (2018). *Partisans, antipartisans, and nonpartisans: Voting behavior in Brazil*. Cambridge University Press.

Sánchez-Talanquer, M., & Greene, K. F. (2021). Is Mexico falling into the authoritarian trap? *Journal of Democracy*, *32*(4), 56–71.

Simonovits, G., McCoy, J., & Littvay, L. (2022). Democratic hypocrisy and out-group threat: Explaining citizen support for democratic erosion. *The Journal of Politics*, *84*(3), 1806–1811. https://doi.org/10.1086/719009

Somer, M., McCoy, J. L., & Luke, R. E. (2021). Pernicious polarization, autocratization and opposition strategies. *Democratization*, *28*(5), 929–948. https://doi.org/10.1080/13510347.2020.1865316

Svolik, M. W. (2019). Polarization versus democracy. *Journal of Democracy*, *30*(3), 20–32. https://doi.org/10.1353/jod.2019.0039

Svolik, M. W. (2020). When polarization trumps civic virtue: Partisan conflict and the subversion of democracy by incumbents. *Quarterly Journal of Political Science*, *15*(1), 3–31. https://doi.org/10.1561/100.00018132

Talanquer, M. S. (2020). Mexico 2019: Personalistic politics and neoliberalism from the left. *Revista de Ciencia Politica*, *40*(ahead), 0–430. https://doi.org/10.4067/s0718-090x2020005000113

Törnberg, P. (2024). Large Language models outperform expert coders and supervised classifiers at annotating political social media messages. *Social Science Computer Review*. https://doi.org/10.1177/08944393241286471

Uribe, A. (2025). Replication data for: Building tolerance for backsliding by trash-talking democracy: Theory and evidence from Mexico. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/4C8VGG

Voeten, E. (2016). Are people really turning away from democracy? Available at SSRN 2882878.

Author Biographies

Lautaro Cella is a PhD Candidate in Political Science at the University of Chicago.

Ipek Çınar received her PhD in Political Science from the University of Chicago in 2024.

Susan Stokes is the Tiffany and Margaret Blake Distinguished Service Professor of Political Science at the University of Chicago and Director of the Chicago Center on Democracy.

Andres Uribe is an Anna Julia Cooper Fellow and incoming Assistant Professor of Political Science at the University of Wisconsin-Madison.