Digital Repression in Autocracies

Erica Frantz, Andrea Kendall-Taylor, Joseph Wright

March 2020

Users Working Paper
SERIES 2020:27
THE VARIETIES OF DEMOCRACY INSTITUTE

UNIVERSITY OF GOTHENBURG
DEPT OF POLITICAL SCIENCE

Varieties of Democracy (V-Dem) is a new approach to the conceptualization and measurement of democracy. It is co-hosted by the University of Gothenburg and University of Notre Dame. With a V-Dem Institute at University of Gothenburg that comprises almost ten staff members, and a project team across the world with four Principal Investigators, fifteen Project Managers, 30+ Regional Managers, 170 Country Coordinators, Research Assistants, and 2,500 Country Experts, the V-Dem project is one of the largest-ever social science research-oriented data collection programs.

Please address comments and/or queries for information to:

V-Dem Institute

Department of Political Science

University of Gothenburg

Sprängkullsgatan 19, PO Box 711

SE 40530 Gothenburg

Sweden

E-mail: contact@v-dem.net

V-Dem Working Papers are available in electronic format at www.v-dem.net.

Digital Repression in Autocracies

Erica Frantz
Assistant Professor
Department of Political Science
Michigan State University

Andrea Kendall-Taylor
Senior Fellow and Director
Transatlantic Security Program
Center for New American Security

Joseph Wright
Professor
Department of Political Science
Pennsylvania State University

March 2020

Abstract

The rise of the digital age initially spread optimism about the potential for democratic change. Those hopes were soon dashed, however, as autocracies learned how to use digital tools to further their interests. In this paper, we leverage data from the Digital Society Project to illustrate a number of important trends with respect to the use of digital repression in autocracies. First, and not surprisingly, we show that reliance on digital repression in dictatorships is on the rise; more dictatorships use digital repression than ever before. The evidence shows that they are wise to do so. We find that digital repression lowers the risk of protest in dictatorships. Importantly, we also find that dictatorships are using digital repression in tandem with more "high-intensity" forms of repression. This suggests that digital repression is not serving as a substitute for more brute repressive acts, but instead as a complement to them. Lastly, we find that some evidence that digital repression is associated with more durable dictatorship. This finding is both relatively small and statistically insignificant in the relatively short panel series, however.

Introduction

Repression is a hallmark feature of authoritarian rule. It raises the costs of disloyalty and makes it more difficult for groups to mobilize against the regime (Wintrobe, 1998). Though dictatorships vary markedly in the extent to which they rely on repression, all regimes use it to some degree (Frantz and Kendall-Taylor, 2014). This reality of authoritarian politics has not changed over time. What has changed, however, are the tools available to autocratic governments to carry out such repression (Xu, 2019).

With the advent of new technologies, dictatorships can censor and filter the Internet to prohibit the spread of unfavorable information, as exemplified by the Chinese regime's "Great Firewall."[1] They can also use bot-driven information-distortion campaigns on social media to cloud information channels with noise and confuse citizens, a tactic at which the Russian government is particularly adept.[2] And they can use artificial intelligence (AI) to surveil their citizens, making it easy to identify, monitor, and target those who oppose them. Saudi Arabia, for example, reportedly hacks into the online accounts of its dissidents using commercially available surveillance technology.[3] In other words, opportunities for leveraging new technologies to carry out repression in new ways – what we refer to as digital repression – are vast.

In this way, digital repression is the new frontier of the autocratic survival toolkit. Yet, beyond vivid anecdotes, particularly those based on the Chinese Communist Party's dystopian tactics, we know very little about what the digital repression landscape looks like in autocracies. This paper seeks to fill this void. It provides insight into how dictatorships are using digital repression, how this has changed over time, and what the consequences are for authoritarian politics.

Specifically, we show that reliance on digital repression lowers an autocracy's risk of protest. This message is particularly important in light of evidence (we offer here) that mass mobilization has become the most frequent and destabilizing threat contemporary dictatorships face. Moreover, we find that reliance on digital repression increases reliance on more "high intensity" forms of repression, such as the use of torture and imprisonment. This suggests that dictatorships are not substituting a new tool for their old ones. Rather, they are using it to make their existing methods more effective. Finally, we offer evidence that reliance on digital repression is associated with longer-lasting authoritarian rule. Though, we cannot establish whether this is a causal relationship, we do unearth evidence that greater digital repression and more durable authoritarianism go hand in hand.

What digital repression is and how it differs from traditional repression

Digital repression refers to the use of new technologies – primarily the Internet, social media, and Artificial Intelligence (AI) – to repress citizens and maintain political control.[4] It can range from relatively rudimentary tactics, such as Internet shutdowns, to the use of misinformation campaigns

[1] See "The Great Firewall of China,"' Bloomberg News, 05 November 2018, `https://www.bloomberg.com/quicktake/great-firewall-of-china` (accessed 24 February 2020).

[2] "How Russia and Other Foreign Actors Sow Disinformation in Elections," National Public Radio, 21 February 2020 `https://www.npr.org/2020/02/21/808275155/how-russia-and-other-foreign-actors-sow-disinformation-in-elections` (accessed 24 February 2020).

[3] "Saudi Arabia: Change Comes with Punishing Cost", Human Rights Watch, 04 November 2019, `https://www.hrw.org/news/2019/11/04/saudi-arabia-change-comes-punishing-cost` (accessed 24 February 2020).

[4] Our focus in this study is on dictatorships, but it is important to emphasize that democracies use digital repression too (albeit to a lesser degree on average than their autocratic counterparts).

on social media to discredit the opposition and AI-powered surveillance to monitor and even predict the actions of potential dissidents.[5]

In many ways digital repression is similar to traditional repression. Like traditional repression, digital repression raises the costs of disloyalty, enables leaders to identify their opposition, and restricts the ability of groups to mobilize against the regime.[6] Yet despite these fundamental similarities, digital repression differs from traditional repression in a number of key ways. Most importantly, digital repression lowers the costs and increases the effectiveness of longstanding repressive tactics. It supercharges established authoritarian methods of control. Take censorship, for example. Rather than having to rely on human operators to monitor the vast amounts of information online, AI can sift through massive amounts of images and text, filtering and blocking content to identify information that is unfavorable to the regime (Feldstein, 2019*b*). In this way, digital tools enable autocracies to cover greater ground so that they can go after their opponents with greater precision.

This is valuable because repression, including censorship, brings with it a number of risks for authoritarian regimes (Gartner and Regan, 1996). It can sow discontent and reduce the regime's political legitimacy, elevating the possibility of civil unrest (Lichbach, 1987; Moore, 1998). Indiscriminate repression, in particular, can raise the chance of a backlash against the regime that strengthens the opposition (Francisco, 1995; Kalyvas, 2006; Rozenas and Zhukov, 2019). Digital surveillance, in contrast, achieves the same goals as traditional repression, but without the collateral damage.

Digital repression also differs from traditional repression with respect to the ease with which regimes can develop the capacity to carry it out. Historically, building an effective repressive apparatus required that a regime cultivate the loyalty of thousands of cadres, train them, and arm them with the tools needed to engage in widespread boots-on-the-ground surveillance. In the case of the East German Stasi, for example, the evidence suggests that there was one East German spy for every 66 citizens (Koehler, 1999). Most dictatorships simply do not have the capacity to create such a vast operation. In the digital age, however, such extensive manpower is no longer required for dictatorships to effectively surveil and monitor their citizens.

Moreover, autocracies increasingly have the possibility of importing the capacity to digitally repress. Aspiring dictatorships can simply purchase their desired new technologies, train a small group of officials in how to use them – often with the support of external actors such as China, which sponsors such seminars – and they are ready to go.[7] This sort of market is already active. For example, Huawei, a Chinese state-backed telecommunications firm, has deployed its digital surveillance technology in over a dozen authoritarian regimes (Greitens, 2019). Similarly, reports suggest that private Israeli companies have sold espionage and intelligence-gathering software to a number of authoritarian regimes across the world, including those in Angola, Bahrain, Kazakhstan, Mozambique, Nicaragua, and the UAE.[8]

These are just a few ways in which digital repression differs from repression in its traditional form. In what follows, we offer greater information on digital repression and its role in the autocratic

[5]See Deibert (2015) for a brief outline of three generations of digital technologies deployed by authoritarian regimes.

[6]See Davenport (2007) for a review of the literature on traditional repression.

[7]Michael Abramowitz and Michael Chertoff, "The Global Threat of China's Digital Authoritarianism," The Washington Post, 01 November 2018, `https://www.washingtonpost.com/opinions/the-global-threat-of-chinas-digital-authoritarianism/2018/11/01/46d6d99c-dd40-11e8-b3f0-62607289efee/_story.html` (accessed 24 February 2020).

[8]Hagar Shezaf and Jonathan Jacobson, "Revealed: Israel's Cyber-spy Industry Helps World Dictators Hunt Dissidents and Gays," Haaretz, 20 October 2018, `https://www.haaretz.com/israel-news/.premium.MAGAZINE-israel-s-cyber-spy-industry-aids-dictators-hunt-dissidents-and-gays-1.6573027` (accessed 24 February 2020).

toolkit.

Digital repression and the autocratic toolkit

The spread of the Internet and social media was originally accompanied by significant enthusiasm that it would usher in greater democracy across the globe. Many analysts believed that this new technology would empower citizens by giving them greater access to information and improve their ability to coordinate, such that autocrats would no longer be able to preserve the power concentration their regimes depend on (e.g. Diamond, 2015; Wilson, 2017). This optimism was most pronounced in the early 2010s, as social media helped bring about the overthrow of four of the world's longest-ruling dictators in Egypt, Libya, Tunisia, and Yemen. In each of these cases, the Internet and social media allowed people to make their voices heard and facilitated their efforts to organize to oust non-responsive and repressive regimes.

As time progressed, it became increasingly clear, however, that these new tools would not just be technologies of liberation. Authoritarian regimes were learning to use them to advance their own interests, as well (Tucker et al., 2017).[9] While digital tools gave citizens the means to hold their governments accountable and facilitated their ability to organize, authoritarians were not sitting idly by. Faced with growing pressure and fear of their own people, dictatorships adapted.

In many ways, authoritarianism has always been a story of adaptation. In the immediate aftermath of the Cold War authoritarian regimes learned to mimic features of democracy (Frantz and Kendall-Taylor, 2017). The widespread adoption of elections and other democratic institutions initially destabilized a number of autocracies around the globe. But these regimes soon learned that they could use these institutions in ways that actually prolong their time in power (Kendall-Taylor and Frantz, 2014). Today the story is the same. Authoritarians are learning to harness technological change to advance regime objectives and push back against new the vulnerabilities that these technologies first introduced.

The Chinese Community Party (CCP), in particular, has demonstrated how far technology can go in creating new possibilities for citizen control (Qiang, 2019; Kendall-Taylor, Frantz and Wright, 2020). Although China is the leading player in the digital repression arena, autocrats of all stripes are looking to follow suit and leverage digital tools to more effectively surveil, censor, and manipulate their citizens. China's capacity to implement digital repression is currently unique, but Beijing's tools and tactics likely foreshadow future trends in autocracies across the globe.

In what follows, we highlight how authoritarian regimes are using digital tools to maintain power. In different ways, these dynamics enable authoritarian regimes to reduce their risk of experiencing large-scale protests and enhance their durability, outcomes we examine further later in this study.

Censorship Internet censorship is perhaps the most obvious way that authoritarian regimes employ digital tools to repress. China, for example, operates what is known as the "Great Firewall" – currently the largest system of censorship in the world, a joint operation between government monitors and technology and telecommunication companies that work together to filter any content the regime considers to be harmful.[10] AI is enhancing the ability of autocracies to execute this type of censorship. AI can sift through images and text in sophisticated ways, allowing the regime

[9] See also Richard Fontaine and Kara Frederick, "The Autocrat's New Tool Kit," Wall Street Journal, 15 March 2018, `https://www.wsj.com/articles/the-autocrats-new-tool-kit-11552662637` (accessed 24 February 2020).

[10] See "The Great Firewall of China," Bloomberg News, 05 November 2018, `https://www.bloomberg.com/quicktake/great-firewall-of-china` (accessed 24 February 2020).

to filter and block content that is unfavorable to it. Meanwhile, bot-driven information-distortion campaigns allow dictators to produce a flurry of misleading posts to blur opponents' messaging and overwhelm information channels with noise. And even if such censorship fails and protests mount, digital autocracies have an added line of defense: they can shut down the Internet– as a whole or in parts – to prevent protesters from communicating, organizing, or broadcasting their messages. Examples of this tactic are plentiful, such as when the Russian government used targeted mobile Internet shutdowns during anti-government protests in Moscow in summer2019 or when the Iranian government successfully shut down the Internet across the country amid widespread protests there in November 2019 (Kendall-Taylor, Frantz and Wright, 2020).

Identifying regime opponents The Chinese pathway suggests that the advancement of AI-powered surveillance is perhaps the most significant evolution in digital repression (Qiang, 2019; Feldstein, 2019*a*). High-resolution cameras, facial recognition, spying malware, automated text analysis, and big data processing are opening up a myriad of new opportunities for citizen control. The CCP, for example, collects an incredible breadth and volume of data on individuals from things such as tax returns, bank statements, and criminal and medical records. It then analyzes this information using AI-powered tools to spot dissenters – just as doctors are seeking to use AI to find data patterns amongst the seemingly healthy to predict disease before it emerges (Wright, 2019). The predictive power of AI allows dictators to pre-empt their opposition and neutralize potential dissidents through targeted detentions and preventive arrests.

Monitoring regime insiders The case of China has also shown us that regimes can use digital repression to monitor their underlings and other political elite. This is valuable given that regime insiders pose an ever-present threat to authoritarian incumbents (Svolik, 2012). Research shows that the CCP, for example, often does not censor citizens' posts about local corruption on Weibo – the Chinese equivalent of Twitter – so that it can keep an eye on the performance of local officials (Qin, Stromberg and Wu, 2017). In this way, digital repression can help authoritarian regimes monitor government officials, gauge their performance, and root out those cadres whose underperformance is harming public perceptions of the regime.

Gauging public sentiment New technologies are also useful in improving dictatorships' access to information about their citizens – historically a critical vulnerability in authoritarian systems (Guvitsky, 2015). Dictatorships typically have limited insight into the sentiment and views of their citizens because their use of repression reduces people's willingness to communicate their beliefs. New technologies can ease this dilemma by improving dictatorships' ability to identify and respond to (even cosmetically) sources of discontent before they spiral into something more threatening. Research shows, for example, that the Chinese government uses digital tools to anticipate events that could create focal points for unrest and then preventatively applies repression to reduce dissent before it spreads (Qin, Stromberg and Wu, 2017). And while digital autocracies may not currently be able to use all the data they collect, advances in big data analysis and decision-making technologies will enhance their ability to read the public mood and respond accordingly (Hoffman, 2019).

Manipulating the information environment Beyond censorship, authoritarian regimes employ a number of more proactive tactics to manipulate their information environments (Roberts, 2018). New technologies augment dictatorships' capacity to shape public perceptions of the regime and its legitimacy. In the authoritarian universe, the ability to control the information environment

is critical because it shapes citizens' willingness to join opposition groups, participate in protests, and engage in anti-regime activity. This is an area where Russia has played a leading role. In addition to the Kremlin's use of surveillance and vague laws to target opposition, Moscow has demonstrated its talents in manipulating the information environment. The Kremlin floods the Internet with pro-regime narratives, diverting attention from negative news, and plants confusion and uncertainty through the dissemination of alternative narratives about events.[11] In this way, new technologies increase the efficiency of dictatorships' efforts to drown out criticism and inflate perceptions of regime support, increasing regime resilience.

Maturing technologies like micro-targeting and "deep fakes" are likely to further boost authoritarians' capacity to manipulate the information environment and boost regime support.[12] Micro-targeting will allow autocrats to tailor content for specific individuals or segments of society, just as the commercial world uses demographic and behavioral characteristics to customize advertisements. AI-powered algorithms will allow autocrats to pin-point persuasion by using individual weaknesses and vulnerabilities to manipulate their citizens. Likewise, the production of "deep fakes" – digital forgeries impossible to distinguish from authentic audio, video or images – will make it increasingly difficult to distinguish the truth. Autocrats will be able to weaponize images, for example, producing unflattering videos or images to undermine their opponents.

Compelling compliance New technologies have also enhanced authoritarians' ability to use cooptation, the other traditional authoritarian survival tactic. Cooptation enhances the durability of authoritarian rule by increasing the benefits of loyalty to the regime (Frantz and Kendall-Taylor, 2014). China's smart cities, for example, do not just facilitate surveillance and control.[13] Tech-powered integration between government agencies allows the CCP to more efficiently solve problems and provide public services, ultimately improving governance and enhancing perceptions of regime performance (Hoffman, 2019). Smart cities also enable the CCP to more precisely control access to government services, calibrating its distribution – or denial – of everything from bus passes and passports to jobs and access to education. An individual that posts information that is critical of the regime, for example, could be deemed "untrustworthy" and find themselves excluded from state-sponsored benefits, such as deposit-free apartment rentals (Ahmed, 2019).[14] In this way, authoritarian regimes can leverage new technologies to fine-tune their use of reward and refusal, blurring the line between cooperative and coercive control (Hoffman, 2019).

Mimicking democracy to improve perceptions of the regime As a final point, in some cases, new technologies are enabling authoritarian regimes to mimic components of democracy, such as participation and deliberation, which research shows is associated with longer-lasting au-

[11]Seva Gunitsky, "The Great Online Convergence: Digital Authoritarianism Comes to Democracies," War on the Rocks, 19 February 2020, https://warontherocks.com/2020/02/the-great-online-convergence-digital-authoritarianism-comes-to-democracies/ (accessed 24 February 2020).

[12]Richard Fontaine and Kara Frederick, "The Autocrat's New Tool Kit," *Wall Street Journal*, 15 March 2018, https://www.wsj.com/articles/the-autocrats-new-tool-kit-11552662637 (accessed 24 February 2020).

[13]Jamil Anderlini, "How China's smart-city tech focuses on its own citizens," *Financial Times*, 04 June 2019, https://www.ft.com/content/46bc137a-5d27-11e9-840c-530737425559 (accessed 24 February 2020).

[14]See, also, Alexandra Ma, "China has started ranking citizens with a creepy 'social credit' system," *Business Insider*, 29 October 2018, https://www.businessinsider.com/china-social-credit-system-punishments-and-rewards-explained-2018-4 (accessed 24 February 2020); Charlie Campbell, "How China Is Using 'Social Credit Scores' to Reward and Punish Its Citizens," *Time*, 16 January 2019, https://time.com/collection/davos-2019/5502592/china-social-credit-score/ (accessed 24 February 2020).

thoritarian rule (Kendall-Taylor and Frantz, 2014). Some local-level Chinese officials, for example, use the Internet and social media to allow citizens to voice their opinions and participate in online deliberative polls or other digitally-based participatory channels so that they can better gauge citizen preferences. Research shows that these portals have enhanced public perceptions of the CCP among less educated citizens (Truex, 2017). In this way, dictatorships can use digital tools to emulate elements of democracy and improve their attractiveness to citizens.

The digital repression landscape in autocracies

To gain a sense of the extent to which dictatorships use digital repression, we make use of data from the Digital Society Project (Mechkova et al., 2019), a component of the Varieties of Democracy project (Coppedge et al., 2019a; Coppedge et al., 2019b; Pemstein et al., 2019), that capture various facets of digital repression from 2000 to 2018.[15] Our measure of digital repression relies on six variables from this data set. Each of these variables measures a concept related to the state's behavior in the digital sphere in practice: *social media censoring, social media monitoring, social media shut down, Internet shut own, Internet filtering*, and *social media alternatives*.[16] In essence, these variables capture a government's ability to monitor, censor, and shut down social media; filter and shut down the Internet; and create social media alternatives that are wholly controlled by either the government or its agents.

Items	Item-test correlation	α
(a) Digital *repression*		
Government social media censorship in practice	0.908	0.928
Government social media monitoring	0.910	0.928
Government social media shut down in practice	0.855	0.937
Government Internet shut down in practice	0.919	0.926
Government Internet filtering in practice	0.898	0.930
Government social media alternatives	0.806	0.945
(b) Digital *capacity*		
Government cyber security capacity	0.784	0.786
Government Internet shut down capacity	0.707	0.838
Government Internet filtering capacity	0.875	0.714
Government capacity to regulate online content	0.851	0.734

Table 1: *Item-test correlations for each item in the digital repression and capacity indices*

Using these six variables, we construct an index of digital repression. To do so, we treat each of these variables as one measure of a latent concept of state-led digital repression and combine them into a single scale using Cronbach's α (which is a test of scale reliability that produces a

[15]The data on dictatorships we describe shortly runs through 2017, so our sample runs from 2001 to 2017, given lags.

[16]Exact question wording for these items and those listed in Table 1 are in Appendix Table A-1.

standardized linear combination of the items similar to a principal component). This scaled index, which combines information from all six variables, we refer to as the *digital repression* index. Its overall scale reliability is 0.943, indicating that on average the items are highly inter-correlated.

The second column of the top panel (a) in Table 1 shows the extent to which each item is correlated with the scaled index, which is akin to measuring how much information the item contributes to the test scale. It shows that *social media shut down* and *social media alternatives* are the least correlated with the *digital repression* index; the other four items are all correlated at roughly 0.90 or more. Overall, each item is relatively highly correlated with the index, indicating that all six items are appropriately included.[17]

To evaluate whether governments differ in their reliance on digital repression simply on account of differences in their capacity to implement digital repression, we also measure digital capacity. We do so using additional variables from the Digital Society Project data set (Mechkova et al., 2019), which tap into concepts related to the state's capacity to intervene in the digital sphere: *government cyber security capacity, Internet shut down capacity, Internet filtering capacity*, and *capacity to regulate online content.* Importantly, these measures are conceptually distinct from a state's willingness to pursue digital repression in practice. Expert assessment of these variables comes from questions that explicitly ask for information about state digital capacity "independent of whether [the government] actually does so in practice" (Coppedge et al., 2019a, 286). As with the *digital repression* index, we treat each of these variables as one manifestation of a latent concept of state digital capacity, and thus combine them into a single scale using Cronbach's α. We refer to this scaled index (which combines information from all four variables) as the *digital capacity* index.

The second column of the bottom panel (b) in Table 1 shows the extent to which each variable is correlated with the scaled index. *Internet filtering capacity* is the most correlated (0.875) with the index, while *Internet shutdown capacity* is the least correlated (0.707). Overall, each variable has a relatively high correlation, however, indicating that it is appropriate to include all four of them.[18] The scale reliability of the *digital capacity* index is 0.818, which suggests that (on average) the variables are inter-correlated, though not to the degree as with the *digital repression* index.[19]

Digital repression and capacity in democracies and dictatorships

To measure dictatorships in this study, we use the Autocratic Regimes Data Set, which captures the start and end dates of authoritarian regimes (Geddes, Wright and Frantz, 2014). The original data run from 1946 to 2010; our updates extend it through 2017. The data set excludes countries with small populations (e.g. Brunei) as well as those where the state does not control the majority of territory due to civil conflict (e.g. Libya after the fall of the Gaddafi regime in 2011). For the purposes of comparing levels of digital repression and capacity across dictatorships and democracies, we exclude those years in which a transition between democracy and dictatorships took place (e.g. Cote d'Ivoire and Tunisia in 2011) because we do not know whether coders assessed the digital repression variable questions before or after the transition event during the calendar year. We use this data set to measure democracies, as well, so that we can compare trends in digital repression in both types of political systems.

[17] Explanatory factor analysis of these six items yields an eigen value of greater than four for the first factor and an eigen value of less than 0.15 for the second factor, indicating very strong uni-dimensionality.

[18] Exploratory factor analysis of these four variables yields an eigen value of greater than 2 for the first factor and less than 0.4 for the second factor, indicating very strong uni-dimensionality.

[19] A factor analysis of all ten variables – those in the *digital repression* index and those in the *digital capacity* index – suggests there are two dimensions in these items: a first factor, which is strongly correlated with items in the *digital repression* index, has an eigen value of over 5, while a second factor, which is strongest for items in the *digital capacity* index, has an eigen value of 1.63.

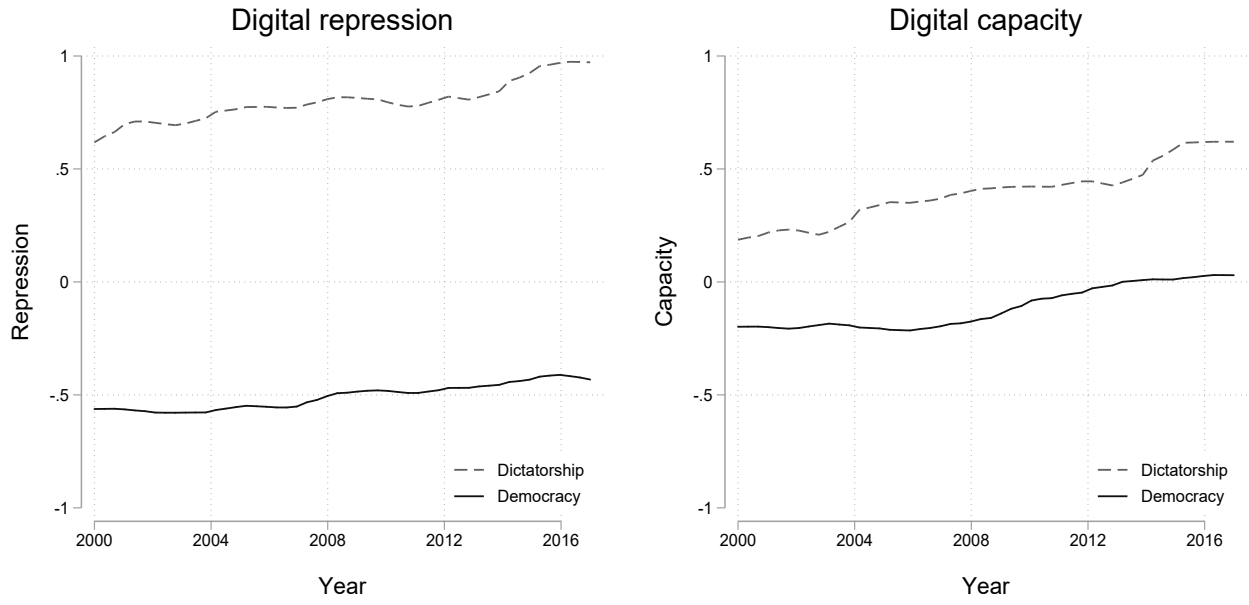

Figure 1: Time trend in digital repression and capacity, by political regime type

Figure 1 illustrates trends over time in *digital repression* (left plot) and *digital capacity* (right plot), with democracies (shown with solid lines) and autocracies (shown with dashed lines). With respect to digital repression, there are two key takeaway messages. The first is that regardless of the type of political system, digital repression has increased over time. This is to be expected given the expansion and growth of new technologies during this period. Second, dictatorships on average use more digital repression than democracies do. This is also unsurprising in light of the fact that autocratic governments tend to use other forms of repression (including restrictions on civil liberties and violations of physical integrity rights, such as imprisonment, torture, and execution) more frequently than their democratic counterparts (Davenport, 2007).

In terms of *digital capacity* (right plot), we see very similar trends. Digital capacity has increased over time in all political systems, and levels of digital capacity are higher in dictatorships than in democracies. That said, the differences are less pronounced than with digital repression, in that dictatorships have a slight edge over democracies in their digital capacity but use digital repression to a far greater extent than democracies do. We presume that actual digital repression is greater that capacity in dictatorships because autocracies often use rudimentary digital repressive tactics, such as Internet shutdowns, that do not require substantial capacity. Examples include Eritrea and Tajikistan, which have higher levels of digital repression than would be expected given their capacity for employing it. Low capacity for digital repression does not mean its absence, in other words; it simply means that governments will be more restricted in the ways they can exercise it, relying primarily on more basic forms of digital repression.

Finally, looking at each type of political system in isolation, democracies use less digital repression than they could given their capacity. This makes sense in light of the fact that new technologies are useful to governments for a wide variety of reasons, beyond just repressing citizens. By contrast, dictatorship use more digital repression than their levels of digital capacity would indicate. This is likely to due to the fact that digital repression can be intense even if the way in which it is exercised is basic (e.g., frequent and widespread Internet shutdowns).

If we continue to narrow the focus on dictatorships, the data reveal that there is substantial variation across regimes in digital repression and digital capacity. Figure 2 illustrates this, showing the average levels of digital repression in dictatorships during the period from 2000 to 2017. Even among these regimes, there are wide-ranging differences in their levels of digital repression. Some regimes, such as North Korea's, exhibit very high digital repression; other regimes, such as those in Angola and Russia, reveal moderate digital repression; and yet other regimes, such as those in Belarus and Mozambique, have very low digital capacity.

Taken together, these patterns give us some basic insight into the digital repression landscape in authoritarian regimes. Some of our intuitions are supported by the data (e.g., that autocracies digitally repress more than democracies do), but other trends that emerge are somewhat less obvious (e.g., that digital repression is higher in dictatorships than digital capacity is). With these things in mind, we now turn to an examination of how digital repression in dictatorships influences outcomes of interest.

Digital repression and protest

We begin by empirically evaluating the impact of digital repression on protests. To motivate the choice of examining protests, we first take a moment to illustrate their importance to the 21st century dictatorship.

Basic statistics about how dictatorships fall from power are telling. Data from the Autocratic Regimes Data Set, which captures the method in which autocratic regimes exit (Geddes, Wright, and Frantz 2014) reveal that, historically, coups have posed the greatest threat to dictatorships. From 1946 to 2000, roughly a third of the 198 regimes that fell were ousted via coup. Protests during this period accounted for only 16 percent of regime failures. Looking to the 2000s, however, a different picture surfaces: from 2001 to 2017 coups only unseated about nine percent of the dictatorships that fell, compared to protests, which toppled 23 percent of them, more than doubling the frequency of coups. In fact, protests are now tied with elections as the most common method through which today's dictatorships collapse. Given that many of these elections are in response to and amid mass protests campaigns (such as Guinea in 2010, Mali in 2013, and Egypt in 2012), the basic message is that protests are now the primary threat to contemporary authoritarians.

Protests, more generally, are on the rise, as well. Data from the Mass Mobilization Project (Clark and Regan, 2016),[20] which capture anti-government protests of 50 participants or more, suggest that from 2000 to 2017, 60 percent of all dictatorships faced at least one protest event.[21] Many of these efforts are small and not too destabilizing, but their ever presence highlights that – at a minimum – they are a persistent nuisance for today's dictatorships. This is particularly true for regimes such as those in Zimbabwe and Iran, which have encountered protests each year since 2000. For regimes like these, protests are more than just an annoyance; they represent a serious challenge that the regime must confront.

New technologies likely account for some of this increase in mass mobilization.[22] The spread of the digital age has changed the context in which opposition groups operate, reducing barriers to coordination and making it easier for citizens to mobilize.[23] As we show here, however, many

[20]Data downloaded on 27 August 2019 from `https://doi.org/10.7910/DVN/HTTWYL`.

[21]Further, while regime collapse is a relatively rare event, on a yearly basis, anti-government protests are much more common.

[22]Gideon Rachman, Benedict Mander, Daniel Dombey, Sue-Lin Wong, and Heba Saleh, "Leaderless Rebellion: How Social Media Enables Global Protests," Financial Times, 25 October 2019, `https://www.ft.com/content/19dc5dfe-f67b-11e9-a79c-bc9acae3b654` (accessed 24 February 2020).

[23]The ease with which social media facilitates protests also helps explain why protests are increasingly

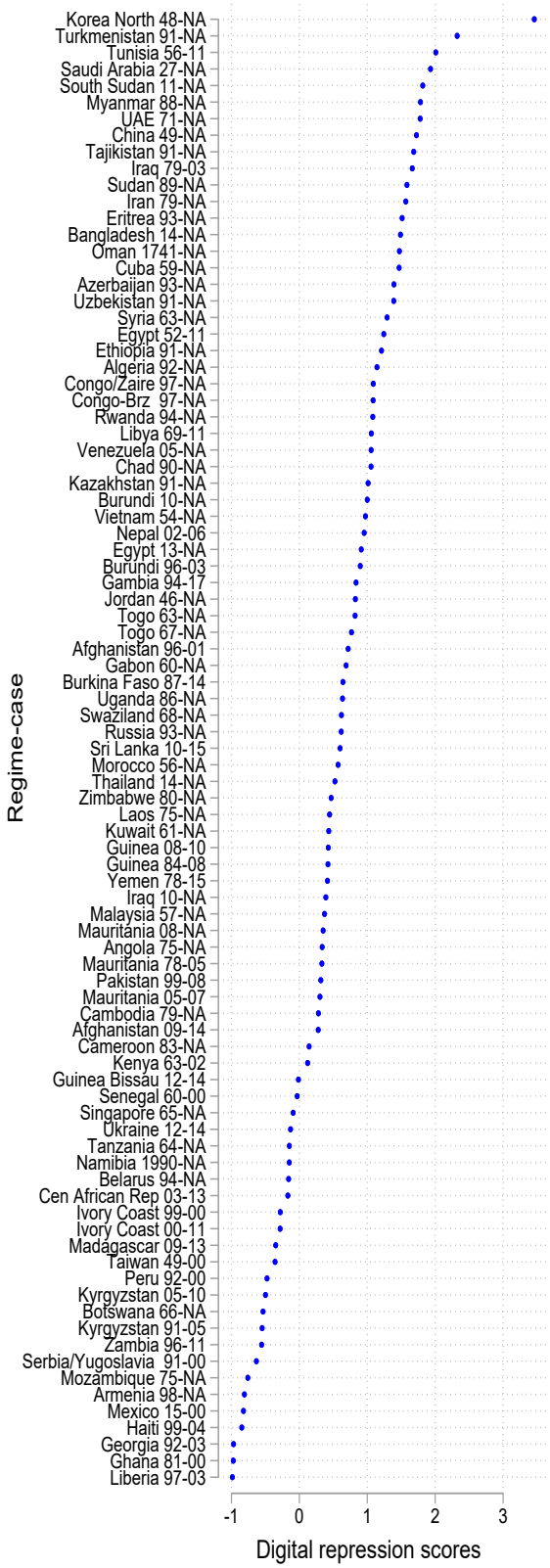

Figure 2: Digital repression, for autocratic regime cases

unsuccessful. Today's movements do not require the broad organization efforts that those in the past did, a feature that is critical for sustaining them.[10] See Antonia Malchik, "The Problem with Social-Media Protests," The Atlantic, 06 May 2019, `https://www.theatlantic.com/technology/archive/2019/05/in-person-protests-stronger-online-activism-a-walking-life/578905/` (accessed 24 February 2020).

authoritarian governments appear to be adapting in response, using these same technologies in ways that help them counter the threat of protest.

An example from Cambodia is illustrative.[24]. The regime that has governed there since 1979, currently led by Hun Sen, held in elections in July 2013. In the period leading up the contest, the regime basically barred the opposition from accessing traditional media outlets. In response, opposition groups turned to digital tools to spread their message and mobilize supporters. The elections were fraudulent, and thousands of citizens protested demanding for a new election. More and more groups joined the protest movement, notably garment workers and Tuk-tuk drivers. The regime hit back with brute force to counter the protests, while also escalating its use of digital repression, such as temporarily blocking Facebook in August 2013 and shutting down Internet cafes in Siem Reap province in December of that year. It created a "Cyber War Team" in late 2014, a group charged with monitoring the Internet and flagging anti-government online activity. And in late 2015, the regime passed a law to give it substantial influence in the telecommunications industry, with an enforcement body in place that can suspend individual firms from providing services and dismiss their staff. These acts of digital repression over the course of 2013 through 2015 paid off for the regime. The protest movement that had been intensifying following the elections began to die off by July 2014. The protest data described earlier corroborate this: anti-government protests peaked in 2014 with 36 that year, declining in 2015 and again 2016, such that there was only one protest event by 2017. Cambodia's ratcheting up of digital repression corresponded with a decline in the incidence of protests.

To evaluate whether dynamics such as these are unique to Cambodia or more widespread, we next use cross-national empirical tests to look at the impact of digital repression on the chance of protest. In these tests, we lag our measure of digital repression one year. We do so because we believe that most country experts who rate the level of digital repression in a country-year incorporate information in their assessments from political events that occur throughout the duration of the calendar year. For example, after protestors topped Ben Ali and his regime in Tunisia in 2011, digital repression decreased substantially. Reflecting this, the country experts who coded digital repression for Tunisia indicate a large decrease in the variables related to digital repression from 2010 to 2011. This is problematic because the observed treatment (digital repression) might be causally prior to the observed outcome (protest). Ideally, we want to model the observed event (protest) using a treatment variable that occurs prior to it, and for this reason lagging the treatment variable is appropriate.

We model protest in two ways. First, we use a binary indicator of protest that accounts for the time (in years) since the last protest event targeting a regime. This is akin to a survival model, helping to answer the question of whether digital repression reduces the onset of anti-government protest. Second, we use a transformed count of the number of protests in a given year.[25] Using this measure, dictatorships in Egypt, Thailand, and Venezuela faced the most protests, even if these protests took place across fewer years. We posit that while the binary annual indicator of protest helps us understand whether digital repression deters or prevents protests from starting, this continuous measure is more likely to capture whether it influences the emergence, escalation, or decline of large, sustained protests movements. There is substantial variation over time within dictatorships in the presence and level of protests, according to both of these measures.

We concentrate on approaches here that isolate the influence of changes over time within dic-

[24]We draw here from three Freedom House Freedom on the Net reports (in 2014, 2015, and 2016). See Freedom House, `www.freedomhouse.org`.

[25]Given the high-skew in the count data, we use the square root of the natural log of the count. We also test negative binomial count models designed to account for the skewed distribution in count data. These yield similar results.

tatorships in their use of digital repression on the two protest outcomes (though we confirm our findings in pooled models too).[26] This helps inform the question of whether changing levels of digital repression within dictatorships deters, prevents, or breeds anti-government mass mobilization.

We adjust for a number of variables that could potentially influence the relationship between digital repression and protest.[27] First, it is possible that some level of digital capacity is needed before digital repression can be exercised, we therefore adjust for digital capacity.[28] We also include a time trend, given strong evidence of this in the data (in terms of rising protests over time, as well as rising digital repression), and a measure of the time (in years) since the last protest. In light of evidence that protests are more common in larger countries and may be more likely when autocratic leaders have been in power for many years,[29] we adjust for population size and leader tenure (both logged). Finally, we add a measure of the extent to which people consume domestic online media, from the Digital Society Project (Mechkova et al., 2019), to proxy for technological advances citizens have adopted that might lead to government censorship (which constitutes one aspect of digital repression) by making protest organization easier.[30]

The left plot in Figure 3 shows the results of the first test using the binary indicator of anti-government protest.[31] The horizontal axis depicts years, while the vertical axis displays the estimated marginal effect of a one-unit increase in digital repression (a value that is slightly larger than the standard deviation of digital repression in the sample – 0.82). Digital repression, according to these estimates, decreases the probability of protest in a given year by between 4 and 6 percent. To put this number in perspective, note that the baseline probability of observing protest in any given year in the sample is 58 percent. Finally, this size of this effect (of digital repression decreasing protest) appears to be declining over time: it was strongest in the earlier part of the sample period in the 2000s but slightly weaker in more recent years.

The right plot in Figure 4 shows a similar result when using the continuous (count) measure of protest, which may be a better proxy for the intensity and/or resilience of protest movements than the binary indicator that only captures whether at least one protest event was observed during the year. Here we find that a one-unit increase in digital repression is associated with a roughly 6 to 10 percent decrease in protest, with the effect getting larger over time.[32] In short, both tests suggest that increases in digital repression within dictatorships over time reduces anti-government protests.

The key message to emerge from the tests we present here is that digital repression in dictatorships lowers the risk of protests. This result holds in a variety of conditions and – given the use of fixed effects – suggests that as levels of digital repression increase within dictatorships, the chance

[26]The results in the main text come from a kernel least squares estimator. Appendix Tables B-1 and B-2 report results from correlated random effects and linear probability estimators. For the count data, we report a kernel estimator but check results with a fixed effects negative binomial model, reported in Appendix Table B-2.

[27]See Appendix Figure B-2 for additional results when adjusting for other potential confounders. Importantly, digital repression may simply be a proxy for state-led violations of human integrity rights, such as the imprisonment, torture, and execution of political dissidents. But, as we show below, digital repression may also shape the state's use of high-intensity repression, which in turn influences protest, making high-intensity repression a post-treatment phenomena. In the main specifications, we therefore do not adjust for this factor.

[28]Similar to the treatment variable, digital repression, we lag digital capacity and physical repression one year to ensure a possible causal relationship is chronologically ordered correctly.

[29]For example, Chenoweth and Ulfelder (2017) find that population-only models of the onset of mass uprisings perform similarly to more complex models. They also note that "[w]hen a single leader has occupied office for an abnormally long period of time... such regimes become increasingly unpopular over time, particularly as new generations begin to challenge the status quo and question the government's legitimacy" (Chenoweth and Ulfelder, 2017, 303).

[30]Appendix Figure B-2 also shows reported results are robust to adjusting for mobile phone penetration.

[31]Reported results from a kernel estimator with unit means of all explanatory variables to proxy for unit fixed effects, thus isolating over time variation in digital repression and protest.

[32]The mean value of this measure of protest is 0.82 with a standard deviation of 0.77.

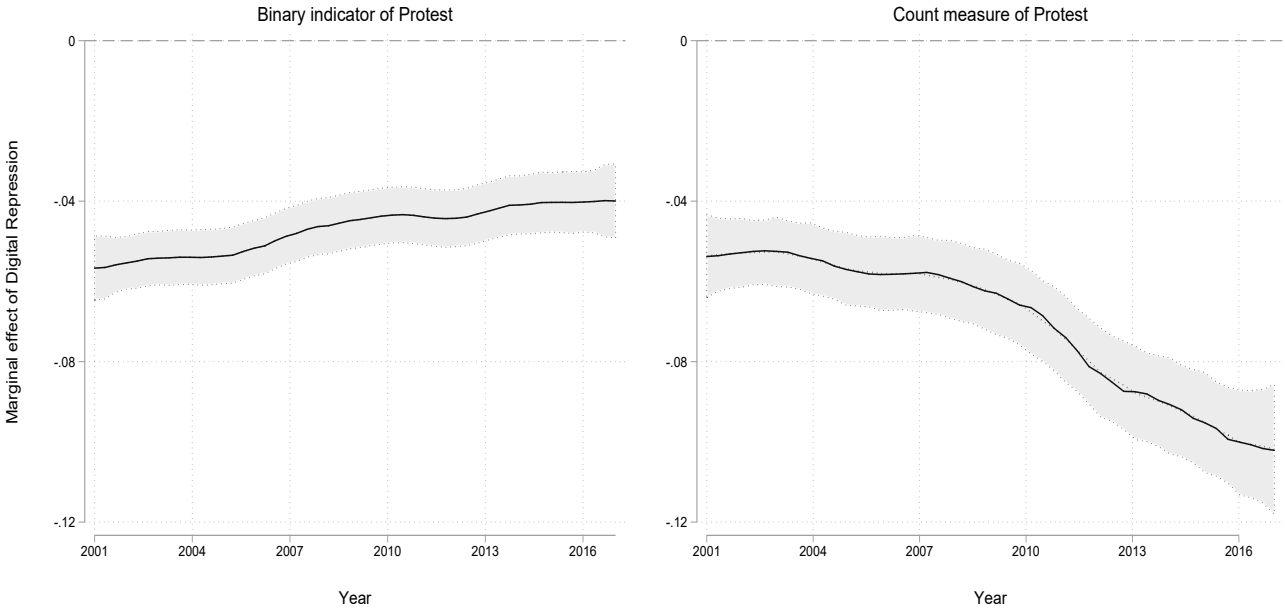

Figure 3: *Digital repression and protest.* Estimates from kernel least squares estimator with units means to proxy for fixed effects; marginal effects of a one standard deviation increase in change in digital repression.

of protest declines. Taken together, the evidence offered in this section reveals that in escalating their use of digital repression, today's authoritarian regimes are countering the most significant threat they now face to their rule.

Digital repression and the broader autocratic repression strategy

Next we look at how digital repression in dictatorships influences the broader repression strategy. Repression comes in many forms but can be grouped into two loose categories based on the intensity of the tactics used.[33] High-intensity repression captures easily observable acts of violence, which usually target well-recognized individuals or groups, such as the Tiananmen Square massacre in 1989 where the Chinese Communist regime killed of hundreds of student demonstrators. Low-intensity repression is far subtler in nature, by comparison. It includes tactics such as surveilling the opposition, detaining government critics for short periods, and using lawsuits to punish opponents. An example would be the Singaporean regime's use of defamation lawsuits to force its opponents into bankruptcy. Most dictatorships use both forms to some degree.

Digital repression falls in the category of low-intensity repression; therefore, we look here at whether its use is impacting reliance on high-intensity repression. Because these two forms of repression contrast each other, comparing them can give us insight into how digital repression is shaping the overall autocratic repression strategy.

There are two contrasting expectations for how digital repression might influence high-intensity repression in dictatorships. On the one hand, if digital repression enables autocratic governments to deter the type of opposition mobilization that destabilizes them (as the prior section on protest

[33]See Frantz (2018) for a more in-depth discussion of these two forms of repression.

suggests), and if high-intensity repression is usual response to this type of mobilization, then we might expect digital repression to serve as a "substitute" for high-intensity repression. This logic presumes that there are political costs – in the form of protestor backlash against observable state-sponsored violence or agency loss in the security sector tasked with carrying out such acts – that autocratic governments would prefer to avoid. If digital repression allows dictatorships to forestall the threats that require high-intensity repression to quell, then they should seek to use digital repression instead, which has fewer political costs. If this logic is true, then we should observe less high-intensity repression in response to an increase in digital repression. According to the substitution argument, in other words, digital repression should decrease high-intensity repression.

An alternative logic, however, suggests the opposite relationship. If digital repression helps governments better identify their opponents, it might enhance their pursuit of high-intensity repression by lowering the costs of the latter. For example, some forms of low-intensity repression, such as the Internet firewall in China, that raise the costs for citizens of gathering information (Roberts, 2018), may enhance the capacity of the regime to identify those dissenters who are willing to incur these costs. Similarly, digital surveillance may lead to better information on dissidents (Xu, 2019), particularly opposition leaders most likely to mobilize against the regime. In these scenarios, governments have better information about which dissidents to target without having to incur the costs of citizens backlash or agency loss. It is reasonable to expect that such costs are more likely to occur when there is indiscriminate state-led violence as opposed to discriminate targeting of the most dangerous dissidents. If digital repression helps mitigate the costs of high-intensity repression, the state may be more likely to ratchet up the latter, albeit in a more methodical and efficient manner. Indeed, this reasoning is consistent with Gohdes (2020), who finds that Internet surveillance by the Syrian government increased targeted repression against the opposition. But where this surveillance is not possible, she finds that the government pursues more indiscriminate or 'un-targeted' repression. According to this logic, then, digital repression "complements" high-intensity repression by making the latter less political costly. The complement argument therefore implies that digital repression should increase high-intensity repression.

To test these alternative expectations, we again use cross-national empirical tests. We measure high-intensity repression using the physical violence index from the Varieties of Democracy project. This index combines information from scaled measures of freedom from torture and freedom from political killings by the government, which we see as a reasonable proxy for high-intensity repression.[34] We check the reported results using a second measure of high-intensity repression from Fariss (2014) and Schnakenberg and Fariss (2014), though there is debate about the extent to which this measure and the Varieties of Democracy measure capture the same concepts. We flip the scales of both measures so that higher values correspond to higher high-intensity repression.

Because we want to know whether digital repression increases, decreases, or has no effect on high-intensity repression, we utilize a differences estimator. This approach, which estimates whether a one-year change in digital repression is associated with a one-year change in high-intensity repression, tests whether there is a short-term relationship between the two within dictatorships, while accounting for all differences between regimes.[35]

Figure 5 shows the results. The first test, with the result shown on the left, is simply a

[34]The question the violence index (`v2x_clphy`) measures is: "To what extent is physical integrity respected?" where physical integrity is "understood as freedom from political killings and torture by the government" (Coppedge et al., 2019a, 263). As the Varieties of Democracy codebook states, "The index is based on indicators that reflect violence committed by government agents" (Coppedge et al., 2019a, 263).

[35]We exclude autocratic regime collapse years from the analysis because this because the Varieties of Democracy data likely capture coders' assessments of concepts using information from after the regime collapse event in those years.

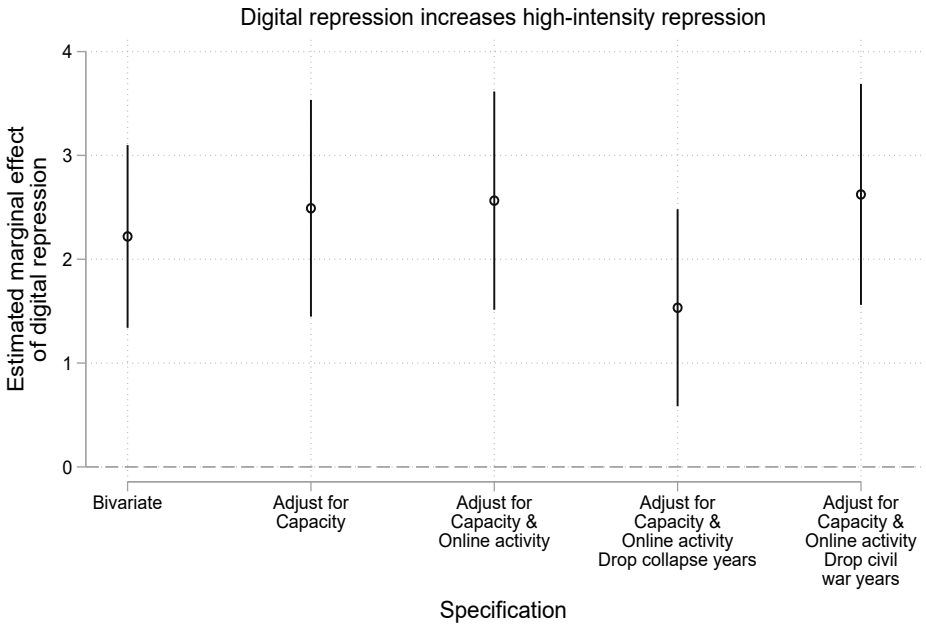

Figure 4: *Digital repression and high-intensity repression.* Estimates based on differenced model; marginal effects of a one standard deviation increase in change in digital repression on standard deviations of change in high-intensity repression.

bivariate differences model. The second test adjusts for digital capacity; and the third adjusts for the existence of online citizen activity. All three specifications yields results that indicate increases in digital repression are associated with subsequent increases in high-intensity repression.[36] Next, we test this model but exclude autocratic regime collapse years from the analysis because the Varieties of Democracy data may capture coders' assessments of concepts using information from after the regime collapse event in those years. Finally, we drop civil war years from the sample. Adjusting the sample in these ways yields similar results.

The evidence presented here is consistent with the "complement" argument. As dictatorships increase their reliance on digital repression, they consequently increase their use of high-intensity repression, as well. Not only are dictatorships leveraging new technologies to suppress protests, but they are also using it to hone their broader repressive strategy. This suggests that digital repression is not replacing high-intensity repression, but instead improving authoritarian's ability to execute it. By making it easier for authoritarian regimes to identify their opposition, digital repression is allowing them to more precisely apply their repressive tactics. Digital repression provides dictatorships with more information about dissidents, lessening their need to apply violence indiscriminately and trigger political costs.

An example from Egypt illustrates these dynamics. In 2019, reports surfaced that the dictatorship there governed by Abdel Fattah el-Sisi had launched a number of sophisticated cyberattacks against regime dissidents.[37] It installed software on the phones of dissidents, which allowed it to read their files and email, monitor their locations, and determine who they had been in touch

[36]Error-correction estimators yield similar findings for the short-term effect. After adjusting for state digital capacity and political civil liberties, we find no support for reverse causation. Further, we find that digital repression granger causes high-intensity repression, while high-intensity repression does not granger cause digital repression.

[37]Ronen Bergman and Declan Walsh, "Egypt Is Using Apps to Track and Target Its Citizens, Report Says," The New York Times, 04 October 2019, `https://www.nytimes.com/2019/10/03/world/middleeast/egypt-cyber-attack-phones.html` (accessed 24 February 2020).

with and when. The regime then used the information gained from the attacks to identify particularly threatening individuals, who were then arrested in government crackdowns on the opposition. Reports indicate that in total at least 33 opposition figures were arrested on account of the cyber-attacks since they began in 2016.

As the Egyptian experience demonstrates and the cross-national empirical tests here show, digital repression facilitates dictatorships' ability to identify opponents, making it easier for them to carry out high-intensity repression. In this way, greater reliance on digital repression in dictatorships brings with it the escalation of state-sponsored violence as a consequence.

Digital repression and autocratic survival

Given that digital repression in autocracies appears to lower the risk of protest and increase the use of high-intensity repression, we next examine whether it contributes to longer-lasting authoritarian rule. Basic summary statistics seem to support this. From 1946 to 2000, the typical dictatorship governed for about a decade. Since 2000 – about the time period when new technologies began to spread worldwide – this number has grown considerably to almost 25 years. Moreover, the evidence suggests that digital repression has played a role in the greater autocratic durability we are witnessing since the turn of the century. From 2000 to 2017, roughly 40 percent of the 91 autocratic regimes that lasted more than a year in power fell. Comparing the regimes that lost power with those that remained in office, we find that average levels of digital repression were substantially higher in the latter group. Instead of withering in the face of the emergence and spread of new technologies, many dictatorships appear to have adapted these tools in ways that strengthen them.

Because these relationships could be due to confounding factors, we next use cross-national empirical tests to evaluate the relationship between digital repression and autocratic durability. We adjust for digital capacity, a measure of the extent to which citizens participate in online activity, and year effects, given evidence of time trends in the data, as we did in our models of protest. We also adjust for two indicators of autocratic regime categories (those led by a high-ranking military officer and those led by a pre-existing political party). These latter two indicators model important political differences across different autocratic regimes (Geddes, Wright and Frantz, 2018). Finally, we adjust for GDP per capita, a quasi-exogenous factor that may be related to both regime stability and digital repression. To test the relationship, we use a survival framework. Our dependent variable is a binary measure of autocratic regime collapse, as measured by Geddes, Wright, and Frantz (2014), excluding those instances in which a foreign power toppled the regime in a military invasion (e.g. Iraq in 2003). From 2001 to 2017 (our sample years), there were 41 regime collapse events, occurring 3.8 percent of the time on average. We lag our measure of digital repression here, as we did earlier in the models of protest; and include polynomials of regime duration.[38]

First, we test a pooled model that treats digital capacity and online citizen activity as potential confounders. This pooled approach, while adjusting for confounders, helps answer the question of whether governments that employ more digital repression are less likely to fall from power than those that employ less. Here, the comparison is across autocratic regimes, not within them over time. The results suggest a negative relationship between digital repression and autocratic regime collapse: a one-unit increase in digital repression decreases the probability of regime breakdown by roughly 1.3 percent on average in a given year. This finding is illustrated in the left panel of Figure 5. The horizontal axis depicts years during the sample, while the vertical axis shows the estimated marginal effect, smoothed across the panel years. While the average marginal effect is

[38]Reported results are from a kernel least squares estimator.

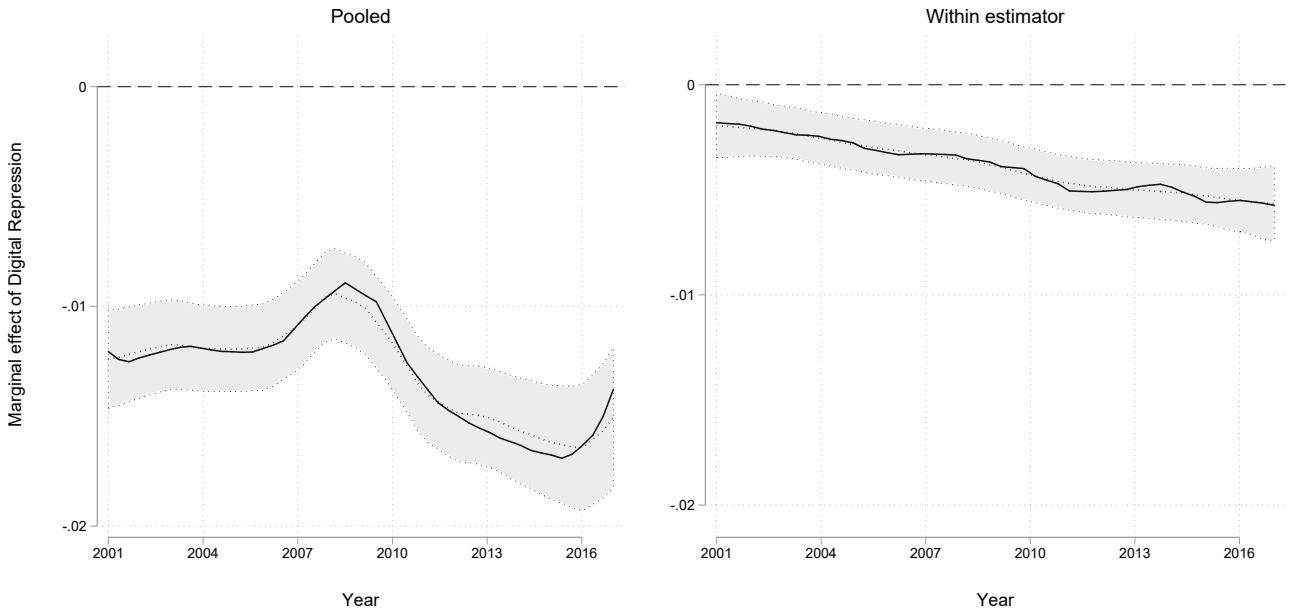

Figure 5: *Digital repression and autocratic regime collapse.*

between -1 and -2 percent, the plot shows that this estimate is closer to 1 percent in the earlier part of the sample period and drops to closer to 2 percent in the latter part of the sample period. While lowering the risk of regime collapse by 1 to 2 percent may seem small, it is important to highlight that the baseline risk of regime collapse is 3.7 percent. In short, when pooling differences across countries in their level of digital repression, the evidence suggests that digital repression lowers the chance of autocratic breakdown.[39]

We next look at whether this relationship holds when isolating changes over time within regimes in digital repression instead. Rather than asking whether more digitally repressive autocracies are less vulnerable to collapse than those that digitally repress less, here we are asking whether increases in digital repression over time within regimes influences their chance of falling from power. In other words, if China or Ethiopia were to increase their reliance on digital repression, how would this influence their prospects for regime survival?

The right plot in Figure 3 illustrates the results. It shows that the estimate of the effect is close to zero; and while visually the polynomial plot of pointwise marginal effects shows average effects shows they are different from zero, the standard error estimate does not. In short, these are substantively small and statistically insignificant results. This indicates that, at least given these data and time period, changes in the level of digital repression within autocratic regimes are not associated with more or less vulnerability to regime collapse.

To summarize, the results offered in this section are mixed. On the one hand, when autocratic regimes increase their use of digital repression, this does not appear to influence their risk of breaking down. On the other hand, those autocratic regimes that rely on digital repression more, are less vulnerable to falling from power.

[39]These results, however, should be interpreted with some caution because we do not find statistically significant pooled estimates when using probit or linear probability estimators.

Conclusion

To date, academics have lagged behind the digital revolution in terms of understanding systematically how the emergence of new technologies is altering the autocratic landscape. In this study, we seek to fill this void by providing some insights into the extent to which dictatorships are using digital repression, how its use has changed over time, and how reliance on digital repression affects other outcomes of interest. In particular, we show that digital repression lowers a dictatorship's risk of experiencing protest, an especially important finding in light of the fact that protests are now the most serious threat to autocratic survival. In addition, reliance on digital repression escalates high-intensity repression, such as the targeted use of torture and imprisonment. Digital autocracies are not simply trading in one repressive tool for another; instead they are using digital tools to fine-tune their broader repressive approach. Lastly, we find that some evidence that digital repression is associated with more durable dictatorship. This finding is both relatively small and statistically insignificant in the relatively short panel series, however. In these ways, this study provides a starting point for future research to more thoroughly examine the myriad ways that digital repression is altering contemporary authoritarian rule.

References

Ahmed, Shazeda. 2019. "Credit Cities and the Limits of the Social Credit System.". https://www.airuniversity.af.edu/Portals/10/AUPress/Books/B_0161_WRIGHT_ARTIFICIAL_INTELLIGENCE_CHINA_RUSSIA_AND_THE_GLOBAL_ORDER.PDF. Accessed 25 February 2020.

Beck, Nathaniel. 2018. "Estimating grouped data models with a binary dependent variable and fixed effects: What are the issues." *arXiv preprint arXiv:1809.06505* .

Chenoweth, Erica and Jay Ulfelder. 2017. "Can structural conditions explain the onset of nonviolent uprisings?" *Journal of Conflict Resolution* 61(2):298–324.

Clark, David and Patrick Regan. 2016. "Mass Mobilization Protest Data.". https://doi.org/10.7910/DVN/HTTWYL. Accessed 25 February 2020.

Cook, Scott J, Jude C Hays and Robert J Franzese. 2018. "Fixed effects in rare events data: a penalized maximum likelihood solution." *Political Science Research and Methods* pp. 1–14.

Cope, Kevin L, Charles Crabtree and Christopher J Fariss. 2020. "Patterns of disagreement in indicators of state repression." *Political Science Research and Methods* 8(1):178–187.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Agnes Cornell, Lisa Gastaldi, Haakon Gjerlow, Valeriya Mechkova, Johannes von Römer, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig and Daniel Ziblatt. 2019a. "V-Dem Codebook v9." *V-Dem Working Paper* .

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Steven Wilson, Agnes Cornell, Lisa Gastaldi, Haakon Gjerlow, Nina Ilchenko, Joshua Krusell, Laura Maxwell, Valeriya Mechkova, Juraj Medzihorsky, Josefine Pernes, Johannes von Römer, Natalia Stepanova, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig and Daniel Ziblatt. 2019b. "V-Dem Dataset v9." *V-Dem Working Paper* .

Davenport, Christian. 2007. "State repression and the political order." *Annual Review of Political Science* 10:1–23.

Deibert, Ron. 2015. "Authoritarianism goes global: Cyberspace under siege." *Journal of Democracy* 26(3):64–78.

Diamond, Larry. 2015. Liberation technology. In *In Search of Democracy*, ed. Larry Diamond. Routledge chapter Chapter 7, pp. 132–146.

Fariss, Christopher J. 2014. "Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability." *American Political Science Review* 108(2):297–318.

Feldstein, Steven. 2019*a*. "The Global Expansion of AI Surveillance.". https://carnegieendowment.org/files/WP-Feldstein-AISurveillance_final.pdf. Accessed 25 February 2020.

Feldstein, Steven. 2019*b*. "The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression." *Journal of Democracy* 30(1):40–52.

Francisco, Ronald A. 1995. "The relationship between coercion and protest: An empirical evaluation of three coercive states." *Journal of Conflict Resolution* 39(2):263–282.

Frantz, Erica. 2018. *Authoritarianism: What Everyone Needs to Know*. Oxford University Press.

Frantz, Erica and Andrea Kendall-Taylor. 2014. "A Dictator's Toolkit: Understanding How Co-optation Affects Repression in Autocracies." *Journal of Peace Research* p. 0022343313519808.

Frantz, Erica and Andrea Kendall-Taylor. 2017. "The Evolution of Autocracy: Why Authoritarianism Is Becoming More Formidable." *Survival* 59(5).

Gartner, Scott Sigmund and Patrick M. Regan. 1996. "Threat and repression: The non-linear relationship between government and opposition violence." *Journal of Peace Research* 33(3):273–287.

Geddes, Barbara, Joseph Wright and Erica Frantz. 2014. "Autocratic breakdown and regime transitions: A new data set." *Perspectives on Politics* 12(2):313–331.

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

Gohdes, Anita R. 2020. "Repression Technology: Internet Accessibility and State Violence." *American Journal of Political Science* forthcoming.

Greitens, Sheena. 2019. "Explaining the Diffusion of Repression Technology: The Global Adoption of Chinese Public Security Technology." *Paper prepared for the Annual Meeting of the American Political Science Association* .

Guvitsky, Seva. 2015. "Corrupting the Cyber-Commons: Social Media as a Tool of Autocratic Stability." *Perspectives on Politics* 13(1):42–54.

Hoffman, Susan. 2019. "China's Tech-Enhanced Authoritarianism.". `https://docs.house.gov/meetings/IG/IG00/20190516/109462/HHRG-116-IG00-Wstate-HoffmanS-20190516.pdf`. Accessed 25 February 2020.

Howard, Phil. 2013. "When Do States Disconnect Their Digital Networks, 1985-2011." *Original Data Set* . Available at: `http://philhoward.org/original-datasets-and-archives/event-history-data-sets/`.

Howard, Philip N, Sheetal D Agarwal and Muzammil M Hussain. 2011. "When do states disconnect their digital networks? Regime responses to the political uses of social media." *The Communication Review* 14(3):216–232.

Kalyvas, Stathis. 2006. *The Logic of Violence in Civil Wars*. Cambridge University Press.

Kendall-Taylor, Andrea and Erica Frantz. 2014. "Mimicking Democracy to Prolong Autocracies." *Washington Quarterly* 37(4):71–84.

Kendall-Taylor, Andrea, Erica Frantz and Joseph Wright. 2020. "The Digital Dictators: How Technology Strengthens Autocracy." *Foreign Affairs* Mar/Apr.

Koehler, John O. 1999. *Stasi: The untold story of the East German secret police*. Westview Press.

Lichbach, Mark Irving. 1987. "Deterrence or escalation? The puzzle of aggregate studies of repression and dissent." *Journal of Conflict Resolution* 32(2):266–297.

Mechkova, Valeriya, Daniel Pemstein, Brigitte Seim and Steven Wilson. 2019. "Digital Society Project Dataset, v1." *Digital Society Project* 1. Available at `https://www.digitalsocietyproject.org`.

Moore, Will H. 1998. "Repression and dissent: Substitution, context, and timing." *American Journal of Political Science* 42(3):851–873.

Mundlak, Yair. 1978. "On the pooling of time series and cross section data." *Econometrica: journal of the Econometric Society* pp. 69–85.

Pemstein, Daniel, Kyle Marquardt, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri and Johannes von Römer. 2019. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." *V-Dem Working Paper* .

Qiang, Xiao. 2019. "The Road to Digital Unfreedom: President Xi's Surveillance State." *Journal of Democracy* 30(1):53–67.

Qin, Bei, David Stromberg and Yanhui Wu. 2017. "Why Does China Allow Freer Social Media? Protests versus Surveillance and Propaganda." *The Journal of Economic Perspectives* 31(1):117–140.

Roberts, Margaret E. 2018. *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton University Press.

Rozenas, Arturas and Yuri M. Zhukov. 2019. "Mass Repression and Political Loyalty: Evidence from Stalin's 'Terror by Hunger'." *American Political Science Review* 113(2).

Rydzak, Jan Andrzej. 2018. "A Total Eclipse of the Net: The Dynamics of Network Shutdowns and Collective Action Responses." *PhD Dissertation, The University of Arizona* .

Schnakenberg, Keith and Christopher J. Fariss. 2014. "Dynamic Patterns of Human Rights Practices." *Political Science Research and Methods* 2(1):1–31.

Svolik, Milan. 2012. *The Politics of Authoritarian Rule*. Cambridge University Press.

Truex, Rory. 2017. "Consultative Authoritarianism and Its Limits." *Comparative Political Studies* 50(3):329–361.

Tucker, Joshua A., Yannis Theocharis, Margaret E. Roberts and Pablo Barbera. 2017. "From Liberation to Turmoil: Social Media and Democracy." *Journal of Democracy* 28(4).

Wilson, Steven Lloyd. 2017. "Information and Revolution." *V-Dem Working Paper* 50.

Wintrobe, Ronald. 1998. *The Political Economy of Dictatorship*. Cambridge University Press.

Wooldridge, Jeffrey M. 2002. *Econometric analysis of cross section and panel data*. MIT press.

Wright, Nicholas D. 2019. "Artificial Intelligence, China, Russia, and the Global Order.". https://www.airuniversity.af.edu/Portals/10/AUPress/Books/B_0161_WRIGHT_ARTIFICIAL_INTELLIGENCE_CHINA_RUSSIA_AND_THE_GLOBAL_ORDER.PDF. Accessed 25 February 2020.

Xu, Xu. 2019. "To Repress or To Co-opt? Authoritarian Control in the Age of Digital Surveillance." *American Journal of Political Science* forthcoming.

Contents

1 Appendix A: Digital Repression Latent Estimate

1.1 Latent linear index estimates

(a) Digital **repression**

Item	*Variable name*	*Question wording*
social media censorship in practice	v2smgovsmcenprc	To what degree does the government censor political content (i.e., deleting or filtering specific posts for political reasons) on social media in practice?
social media monitoring	v2smgovsmmon	How comprehensive is the surveillance of political content in social media by the government or its agents?
social media shut down in practice	v2smgovsm	How often does the government shut down access to social media platforms?
Internet shutdown in practice	v2smgovshut	How often does the government shut down domestic access to the Internet?
Internet filtering in practice	v2smgovfilprc	How frequently does the government censor political information (text, audio, images, or video) on the Internet by filtering (blocking access to certain websites)?
social media alternatives	v2smgovsmalt	How prevalent is the usage of social media platforms that are wholly controlled by either the government or its agents in this country?

(b) Digital **capacity**

Item	*Variable name*	*Question wording*
cyber security capacity	v2smgovcapsec	Does the government have sufficiently technologically skilled staff and resources to mitigate harm from cyber-security threats?
Internet shut down capacity	v2smgovshutcap	Independent of whether it actually does so in practice, does the government have the technical capacity to actively shut down domestic access to the Internet if it decided to?
Internet filtering capacity	v2smgovfilcap	Independent of whether it actually does so in practice, does the government have the technical capacity to censor information (text, audio, images, or video) on the Internet by filtering (blocking access to certain websites) if it decided to?
capacity to regulate online content	v2smregcap	Does the government have sufficient staff and resources to regulate Internet content in accordance with existing law?

Table A-1: *Question wording for each item in the digital repression and capacity indices*

Table A-1 lists the items (i.e. variables), variable names, and question wording for the six items in the *digital repression* index and the four items in the *digital capacity* index. We construct the respective linear indexes by using the standardized (mean 0, variance 1) values of the individual items for each index. Table 1 in the main text shows the item-test correlation for each item used to construct the respective indices. Overall the items in each index are highly inter-correlated, suggesting that they can be reliably combined into one scaled index. The digital repression measure is positively correlated with digital capacity (0.42) and infant mortality rates (0.30) but negatively correlated with urban population (-0.16). There is almost no correlation between digital repression

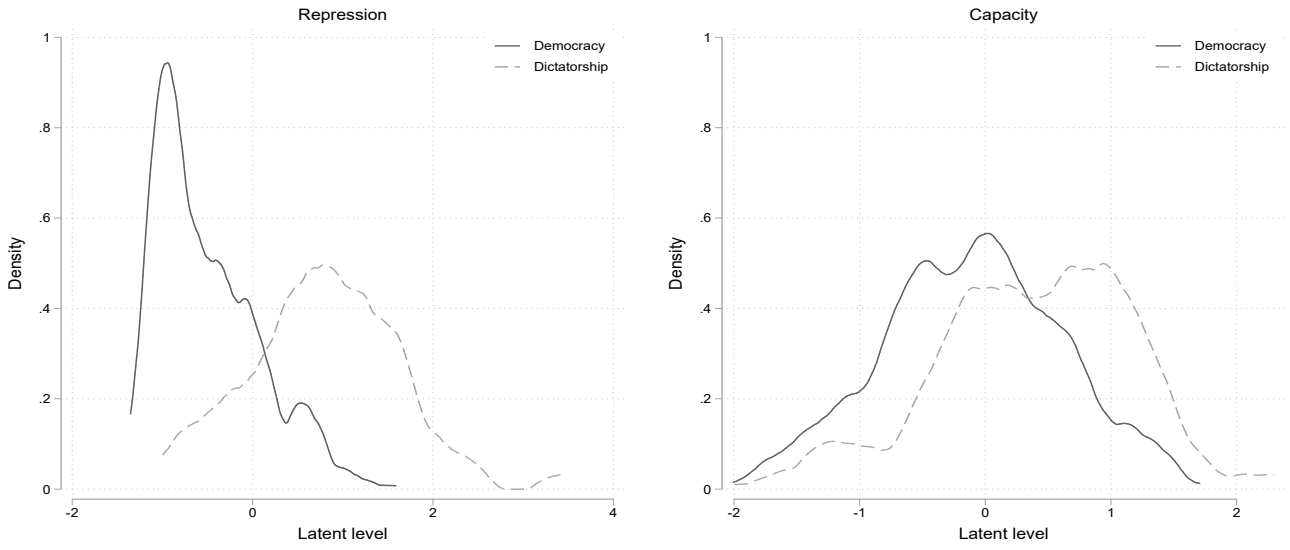

Figure A-1: Distributions of digital repression and capacity measures, by political regimes

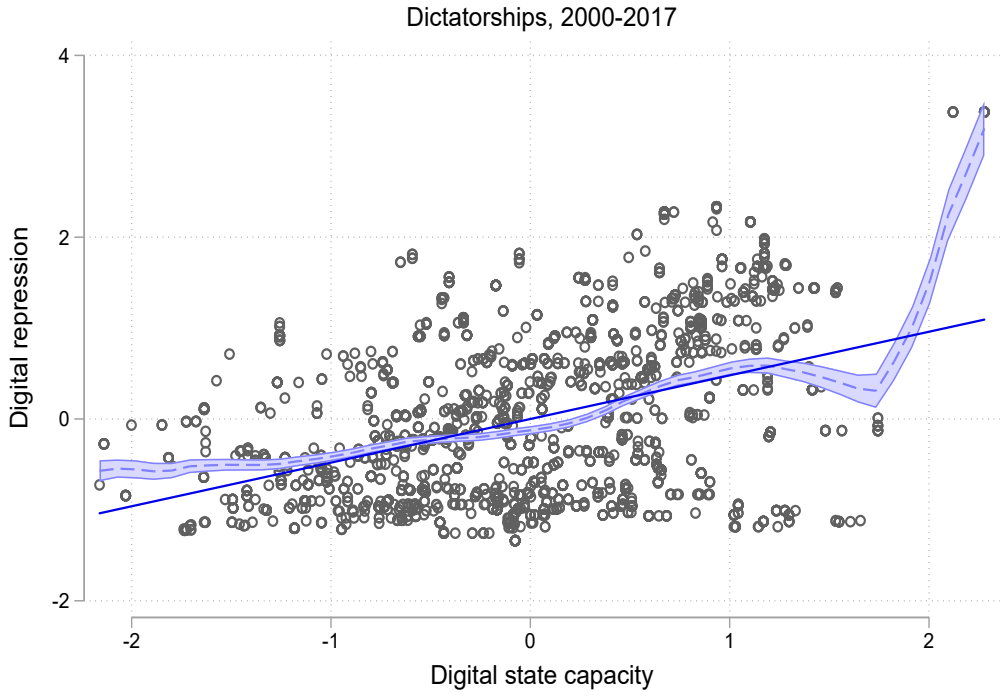

Figure A-2: Digital repression and capacity measures

and GDP per capita (0.01).

Figure A-1 shows the distribution of the repression and capacity indexes, by regime type. The left plot shows the repression distribution: on average, dictatorships (0.79) have substantially more observed digital repression than democracies (-0.51). The right plot shows the capacity distribution:

while autocratic governments (0.39) have, on average, more digital capacity than their democratic counterparts (-0.11), the difference is not as large. This suggests that autocratic governments use more of their repressive capacity than democratic governments.

Figure A-2 plots the digital capacity score (horizontal axis) against the digital repression score (vertical axis). There is a strong, positive correlation between the two scores, especially through the middle of the distribution of capacity (roughly -1 to 1 on the horizontal axis). The outlying observations in the upper-right of the plot all come from North Korea, which appears to have substantially more digital repression that expected given its (high) level of digital capacity.

Finally, Figure A-3 shows the digital repression scores by regime cases (listed on the vertical axis). Some countries have more than one regime (for example Guinea and Mauritania) but most countries have only one. The horizontal axis depicts the latent level of digital repression, where higher values mean more repression. The blue dots are the regime-case mean levels, while the horizontal gray lines show the minimum and maximum values during the period from 2000 to 2017. Many cases do not appear to have gray lines simply because the mean value reflects (roughly) the minimum and maximum values; this indicates there is very little variation in the level of digital repression in these cases. The regime with the highest score is North Korea.

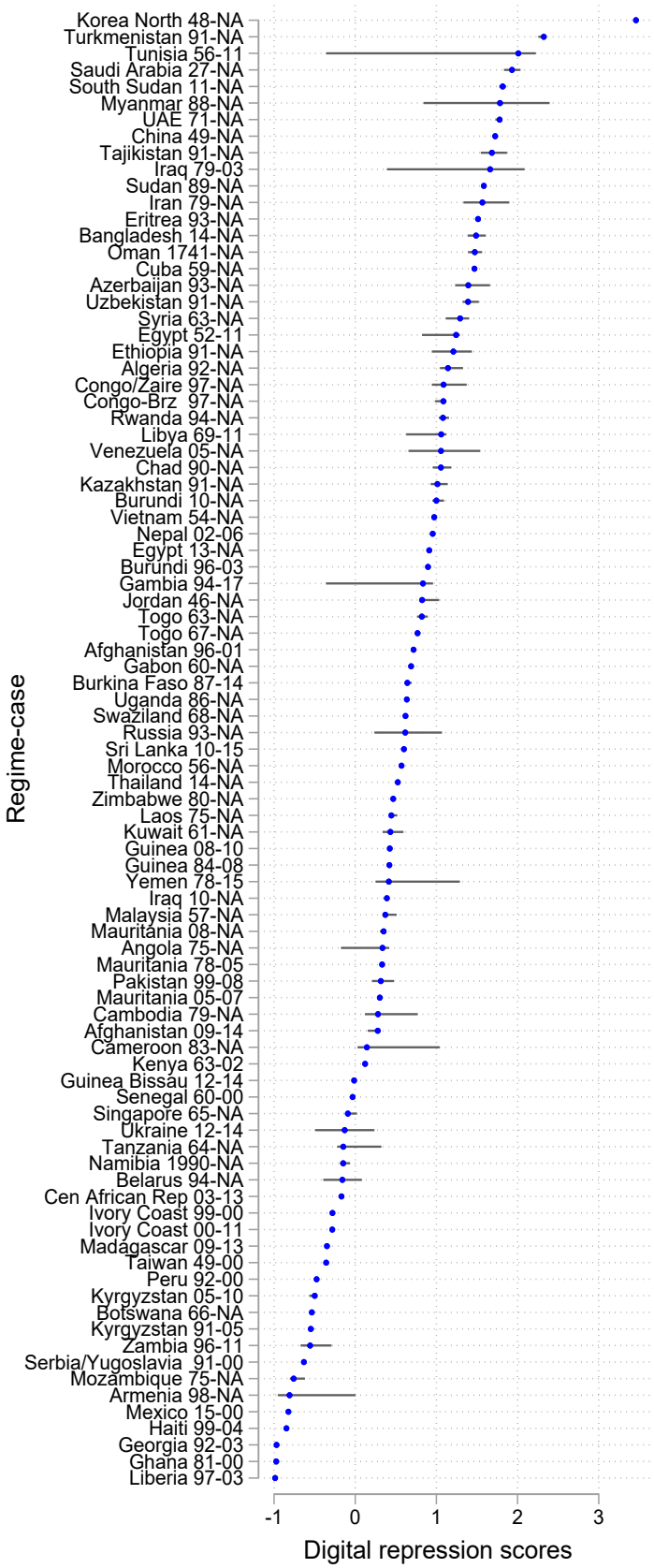

Figure A-3: Digital repression for autocratic regime cases

1.2 Reliability and validity

Internal reliability To assess internal reliability of the *digital repression* measure, we re-estimate the linear scaling model four different ways. First, we divide the sample randomly into two groups and re-estimate the scaled index for each group (*random partition*). Second, we divide the sample by time period, grouping the years 2000 to 2008 into one bin and the years 2009 to 2017 into another and re-estimate the scaled index for each group (*time partition*). Third, we divide the sample into three geographic groups (Americas + Europe; Africa + Middle East; and Asia) and re-estimate the scaled index for each group (*geographic partition*). Finally, we divide the six items into two groups of three items and re-estimate the scaled index for each group (*item partition*). We then compare how these re-estimated measures compare with the original scaled index. For the first three partitions, the correlation is greater than 0.99; for the *item partition* the correlation is over 0.97. This suggests that the linear index measure is reliable across time and space and even across different sets of items.

External validity This section examines how the digital repression measure we construct from the six items in the VDem data set matches up with data from three external measures of Internet shut downs and Internet filtering. We look at two extant data collection efforts for Internet shut downs, from Howard, Agarwal and Hussain (2011) and Rydzak (2018), and the Internet filtering score for political content from the OpenNet Initiative (ONI).[40] The ONI data, however, only cover selected countries. The Internet shut down data record government shut down events, which produces count data, while the ONI data is a five-point ordinal scale measuring the extent to which the government filters and blocks political content on the Internet.[41]

First, we assess how the digital repression index compares with the shut down data and the ONI score. We also include, as a separate measure, the VDem variable for Internet shut downs. Figure A-4 shows the pairwise correlations. The digital repression measure is highly correlated with the VDem measure of Internet shut downs (recall that the latter is part of the former). But neither the digital repression measure nor the VDem Internet shut down variable is closely correlated with Internet shut down data from Howard and Rydzak. Further, the Howard and Ryzdak data are not particularly highly correlated (0.24). Finally, the ONI political score is not particularly highly correlated with any of the other measures. In short, while the VDem data is highly inter-correlated, neither the full index of digital repression nor the individual indicator of Internet shut downs matches well other data sets measuring the same phenomena. One reason for the apparent mismatch is that the event data often records multiple local Internet shut downs (e.g. in Kashmir in India) whereas the VDem is measuring national-level phenomena and therefore more likely to be picking up what occurs in the main cities and the capital city – the very places protests are most likely to destabilize autocratic governments.

Next, we attempt to construct a scaled index of the concept of "Internet shut down" by combining the VDem measure of this concept with the Howard, Rydzak, and ONI data into a single linear index. The left panel of Figure A-5 shows that each of these items – VDem, Howard, Ryzdak and ONI – are correlated with the scaled index at between 0.6 and 0.8, but the overall alpha score is relatively low – 0.48. This indicates that combining these items into a single index is likely an unreliable approach. To contrast this, the left panel of Figure A-5 shows item-test correlations for six items that comprise the digital repression index; all are correlated with the index at 0.85 of

[40] See Howard (2013) for the data.

[41] The five categories are: pervasive; substantial; selective; suspected; and no evidence. Given the paucity of scores at the high end, we combine pervasive, substantial, and selective into one category, producing a three-value ordinal scale.

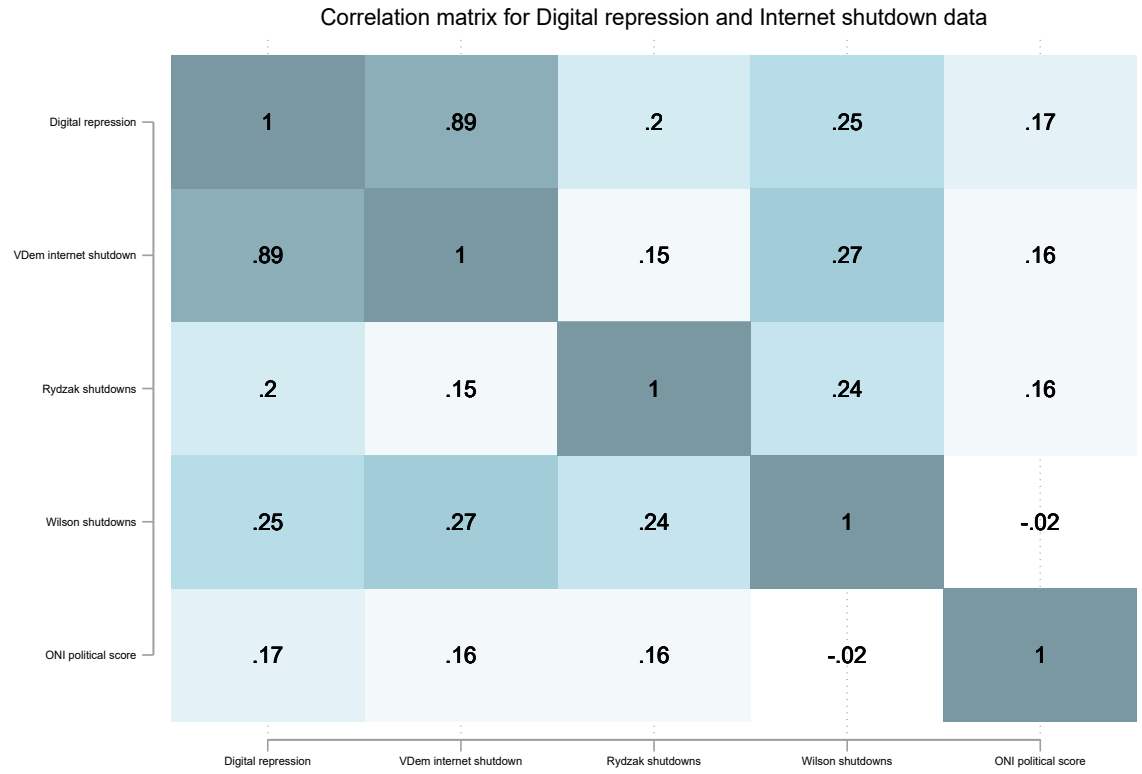

Figure A-4: Digital repression and Internet shutdown correlation matrix.

higher and the alpha score is 0.96. This suggests a highly reliable latent measure.

Next, we explore the extent to which the six VDem items (listed in Table A-1) and the three external items (Howard, Rydzak, and ONI) inter-correlate using factor analysis. The eigenvalues from this exercise are shown in the left panel in Figure A-6. The first factor – or dimension – has an eigenvalue of roughly 5, indicating strong loading on this dimension. The second factor has an eigenvalue of roughly 0.5, indicating a weak second dimension in the data. The right panel of Figure A-6 plots the (rotated) first two factors, placing the loadings for each item (there are nine items) in the two-dimensional space. The six VDem items cluster in the lower right corner, all with relatively high values on the first factor and low values on the second. The Howard measure and the ONI score cluster in upper left corner, with low values on the first factor and high values on the second. This indicates that the VDem items likely belong together in a single index, while the Howard and ONI measures might constitute a second dimension. Finally, the Rydzak data do not correlate with either cluster. Of note, this plot also indicates that the three external data series (Howard, Rydzak and ONI) should not be part of measure of digital repression using the VDem data: they pick up different variation.

That said, we still produce an index that aggregates information from the six VDem items plus the two external items that correlate (Howard and ONI). Figure A-7 shows the resulting index on the vertical axis, plotted against the digital repression index we use (constructed just from the six VDem items) on the horizontal axis. While the pattern shows the two measures are highly correlated, adding the external data simply increases the latent score for some observations when one of the two external data series registers an Internet shut down event.

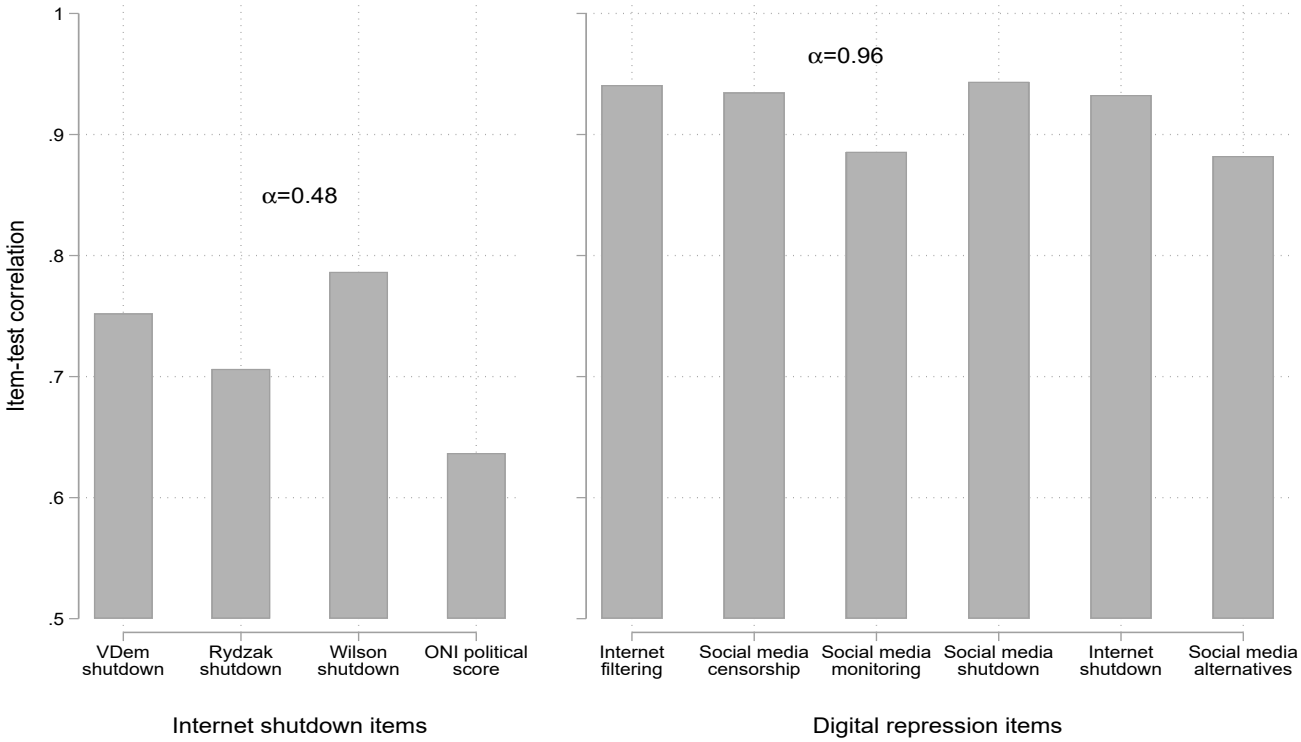

Figure A-5: Item-test correlation for Internet shut down index and digital repression index

Overall, we find that the digital repression index constructed from the six VDem items does not correlate highly with extant, external data series measuring Internet shut downs. This means we should *not* be adding these data series to measures we construct from the VDem items.

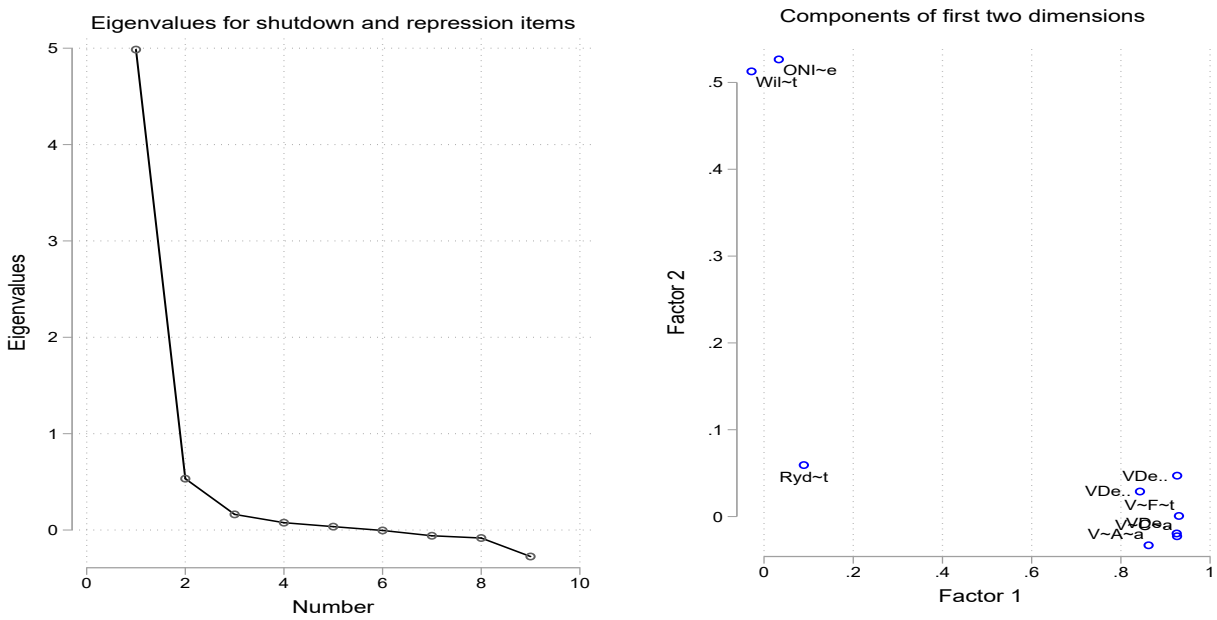

Figure A-6: Eigenvalues and Item loadings.

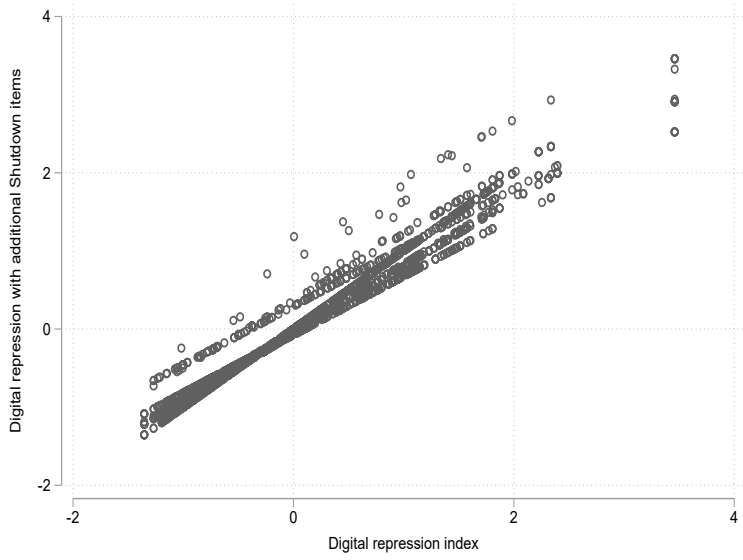

Figure A-7: Digital repression scale with added Internet shut down items.

1.3 Item-response theory estimates

This section introduces and discusses an IRT model for estimating the latent level of digital repression. In the main text we employed a linear combination of six items used to measure various aspects of digital repression. The items in this index, listed in Table A-1, are the mean point estimates from latent models that aggregate responses from multiple country experts who rate each country-year for each item. We combined these six continuous variables – which reflect mean point estimates – using a linear index. This provides the *digital repression* variable used throughout. VDem also provided ordinal measures of all of the concepts used in the digital repression index (i.e. censor social media; monitor social media; shut down social media; shut down Internet; filter Internet; alternative social media). As a robustness test, we construct a measure of digital repression using an IRT graded response model that combines information from ordinal items for the set of underlying concepts.

Figure A-8 plots the item information functions (IIF) for the six items in the IRT latent estimate of *digital repression*, or θ. The vertical axis measures the item discrimination parameter: higher values indicate more information in the latent estimate over a smaller range of θ values. The horizontal axis corresponds to the "difficulty" parameter: larger values indicate items for which observations have a higher estimate of θ. If the model accurately estimates latent *digital repression*, more "difficult" items are those for which an observation must be more repressive to observe a larger ordinal value for this item. This parameter captures how well an item splits high and low *repression* cases at a particular point in the latent space.

The item *Shut down Social Media* provides the most information for separating cases along the middle part of the latent distribution, while *Monitor Social Media* and *Alternative Social Media* provide the least information. Second, *Filter Internet* is a high "difficulty" item, which helps place cases along the high end of the estimated latent space. That is, to be coded as having high *digital repression*, the observation likely has a high value on this item. In contrast, *Monitor Social Media* and *Censor Social Media* both have low "difficulty", indicating that these items help sort cases along the low end of the estimated latent space.

This IRT model produces predicted values for each country-year observation as the empirical Bayes means of latent variables; we call this estimate of digital repression the "IRT estimate" going forward. Importantly, it is correlated with the measure we use in the main text, which is derived from a linear combination of mean point estimates, at 0.98. The IRT measure, however, has slightly more (20.0 percent) "within" variation than the linear scale (18.5 percent).

1.4 Results with alternative measures of digital repression

This section reports the main results for three sets of analyses when using alternative measures of digital repression. For each of three outcomes – binary protest, count protest, and high-intensity repression – we re-estimate the main models but substitute alternative measures of digital repression. For each outcome we report results from five tests. First, we report the original test using the scaled linear index introduced in the main text. Because this measure is a linear index derived from latent model estimates provided by VDem, we call this measure "Model estimate, linear index". As a first alternative measure, we add Internet shut down data from Wilson and ONI as items alongside the original six items from VDem. The resulting linear index is called "Add shut down data, linear index". Third we construct a linear index from the ordinal values provided by VDem for each item (i.e. not the model estimate means); in VDem terminology these are *ord* variables. We call this measure "Ordinal scale, linear index". Fourth, we report results from a linear index constructed from the six VDem items placed on their original scale (*osp* in VDem terminology);

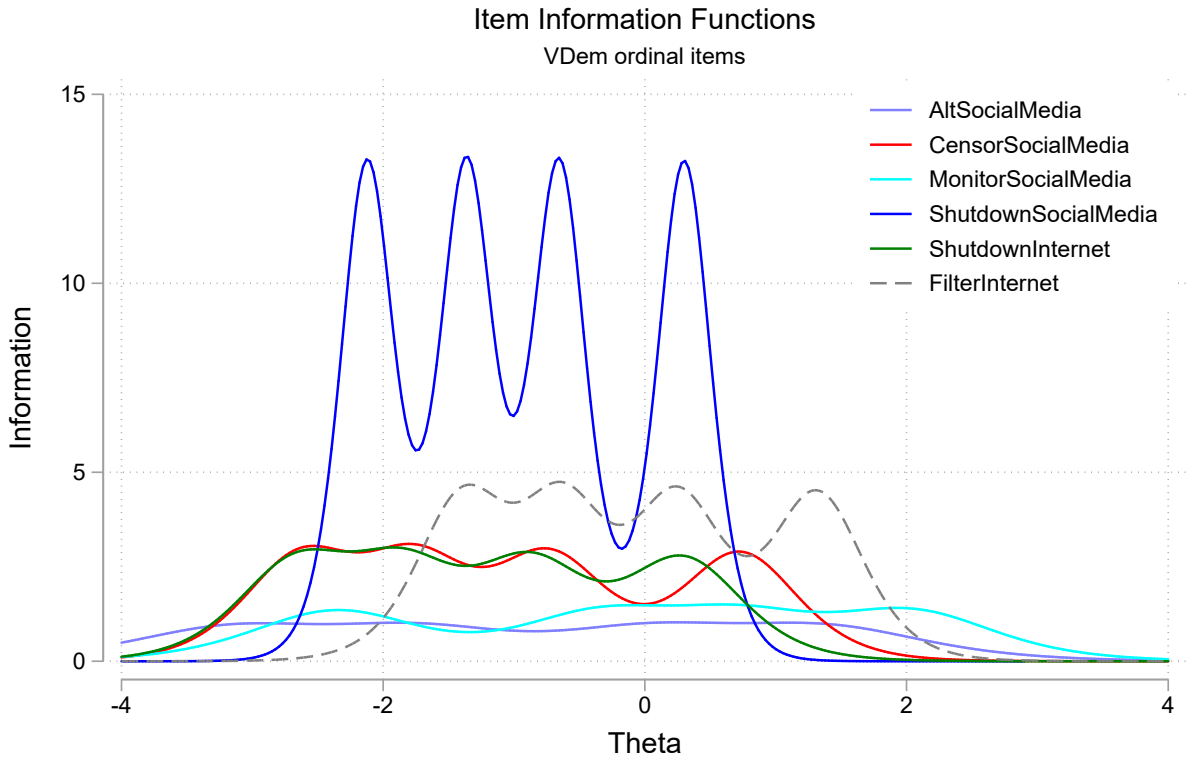

Figure A-8: Item information function, IRT-GRM estimate of digital repression.

we call this variable "Original scale, linear index". Finally, we report results when using the "IRT estimate" discussed in the prior section. Note that the items in this index are the ordinal VDem variables.

Figure A-9 reports results from a 2-way fixed effects linear probability models where the outcome is a binary indicator of protest ($\mu = 0.58$). The first test reports the estimate for *digital repression* used throughout, which we call "Model estimate, linear index". The coefficient estimate is negative and significant. The next four results substitute different, alternative measures of *digital repression*: linear indexes that add Internet shut down data, utilize original scale and ordinal items, and use the ordinal items in an IRT model. All estimates of *digital repression* – irrespective of how this is measured – are negative and significant. This indicates that the main reported findings for binary protest are robust to alternative measures of digital repression.

Figure A-10 reports results from fixed effect negative binomial models where the outcome is a count measure of protest. All the estimates for *digital repression*, irrespective of how this is measured, are negative and significant. Figure A-11 reports results from differenced models of high-intensity repression, similar to those reported in the main text. All the estimates for *digital repression* are positive and significant. Overall, these findings suggest that the exact method of measuring *digital repression*, at least when using the six items from VDem listed in Table A-1, does not alter the main results.

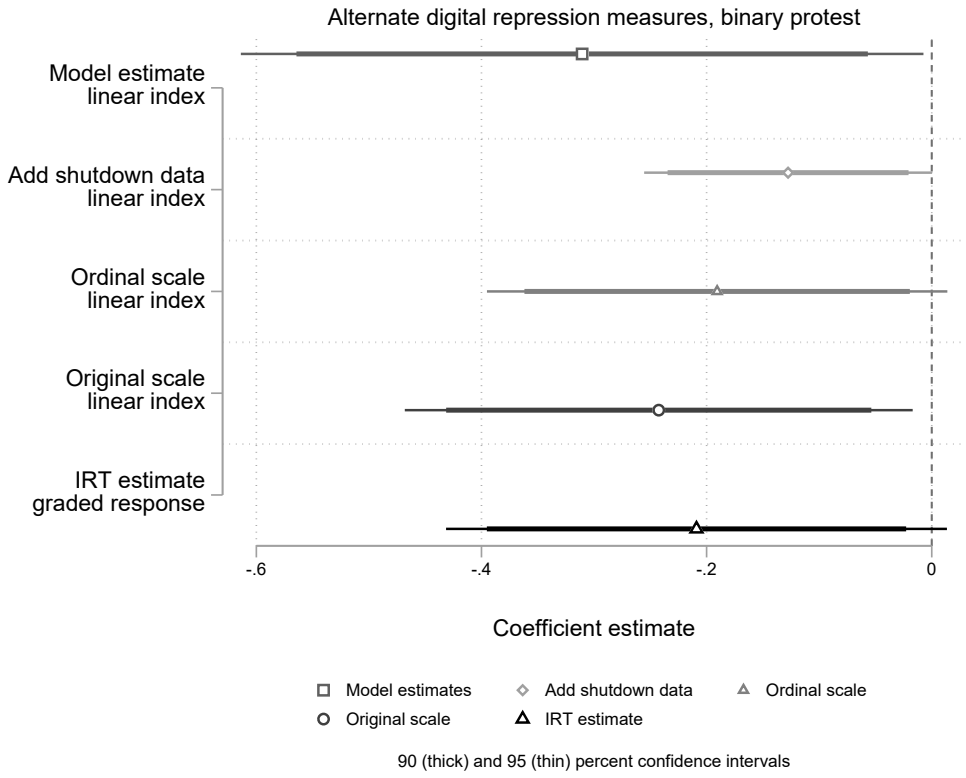

Figure A-9: Binary protest indicator, alternative digital repression measures

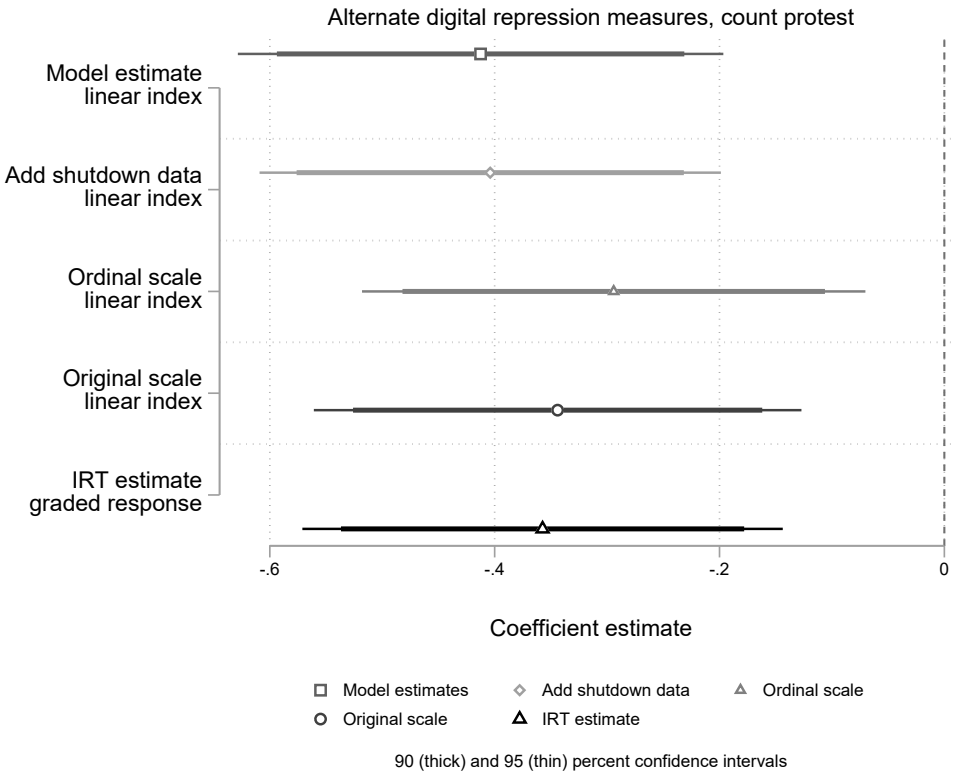

Figure A-10: Count protest data, alternative digital repression measures

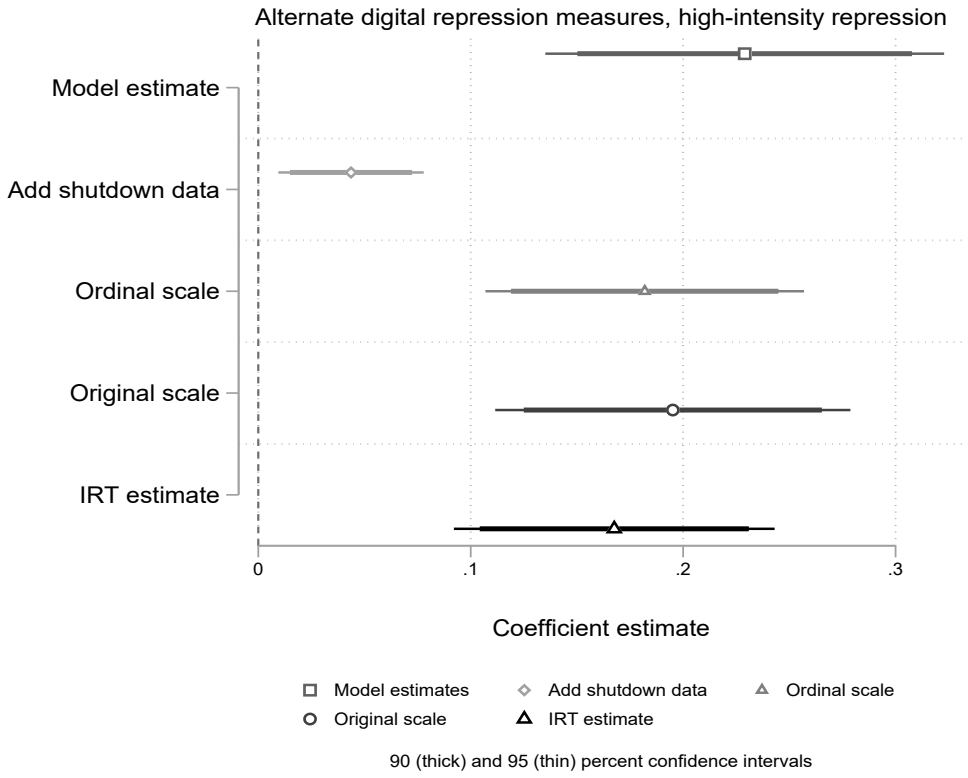

Figure A-11: High-intensity repression, alternative digital repression measures

2 Appendix B: Additional Results for Protest

This section discusses additional results for the protest finding. We start by showing the distribution of the protes count data. While 58 percent of country-year observations have at least one protest event, nearly all regimes in the sample period (2001–2017) face a protest at some point (94 percent of regimes). Because the distribution of the raw counts of protest are highly skewed, we transform the data as following, where T is the transformed count and C is the raw count: $T = (log(C+1))^{0.5}$. The left plot in Figure B-1 shows the distribution of the raw count of protest plotted against the tranformed count; and the right plot shows the distribution of the transformed count. The mean raw count is 23 protests, with a standard deviation of 155; the transformed count has mean of 0.82 with a standard deviation of 0.77. We use a binary ($\mu = 0.58$) indicator of protest for binary dependent variable models and linear probability models; we use the transformed count for linear models; and we retain the raw count for negative binomial models.

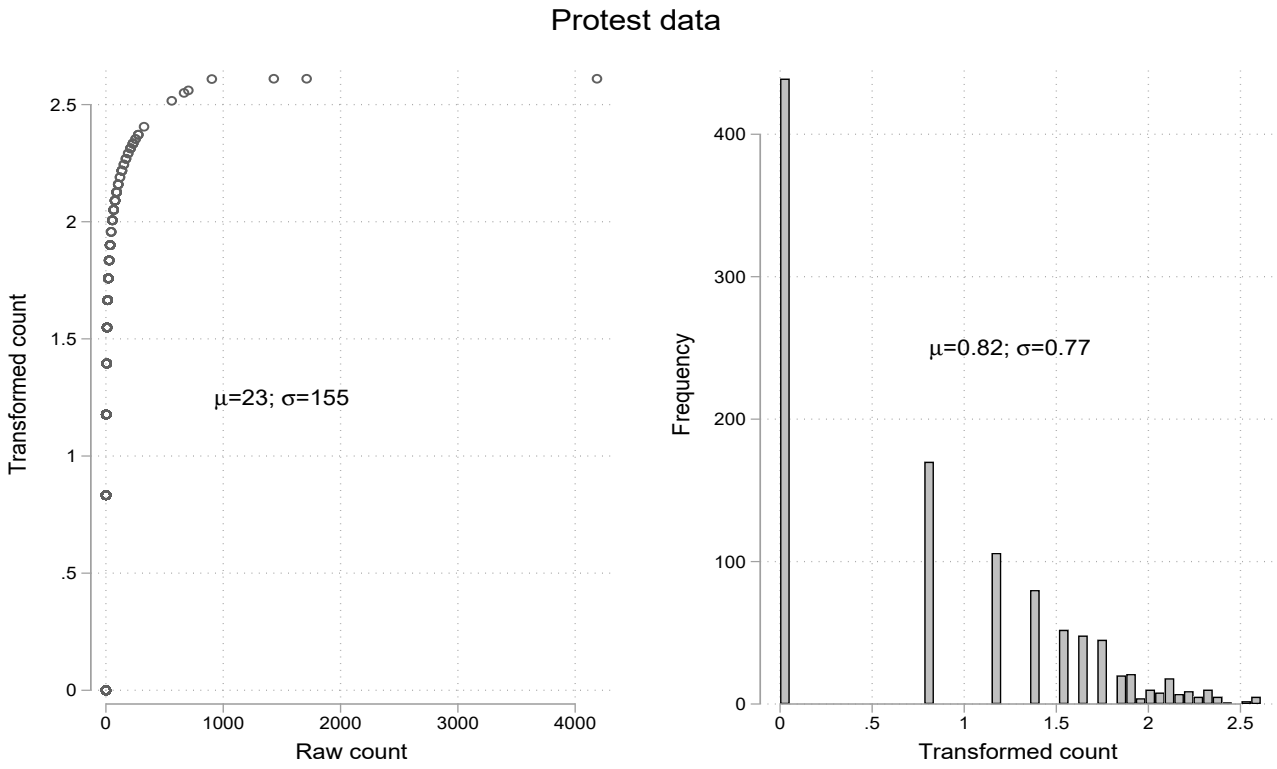

Figure B-1: Protest count data.

The main text reports results using kernel least squares (krls) estimators with unit means of all explanatory variables as proxies for regime-case fixed effects. The krls estimator is a machine learning method to fit multidimensional functions for regression and classification problems without relying on linearity or additivity assumptions. These results are reported in columns (2) and (4) in Table B-1, for the binary protest indicator and for the transformed protest count, respectively. In addition, we test a correlated random effects (CRE) probit for the binary protest indicator and a similar negative binomial for count data, as described next. These are reported in columns (1) and (3) of Table B-1, respectively.

Table B-1: Digital repression and protest

	Binary protest		Count protest	
Estimator	CRE- Probit	KRLS	CRE-NegBin	KRLS
	(1)	(2)	(3)	(4)
Digital repression	-1.1352	-0.0478*	-0.4797	-0.0709*
	(0.5876)	(0.016)	(0.4705)	(0.029)
Digital capacity	0.4545	-0.0007	0.1043	0.0003
	(0.5314)	(0.016)	(0.5902)	(0.029)
Online	-0.2805	-0.0088	-0.2796	-0.0219
	(0.2164)	(0.015)	(0.2180)	(0.026)
Time since last protest (log)	-0.2197*	-0.0414*	-0.3448*	-0.1049*
	(0.0911)	(0.017)	(0.1115)	(0.028)
Regime duration	-0.5806	-0.0108	-0.3249	-0.0280
	(0.3234)	(0.010)	(0.2022)	(0.016)
Population size	-0.0487	0.0120*	-0.5031	0.0250*
	(1.3298)	(0.005)	(0.8981)	(0.009)
(Intercept)	0.5909		-2.1278	
	(0.7716)		(0.6482)	
Time trend	✓	✓	✓	✓
Random intercepts	✓		✓	
Unit means	✓	✓	✓	✓

Dependent variable is protest: binary in (1) and (2); raw count (3); and log count (4). NxT=960; 88 regime-cases in 75 countries; 2001-2017. * $p < .05$. Estimate for digital repression in column (1) is statistically significant at the 0.053 level. Cluster-robust errors in (1); bootstrapped errors in (3).

One way to account for unit heterogeneity in a probit model is with random intercepts (RE). This approach assumes that the unit effect is not correlated with the explanatory variables. If this assumption is not met, RE estimates may be biased. A fixed effects estimator with limited dependent variable, however, has drawbacks as well because some regimes (panel units, i) are short-lived and have a small t. Further, protest is absent in a handful of regimes; a fixed-effects estimator will not draw inferences about marginal effects from these regimes. There are many approaches to dealing with this issue (e.g. Mundlak, 1978; Cook, Hays and Franzese, 2018; Beck, 2018). Our preferred estimator follows spirit of the Mundlak-Chamberlain approach, employing the correlated random effects estimator (CRE). Instead of estimating separate intercepts for each panel, we include the unit-means of explanatory variables in the specification (Wooldridge, 2002, 488):

$$Pr(Protest_{i,t} = 1) = \alpha_{j[i]} + \beta_1 D_{i,t} + \delta_1 \bar{D}_i + \beta_2 X_{i,t} + \delta_2 \bar{X}_i + \varepsilon_{i,t};\ \alpha_j \sim N(0, \sigma_\alpha^2) \quad \varepsilon \sim N(0,1) \quad (1)$$

$D_{i,t}$ is the treatment variable; $X_{i,t}$ are time-varying confounders; and $\bar{D}_i$ and $\bar{X}_i$ proxy for fixed unit effects. The estimate of β_1 adjusts for the unit means of all RHS variables for all regimes (panel units) and not just regimes that experience mass uprisings. The marginal effects estimates also draw information from cases where no protest onset has occurred (yet) while still accounting for unobserved time-invariant unit effects. The result reported in column (1) of Table B-1 is this CRE-probit model, where $\alpha_{j[i]}$ are the random intercepts and $\bar{D}_i$ and $\bar{X}_i$ are the unit means. This is a "within" estimator because it isolates variation over time within regimes to draw inferences. The estimate for *Digital repression* is negative and statistically significant at the 0.053 level.[42] We then

[42]In an unreported test, we find that the estimate from a RE-probit is -0.505 (0.111), which is statistically signficant at the 0.001 level. The CRE estimate is therefore substantially larger in absolute size but has higher variance. The

extend this approach to a negative binomial model for count data. We again include unit means as proxies for fixed unit effects and model the dispersion parameter as random variable (fixed for each unit). The results from this model are reported in column (3) of Table B-1. Note here that while the estimate is negative, the estimated standard error is quite large. However, as we show in the next table, this estimate is nearly identical to a standard fixed effects negative binomial estimate.

Table B-2: Digital repression and protest, various estimators

Estimator	Logit	RE-logit	Cond. Logit	OLS	RE-OLS	FE-OLS	FE-NegBin	Within RE-NegBin
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Digital repression	-0.7133*	-0.8732*	-1.7013	-0.1303*	-0.1303*	-0.3461*	-0.4641*	-0.4797
	(0.1431)	(0.1993)	(0.9715)	(0.0264)	(0.0279)	(0.1703)	(0.1222)	(0.4705)
Digital capacity	0.1999	0.2413	0.4453	0.0341	0.0341	0.0915	-0.0770	0.1043
	(0.1543)	(0.2084)	(0.9101)	(0.0273)	(0.0315)	(0.1605)	(0.1455)	(0.5902)
Online	0.0969	0.0832	-0.5408	0.0184	0.0184	-0.0711	0.0714	-0.2796
	(0.0944)	(0.1194)	(0.3690)	(0.0168)	(0.0178)	(0.0630)	(0.0858)	(0.2180)
Time since last	-0.9762*	-0.8025*	-0.3345*	-0.1931*	-0.1931*	-0.0726*	-0.5406*	-0.3448*
protest (log)	(0.1149)	(0.1581)	(0.1560)	(0.0219)	(0.0227)	(0.0314)	(0.0890)	(0.1115)
Regime duration	-0.3501*	-0.3814*	-1.0482	-0.0556*	-0.0556*	-0.1165	-0.0253	-0.3249
	(0.1177)	(0.1415)	(0.8078)	(0.0176)	(0.0185)	(0.0753)	(0.0906)	(0.2022)
Population size	0.3580*	0.4285*	-0.6646	0.0637*	0.0637*	-0.0402	0.2207*	-0.5031
	(0.0805)	(0.1200)	(2.0721)	(0.0138)	(0.0154)	(0.3484)	(0.0691)	(0.8981)
(Intercept)	-3.5548*	-4.6028*		-0.1233	-0.1233		-4.7272*	-2.1278*
	(1.3925)	(2.0018)		(0.2462)	(0.2724)		(1.1673)	(0.6482)
Year effects	✓	✓	✓	✓	✓	✓		
Non-linear time trend							✓	✓
Random intercepts		✓			✓			
Fixed dispersion							✓	
Random dispersion								✓
Unit means								✓

Dependent variable is protest; NxT=960; 88 regime-cases in 75 countries; 2001-2017. * $p < .05$. Estimate for digital repression in column (3) is statistically significant at the 0.080 level. Cluster-robust errors in (1)-(6); bootstrapped errors in (7)-(8).

B-2 reports results from various estimators used to model the binary indicator of protest. We begin with an ordinary logit that does *not* account for unit heterogeneity. We then test a random effects logit and finally a conditional logit. The latter accounts for all cross-section differences between regimes, but drops cases that never experience a protest (roughly 6 percent of regimes). In all three tests, the estimate for *Digital repression* is negative and statistically significant, though only at the 0.080 level for the conditional logit estimate. That said, the estimate size in the conditional logit is much larger (absolutely) than in the other two models. This suggests that, if anything, the random effects estimator is biased towards zero. Next we test the same set of models using linear estimator: OLS, RE-OLS, and FE-OLS. Note that all specifications include year fixed effects to account for common time trends. Again all estimates for *Digital repression* are negative and statistically significant. The marginal effect estimate in the FE model is over twice as large as the estimates in teh OLS and RE-OLS models. Finally, column (7) in Table B-2 reports a FE negative binomial model with count data and bootstrapped standard errors. The estimate for *Digital repression* is negative and statistically significant. Note that the fixed effects in this approach is a fixed dispersion parameter not a "within" transformation of the conditional mean. That said, the within estimator outlined above and shown again in column (8) yields an almost

RE-probit margin effect is half the size of the CRE-probit marginal effect.

identical estimate for *digital repression* but with a much larger variance.

Figure B-2 reports results when adding potential confounders to the specification. Recall that baseline specification includes the following covariates: digital repression, online existence, time since last protest (log), regime duration (log), and population size. We posit that none of these covariates are post-treatment phenomena: digital repression should have no effect on these variables. We show results for two approaches, both linear probability models (RE and 2-way FE) when we include additional potential confounders, one at a time, for each of 20 variables. The top plot shows that the RE estimate is robust to including any one of these variables. The bottom plot shows that the 2-way FE result is robust to including any of 18 variables. The two variables for which the estimate of interest is no long significant at the 0.10 level are: judicial independence and trade. Both estimates, however, are larger in absolute size than the RE estimates.

Finally, Figure B-3 reports results from linear probability models with the lagged dependent variable in the specification. For reference, we first plot the 2-way FE estimate, which is roughly 0.34 and statistically significant. Next, we report the Lag DV estimate (without year effects); it is roughly -0.13 and signficant. Next, we report a lag DV estimate with year fixed effects and the estimate is roughly -0.12. Finally, we report a test with 2-way FE and a lagged DV; the estimate is -0.32 and significant at the 0.056 level. If we believe these two approaches bracket the true effect, then it would fall between roughly -0.3 and -0.1. Note that the RE-OLS estimate is -0.13 and statistically significant.

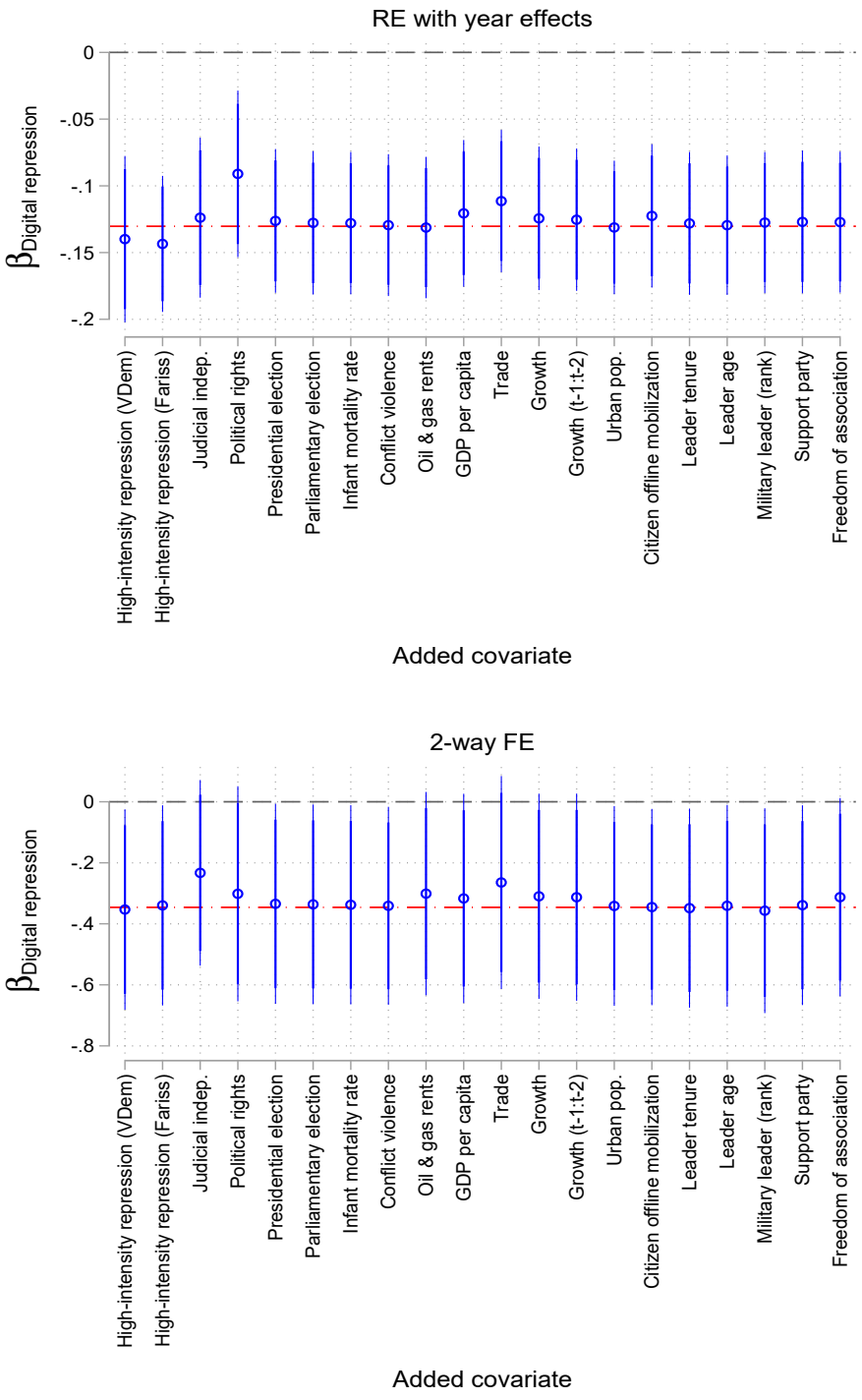

Figure B-2: Digital repression and protest, additional confounders. OLS estimators: RE and 2-way FE. Dependent variable is a binary indicator of protest. Cluster-robust errors.

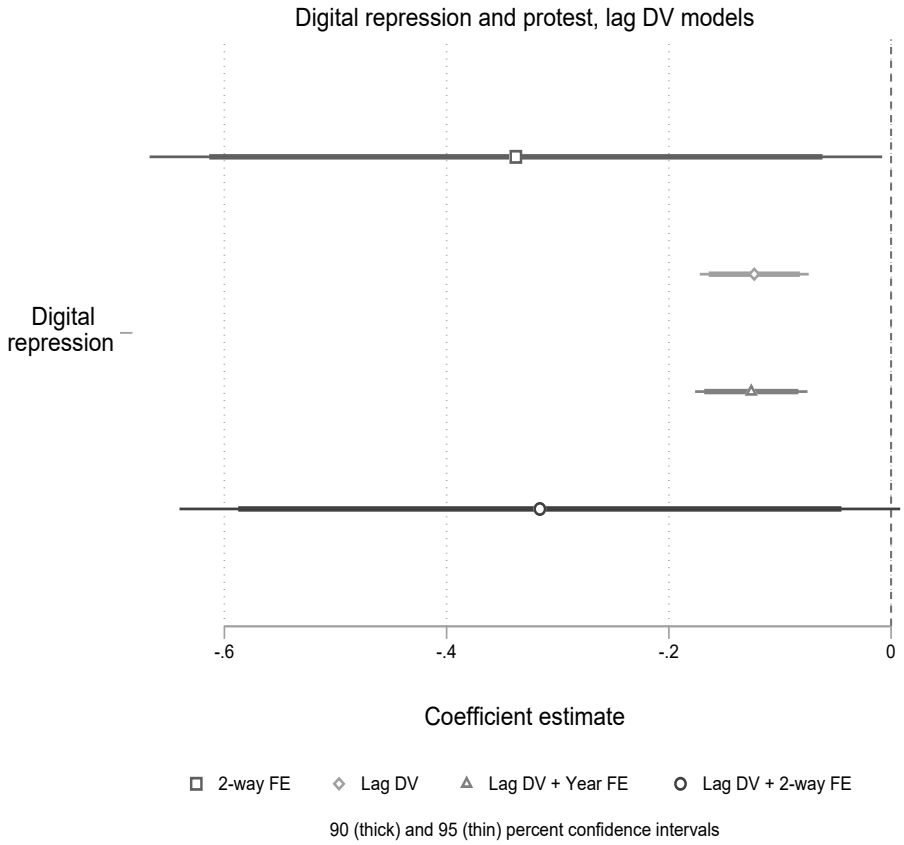

Figure B-3: Digital repression and protest, lagged dependent variable models. OLS estimators. Dependent variable is a binary indicator of protest. Cluster-robust errors.

3 Appendix C: Additional Results for High-Intensity Repression

This section reports additional test for the violent (or "high-intensity" repression) analysis. Figure C-1 shows the average level of high-intensity repression and digital repression in dictatorships during the sample period (2001–2017). The histograms in the background depict the distribution of each variable, both scaled on (0,1) for visual comparison. While there is a slight decline in high-intensity repression during the first part of the period, the average level appears to increase slightly after 2013. In contrast, the average level of digital repression is increasing over time, though not by very much.

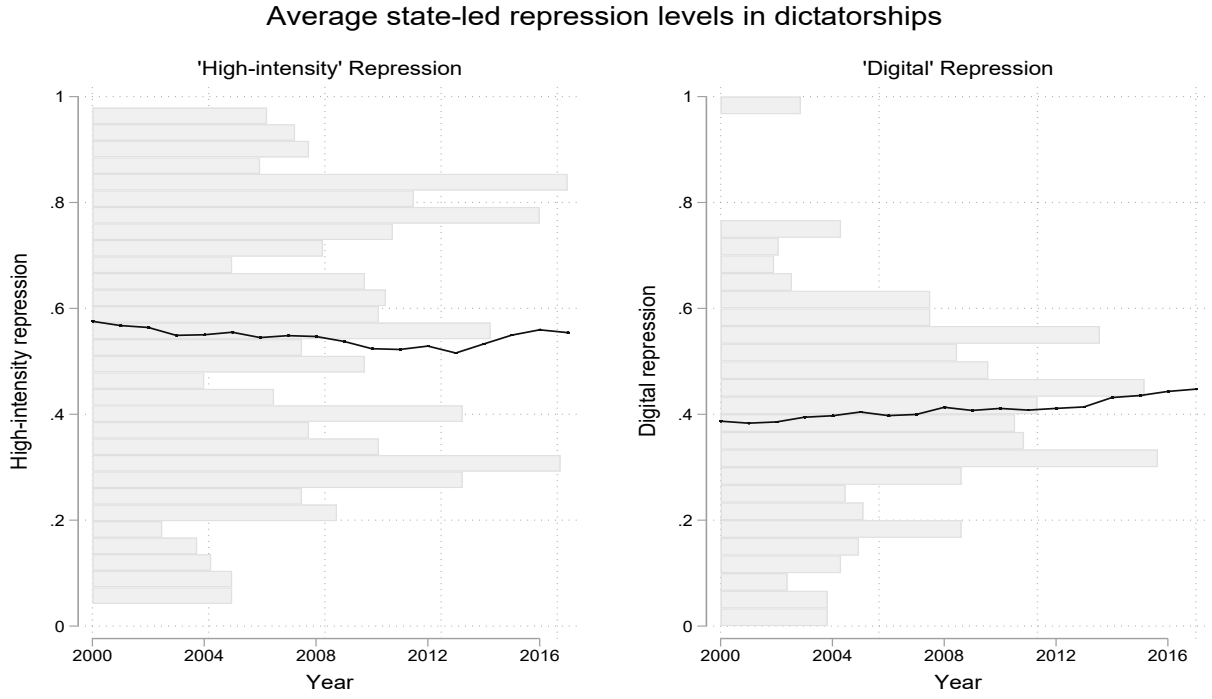

Figure C-1: 'High-intensity' repression and digital repression trends over time

In the analysis we modeled high-intensity repression using difference models because the data have very high serial (or auto-) correlation. A Wooldridge test for panel data serial correlation in levels yields a test statistic of 77, indicating a high level of auto-correlation. In differences, this test statistic is just over 2 and not statistically significant, indicating that differencing the data yields a panel data series that is unlikely to suffer from auto-correlation. The models reported in the main text are therefore linear regressions in differences.

Figure C-2 reports results from tests of error-correction models (ECMs), which allow tests of both short-term and long-run effects in panel data series by including both differences and levels of explanatory variables. We test models with no effects; year effects; unit effects; and both types of effects. The left plot shows estimates for the short-term, year-on-year, marginal effect of *digital repression*: irrespective of the exact model, the estimated short-run effect is positive and statistically significant. The size of these estimates indicates that once we adjust for the lagged level of high-intensity repression (and digital repression), the short-term effect is roughly one-third of the estimate reported in the main text. The right plot shows the estimated long-run effects,

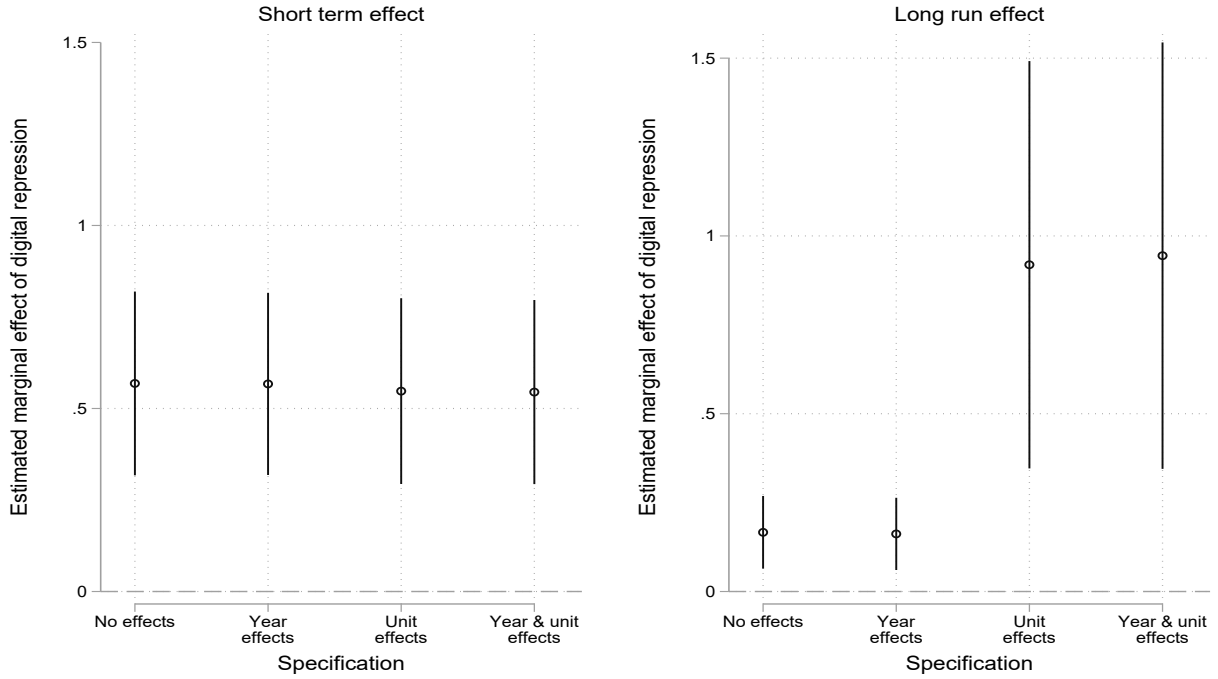

Figure C-2: Error correction models

which are all positive and significant. This suggests that the total long-run effect – which combines both short- and long-term effects – should be positive. Nonetheless, we are skeptical that there should be observed long-run effects because after an initial uptick in high-intensity repression, this tactic should be less necessary if digital repression reduces anti-government mobilization.

Figure C-3 reports results from tests that use alternative measures of high-intensity repression. The first alternative is a variable from Fariss (2014) and Schnakenberg and Fariss (2014), which is a latent estimate of human rights protection that combines information from multiple observed measures of human rights protections and state-led repression.[43] The second measure is from VDem but only contains information on government killings and torture, not other forms of high-intensity repression, such as imprisonment. The latter two measures are derived from State Department human rights reports and Amnesty International reports. Both of these scales are also incorporated into the latent measure from Fariss. To facilitate visual comparison of estimates, we standardize and rescale all of these variables so that larger values reflect more repression; all outcomes are centered at `0` with a standard deviation of `1`. While all the estimates are positive and statistically significant, the size of the estimates varies widely.

Finally, Table C-1 reports results of Granger causality tests, which take the following form, where $D_{i,t} - D_{i,t-1}$ is the differenced treatment and $Y_{i,t} - Y_{i,t-1}$ is the differenced outcome:

$$Y_{i,t} - Y_{i,t-1} = \alpha_0 + \beta(D_{i,t} - D_{i,t-1})_{t-1} + \delta(Y_{i,t} - Y_{i,t-1})_{t-1} + \varepsilon_{i,t} \tag{2}$$

[43] See Cope, Crabtree and Fariss (2020) for more information on how the VDem and Fariss measures compare. We find a positive correlation in levels (0.31) but a very weak correlation in differences (0.05).

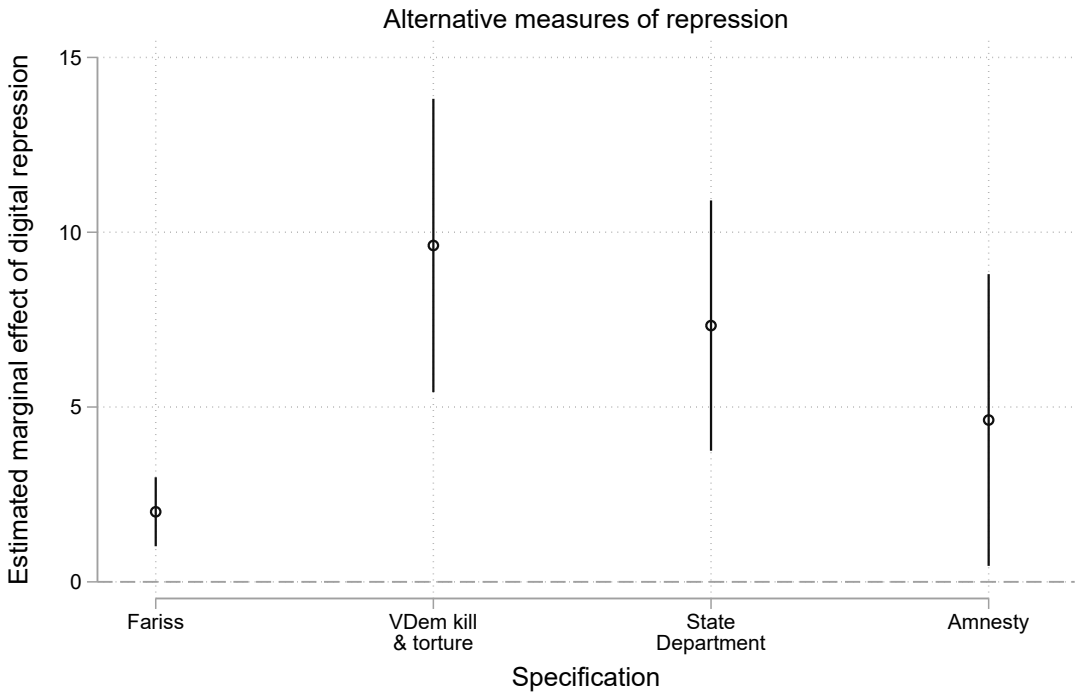

Figure C-3: Alternative measures of 'high-intensity' repression

That is, we regress the lag of the differenced outcome and the lag of the differenced treatment on the differenced outcome. First, we treat violent (or "high-intensity") repression as the outcome and test three linear models: no covariate adjustments; covariate adjustment; and covariate adjustments with FE. Note that the FE model means that the estimator (weighted) averages the 'within' estimates for all the distinct panels. These are reported in columns (1)-(3). Then we estimate the model with covariate adjustment but with a kernel estimator, reported in (4). In each of these four tests, where the lag of the outcome is treated as a conditional mean adjustment, the lag of the treatment ($D.Digital\ repression_{t-1}$) is positive and significant. This suggests that digital repression (in differences) may granger cause high-intensity repression (in differences). Next we repeat this exercise with digital repression as the outcome and high-intensity repression as the treatment ($D.Digital\ repression_{t-1}$ is interpreted as a conditional mean adjustment). The lag of the treatment, in these models, while positive is not significant in any model. This suggests that high-intensity repression (in differences) does not granger cause digital repression (in differences).

Table C-1: Granger causality tests

Outcome variable	*D.High-intensity repression*				*D.Digital repression*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
D.High-intensity repression$_{t-1}$	-0.0415	-0.0406	-0.1289*	0.0005	0.1050	0.1295	0.1275	0.0291
	(0.0630)	(0.0632)	(0.0571)	(0.019)	(0.1526)	(0.1550)	(0.1468)	(0.003)
D.Digital repression$_{t-1}$	0.0672*	0.0650*	0.0688*	0.0213*	-0.0291	-0.0894	-0.1576	0.0071
	(0.0183)	(0.0205)	(0.0216)	(0.006)	(0.1206)	(0.1295)	(0.1145)	(0.010)
Adjust for covariates		✓	✓	✓		✓	✓	✓
Fixed effects			✓				✓	
OLS	✓	✓	✓		✓	✓	✓	
kernel regression				✓				✓

NxT=883; 77 regime-case; 2001-2017. * $p < .05$. Cluster-robust errors in (1)-(3) and (5)-(7).

4 Appendix D: Additional Results for Regime Collapse

This section reports results for models that assess the empirical relationship between digital repression and autocratic regime collapse. There are 37 regime collapse events in the sample period (2001 to 2017) with information on the lagged covariates. However, we treat the the U.S. invasions of Afghanistan (2001) and Iraq (2003) as a right-censored because these are events in which regime collapse is caused by foreign intervention as opposed to domestic political actors ousting the incumbent regime. This leaves 35 regime collapse events from 2001 to 2017.[44] Of these 35 regime collapse events, 26 are coded as democratic transitions and the other 9 as transitions to either a failed state (e.g. Libya 2011 and Yemen 2015) or a new autocracy (e.g. 2008 coup in Guinea).

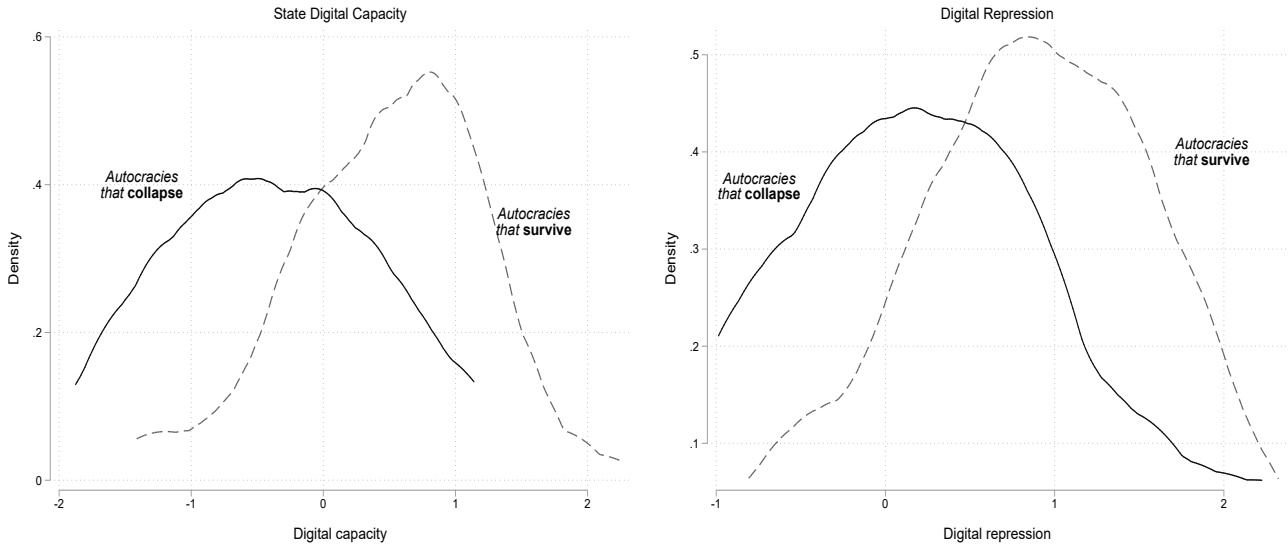

Figure D-1: Digital repression and capacity, by collapsed and surviving regimes

The left plot in Figure D-1 shows the average level of digital capacity for autocracies that survive and those that collapse. Those that survive tend to have higher capacity than those that collapse. The right plot shows the same but for digital repression: autocracies that survive use more digital repression than those that collapse. In broad strokes, the pattern in the left plot would suggest that digital repression may help stabilize autocracies by lowering the risk of regime collapse. But the pattern in the left plot indicates that digital capacity likely confounds this relationship because we know that digital capacity and repression are highly correlated.

Table D-1 reports results from a series of probit and kernel regression models – all with the same baseline specification. Column (1) reports results from a pooled probit. The estimate for *digital repression* is negative but not statistically significant. Note that if the sample size were twice as large and the data patterns were to hold, then the size of the error estimate would be roughly half of the reported value and the estimate for *digital repression* would be statistically significant at conventional levels. Column (2) reports results from a random intercept probit; the estimate for *digital repression* is again negative but not significant. In these two models, estimates for *digital*

[44] Afghanistan 2001 and Taiwan 2000 are also dropped due to missing data on the World Bank's measure of GDP per capita. And all regime collapse events in 2000 are dropped because there is no information on lagged digital repression for that year. Thus Afghanistan 2001 and Taiwan 2000 are dropped from the sample irrespective of missing GDP data.

Table D-1: Digital repression and regime collapse

Estimator	Probit	RE Probit	CRE Probit	KRLS	KRLS
	(1)	(2)	(3)	(4)	(5)
Digital repression	-0.1694	-0.2138	0.7687	-0.0124*	-0.0038
	(0.1594)	(0.2379)	(0.8153)	(0.006)	(0.004)
Digital capacity	-0.3953*	-0.4895	-3.3945	-0.0199*	-0.0086*
	(0.1555)	(0.3328)	(1.9919)	(0.007)	(0.004)
Online	-0.0012	-0.0120	0.2581	-0.0020	0.0019
	(0.0842)	(0.1074)	(0.3798)	(0.005)	(0.003)
Military leader	0.5234*	0.6653	-0.0725	0.0318*	0.0214*
	(0.2223)	(0.5032)	(0.7883)	(0.016)	(0.010)
Supporting political party	-0.3109	-0.3485	-1.1276*	-0.0347*	-0.0327*
	(0.2052)	(0.2763)	(0.4868)	(0.016)	(0.012)
GDP per capita	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(Intercept)	-1.8966*	-2.2406	2.1618		
	(0.5634)	(1.3236)	(1.7513)		
Year effects	✓	✓		✓	
Non-linear time trend			✓		✓
Regime duration polynomials	✓	✓	✓	✓	✓
Random intercepts		✓			
Unit means			✓		✓

Dependent variable is regime collapse; NxT=929; standard errors clustered on 82 regime-cases in 70 countries; 2001-2017. * $p < .05$.

capacity are also negative but only significant in the pooled model in (1). Further, estimates for *military leader* and *supporting political party* are in the expected directions, positive and negative respectively.

Column (3) reports results from a correlated random effects (CRE) probit, where unit means of all explanatory variables are added to the specification as proxies for unit (regime-case) fixed effects (Wooldridge, 2002, 488). The estimate for *digital repression* flips signs and is now positive but is again insignificant. This CRE approach likely over-fits the data given the relatively low number of regime collapse events and the short time span on the sample. For this reason, some coefficient estimates (and error estimates) are extremely large, reflecting the relatively low 'within' variance of many of these variables.

Columns (4) and (5) report results from kernel least squares (KRLS) estimators. Column (4) reports a pooled model and column (5) a model with unit means as proxies for fixed effects. The marginal effects for *digital repression* from these models are reported in the main text. In both models the *average* marginal effect is negative but not statistically significant. Again, if the sample size were twice as large and the data patterns remained the same, the pooled estimate would likely be statistically significant at conventional levels.

The results from Table D-1 offer little support for the expectation that digital repression stabilizes autocracies by lowering the risk of regime collapse. Given the relatively short time series and relative dearth of regime collapse events (35), however, it should not be surprising that many estimates are not statistically significant. If the sample period were from 1990 to 2020 (and the data patterns remained the same), for example, many of the estimates in the pooled models would be significant. That said, the models that adjust for unit effect yield null results. Thus we cannot rule out possible confounding from unobserved regime features is biasing the estimates in the pooled model. Simply put, there are too few regime collapse events in a 17-year panel data set to pick up consistent correlation patterns linking digital repression and regime collapse.

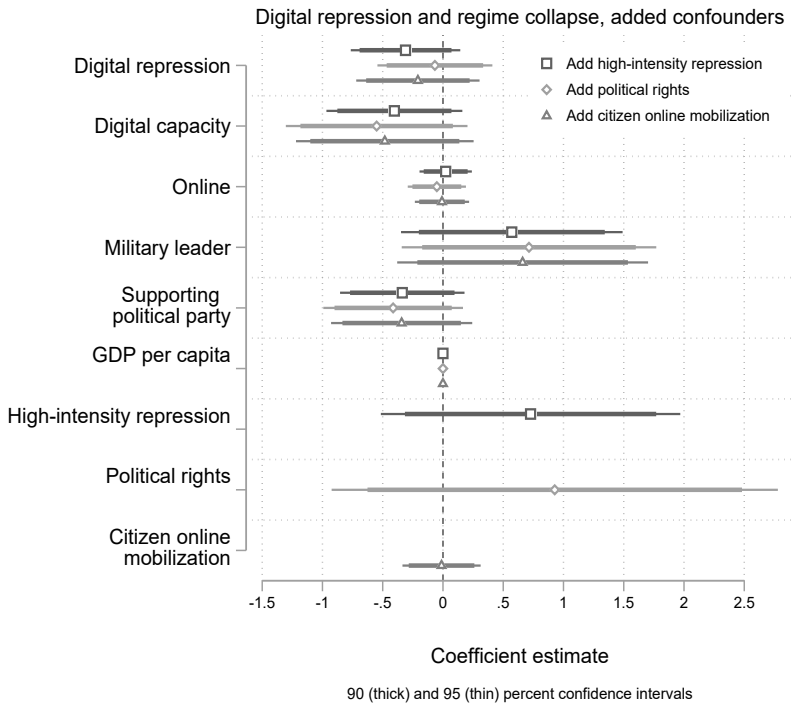

Figure D-2: Digital repression and regime collapse, additional covariates. Random effects probit.

Figure D-2 shows results from RE probit models that add covariates to the specification. The first is a measure of violent (or "high-intensity") repression. This is likely a post-treatment phenomena that often backfires to result in raising the risk of regime collapse. For this reason, adding it to the specification makes the estimate for *digital repression* slightly stronger but still not statistically significant. The second model adds a measure of political rights to the specification. This measure likely incorporates information on *digital repression* and should not be in a specification because it is highly correlated with digital repression and conceptually similar. The final model adds a measure of citizen offline mobilization via social media. Again, this is likely a post-treatment phenomena that is associated with increased risk of regime collapse; including this measure in the specification makes the estimate for *digital repression* slightly stronger but still not statistically significant.

Figure D-3 reports results from models of democratic transition, a subset of all regime collapse events. The results generally show a negative but insignificant association between *digital repression* and democratic transition. Notably, however, the CRE probit result is negative and slightly stronger (in absolute size) than the pooled probit or the RE probit result. These results, while statistically insignificant, point in the same direction and provide some evidence consistent with the contention that digital repression stabilizes autocracies by reducing the risk of democratic transition.

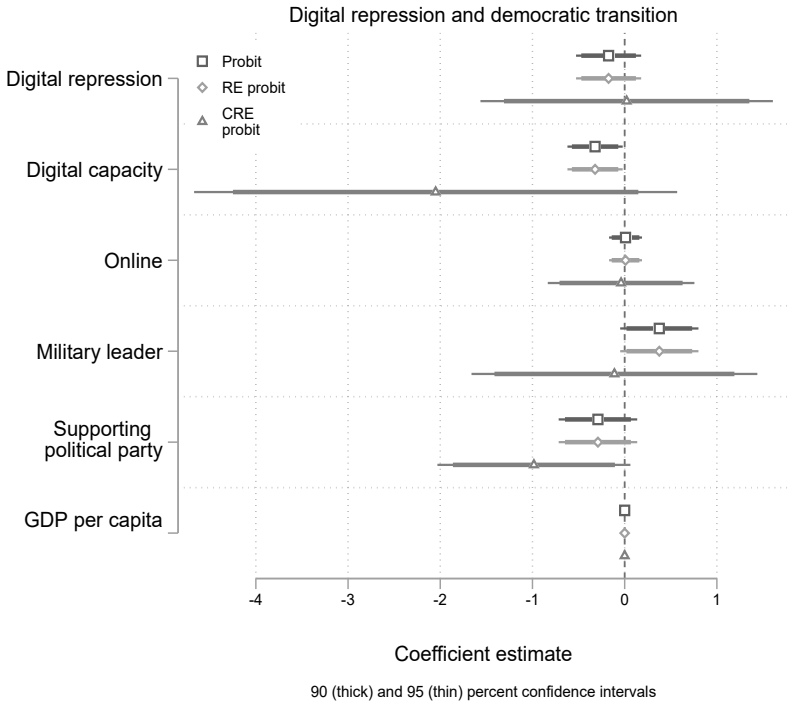

Figure D-3: Digital repression and democratic transition