NBER WORKING PAPER SERIES

SURVEILLANCE OF REPRESSION: THEORY AND IMPLEMENTATION

Veli Andirin
Yusuf Neggers
Mehdi Shadmehr
Jesse M. Shapiro

Working Paper 30167
http://www.nber.org/papers/w30167

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2022, Revised December 2024

This paper supersedes our working papers, "Measuring the Tolerance of the State" and "Real-time Surveillance of Repression." We acknowledge funding from the Data Science Initiative, the Population Studies and Training Center, the Eastman Professorship, and the JP Morgan Chase Research Assistant Program at Brown University; and the Kenneth C. Griffin Economics Research Funds at Harvard University. We are grateful to the United Nations Group of Experts on Geographical Names for sharing their database of country names and official languages with us in a convenient form, to the Global Sanctions Database for answering our questions, to Mohammad Ali Kadivar and Yao Li for sharing data on protests, and to Audrey Sacks for sharing data on commodities. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the funding or data sources. We thank Tariq Ahmad, Isaiah Andrews, Chris Blattman, Dan Björkegren, Ray Fair, Ben Feigenberg, Fred Finan, Terry Fletcher, Bård Harstad, Lihua Lei, David I. Levine, Eduardo Montero, David Yang, audiences at Brown University, Georgetown University, Reichman University (IDC Herzliya), Harvard University, ETH Zurich, the Cowles Foundation, SSE/SITE, Stockholm University, the NBER, UNC Chapel Hill, NYU, the GLPET, the AEA, the University of Chicago, the University of Miami, the Wallis Institute Annual Conference, the Autocracies and Conflict Conference, the Northeastern Universities Development Consortium, the Kelley School of Business, Stanford University, and especially discussants Brian Knight, Jorge Lemus, and Scott Gehlbach for comments and suggestions. We thank our many dedicated research assistants for their contributions to this project. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w30167

Surveillance of Repression: Theory and Implementation
Veli Andirin, Yusuf Neggers, Mehdi Shadmehr, and Jesse M. Shapiro
NBER Working Paper No. 30167
June 2022, Revised December 2024
JEL No. C55, D74

ABSTRACT

A repressive regime can suppress dissent (e.g., protest) after it manifests (e.g., by beating or arresting protesters) or prevent it beforehand (e.g., by obstructing public spaces). We separately study and measure these two forms of repression, focusing on the case of political protest. We show in a model that preventive repression makes protest less predictable. We introduce a new database of security alerts that includes advance warning of protest. We validate the database against objectively measured events and use it to produce global, monthly indices of repression grounded in our model. We show that these indices track salient episodes of repression and predict public expressions of human rights concerns. We find that additional government resources increase prevention but not suppression. We argue that these indices can be a valuable tool for surveillance by the international community and analysis by the research community.

Veli Andirin
Department of Economics
Yale University
New Haven, CT 06511
velimurat.andirin@yale.edu

Yusuf Neggers
Gerald R. Ford School of Public Policy
University of Michigan
735 S. State St. #4216
Ann Arbor, MI 48109
yneggers@umich.edu

Mehdi Shadmehr
Department of Public Policy
University of North Carolina at Chapel Hill
Abernethy Hall, 131 S Columbia St
Chapel Hill, NC 27514
United States
mshadmeh@gmail.com

Jesse M. Shapiro
Department of Economics
Harvard University
Littauer Center
Cambridge, MA 02138
and NBER
jesse_shapiro@fas.harvard.edu

Surveillance of Repression: Theory and Implementation

Veli M. Andirin, *Yale University*[*]
Yusuf Neggers, *University of Michigan*
Mehdi Shadmehr, *University of North Carolina at Chapel Hill*
Jesse M. Shapiro, *Harvard University and NBER*

December 2024

Abstract

A repressive regime can suppress dissent (e.g., protest) after it manifests (e.g., by beating or arresting protesters) or prevent it beforehand (e.g., by obstructing public spaces). We separately study and measure these two forms of repression, focusing on the case of political protest. We show in a model that preventive repression makes protest less predictable. We introduce a new database of security alerts that includes advance warning of protest. We validate the database against objectively measured events and use it to produce global, monthly indices of repression grounded in our model. We show that these indices track salient episodes of repression and predict public expressions of human rights concerns. We find that additional government resources increase prevention but not suppression. We argue that these indices can be a valuable tool for surveillance by the international community and analysis by the research community.

JEL Codes: C55, D74, F51
Keywords: preventive repression, reactive repression, international sanctions

[*]This paper supersedes our working papers, "Measuring the Tolerance of the State" and "Real-time Surveillance of Repression." We acknowledge funding from the Data Science Initiative, the Population Studies and Training Center, the Eastman Professorship, and the JP Morgan Chase Research Assistant Program at Brown University; and the Kenneth C. Griffin Economics Research Funds at Harvard University. We are grateful to the United Nations Group of Experts on Geographical Names for sharing their database of country names and official languages with us in a convenient form, to the Global Sanctions Database for answering our questions, to Mohammad Ali Kadivar and Yao Li for sharing data on protests, and to Audrey Sacks for sharing data on commodities. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the funding or data sources. We thank Tariq Ahmad, Isaiah Andrews, Chris Blattman, Dan Björkegren, Ray Fair, Ben Feigenberg, Fred Finan, Terry Fletcher, Bård Harstad, Lihua Lei, David I. Levine, Eduardo Montero, David Yang, audiences at Brown University, Georgetown University, Reichman University (IDC Herzliya), Harvard University, ETH Zurich, the Cowles Foundation, SSE/SITE, Stockholm University, the NBER, UNC Chapel Hill, NYU, the GLPET, the AEA, the University of Chicago, the University of Miami, the Wallis Institute Annual Conference, the Autocracies and Conflict Conference, the Northeastern Universities Development Consortium, the Kelley School of Business, Stanford University, and especially discussants Brian Knight, Jorge Lemus, and Scott Gehlbach for comments and suggestions. We thank our many dedicated research assistants for their contributions to this project. E-mail: velimurat.andirin@yale.edu, yneggers@umich.edu, mshadmehr@unc.edu, jesse_shapiro@fas.harvard.edu.

1 Introduction

In March 2011, according to international media reports, Zimbabwean authorities mobilized riot police and arrested organizers in response to indications that the opposition planned to march against the government (Mavhunga 2011; see also Human Rights Watch 2011). No such march occurred. Contrast this episode with one in August 2018, when international media published video and photo documentation of Zimbabwean military firing on crowds, and beating individual protesters, following a disputed election (Burke 2018; Bearak 2018; see also Human Rights Watch 2018).

These two episodes illustrate two forms of repression. In the first episode, the regime prevented protest altogether, thus avoiding the need to suppress it publicly. In the second episode, the regime suppressed protest and avoided revolution, but at the cost of public and brutal violence. The first form of repression, which we call *prevention*, is inherently less public, and less visible to those outside the country, than is the second form, which we call *suppression*.

This imbalance affects both the international community and the research community. For the international community, it is practically and politically easier to sanction visible acts of suppression than less visible and more deniable acts of prevention. In the case of the two episodes in Zimbabwe, for example, the US only expressed concern in response to the episode of prevention (Crowley 2011; CNN Wire Staff 2011), whereas it issued new sanctions in response to the episode of suppression (U.S. Department of the Treasury, 2020).[1] Multiple international bodies have noted the differences in scrutiny these two forms of repression bring.[2]

For the research community, a number of cross-national datasets track broad measures of freedom or specific forms of visible conflict between states and citizens. These and other datasets have enabled a large literature studying both individual countries and broad cross-national trends (see, e.g., Lyall 2009; Conrad and Moore 2010; Fariss 2014, 2019; Rozenas et al. 2017; Cingranelli and Filippov 2018; Carey and Gohdes 2021; Arnon, Haschke, and Park 2023), and contributing to our understanding of state-citizen interactions (Tilly 1978, 2006; Fording 2001 ; McAdam, Tarrow, and Tilly 2001; Bueno de Mesquita et al. 2003; Guriev and Treisman 2022). By contrast, there is no systematic ongoing measurement of preventive

[1]For another example, see the case of Cuba in 2021, in which the US imposed sanctions in response to regime security forces violently suppressing protesters in July (U.S. Department of State 2021; U.S. Department of the Treasury 2021a; U.S. Department of the Treasury 2021b), but not to regime security agents and supporters subsequently confining dissidents to their homes to prevent protest in November (Robles 2021).

[2]The International Network of Civil Liberties Organizations notes that, "perhaps because of the fact that its stifling effect on protest is less 'visible' than illegal or excessive use of force, this issue [of legal and administrative limitations on the right to protest] receives less attention in international discussion fora" (2013, p. 63). See also Głowacka et al. (2021, Section 4.2.1) and National Intelligence Council (2022, p. 3).

repression at global scale, and while, as we discuss later, there is a significant body of work studying preventive repression in particular forms and contexts, much less is known about its broad determinants.

In this paper, we separately study and measure prevention and suppression. We show in a model that preventive repression makes protest less predictable. We introduce a new database of security alerts that includes advance warning of protests. We validate the database against objectively measured events, and we use it to develop global, monthly indices of prevention and suppression. We show that these indices track salient episodes of repression and predict the expression of human rights concerns by the international community and the onset of sanctions. And we use the indices to study how greater resources affect both the amount and type of repression favored by regimes.

We begin by laying out our model. In the model, grievances arise stochastically, and an opposition decides whether to plan protest in public, in secret, or not at all. A regime monitors society for signs of protest and decides whether to take action to block it, for example by closing public spaces or shutting communications channels. For the regime, blocking protest carries a cost, for example because it curtails economic activity. For the opposition, secret planning is more costly than public planning, but less likely to be thwarted by the regime.

The equilibrium of the game features a cutoff such that the regime blocks protest whenever it thinks protest is sufficiently likely. As a result, in a highly preventive regime, protest is rarely very likely *ex ante*; i.e., it is hard to predict, because if it were easy to predict, the regime would block it. Or, as the late Zimbabwean political scientist John Makumbe said in the wake of the 2011 march that wasn't, "I am sure it [protest] will happen. When it happens, the state structure will be caught napping" (Mavhunga, 2011).

To check whether protest is predictable in advance, or must instead "catch the regime napping," we introduce a new global database of security alerts from Crisis24, a global risk management firm, covering 2010 through 2019. These alerts are sourced from official channels, local, international, and social media, and event participants, and are curated by teams of analysts. Clients use these alerts to plan travel and other business activities, so the alerts endeavor to give advance warning of events. Protests are therefore frequently flagged ahead of their occurrence, an unusual feature that makes these data especially suited to our purposes.

We validate the security alerts in two main ways. First, we compare to global databases of objectively measured events such as natural disasters. Such events are not our focus, but are an appealing point of comparison because they can be measured objectively (e.g., from weather stations or seismology) and without reporting biases. The security alerts perform well, for example capturing more than 89 percent of severe natural disasters, much larger than the 38 percent captured by the most comparable open-source database we are aware of. Second, we compare the security alerts to curated academic databases of protests produced by area experts. Unsurprisingly, the security alerts omit many protests that experts identify, especially those that are smaller or less prominent. But, the security alerts tend to capture as many or more protests than the other global sources we compare to.

We use the security alerts to produce a monthly, global index of preventive repression. Our main index takes an extremely simple form: it is (one minus) an indicator for whether the alerts flag a protest in advance. In our model, this corresponds to an indicator for whether the opposition plans protest in public, and provides a bound on the regime's cutoff for blocking protest. If the opposition typically plans in secret, this bound may be very slack. We therefore produce an alternative index based on a forecasting model that combines multiple signals, including from search queries, news media, and social media, to predict the onset of protest. Because regimes routinely monitor such sources, it is plausible that if we can use these sources to predict protest, so can the regime. We explore a number of approaches to incorporating this additional information into our index, and find that the resulting improvement relative to the main index is typically small.

We validate our index by comparison to salient episodes of repression. We attempted to find all episodes of martial law or states of emergency during our sample period. We focus on those with a clear timeline and evidence of preventive repression. We find that our index tracks these episodes well from month to month. Importantly, our index often (correctly) indicates the occurrence of prevention even in months when protest does occur, showing that the index does not simply reflect the (non-)occurrence of protest. Also importantly, our index correctly flags periods when repression is known to have continued even after the formal end of martial law or emergency, showing the value of tracking *de facto* rather than *de jure* repression.

We use the security alerts also to produce a monthly global index of suppression, formed by parsing the text of the alerts and looking for indications of violence or arrests. This index is less original than our index of preventive repression, as there are many large-scale databases that track conflict of this kind. The advantage of producing an index of suppression using the security alerts is that it allows us to measure both prevention and suppression using a common platform.

Our model shows the importance of separately measuring prevention and suppression. If protest occurs, the regime chooses whether to suppress protest in order to prevent revolution. The more costly is suppression—for example, due to sanctions or opprobrium from the international community—the more the regime leans on prevention. If the international community finds it easier to sanction suppression than prevention, this can therefore lead to perverse incentives, where regimes use prevention to avoid the gaze of outsiders, or, in the words of a Department of State report about Taiwan in the 1960s, "to avoid any political incident that can be avoided" (quoted in Greitens, 2016, p. 200).

To study the reactions of the international community, we introduce a new database of news alerts by Human Rights Watch, a global nongovernmental organization, and a new extension of the Global Sanctions Database. We find that mentions of human rights concerns in Crisis24 protest alerts, and mentions of these and other signs of preventive actions by the regime in Human Rights Watch alerts, are positively correlated with our index of preventiveness. Consistent with the greater visibility of suppression, we find that Human Rights Watch alerts mentioning suppression are even more strongly correlated with our measure of suppression, and that only suppression, and not prevention, predicts the onset of sanctions.

Lastly, we study the effect of economic resources on repression. We find that preventive repression increases in response to increases in government resources, but not in response to general increases in economic resources. Following a large literature, we use variation in commodity export prices to isolate variation in resources that is both plausibly exogenous to domestic events and especially amenable to extraction by the regime. We find strong evidence that commodity export price improvements increase prevention. This pattern is particularly sharp at the monthly level, highlighting the value of subannual measurement. And we find no similar pattern for suppression, highlighting the value of unbundling these two forms of repression. Interestingly, we find suggestive evidence that the effect of commodity prices is larger in regimes that face a stronger opposition.

Because a preannounced protest can only occur if there is protest, our monthly index of preventive repression is mechanically related to the occurrence of protest. We show direct evidence that the patterns we find are not driven by this mechanical relationship. In our analysis of episodes of martial law or emergency, we find that our index often (correctly) indicates high preventiveness even in months when protest does occur. In our analysis of the reactions of the international community, we find that preventiveness is correlated with a measure of the expression of human rights concerns that considers only those concerns raised in the context of protest alerts. In our analysis of the role of government resources, we show that the connection between commodity export prices and prevention is robust to controlling for the occurrence of protest and to excluding countries in which protest is especially rare.

The remainder of this section discusses connections to the literature. Section 2 presents the model and formal results, including on the identification of preventiveness. Section 3 describes our data and evidence on its validity, and introduces our indices of repression. Section 4 applies our indices to surveillance by the international community. Section 5 applies our indices to studying the effect of economic resources on repression.

Relationship to the Literature

Our main contribution is to show how to measure preventive repression and to study its determinants at global scale. The building blocks of our approach contribute to research on the theory and measurement of protest and repression.

We develop a new model of protest and repression that shows how preventive repression can be inferred from the predictability of protest. A large prior literature, reviewed for example in Gehlbach, Sonin, and Svolik (2016), studies the dynamics of protest, dissent, and repression, especially in autocracies (see also Davenport 2007; Earl 2011; Acemoglu, Egorov, and Sonin 2015; Davenport et al. 2019; Cantoni et al. 2024). Within this literature there are relatively few papers featuring a dynamic equilibrium model with protest prevention. Recent exceptions include De Jaegher and Hoyer's (2019) analysis of the backlash effects of preventing dissidents from participating in protest, and Gibilisco's (2021) analysis of center-periphery relations.[3] The distinction we study between public and private coordination of protest relates to work by Chau et al. (2023), and others, on the influence of communication technology on protest coordination. To our knowledge, the idea that sanctioning suppression can encourage prevention has not previously been formalized within the large literature on the political economy of sanctions (see, e.g., Baliga and Sjöström 2023).

We introduce a new database of security alerts that features advance notice of protests. The closest public counterpart that we are aware of is a database of diplomatic postings that we discuss in more detail in the paper. Although our main index of preventiveness uses only whether alerts flag protests in advance, our alternative index uses a forecasting model that connects our work to an active literature on the prediction of civil unrest (e.g., Ramakrishnan et al. 2014, Hoegh et al. 2015, Hoegh, Ferreira, and Leman 2016, Qiao et al. 2017, Wu and Gerber 2018).[4] Much of this literature focuses on the predictive task itself, whereas we use the prediction of protest as an input to measuring preventive repressiveness.[5] Kuran (1991) studies the predictability of revolution; see also Langørgen (2016) and Sonin and Wright (2023).

We use the security alerts to develop a global, monthly index of preventive repression, which to our knowledge is the first of its kind. Fariss (2014) measures trends over time in particular forms of repression such as killing by the state (see also Guriev and Treisman 2022).[6] Recent empirical studies of

[3] Aktan (2022a) studies a model featuring asymmetric information and both prevention and suppression of dissent in a stage game that is not repeated (see also Dragu and Przeworski 2019; Aktan 2022b).

[4] See also Bagozzi, Chatterjee, and Mukherjee (2019), Deng, Rangwala, and Ning (2019), Hoegh (2019), and Ross et al. (2019). Mueller and Rauh (2018), among others, study prediction of related outcomes such as armed conflict. Zhao (2022) provides a survey.

[5] For recent work on the measurement of social unrest and its relationship to the macroeconomy, see, e.g., Barrett et al. (2022).

[6] A large literature studies the determinants of these and other repressive acts (e.g., Franklin 2008; Licht and Allen 2018;

preventive repression include Truex's (2019) study of dissident arrests, Carter and Carter's (2022) study of propaganda-based threats in China, and Esberg's (2021) study of the treatment of politicians in Pinochet's Chile. These studies focus on specific, measurable forms of preventive repression in a particular context. By contrast, our index is global, and allows that prevention may take many forms. Several academic and nongovernmental projects publish annual indices of freedom based on expert ratings, but these are typically updated no more than annually, and do not distinguish between prevention and suppression.[7] Such measures nevertheless play a prominent role in public policy, suggesting the value of improving them (Kelley and Simmons, 2015).[8]

Among cross-national studies of preventive repression, Danneman and Ritter (2014) show in a country-year panel that human rights abuses in a given country increase in response to civil war in neighboring ones, and Carey et al. (2022) show in a country-year panel that a news-based measure of disruptive actions by the state increases following oil discoveries in low-capacity states. Ritter and Conrad (2016) show in a province-day panel for Africa that responsive repression (similar to our concept of suppression) increases following mobilization of dissent, and that this effect depends on the strength of democracy in the country (a proxy for preventive repression). These studies all measure repression either with broad indices or with data on specific reported acts, and none of them measures preventive repression at the subannual level.[9]

2 Model of Protest and Repression

In this section we lay out a game of protest and repression that plays out in a few stages, for example over a single day. We begin with a game that features only preventive repression, and then add suppression in a second step. Supplemental Appendix S.A.1 generalizes the model to a dynamic setting with time-varying primitives, and relaxes many of the simplifying assumptions on primitives that we adopt here.

Scharpf et al. 2023).

[7]For indices related to repression of protest, see, for example, the *Freedom in the World* report's Freedom of Assembly index (Freedom House 2021c), the Varieties of Democracy Project's Freedom of Peaceful Assembly index (Pemstein et al. 2023), and the Cingranelli-Richards (CIRI) Human Rights Data Project's Freedom of Assembly and Association index (Cingranelli, Richards, and Clay 2014b). Some indices use survey data to measure freedom at a point in time (e.g., Logan and Mattes 2012; Pickel, Breustedt, and Smolka 2016). Some indices incorporate information on *de jure* freedoms including those guaranteed by constitutions (e.g., Merkel et al. 2020). There is mixed evidence on whether such guarantees are associated with greater *de facto* freedoms (see, e.g., Keith and Poe, 2004; Keith, Tate, and Poe 2009; Chilton and Versteeg 2015). See also the typology of human rights measurement in Landman (2004).

[8]The US Millennium Challenge Corporation incorporates Freedom House's indices into its criteria for determining a country's eligibility for assistance (Millennium Challenge Corporation 2020). Canada's Country Indicators for Foreign Policy project integrates Freedom House indicators into data aimed at providing guidance to development-agency staff (Carment 2010). The Open Government Partnership Global Report cites Freedom House data in the context of identifying potential areas for future work and improvement (Open Government Partnership 2019, pp. 72, 78, 96).

[9]Gueorguiev (2017) and Esberg (2021), among others, consider the strategic distinction between private and public repression. For other work on the empirical dynamics of protest, dissent, and repression, see, for example, Moore (1998), Carey (2006, 2009), and Archibong, Moerenhout, and Osabuohien (2022).

2.1 Preventive Repression

In the game, nature first determines whether citizens are aggrieved, which we capture by an indicator $\omega \in \{0,1\}$, and which we may think of as reflecting factors such as the state of the economy (e.g., Tilly 1975, 1995; The Economist 2022a,b) or the progress of a foreign war (e.g., Lewis 2003; Ascher 2004; Adamovsky 2024).

If citizens are aggrieved they would like to protest, but only if others will protest with them. We assume that only the opposition, which is informed of ω, can coordinate protest. Supplemental Appendix S.A.2 microfounds that assumption as the outcome of a game in which the opposition communicates the location of protest to citizens (see also Chau et al. 2023).

The opposition can coordinate protest in one of two ways. The first way, denoted by $a \in \{0,1\}$, is by making a public announcement, for example via broadcast media or a public social media feed. The second way, denoted by $m \in \{0,1\}$, is by making a secret plan, for example using private messaging platforms (e.g., Jordan 2006; Ketchley 2017, p. 154) or by planting activists at a public gathering (e.g., Muradiniya 1998, pp. 216-231). Making a public announcement is costless to the opposition. By contrast, making a secret plan entails costly effort to avoid detection, such as coordinating across multiple locations (e.g., Muradiniya 1998, pp. 208-11) or restricting modes of communication (e.g., Ketchley 2017, p. 154). We denote the cost of secret planning by $\mu > 0$, and assume that μ is privately known to the opposition and distributed according to a commonly known continuous distribution with full support on $(0,\infty)$. For simplicity, we assume that when indifferent the opposition does not coordinate protest.[10]

After the opposition makes its choice, the regime learns some information $x \in \mathscr{X}$, and must then decide whether to take a costly action to prevent protest, where we denote the cost by $\rho > 0$ and the action by $r \in \{0,1\}$. We may think of the action as blocking public spaces (e.g., Alimagham 2020, pp. 92-93), canceling events (e.g., Leung 2019; Kryzhanouski 2022), forcing an internet outage (e.g., Freedom House, 2017, p. 168; Earl et al. 2022; Bajoria 2023), or deploying police to potential hotspots (e.g., Ketchley 2017, Ch. 6). These actions are costly to the regime because they disrupt economic activity and require the application of personnel or other resources. We assume that, when indifferent, the regime does not prevent protest. For now, we also assume that prevention always works; the more general dynamic model allows that it may fail (e.g., Ketchley 2017).

We assume that the regime's information concerns three things. First, the regime is at least partially

[10]Of course, protests are sometimes sparked by spontaneous events—the beating, arrest, or killing of an innocent citizen (e.g., Mahsa Amini in Iran), a public display of deviance (e.g., the self-immolation of Mohamed Bouazizi in Tunisia)—that do not require costly planning (see, e.g., Kuran 1989; Kuran 1991; Alimagham 2020). We will think of these cases as ones with a very low cost of coordination, and so will not model them separately.

informed of grievances ω, in the sense that if citizens are aggrieved, there is a positive probability that the regime knows this with certainty, for example from its monitoring of public social media channels (e.g., Gunitsky 2015). Second, the regime is fully informed of any public announcement, i.e., of $a \in \{0,1\}$. Third, the regime receives clues about the opposition's secret plans m, for example through infiltration or digital surveillance of public and private channels (e.g., Earl 2011; Gunitsky 2015; Jones 2015; Kadivar 2015; Greitens 2016; Sullivan 2016; Nikolayenko 2017; Freedom House 2019b; Jones 2020; Nugent 2020; Earl et al. 2022; Beraja et al. 2023; Freedom House 2023).

We formalize these assumptions as follows. First, there is a nonempty subset $\mathscr{X}_\omega$ of the sample space $\mathscr{X}$ that is reached with strictly positive probability if and only if $\omega = 1$, regardless of the opposition's choices. Second, the sample space $\mathscr{X}$ is partitioned according to whether there is a public announcement a. Third, when $\omega = 1$ and $a = 0$, the support of x does not depend on m, so that we can write $\lambda(x) \in (0,\infty)$ as the likelihood ratio of the information on the opposition's plans, viz., its density if the opposition has made a secret plan, divided by its density if the opposition has not made a secret plan.[11] We assume that the regime's information is rich in the sense that the image of the support of x in $\lambda(\cdot)$ on $\mathscr{X}_\omega$ is $(0,\infty)$, so that some realizations of x strongly indicate the presence of a secret plan, and some strongly indicate its absence.[12]

If protest occurs, the regime receives a payoff of $-L^* < 0$ and the opposition a payoff of $B^* > 0$. Later we will endogenize these objects as continuation payoffs reflecting the possibility of revolution, and other future events. For now we take them as given.

We solve the game via backwards induction. Begin with the regime's strategy. If citizens are aggrieved and the opposition coordinates protest, then protest occurs if and only if the regime does not prevent it. The regime therefore prevents protest if (and only if) protest is sufficiently likely. To state this more precisely, let $\hat{p}(x) \in [0,1]$ be the probability the regime assigns to protest occurring, absent prevention, given the regime's information $x \in \mathscr{X}$.[13] Then in any equilibrium, the regime prevents protest if and only if $\hat{p}(x) > \overline{p}^* = \min\{\rho/L^*, 1\}$.

Next consider the opposition's strategy. If the citizens are not aggrieved, then the opposition does not coordinate protest. So suppose that citizens are aggrieved, $\omega = 1$. If $\overline{p}^* \geq 1$ then prevention never occurs and the opposition makes a public announcement, $a^* = 1$. So suppose instead that $0 < \overline{p}^* < 1$. In this case, the opposition does not make a public announcement, $a^* = 0$, but may make a secret plan. Specifically,

[11] That is, $\lambda(x) = \Lambda(x|m=1,\omega=1,a=0)/\Lambda(x|m=0,\omega=1,a=0)$ for $\Lambda(x|m,\omega,a)$ the conditional density of x, which we assume to exist for all $m,\omega,a \in \{0,1\}$.

[12] As a technical condition we also assume that the distribution of the random variable $\Pr(m=1,\omega=1|x,a=0)$ has no mass points given any non-degenerate prior on m.

[13] That is, $\hat{p}(x) = \Pr(\omega(a+m)|x)$.

let $\hat{\mu} = B^* \Pr(\hat{p}(x) \le \overline{p}^* | m = 1, \omega = 1)$ be the expected benefit of secretly planning to coordinate protest. Then the opposition makes a secret plan if and only if $\mu < \hat{\mu}$.

Because a higher threshold $\hat{\mu}$ increases the incentive for the regime to prevent protest, which in turn dampens the incentive to plan protest, our assumptions suffice to ensure the existence of an equilibrium threshold $\overline{\mu}^*$, and the uniqueness of the outcome. The Appendix proves the following.

Proposition 1. (Cutoff for prevention) *There is an equilibrium. In any equilibrium, there is a function* $p^* : \mathscr{X} \to [0,1]$*, and a cutoff* $\overline{p}^* \in (0,1]$*, such that if the regime does not prevent protest, protest occurs with probability* $p^*(x)$ *with respect to the regime's information set, and the regime prevents protest if and only if* $p^*(x) > \overline{p}^*$.

2.2 Identification of Preventiveness

In the model, the possibility of prevention affects the likelihood of protest even when prevention is not observed. As Greitens (2016, pp. 43-4) writes, "If protesters or opposition activists learn that they are virtually guaranteed to be detected in advance and prevented from acting, they will be deterred from organizing and therefore no state response will be required" (see also Li 2018, p. 25). We therefore focus on measuring the regime's willingness to prevent protest, summarized by its *strategy* $\overline{p}^*$, rather than the occurrence of acts of prevention, summarized by the regime's *action* r.

Specifically, we define the value $1 - \overline{p}^*$ as the regime's *preventiveness*. The preventiveness can be represented as a direct function of model primitives, $1 - \overline{p}^* = 1 - \min\{\rho / L^*, 1\}$. Because it is difficult to measure primitives such as the cost of prevention ρ, we focus on direct identification of preventiveness. In focusing on direct recovery of player strategies in a game, we follow a tradition in industrial organization (e.g., Pakes et al., 2007). In contrast to that literature, which often uses recovered policy functions to infer other primitives (though see Benkard et al. 2018), our interest is in preventiveness itself.

We imagine that an econometrician observes the occurrence of protest, $z \in \{0,1\}$, and some function $\chi(\cdot)$ of the regime's information x. When the function $\chi(\cdot)$ is one-to-one, the econometrician has, in effect, the same information as the regime; when the function $\chi(\cdot)$ is many-to-one, the econometrician observes a coarsening of the regime's information, akin to observing the regime's information with noise. We do not assume that the econometrician knows the function $\chi(\cdot)$; even when it is one-to-one, the econometrician may not be able to recover the information as it was received by the regime.

We imagine that the econometrician observes the data $(z, \chi(x))$ over repeated play of the game, such that the econometrician learns the joint distribution of $(z, \chi(x))$, and hence the marginal distribution Φ^* of $\Pr(z|\chi(x))$. Treating the population distribution as "data" allows us to separate the analysis of identification from the analysis of estimation (Lewbel 2019). We return later to our approach to estimation.

Our approach to identification depends on an object $\overline{p} = \inf\{p \in [0,1] : \Phi^*(p) = 1\}$, which is the largest value that $\Pr(z|\chi(x))$ reaches in equilibrium. If $\chi(\cdot)$ is one-to-one, then $\Pr(z|\chi(x)) = \Pr(z|x)$. In that case, because we have assumed that the signal structure is rich, it follows by Proposition 1 that $\overline{p} = \overline{p}^*$. More generally, if $\chi(\cdot)$ is many-to-one, then $\Pr(z|\chi(x))$ averages over several values of $\Pr(z|x)$. It follows from the law of total expectation that $\overline{p} \leq \overline{p}^*$. Intuitively, if the probability of protest is sometimes very high, then the equilibrium threshold for prevention, $\overline{p}^*$, must also be high. The following proposition, whose proof is completed in the Appendix, summarizes these arguments.

Proposition 2. (Identification of preventiveness) *In any equilibrium, the preventiveness* $1-\overline{p}^*$ *is partially identified from the marginal distribution* Φ^* *of* $\Pr(z|\chi(x))$*, and is point identified if* $\chi(\cdot)$ *is one-to-one. More specifically,* $1-\overline{p}^* \leq 1-\overline{p}$*, for* $\overline{p} = \inf\{p \in [0,1] : \Phi^*(p) = 1\}$*, with equality when* $\chi(\cdot)$ *is one-to-one.*

The reasoning behind Proposition 2 shows why information on the *distribution* of $\Pr(z|\chi(x))$ is important for identification. Suppose instead that the econometrician observes only the marginal distribution of $z \in \{0,1\}$, which can be summarized by $\Pr(z=1)$. Consider an example where $\Pr(z=1) = 0.5$, so that protest is observed half of the time the game is played. This situation might be one in which preventiveness is high, so that the regime will prevent if the probability of protest is much above 0.5, and where on most days the opposition nevertheless attempts to plan in secret, perhaps because the cost of planning is often low. The situation might also be one in which preventiveness is low, so that the regime will prevent only if the probability of protest is nearly 1, but the opposition plans only about half of the time, for example because the rest of the time planning is too costly. This example illustrates why data on $\Pr(z=1)$ alone is not sufficient to identify the regime's preventiveness.

2.3 Suppression and the International Community

To incorporate the possibility of suppression, we now add an additional stage to the game in which protest may escalate to revolution if it is left unchecked. Given our focus on preventive repression, we keep this stage simple.

If protest begins, the regime observes information $y \in \mathscr{Y}$ that indicates the likelihood that protest may escalate into revolution. This information may concern, for example, the weather (e.g., Collins and Margo 2007) or the reaction of foreign powers (e.g., Milani 2011, pp. 391-392). We also subsume information about any actions by the opposition to foment revolution into the regime's information $y \in \mathscr{Y}$, and do not explicitly model the opposition's information and actions. We assume that the regime's information y is independent of its information x in the previous stage, that the regime's prior probability of revolution given protest is interior, and that the image of y in the likelihood ratio for revolution (the analogue of $\lambda(\cdot)$) is $(0,\infty)$.

After receiving the information $y \in \mathscr{Y}$, the regime decides whether to suppress protest, denoted by $s \in \{0,1\}$, for example through arrests or violence (e.g., Ritter and Conrad 2016). We assume that suppression has a cost $\sigma > 0$ to the regime and allow that it has a cost α to the opposition, but we do not require that α is nonzero, or even that it is positive. For simplicity we suppose that when indifferent the regime does not suppress protest. For now, we also suppose that suppression prevents revolution with certainty; our more general dynamic model relaxes this assumption.

Revolution entails a loss $-L < 0$ to the regime and a benefit $B > 0$ to the opposition, where L exceeds the cost σ of suppression for the regime, and B exceeds any cost α of suppression to the opposition, reflecting the large stakes associated with revolution and regime change (e.g., Abrahamian 1999; Cole and McQuinn 2015; Fitzpatrick 2017).

The equilibrium of this stage of the game is analogous to that of the preceding stage. Let $q^* : \mathscr{Y} \to [0,1]$ describe the regime's belief about the probability of revolution absent suppression. Then in any equilibrium there is some $\overline{q}^* \in (0,1]$ such that the regime suppresses if and only if $q^*(y) > \overline{q}^*$. Any equilibrium also features some *ex ante* probability $S^* \in [0,1)$ of suppression and probability $Q^* \in (0,1)$ of revolution, given protest.[14]

We can think of the payoffs (L^*, B^*) in the previous stage game of protest and prevention as continuation payoffs that reflect the outcome of the stage game of suppression and revolution. To see this, let $L^* = Q^* L + \sigma S^* > 0$ denote the expected loss to the regime given protest, taking account of the equilibrium probabilities of revolution and suppression, and let $B^* = Q^* B - \alpha S^*$ denote the expected benefit to the opposition, defined analogously. If $B^* > 0$, the characterization in Proposition 1 is unchanged by the addition of the stage we define here.

Incorporating this stage of the game allows us to characterize the effect of sanctions on repression. We can think of a change in the likelihood of sanctions as a change in the cost of repression. Increasing the cost of suppression σ makes it more desirable for the regime to prevent protest. Likewise, increasing the cost of prevention ρ makes it harder to prevent protest, increasing the likelihood that the regime resorts to suppression. To state this result formally, let Z^* be the equilibrium probability of protest.[15] The Appendix proves the following.

Proposition 3. (Substitution in repression) *The equilibrium preventiveness* $1 - \overline{p}^*$ *is increasing in the cost*

[14] Specifically, $\overline{q}^* = \min\{\sigma/L, 1\} \in (0,1]$, $S^* = \Pr(q^*(y) > \overline{q}^*) \in [0,1)$, and $Q^* = (1 - S^*)\mathrm{E}(q^*(y) | q^*(y) \leq \overline{q}^*) \in (0,1)$.

[15] If $\overline{p}^* = 1$, then $Z^* = \Pr(\omega = 1)$. Otherwise, $Z^* = \Pr(\omega = 1)\Pr(\mu < \overline{\mu}^*)\Pr(p^*(x) \leq \overline{p}^* | m = 1, \omega = 1, a = 0)$.

of suppression σ, strictly so whenever equilibrium cutoffs $\overline{p}^*$, $\overline{q}^*$ are interior. Likewise, the equilibrium probability of suppression Z^*S^* is increasing in the cost of prevention ρ, strictly so whenever the cutoffs $\overline{p}^*$, $\overline{q}^*$, $\overline{\mu}^*$ are interior.

If sanctions focus too much on one form of repression (say, suppression), they risk encouraging the other form (prevention). There are two reasons to expect that international sanctions will focus especially on suppression. First, suppression is more visible to outsiders, something that regimes are keenly aware of (see, e.g., International Network of Civil Liberties Organizations 2013; Greitens 2022). Second, suppression is less deniable: it may be easier for the international community to act in the face of, say, video footage of military shooting protesters, than in the face of, say, the mysterious disappearance of organizers, or restrictions on public gatherings imposed under a pretense of public health, even if in the latter cases the repressive intent is clear.[16]

3 Measuring Protest and Repression

In this section we introduce our main data source, explain its origins, present evidence on its validity, and discuss how we use it to measure protest and repression.

3.1 A Database of Security Alerts

Our main data source is the full text of security alerts from Crisis24 (2022), a global risk management firm.[17] Crisis24 analysts produced these alerts as a service available to clients interested, among other things, in events impacting travel and other business activity. The alerts concern a range of topics including crime, health, natural disasters, disruptions to transportation, conflict, terrorism, and protest or unrest. Events range from the relatively commonplace (e.g., cancellation of ferry services due to a storm, a string of ATM thefts) to the more extreme (e.g., a terrorist attack triggering a nationwide lockdown, a major hurricane). For each alert, the database includes the country, a short title and longer description, and fields indicating the date and time of the alert.

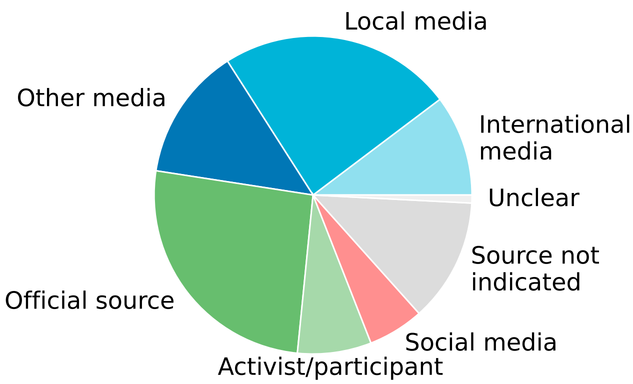
Analysts produced alerts using a range of public and proprietary sources. Analysts typically required at least two "reliable, independent" sources before creating an alert, and avoided sourcing information from social media accounts other than official governmental accounts or accounts of "established labor organizations, activist groups, or corporations reporting on activities they are planning themselves."[18]

[16] Regarding the effect of ambiguity on the enforcement of international norms, see, e.g., Donno (2013, p. 177).

[17] During our sample period, the alerts were produced by Drum Cussac as an independent entity. Drum Cussac was acquired by GardaWorld in November 2019 and combined with GardaWorld Travel Security, NYA, and FAM International Security under the new Crisis24 brand in 2020 (ScaleUp Capital 2019; GardaWorld 2024).

[18] Crisis24 documentation provided via e-mail, March 22, 2024. Other social media accounts were sometimes used as an initial signal leading to additional effort to verify the reported information from reliable sources.

Figure 1: Sources for Security Alerts



Notes: The pie chart shows the frequency of cited data sources in alerts from Crisis24. For each alert, we extract common phrases that indicate an information source along with the appropriate surrounding text, such as "according to...," "...reported that," or "reports from..." We use language cues to classify the source according to the categories in the pie chart, or to classify the source as not indicated by the alert text. Cases labeled as "unclear" are those where our automated parsing procedure could not assign a source and could not determine that a source was not indicated.

Figure 1 shows a breakdown of the sources based on our parsing of the alert text. The most common sources for alerts are official sources and local media, followed by other media, activists/participants, and social media.[19]

We parsed the alerts to identify events of interest using a set of rules described in more detail in Supplemental Appendix S.B. We focus on the period 2010 through 2019, during which alert generation was global in scope, and restrict attention to the 150 countries with a population of at least 1,000,000 in 2010 and for which there is at least one year in the sample period with at least 10 alerts.[20] For the applications of our measures of repression in the following sections, we further restrict attention to non-OECD countries.[21]

To validate the alerts we compared their coverage of natural disasters to that of the Global Disaster Alert and Coordination System (GDACS 2024), a collaboration between the United Nations and the European Commission that provides real-time information on earthquakes, floods, cyclones, volcanoes, droughts, and forest fires. The GDACS is an appealing point of comparison because it includes information from remote sensing and other objective sources, and aims to be comprehensive.

The GDACS rates events in severity from green to red. Less severe events include those affecting less populated areas or with less potential for harm. We expect the security alerts will omit many of these less severe events. As a benchmark for what is a reasonable inclusion rate, we compare to two other sources. One is the database of travel advisories from the US Department of State, Embassies, and Consulates

[19]These findings align with published documentation (Drum Cussac 2018) and our discussions with Crisis24 staff (Crisis24 documentation provided via e-mail, March 22, 2024; and telephone conversation with Crisis24 staff, March 29, 2024).

[20]We define a country to correspond to the definition in the ISO 3166 (International Organization for Standardization 2021). We obtain data on 2010 population from the World Bank (2024). For Taiwan, we obtain population data from National Statistics, Republic of China (2010). Supplemental Appendix S.B shows results where we increase the population threshold to 2,000,000 as well as results where we increase the alerts threshold to 20.

[21]We define OECD membership as of the year 2010 using information from the OECD (2023).

maintained by the US Overseas Advisory Council (OSAC 2024). This database has a similar structure to our main database of security alerts, but is produced independently and is in the public domain. The other is the Global Database of Events, Language and Tone (GDELT) Project (2020; 2023; 2024), a global database that sources events from news text, and that has been used recently in economics and finance research (e.g., Acemoglu, Hassan, and Tahoun 2018; Campante and Yanagizawa-Drott 2018; Manacorda and Tesei 2020).

To measure the inclusivity of each source we calculate, among all country-months with a GDACS alert of given severity, the share in which a natural disaster was flagged by each comparison source. To measure overinclusion, we also calculate this share for country-months without any GDACS alerts.

We begin with a comparison to OSAC. Table 1 shows results on the performance of Crisis24 and OSAC. Crisis24 flags a disaster in 52 percent of country-months with a GDACS alert in the least severe category. This value rises to 89 percent in the most severe category. These values are far larger than for OSAC, which flags a disaster in only 8 percent of the months with a GDACS alert in the least severe category, and only 38 percent of the months with a GDACS alert in the most severe category.

Turning to a comparison across types of countries, we find that Crisis24 includes alerts in a larger share of months with GDACS alerts in OECD than in non-OECD countries, though the share in months with severe alerts remains large at 88 percent even in non-OECD countries. Crisis24 also flags alerts in a larger share of months without GDACS alerts in OECD than non-OECD countries.

We turn next to a comparison to GDELT. Because GDELT only classifies natural disasters beginning in 2015, Online Appendix O.A.1 presents results separately for 2010-2014 (for Crisis24 and OSAC only) and for 2015-2019 (for Crisis24, OSAC, and GDELT). We find evidence that GDELT is overinclusive. For example, GDELT is more likely to flag a natural disaster in a month with *no* GDACS alert, or a month with a less severe alert, than in a month with a severe GDACS alert. GDELT also includes far more alerts per country-month than both Crisis24 and OSAC. Online Appendix O.A additionally reports comparisons of C24 and OSAC to GDACS at the daily level, and comparisons of C24 and OSAC to the World Health Organization's Disease Outbreak News (DONs) database.

3.2 Measuring Protest

We measure protest using structures in text fields. Table 2 shows summary statistics on the incidence of protest. Panel A shows, across countries, statistics of the distribution of the share of months with protest. The average country experiences protest in 49 percent of months, with a 75th percentile of 69 percent and a 25th percentile of 29 percent.

To give a sense of how these protests are clustered, Panel B shows, across country-months, the share

Table 1: Summary of Crisis24 and OSAC Coverage of GDACS Disaster Alerts, Monthly

		Proportion with:		Average # of:		
		C24	OSAC	GDACS	C24	OSAC
Alert Type	Obs.	alerts	alerts	alerts	alerts	alerts
Panel A: All countries						
Green	5,143	0.52	0.08	7.03	1.55	0.11
Orange	484	0.86	0.26	15.36	3.94	0.44
Red	132	0.89	0.38	22.73	4.97	0.67
None	12,857	0.14	0.01	-	0.24	0.02
Panel B: OECD countries						
Green	1,364	0.67	0.09	10.94	2.56	0.17
Orange	72	0.96	0.46	45.24	6.61	1.07
Red	17	1.00	0.76	84.18	7.59	2.06
None	2,356	0.23	0.01	-	0.43	0.01
Panel C: Non-OECD countries						
Green	3,779	0.47	0.07	5.61	1.18	0.09
Orange	412	0.84	0.23	10.14	3.47	0.33
Red	115	0.88	0.32	13.65	4.58	0.47
None	10,501	0.12	0.01	-	0.19	0.02

Notes: The table shows statistics relating to GDACS, Crisis24 (C24), and OSAC coverage of country-months with natural disasters. “Alert Type” is the GDACS natural disaster classification, originally coded as green (least severe), orange (medium), or red (most severe). “Obs” displays the number of country-months with at least one alert of the type given by the column “Alert Type” or a higher severity alert type. These country-months are the sample used for calculating the statistics displayed in other columns. “Proportion with [C24/OSAC] alerts” refers to the proportion of country-months with a GDACS alert which are also flagged as having a disaster by [C24/OSAC]. “Average # of [GDACS/C24/OSAC] alerts” is the average number of total [GDACS/C24/OSAC] alerts in country-months within the row’s sample. Panel A shows results for all countries in our sample. Panels B and C include OECD and non-OECD countries, respectively.

with 0 protests, 1 protests, etc., both in the actual data, and in data where we have permuted the protest indicator across dates within each country. The real data exhibit more months with few and with many protests than the permuted data, showing that protests tend to cluster in particular country-months. Online Appendix O.B shows statistics on the clustering of protests within country-months.

In contrast to events like natural disasters, there is no source of ground truth for protests. Global databases typically rely on media accounts which are necessarily selective, and different sources may consider different events worthy of coverage (e.g., Hendrix and Salehyan, 2015; Jenkins and Maher, 2016; Weidmann and Rød, 2019). We therefore turned instead to curated academic databases studying particular contexts, often for short periods. Our search yielded the following: a database from Kadivar et al. (2023, 2024) covering two periods in Iran; a database from Enikolopov et al. (2020) covering protest in Russia during a single week; a database from Li (2018) covering protest in China over a two-year period; and a database from the Centre for Social Conflict and Cohesion Studies (COES; 2020) covering protest in Chile over the full sample period. These data cover important contexts and are produced by domain experts.

Table 2: Protest Event Summary Statistics, Across Countries

Panel A: Protest Event Incidence

	Share of months with:			Share of protest days with:	
	Protest	Anticipated protest	Suppressed protest	Anticipated protest	Suppressed protest
Maximum	1.000	0.933	0.850	0.927	1.000
75th percentile	0.692	0.465	0.350	0.585	0.365
Mean	0.494	0.321	0.222	0.465	0.271
Median	0.475	0.267	0.167	0.472	0.260
25th percentile	0.292	0.158	0.067	0.354	0.158
Minimum	0.008	0.000	0.000	0.000	0.000

Panel B: Within-Month Clustering of Protests

Mean share of months with:				
0 Protests	1 Protest	2-4 Protests	5-7 Protests	8+ Protests
0.506	0.191	0.206	0.061	0.037
(0.406, 0.413)	(0.238, 0.248)	(0.260, 0.268)	(0.063, 0.068)	(0.017, 0.019)

Notes: Panel A shows summary statistics across countries, for the distribution of the share of months with each listed event type (left panel) and the distribution of the share of protests with each listed attribute (right panel). Panel B reports the mean, across countries, of the share of months with the given number of protests. In parentheses below each mean is the range, from the 2.5th to 97.5th percentiles, of the same statistic across 200 random permutations of the protest indicator.

Online Appendix O.A.3 provides detailed findings. As a broad generalization, we tend to find that OSAC is underinclusive, GDELT is overinclusive, and Crisis24 is in between. For example, among protests in large cities in Iran, OSAC includes *none* during the period covered by Kadivar's (2023; 2024) data. Only 49 percent of the protests in large cities in GDELT match to one in Kadivar's (2023; 2024) data, suggesting that GDELT is overinclusive. Of course, a benefit of overinclusiveness is a higher capture rate, and indeed, GDELT's data captures 76 percent of protests in large cities in Kadivar's (2023; 2024) data. Turning to Crisis24, 89 percent of protests in Crisis24 also appear in Kadivar's (2023; 2024) data, suggesting that protests included in Crisis24 are verifiable. This high quality comes at the expense of some underinclusion: only 24 percent of large-city protests in Kadivar's (2023; 2024) data match to Crisis24. In relative terms, Crisis24's underinclusion is much less severe in Russia and China, where *only* Crisis24 captures any of the protests in the databases of Enikolopov et al. (2020) and Li (2018). In Chile, all comparison sources tend to miss protests outside of the capital city; among capital-city protests; the performance differences are qualitatively similar to those in Iran.

We conclude that the Crisis24 data are far more complete than the OSAC data, which is the most similar public database of which we are aware, and more verifiable than those in the GDELT data, which does not include complete event descriptions. We proceed using Crisis24 data in our analysis.

3.3 Measuring Repression

From the preceding subsection we have an indicator $z_{it} \in \{0,1\}$ equal to one if an alert indicates the occurrence of a protest on date t in country i, and zero otherwise. This indicator proxies for its counterpart in the model. Our remaining task is to measure repression also according to the concepts in the model.

For suppression, our aim is to measure its incidence, which we do by searching the alert text for indications of violence or arrests of protesters. We define an indicator $s_{it} \in \{0,1\}$ equal to one if an alert posted on date t or after indicates the occurrence of a protest on t and use of force, and zero otherwise. This indicator proxies for its counterpart in the model, and is strongly correlated with a similar concept from a standard global database (see Online Appendix O.A.4). Table 2 shows summary statistics on the incidence of suppression, which occurs in 27 percent of protests in the average country in the sample.

For prevention, our aim is to measure the *preventiveness* of the regime in a manner that is consistent with the ideas in Section 2. We focus on measuring preventiveness at the country-month level. Aggregating to the country-month level allows us to capture clusters of protests in a single measurement unit, while, as we will see in the next section, also allowing sufficient resolution to track important events.

Central to our approach to measuring preventiveness is the indicator $a_{it} \in \{0,1\}$ for anticipated or announced protest, which we define to be equal to one if an alert posted at least one day earlier indicates the occurrence of a protest on the given date, and zero otherwise.[22] Table 2 shows that about half of protests are anticipated according to this definition. The indicator $a_{it} \in \{0,1\}$ proxies for its counterpart in the model: if Crisis24 knows about the protest in advance, we assume that the regime does, too.

Neither the curated academic databases we discuss in Section 3.2, nor any other extant global databases of protest that we are aware of, includes information on whether a given protest was anticipated. Capturing this information is one of our main motivations for introducing the Crisis24 database. Still, it is useful to validate the indicator for anticipated protest against an external source. The best external comparison we are aware of is to the OSAC alerts that we have collected, as these follow a structure similar enough to Crisis24 to permit extracting an indicator analogous to a_{it}. Online Appendix O.A.5 shows that Crisis24 captures a large portion of OSAC anticipated protests, whereas OSAC captures a smaller portion of Crisis24 alerts, consistent with the finding in Section 3.2 that OSAC tends to be underinclusive.

Section 2 shows that the occurrence of anticipated protest implies that preventiveness is zero. We

[22]We also set the indicator equal to one in cases where the alert is posted on the day of the protest but where the title of the alert indicates a planned or scheduled protest (e.g., "protests scheduled"). Supplemental Appendix S.B presents results where we instead set the indicator equal to zero in such cases.

therefore define our main index of preventiveness as (one minus) an indicator for whether anticipated protest ever occurs during the given country-month. Importantly, Section 2 also shows that the absence of anticipated protest does *not* imply that the preventiveness is one. We should therefore think of our index as an upper bound on the preventiveness. Following Proposition 2, such a bound is the best an econometrician can hope to achieve without sharing the regime's information set.[23] We turn next to tightening the bound by incorporating additional signals of anticipated protest.

3.4 Tightening Bounds with Additional Signals of Protest

Section 2 showed that we can tighten the measured bound on preventiveness by using more of the information in the regime's information set. We therefore collected additional information that we thought would either be readily available to the regime, or closely correlated with readily available information: social media mentions of protest, media mentions of protest, and search queries about protest (see Supplemental Appendix S.B). We included standardized, lagged values of these variables in a predictor vector $\mathbf{x}_{it}$, along with the indicator a_{it}, and used them to form a predicted probability of protest $\hat{p}(\mathbf{x}_{it})$ for each country i and date t in our sample. Importantly, we trained the predictor $\hat{p}(\mathbf{x}_{it})$ using data from *other* country-months, so that information from country i in the month containing date t is never used to construct the function $\hat{p}(\cdot)$. We tried several types of predictive models, estimators, and sample-splitting schemes; these are detailed in Supplemental Appendix S.C.

Given the estimated probabilities $\hat{p}(\mathbf{x}_{it})$, our goal is to estimate the largest true probability of protest in each country-month.[24] We do this in three ways. First, we simply take the largest value of $\hat{p}(\mathbf{x}_{it})$ in the country-month. This procedure is simple to describe but is known to be biased (e.g., Hirano and Porter 2012; Lei et al. 2021). To address this bias, we take advantage of recent ideas in statistics that relate to inference on black-box predictors (e.g., Lei et al. 2021). A key idea is that, while it is impossible to demonstrate that an outcome is *not* predictable (because the econometrician may simply have the wrong predictors or the wrong model), it is possible to show with statistical confidence that an outcome *is* predictable (by showing that the econometrician can reliably predict it using a function trained out-of-sample).

Accordingly, our second estimate of the largest true probability of protest in a given country-month is an indicator for whether protest occurs on the date with the largest value of $\hat{p}(\mathbf{x}_{it})$. Because we train $\hat{p}(\cdot)$ out of sample, this indicator is a Bernoulli random variable whose expectation is no larger than the largest true probability of protest in the country-month. It therefore yields an unbiased estimate of an upper

[23] To connect the bound back formally to Proposition 2, we can imagine that the econometrician's coarsening is simply $\chi(x) = a$.

[24] In the notation of Section 2, the largest probability of protest with respect to the econometrician's information set is $\overline{p}$; if the econometrician's information set is equivalent to the regime's, then $\overline{p} = \overline{p}^*$.

bound on the preventiveness. Intuitively, if protest often occurs on the day on which it is predicted to be most likely, then there must be some days on which the probability of protest is large, and preventiveness must therefore be low.

A limitation of this second measure is that it only uses information on the incidence of protest on a single day in each month. Intuitively, we may expect that more refined inferences are available if we use information on whether protest occurs on the date with the second-largest, third-largest, etc. probabilities $\hat{p}(\mathbf{x}_{it})$. To incorporate this information, our third estimate uses the procedure in Lei (2023) to construct a finite-sample confidence bound for the largest true probability of protest in the given country-month, from which we compute an indicator for whether this confidence bound includes preventiveness near one. This procedure relies on independent and identically distributed data.

In practice, all of these approaches yield indices of preventivenesss that are closely correlated with one another, and with our main index based only on the anticipation indicator a_{it}.[25] Intuitively, the reason is that most of the predictive information in the predictors $\mathbf{x}_{it}$ comes from the anticipation indicator a_{it}. This is not surprising: Crisis24 analysts use a wide range of sources to produce alerts (including many of those that underlie our predictors $\mathbf{x}_{it}$), and have an incentive to flag anticipated events before they happen. As a result, Crisis24 security alerts do a good job summarizing the information available in other public signals of impending protest. Given these findings, we proceed using the index defined in Section 3.3 for the analysis in the main text, and the additional indices defined in this section for sensitivity analyses shown in Supplemental Appendix S.C.

4 Application to Surveillance of Repression

Our first application is to the surveillance of repression by the international community. We begin by showing how our indices track salient episodes of repression. This serves as both validation of our approach to measurement, and proof-of-concept for use of these indices in surveillance. We then study how the international community responds, both in word and in deed, to different forms of repression.

4.1 Evolution of Protest and Repression During Salient Episodes of Repression

We attempted to collect data on all episodes of martial law and states of emergency reported by the *Freedom in the World* reports during our sample years (Freedom House, 2010-2019). We focused on episodes for which there is clear narrative evidence of preventive repression, and a clear timeline. Supplemental Appendix S.D provides additional details about our data collection procedure.

Our search yielded two episodes of martial law and two states of emergency for which there is a clear

[25]The correlations among the four measures are always above 0.949 in monthly first differences.

timeline. Supplemental Appendix S.D discusses three additional episodes of states of emergency that occur over multiple years and phases.

Figure 2 depicts the evolution of preventiveness (left) and the incidence of protest and suppression (right) for the two episodes of martial law. We plot our indices at the monthly level, where we define the incidence of protest as the occurrence of any protest during the month, and likewise for the incidence of suppression. We annotate salient events with their exact dates.

Panel A of Figure 2 concerns events in Bahrain in 2011 before, during, and after the Pearl Uprising (see, e.g., Wehrey 2014, Chapter 5; Shehabi and Jones 2015; Jones 2020, Chapter 5). The uprising began with protests in February 2011. The protests threatened the regime; among protester demands was the end of the ruling monarchy. Protests continued in the Pearl Roundabout through March 2011. In March 2011, the regime declared martial law and, with the assistance of international forces from the Gulf Cooperation Council, reasserted control of the streets. Martial law ended in June 2011. Importantly, although the regime succeeded in reasserting control of the Pearl Roundabout during the period of martial law, protests and clashes with security forces continued throughout the year.

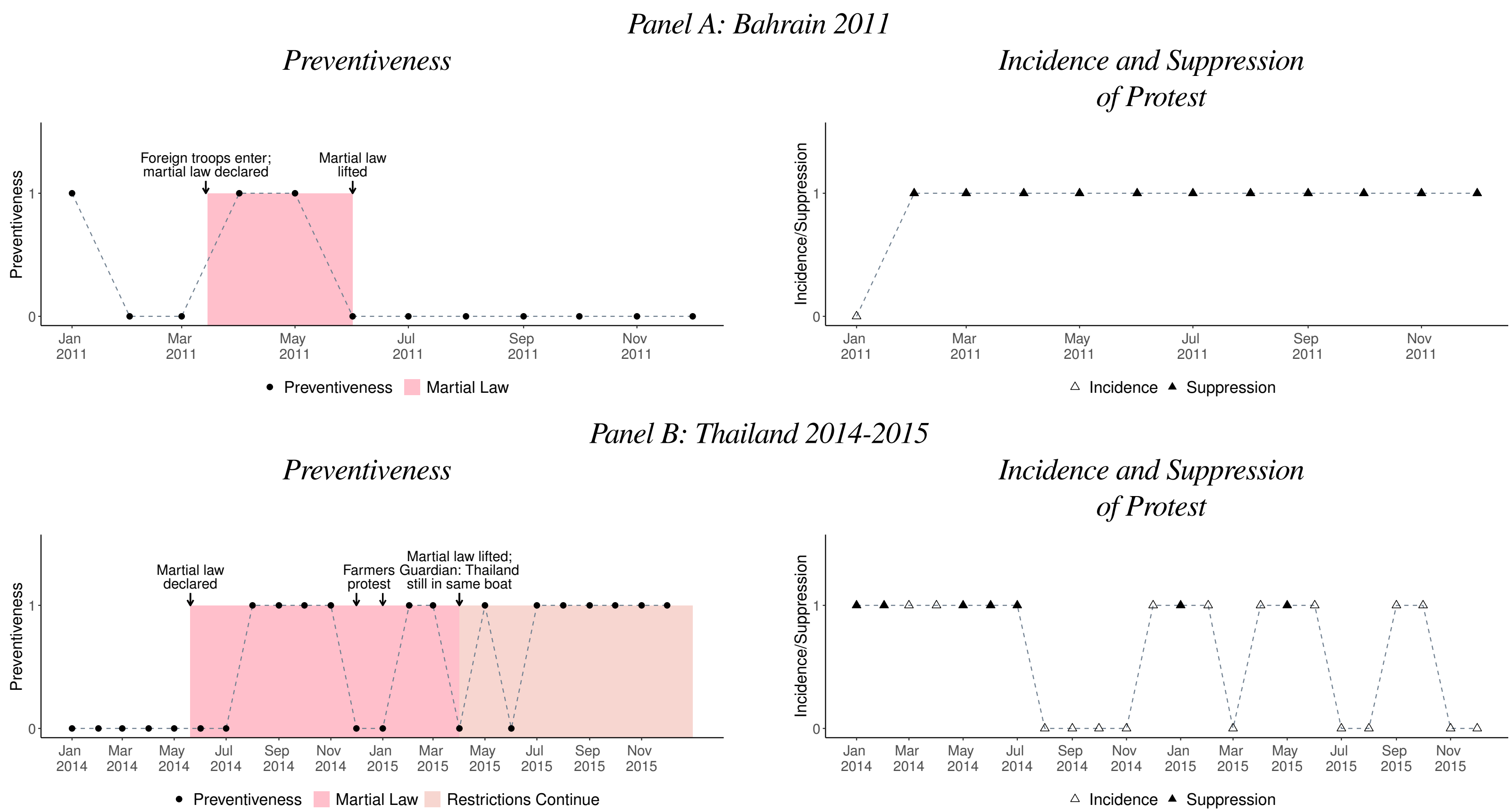
Our indices track this episode well. Our index of preventiveness shows a spike that coincides closely with the period of martial law (left plot). Our indices of protest incidence and suppression, on the other hand, show continued protest and suppression beginning in February and lasting throughout the year (right plot). These findings reinforce that our index of preventiveness is not driven only by the incidence of protest, and that prevention and suppression are distinct forms of repression.

Panel B of Figure 2 concerns events in Thailand before, during, and after the imposition of martial law in 2014. The declaration of martial law occurred in May 2014 after months of political instability, and was followed shortly by a military coup (Harlan and Samuels 2014). While martial law was maintained formally for nearly one year, many accounts indicate that restrictions on protest continued under the new government after the formal end of martial law (Associated Press in Bangkok 2015; Freedom House, 2016, p. 694).

Our indices also track this episode well. The period of martial law coincides with high preventiveness, punctuated by months with planned protests by rubber farmers (left plot). Although some of the months with high preventiveness feature no protest (right plot), other months with high preventiveness do feature protest (right plot), again showing that the two concepts are not simply inverses of one another.

Most importantly, although martial law formally ended in April 2015, our measure of preventiveness indicates episodic but continued prevention throughout the remainder of the calendar year (left plot), including during months with protest (right plot), and closely in line with contemporary accounts. As Thai

Figure 2: Dynamics of Protest and Repression During Episodes of Martial Law



Notes: The plots show the evolution of protest preventiveness, incidence, and suppression during episodes of martial law, with annotations for important dates. Shaded areas correspond to the exact dates of the labeled period. Supplemental Appendix S.D includes additional details about our data collection procedure and about the episodes we consider.

political scientist Thitinan Pongsudhirak noted at the time, "From the outside, the lifting of martial law is good news for business and tourism...But from the inside, we're functionally in the same boat...Similar restrictions are still in place" (Associated Press in Bangkok, 2015). Our index of preventiveness successfully tracks the boat the Thai people were in, rather than the boat the regime portrayed.

Figure 3 depicts the evolution of preventiveness (left) and protest incidence and suppression (right) for the states of emergency. Panel A of Figure 3 concerns events in Thailand in 2010. Beginning in March, the opposition United Front for Democracy Against Dictatorship ("Red Shirts") organized protests against the government. The government declared a state of emergency in Bangkok in April and extended it more widely in May, suppressing the protests and banning future ones. Following the initial lifting of the emergency in July, planned protests resumed, including a large Red Shirt rally in September (BBC, 2010).

Our measures track this episode well, including the initial suppression and subsequent prevention of protest during the state of emergency, and the resumption of permitted protest after its end. Of the three months of the year with high preventiveness, two are during the height of the emergency.[26] This is true even though protest continued during the emergency.

Panel B of Figure 3 concerns events in Mali in 2013. A state of emergency began in January. Although in the context of a French-assisted military operation against radical Islamist groups, the emergency banned large gatherings (Freedom House, 2014, p. 446). The government ended the emergency in July, timing that allowed campaigning in advance of the presidential election later that month, which had been earlier postponed (Al Jazeera 2013a; Freedom House, 2014, p. 446). We measure the period of the emergency as fully preventive and the period after as only sporadically preventive, with the least preventive periods coinciding with second rounds of the presidential and parliamentary elections. Violence, including political violence, continued following the end of the emergency and throughout the election period (Al Jazeera 2013b; International Crisis Group 2014).

4.2 Response of the International Community

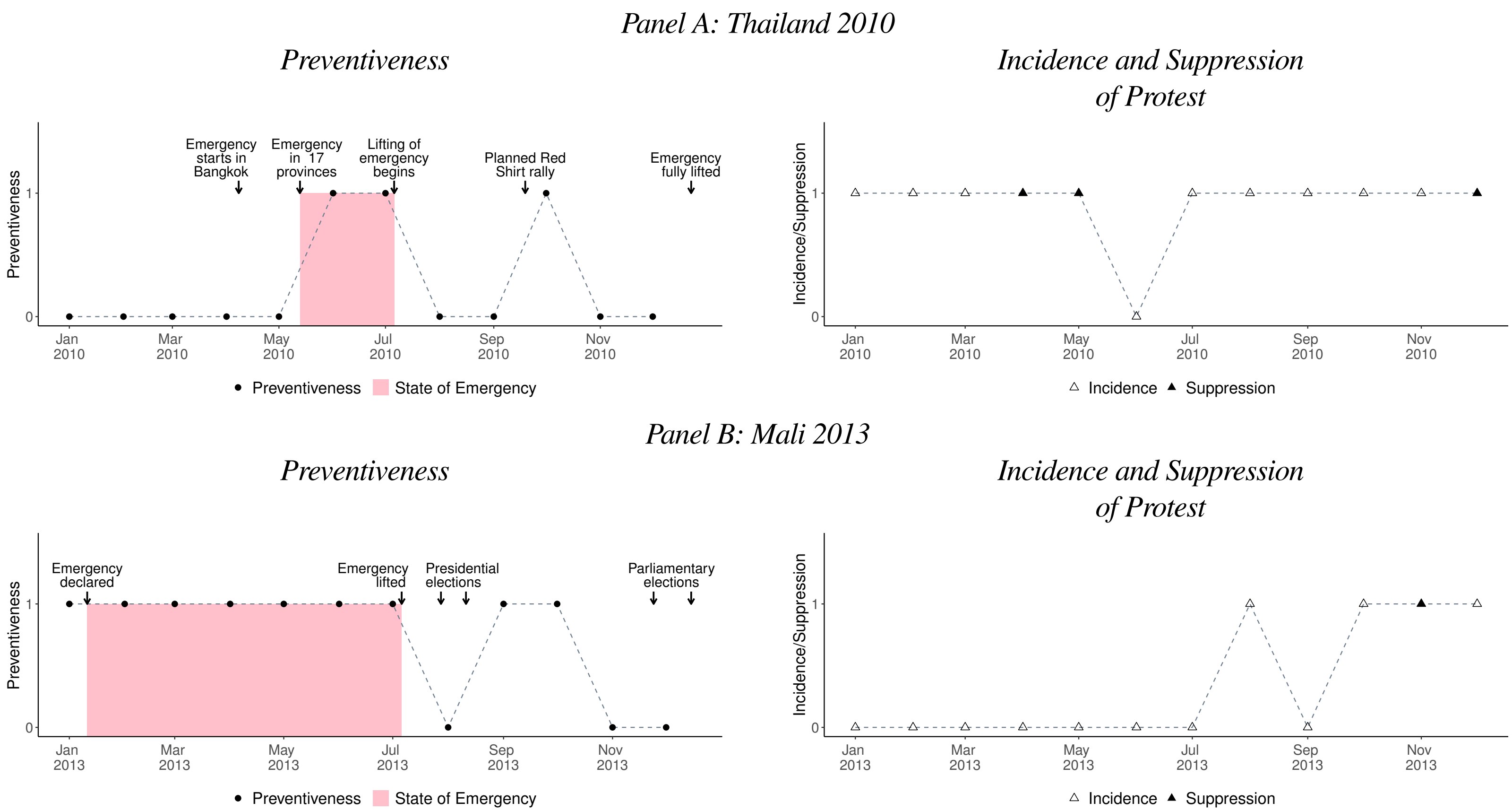
Having shown that our indices can be used to track salient episodes of repression, we now use the indices to study how the international community responds to repression in our sample of non-OECD countries. We begin by studying what the international community says, and then turn to what it does.

4.2.1 Expressions of Concern in Protest Alerts

As a first measure of the international community's response, we measure, in each country i and month k, the share $h_{i,k}$ of C24 protest alerts that mention human rights concerns. We measure this share based on the

[26]The third, October 2010, is a month during which there were restrictions on demonstrations in Bangkok (Kosolpradit 2010; Voice of America News 2010).

Figure 3: Dynamics of Protest and Repression During States of Emergency



Notes: The plots show the evolution of protest preventiveness, incidence, and suppression during states of emergency, with annotations for important dates. Shaded areas correspond to the exact dates of the labeled period. Supplemental Appendix S.D includes additional details about our data collection procedure and about the episodes we consider.

occurrence of terms such as "human rights violations," "human rights abuse," or mentions of human rights groups, as detailed in Supplemental Appendix S.B. Expressions of human rights concerns in C24 alerts have the advantage of being tightly tied to the protests we study while still not mechanically related to our indices of preventiveness or suppression. Because the measure h_{ik} is defined only in country-months with a protest alert, it is also not affected by any mechanical connection between our measure of preventiveness and the occurrence of protest.

We specify and estimate the following model:

$$h_{i,k} = \phi_i^h + \varphi_k^h + \gamma_0^{h,p}\left(1-\overline{p}_{i,k}\right) + \gamma_{-1}^{h,p}\left(1-\overline{p}_{i,k-1}\right) + \gamma_0^{h,s} s_{i,k} + \gamma_{-1}^{h,s} s_{i,k-1} + \varepsilon_{i,k}^h \tag{1}$$

where ϕ_i^h is a country fixed effect, φ_k^h is a calendar month fixed effect, $\left(1-\overline{p}_{i,k}\right)$ is our index of preventiveness, $s_{i,k}$ is our index of suppression, the γ's are coefficients, and $\varepsilon_{i,k}^h$ is a residual. We include both the current and lagged values of preventiveness and suppression to allow for delayed effects, which are plausible for some of the responses we study in this section.

Column (1) of Table 3 presents an OLS estimate of Equation (1). To facilitate interpretation we express these coefficients, and all others in this section, relative to the baseline mean of the dependent variable, defined as the sample mean in country-months in which protest occurs but there is neither preventiveness nor suppression. We report the baseline mean in Table 3 for reference. We perform inference in two ways: first, using asymptotic standard errors, and second, using a 95% credible interval obtained from a Bayes bootstrap, both clustered by country.[27] The Bayes bootstrap retains a useful interpretation even when the usual local asymptotically normal approximation breaks down (Andrews and Shapiro, 2024).

Column (1) of Table 3 reports that the share of alerts mentioning human rights concerns is larger in preventive months by an amount equivalent to 85.7 percent of the baseline mean. There is a smaller, but still economically meaningful, positive association with the incidence of suppression, representing 57 percent of the baseline mean. The share of alerts mentioning human rights concerns in the current month is not consistently related to either preventiveness or suppression in the previous month.

4.2.2 Expressions of Concern in Human Rights Watch Alerts

As a second measure of the international community's response, we turn to alerts published by Human Rights Watch (HRW 2024), a global NGO that advocates for human rights. We collected data on all of the alerts published by HRW during our sample period, and used automated methods described in Supplemental Appendix S.B to classify the country they target and their contents. Relative to our first measure, these data have the drawback that we cannot tie each alert to a specific protest in our database.

[27] Here and below we compute the credible interval using 500 replicates.

But, they have the corresponding advantage of being constructed from a source that is not used in constructing our indices of repression. And, because HRW focuses on human rights, the alerts tend to be granular regarding the behaviors of the regime that are causing concern.

Accordingly, our second measure of the international community's response is the average of the share of HRW alerts mentioning each of the following: general human rights concerns, police deployments, restrictions on press freedom, restrictions on the internet, and torture. Column (2) of Table 3 presents an OLS estimate of an analogue of Equation (1) using this alternative dependent variable. We estimate that this share is larger in preventive months by 33.7 percent of the baseline mean. This estimate is smaller than its counterpart in column (1) but still economically meaningful and clearly statistically significant, showing that our measure of preventiveness is strongly associated with an independently measured proxy for the expression of human rights concerns by the international community. We see a more modest, and statistically insignificant, association with suppression.

One of the motivations for our study is the hypothesis that acts of suppression are more visible than acts of prevention. Column (3) of Table 3 repeats the specification of column (2) but using the average of the share of HRW alerts mentioning each of the following: use of force against protesters, dispersal of protest, and arrests of protesters. Consistent with the greater visibility of these acts, the coefficient on our indicator for incidence of suppression is very large, representing 153.9 percent of the baseline mean.[28] HRW is more likely to mention acts of suppression in months following the incidence of suppression, and is less likely to mention acts of suppression in months with greater preventiveness.

4.2.3 Initiation of Sanctions

As a final measure of the international community's response, we measure whether sanctions against country i are announced in month k. We obtain information on sanctions during our sample period from the Global Sanctions Data Base (GSDB 2023; see also Felbermayr et al. 2020; Kirilakha et al. 2021). We supplement these data with dates of sanction announcements based on our own searches of news archives. We focus on cases where the sanctioning entity includes the UN, EU, US, or at least one other member of the OECD. In column (4) of Table 3, the dependent variable is an indicator for the announcement of sanctions related to human rights or democracy. In column (5) of Table 3, the dependent variable is an indicator for the announcement of sanctions in another category (e.g., territorial conflict). Because we wish to be agnostic about the timing of sanctions in relation to protest, in both columns we include

[28] Also consistent with the greater visibility of acts of suppression, Supplemental Appendix S.E shows that the total number of HRW alerts in a given country-month is positively related to our indicator for the incidence of suppression, and (more weakly) negatively related to our index of preventiveness.

Table 3: Response of the International Community

	Share of C24 protest alerts mentioning human rights	Mean share of HRW alerts mentioning		Announcement of sanctions	
		Prevention	Suppression	Human rights/ Democracy	Other
	(1)	(2)	(3)	(4)	(5)
Preventiveness:					
Current month	0.857	0.337	-0.994	-0.217	-0.383
	(0.306)	(0.134)	(0.345)	(0.313)	(0.465)
	[0.280, 1.708]	[0.074, 0.615]	[-1.635, -0.448]	[-0.873, 0.476]	[-1.143, 0.850]
Previous month	-0.107	-0.024	0.084	-0.348	-0.270
	(0.257)	(0.141)	(0.333)	(0.358)	(0.513)
	[-0.698, 0.307]	[-0.294, 0.257]	[-0.565, 0.791]	[-1.480, 0.312]	[-1.715, 0.731]
Suppression:					
Current month	0.570	0.104	1.539	0.105	0.343
	(0.272)	(0.116)	(0.329)	(0.412)	(0.387)
	[0.087, 1.177]	[-0.099, 0.361]	[0.773, 2.698]	[-0.707, 1.075]	[-0.243, 2.300]
Previous month	-0.038	0.094	0.681	0.799	-0.071
	(0.213)	(0.121)	(0.383)	(0.360)	(0.343)
	[-0.439, 0.395]	[-0.136, 0.323]	[-0.003, 1.500]	[0.178, 1.940]	[-0.930, 0.585]
Baseline mean	0.0139	0.0093	0.0043	0.0065	0.0051
Only months with protest	X	X	X		
Control for protest incidence				X	X
Number of countries	119	119	119	119	119
Number of country-months	6,512	6,512	6,512	14,161	14,161

Notes: Each column presents results from a single linear regression model. The unit of analysis is a country-month. Coefficients are reported as a fraction of the baseline mean of the dependent variable, defined as the sample mean in country-months with protest but without preventiveness or suppression. For each coefficient we report an asymptotic standard error in parentheses and a Bayes bootstrap credible interval in brackets, both clustered by country. All models include country and month fixed effects. In column (1), the dependent variable is the share of C24 protest alerts that mention human rights concerns. In column (2), the dependent variable is the mean of the share of HRW alerts that mention human rights concerns, internet shutdowns, press restrictions, and torture. In column (3), the dependent variable is the mean of the share of HRW alerts that mention use of force against protesters, dispersal of protest, and arrests of protesters. In column (4), the dependent variable is an indicator for the announcement of international sanctions related to human rights or democracy. In column (5), the dependent variable is an indicator for the announcement of other sanctions. In columns (1), (2), and (3) we include only country-months with a protest; in columns (4) and (5) we include all country-months and control for indicators for the occurrence of protest in the current and previous months.

all country-months in the regression sample, and we modify Equation (1) by adding controls for the incidence of protest in the current month and the previous month.

We hypothesize that the greater visibility and lower deniability of suppression makes it more likely that sanctions will respond to suppression than to preventiveness. As sanctions can take time to coordinate (see, e.g., O'Toole and Sultoon 2019; Barigazzi and Vela 2022), we also hypothesize that sanctions will be announced with a lag relative to the acts of repression that they respond to. Column (4) of Table 3 shows that both hypotheses are borne out in our data. We estimate that the international community is more likely to announce sanctions related to human rights or democracy in a given month if the sanctioned country has suppressed protest in the preceding month; the magnitude of this relationship is equivalent to 79.9 percent of the baseline mean. By contrast, the contemporaneous relationship between the announcement of sanctions and suppression is smaller and statistically insignificant, and the relationship between the announcement of sanctions and preventiveness (both in the current and previous months) is negative and statistically insignificant.[29]

We expect a weaker connection between repression and sanctions that are not related to human rights or democracy. Column (5) of Table 3 shows that this expectation is borne out in the data, with none of the estimated coefficients being statistically significant, or as large economically as the coefficient on the previous month's suppression in column (4). It is important to note, though, that the credible intervals in column (5) do not rule out economically large relationships.

In summary, we find that preventiveness predicts the expression of concerns about human rights and preventive tactics. We also find that the occurrence of suppression even more strongly predicts the expression of concern about suppressive tactics, and predicts the later announcement of sanctions related to human rights or democracy. These patterns are consistent with prevention being partly visible to the international community, but less visible, more deniable, and harder to sanction than suppression.

5 Application to the Economic Determinants of Repression

Our next application is to the question of how economic shocks affect repression. A large literature studies the effect of economic change on freedom and democracy (e.g., Egorov et al. 2009; Burke and Leigh 2010; Caselli and Tesei 2016). Our analysis adds two main elements relative to prior work. First, we can separate the repression of dissent into prevention and suppression. Second, we can study repression sub-annually. Both of these elements are enabled by our novel indices of repression. In the remainder of this section, we lay out our data, hypotheses, methods, and findings.

[29] We cannot reject statistically that the coefficient on the incidence of suppression in the preceding month is equal to that on preventiveness in the preceding month.

5.1 Measures of Economic Variables

We obtain data on gross domestic product (GDP) in 2015 US dollars, as well as the share of GDP that is accounted for by government revenue, from the World Development Indicators (World Bank, 2024).[30] We obtain data for an index of the price of commodity exports, weighted by each commodity's export share of the given country's GDP, from the IMF's Commodity Terms of Trade (CTOT) database (Kebhaj and Gruss, 2019).[31]

For the sample of non-OECD countries that we study, the macroeconomic aggregates we use are much more often available annually than sub-annually.[32] We therefore conduct our initial analysis on an unbalanced annual panel of countries, constructing annual measures of preventiveness, incidence of suppression, and incidence of protest for each country by taking the average of these monthly variables across the months of each calendar year. We return later to an analysis of sub-annual patterns using a subset of available variables.

5.2 Hypotheses and Descriptive Findings

As an initial descriptive exercise we specify and estimate the following two-way fixed effects model:

$$1 - \overline{p}_{ic} = \phi_i^p + \varphi_c^p + \gamma^{p,y} y_{ic} + \varepsilon_{ic}^p \tag{2}$$

where $1 - \overline{p}_{ic}$ is the average estimated preventiveness for country i in calendar year c, ϕ_i^p is a country fixed effect, φ_c^p is a calendar year fixed effect, y_{ic} is the log of real GDP, $\gamma^{p,y}$ is a coefficient, and ε_{ic}^p is an error term reflecting, for example, noise in the measure of preventiveness. We estimate this model (and all models below) in annual differences. We again perform inference using both asymptotic standard errors and a Bayes bootstrap credible interval, both clustered by country.

Column (1) of Table 4 presents an OLS estimate of Equation (2). We estimate a positive and statistically insignificant value of the coefficient $\gamma^{p,y}$, indicating no statistical evidence that changes in economic resources are associated with changes in preventiveness.

Equation (2) treats all economic resources symmetrically. Our theoretical model, by contrast, identifies multiple pathways by which resources may affect preventiveness. Greater resources for the regime may increase the incentive to retain power, and improve the technology for preventing dissent, thus increasing

[30] In the few cases where government revenue accounts for less than 0.1 percent of GDP, we treat government revenue as missing.

[31] Export shares are from a backward-looking moving average. We use the original data from Kebhaj and Gruss (2019) where possible, and the data from International Monetary Fund (2024) where this is unavailable. Supplemental Appendix S.B shows results where we instead use the commodity terms of trade, which depends on both import and export shares.

[32] For example, the World Development Indicators (World Bank, 2024) do not include a subannual measure of government revenues.

preventiveness.[33] Greater economic resources for citizens may, on the other hand, reduce the likelihood that protest escalates to revolution, thus lessening the need to prevent protest.

We therefore specify and estimate a second model in which we focus on the resources available to the government:

$$1-\overline{p}_{ic}=\phi_i^p+\varphi_c^p+\gamma^{p,g}g_{ic}+\varepsilon_{ic}^p \tag{3}$$

where g_{ic} denotes the log of government revenue in country i and year c and $\gamma^{p,g}$ is a coefficient. Here and below, we abuse notation by retaining the same notation for the fixed effects and residual as in Equation (2).

Column (2) of Table 4 presents an OLS estimate of Equation (3). We estimate a positive and statistically significant coefficient $\gamma^{p,g}$. Its magnitude implies that an increase of 10 log points (approximately 10 percent) in government revenue is associated with an increase in preventiveness of 1.3 percentage points, equivalent to about 6.1 percent of a standard deviation of the change in preventiveness in the sample.

Resources available to the government are of course correlated with general economic resources.[34] We therefore specify and estimate the following model that includes both government and general resources:

$$1-\overline{p}_{ic}=\phi_i^p+\varphi_c^p+\gamma^{p,g}g_{ic}+\gamma^{p,y}y_{ic}+\varepsilon_{ic}^p \tag{4}$$

where we again abuse notation by retaining the same notation for the coefficients as in the earlier models. In this model, we can think of $\gamma^{p,g}$ as reflecting the association between preventiveness and government resources, after accounting for their linear association with GDP.

Column (3) of Table 4 presents an OLS estimate of Equation (4). We estimate a positive and statistically significant coefficient $\gamma^{p,g}$, similar in magnitude to the estimate in column (2). The coefficient $\gamma^{p,y}$ on the log of GDP is statistically insignificant and less precisely estimated than the coefficient $\gamma^{p,g}$ on the log of real government revenue, reflecting that government revenue is more variable than other components of GDP.[35] Nevertheless, we find it interesting that the point estimate of $\gamma^{p,y}$ is negative, consistent with greater resources for citizens reducing the need to prevent protest.

5.3 Using Commodity Price Shocks to Isolate Exogenous Changes in Government Resources

An important limitation of the model in Equation (4) is that it aggregates all government revenue together. Not all government revenue directly benefits the regime; some will, for example, be used in broad redistri-

[33]These implications are particularly clear in the dynamic extension in Supplemental Appendix S.A.1, where, holding constant the equilibrium probability of revolution and suppression given protest, the equilibrium preventiveness in a given environment is determined by the ratio of the cost of prevention to the continuation value to the regime of remaining in power, both of which may be affected by the regime's access to resources.

[34]In our sample, the correlation of the annual change in the log of real government revenues with the annual change in the log of real GDP is 0.32.

[35]In our sample, the standard deviation of the annual change in the log of real government revenue is 0.115, as compared to 0.042 for the annual change in the log of real GDP excluding government revenue.

Table 4: Economic Determinants of Prevention, Country-Year Panel

	Preventiveness						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(GDP)	0.011		-0.154				
	(0.186)		(0.205)				
	[-0.363, 0.334]		[-0.562, 0.226]				
Commodity export price component					-0.075	-0.537	-0.933
					(1.122)	(0.980)	(0.682)
					[-2.697, 2.844]	[-3.309, 1.496]	[-2.875, 0.308]
Remainder component					-0.143	-0.087	-0.058
					(0.207)	(0.168)	(0.146)
					[-0.547, 0.236]	[-0.426, 0.264]	[-0.359, 0.203]
log(Government revenue)		0.130	0.144				
		(0.048)	(0.054)				
		[0.043, 0.227]	[0.047, 0.257]				
Commodity export price component				0.486	0.419	0.465	0.337
				(0.237)	(0.211)	(0.140)	(0.093)
				[-0.020, 1.274]	[-0.047, 1.060]	[0.178, 0.981]	[0.125, 0.678]
Remainder component				0.110	0.122	0.081	0.101
				(0.046)	(0.051)	(0.051)	(0.044)
				[0.025, 0.200]	[0.032, 0.222]	[-0.008, 0.188]	[0.025, 0.197]
Control for incidence of:							
Suppression						X	X
Protest							X
Number of country-years	729	729	729	729	729	729	729
Number of countries	90	90	90	90	90	90	90

Notes: Columns (1)-(4) present OLS estimates of Equations (2), (3), (4), and (6) respectively. Columns (5), (6), and (7) add additional variables to the specification in column (4). All models are estimated in annual differences. Economic variables are in 2015 US dollars. Annual measures of preventiveness, the incidence of suppression, and the incidence of protest are calculated as averages across months. Below each coefficient we report an asymptotic standard error in parentheses, and a 95% Bayes bootstrap credible interval in brackets, both clustered by country.

bution or public goods provision. An important limitation of the OLS estimate of Equation (4) in column (3) of Table 4 is that government revenue is not econometrically exogenous: changes in repression, for example, could directly affect the economy and the government's ability to extract resources.

Variation in commodity export prices provides a way to address both limitations. Because commodity exports tend to be heavily leveraged for rent-seeking (e.g., Dube and Vargas, 2013; Asher and Novosad, 2023), the resources they yield are disproportionately (though not entirely) available to the regime. And because commodity price shocks can be taken as exogenous with respect to activity in most individual countries (e.g., Aghion et al., 2010; Bazzi and Blattman, 2014; Kebhaj and Gruss, 2019), there is limited scope for reverse causality from preventiveness to commodity prices.

We begin by estimating a model that relates commodity export prices to government revenue. Specifically, we estimate the following model:

$$g_{ic} = \phi_i^g + \varphi_c^g + \gamma^{g,x} x_{ic} + \varepsilon_{ic}^g \tag{5}$$

where ϕ_i^g is a country fixed effect, φ_c^g is a calendar year fixed effect, x_{ic} is the log of the index of commodity export prices, $\gamma^{g,x}$ is a coefficient, and ε_{ic}^g is an error term.

Column (1) of Table 5 presents an OLS estimate of Equation (5). We estimate a positive and highly statistically significant coefficient $\gamma^{g,x}$. Its magnitude implies that a 10 percent increase in commodity export prices leads to an 8.1 percent increase in government revenue.

We use the estimated Equation (5) in column (1) of Table 5 to decompose government revenues into a component $g_{ic}^x = \gamma^{g,x} x_{ic}$ predicted from commodity export prices and a remainder $g_{ic}^{-x} = g_{ic} - g_{ic}^x$. We then estimate the following modification of Equation (3):

$$1 - \overline{p}_{ic} = \phi_i^p + \varphi_c^p + \gamma^{p,g,x} g_{ic}^x + \gamma^{p,g,-x} g_{ic}^{-x} + \varepsilon_{ic}^p. \tag{6}$$

Here $\gamma^{p,g,x}$ estimates the effect on preventiveness of variation in government revenues due to changes in commodity export prices, and $\gamma^{p,g,-x}$ estimates the association with variation in government revenues due to other factors. An OLS estimate of $\gamma^{p,g,x}$ in (6) is algebraically equivalent to a 2SLS estimate of $\gamma^{p,g}$ in (3), treating x_{ic} as an excluded instrument. An advantage of the decomposed formulation in Equation (6) is that it permits a direct comparison between the estimated effect $\gamma^{p,g,x}$ of variation in government revenues due to commodity export prices and the estimated association $\gamma^{p,g,-x}$ with variation in government revenues due to other factors.

Column (4) of Table 4 presents an OLS estimate of Equation (6). We estimate positive values of both $\gamma^{p,g,x}$ and $\gamma^{p,g,-x}$, with the latter statistically significant and the former marginally so. The estimated value of $\gamma^{p,g,x}$ implies that a 10 log-point increase in government revenue due to commodity export price changes increases preventiveness by 4.9 percentage points. The estimated value of $\gamma^{p,g,-x}$ is economically smaller, implying that a 10 log-point increase in government revenue due to factors orthogonal to commodity export price changes is associated with an increase in preventiveness of 1.1 percentage points. Although the difference $(\gamma^{p,g,x} - \gamma^{p,g,-x})$ is economically large, it is not statistically significant according to a 95% Bayes bootstrap credible interval.

Of course, export prices may affect sectors of the economy other than the government. Column (2) of Table 5 shows that the estimated contemporaneous effect of commodity export prices on GDP is small and statistically insignificant. Columns (3) and (4) of Table 5 show, however, that while commodity export prices have an immediate effect on government revenues (column (3)), they affect GDP with a lag (column (4)). The latter finding is in line with prior work that studies the macroeconomic effects of changes in commodity terms of trade (e.g., Bazzi and Blattman 2014, Online Appendix Table 1-4; Kebhaj and Gruss 2019, Table 1).

Column (5) of Table 4 presents an OLS estimate of an analogue of Equation (4) where we decompose both government revenues and GDP into components predicted from commodity export prices (up to two lags) and a remainder component. The magnitudes and statistical significance of the coefficients for government revenue are similar to their counterparts in column (4). The coefficients for GDP, on the other hand, are statistically insignificant and negative.

Our model emphasizes that different forms of repression can be substitutes. Accordingly, column (6) of Table 4 adds to the specification in column (5) a control for the incidence of suppression. Relative to column (5), the estimate of $\gamma^{p,g,x}$ in column (6) is larger, more precise, and statistically significant, whereas the estimate of $\gamma^{p,g,-x}$ is smaller, similarly precise, and only marginally statistically significant. Of course, as we do not have a source of exogenous variation in the incidence of suppression holding economic resources constant, we should interpret this specification with caution.

Our measure of preventiveness is mechanically related to the occurrence of protest. It is therefore important to verify that the patterns we find in Table 4 are not simply driven by an association between economic variables and the incidence of protest. Column (7) of Table 4 investigates this possibility by adding to the specification in column (6) a further control for the incidence of protest. Relative to column (6), the estimate of $\gamma^{p,g,x}$ in column (7) economically smaller but more statistically precise, and highly statistically significant, whereas the estimate of $\gamma^{p,g,-x}$ is economically larger, more precise, and statistically significant. Of course, as we do not have a source of exogenous variation in the incidence of protest holding economic resources constant, we should again interpret this specification with caution.

The effect of economic resources on repression may differ across different forms of repression. For

Table 5: Economic Effects of Commodity Price Shocks, Country-Year Panel

	log(Government revenue) (1)	log(GDP) (2)	log(Government revenue) (3)	log(GDP) (4)
log(Commodity export price)				
Contemporaneous	0.815	-0.029	0.781	-0.003
	(0.261)	(0.026)	(0.261)	(0.030)
	[0.355, 1.228]	[-0.082, 0.022]	[0.273, 1.247]	[-0.060, 0.057]
One-year lag			0.347	0.085
			(0.092)	(0.048)
			[0.167, 0.531]	[-0.011, 0.175]
Two-year lag			0.140	0.102
			(0.114)	(0.032)
			[-0.081, 0.338]	[0.042, 0.159]
Number of country-years	729	729	729	729
Number of countries	90	90	90	90

Notes: Column (1) estimates Equation (5). Column (2) replaces log(government revenue) with log(GDP) as the dependent variable. Columns (3) and (4) add lagged independent variables. All models are estimated in annual differences. The commodity export price is an average of commodity prices weighted by a backward-looking average of each commodity's export share in the country's GDP. This variable is a price index; all other economic variables are in 2015 US dollars. Below each coefficient we report an asymptotic standard error in parentheses, and a 95% Bayes bootstrap credible interval in brackets, both clustered by country.

example, while greater economic resources for the regime might encourage the regime to suppress protest more to protect its rents, the greater use of prevention associated with greater resources might entail less protest and hence less need to suppress, thus leading to a more muted (or even different-signed) relationship between resources and suppression than between resources and prevention. Our indices allow us to explore this possibility directly.

Supplemental Appendix Table 7 presents analogues of columns (1) through (5) of Table 4, but using the incidence of suppression as the dependent variable. For many coefficients of interest, the point estimates have the opposite sign from those for preventiveness, though they are generally not statistically significant. Separating the two forms of repression thus allows us to uncover effects that might be obscured by considering them jointly.

5.4 Improving Precision by Analyzing Relationships at the Monthly Level

Using annual variation limits statistical precision and obscures the important month-to-month dynamics in repression that are visible in Section 4.1. Although for most countries in our sample we cannot observe all of the variables in our analysis subannually, we *can* observe commodity price shocks at the monthly level. Because the preceding analysis shows that commodity export prices are an important driver of government revenues and of prevention, we proceed here to analyze the relationship between commodity prices and prevention at the monthly level.

To study this relationship we specify and estimate the following model:

$$1-\overline{p}_{ik} = \phi^{p}_{im(k)} + \varphi^{p}_{k} + \gamma^{p,x} x_{ik} + \varepsilon^{p}_{ik} \tag{7}$$

where $\phi^{p}_{im(k)}$ is a country-and-month-of-year fixed effect (to allow for seasonality), φ^{p}_{k} is a calendar month fixed effect, x_{ik} is the log of the index of commodity export prices, $\gamma^{p,x}$ is a coefficient, and ε^{p}_{ic} is an error term.

Panel A of Table 6 presents an OLS estimate of Equation (7). We estimate a positive and statistically significant coefficient $\gamma^{p,x}$. Its value implies that a 10 log-point increase in commodity export prices increases preventiveness by 4.5 percentage points. If we assume that the relationship between government revenues and commodity export prices from column (1) of Table 5 applies at a monthly horizon, then our estimate of Equation (7) implies an estimate of the effect $\gamma^{p,g,x}$ of commodity-driven variation in government revenues of 0.553, which is larger than—but not statistically different from—the estimate of 0.486 in column (4) of Table 4.

Panel B of Table 6 depicts an event-study plot in annual differences (Freyaldenhoven et al., 2019, Section III.B). Changes in commodity export prices are estimated to affect preventiveness over the year that they occur, consistent with the contemporaneous effect on government revenues that we observe in Table 5.

Bazzi and Blattman (2014, Section II), among others, note that some countries may play a large enough role in commodity markets that commodity prices can no longer be taken as exogenous with respect to these countries' economic outcomes (though see also Kebhaj and Gruss 2019, Section 4.1.3). Panel C of Table 6 explores the sensitivity of our findings in Panel A of Table 6 to successively dropping those countries that account for a large proportion of the export market in a given commodity. To do this, we use data from the UN Comtrade database (United Nations Statistics Division, 2024) to construct an annual panel of each country's export values in each of 36 commodities over our sample period. We then calculate each country's maximum export share of the global exports across commodities and years, and successively exclude countries from our analysis in descending order of this maximum export share. Although the statistical precision of the estimate necessarily declines as we remove countries from the sample, the point estimate is fairly stable, and the statistical significance is preserved even when a large number of countries is dropped.

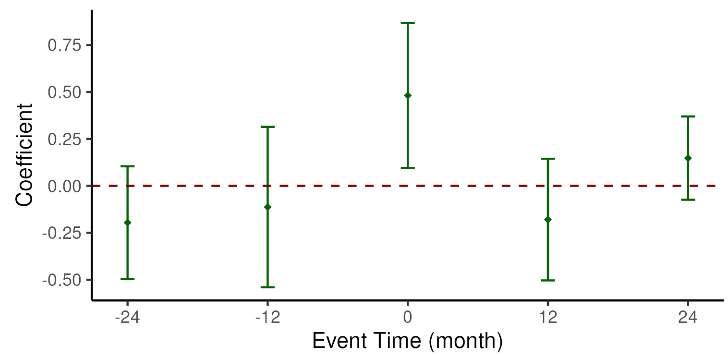
Table 4 showed that the key correlations in the country-year panel survive controlling for protest incidence. The greater precision of the monthly analysis allows us to go further. In Panel D of Table 6, we successively drop countries from the sample in ascending order of the share of sample months with

Table 6: Economic Determinants of Prevention, Country-Month Panel

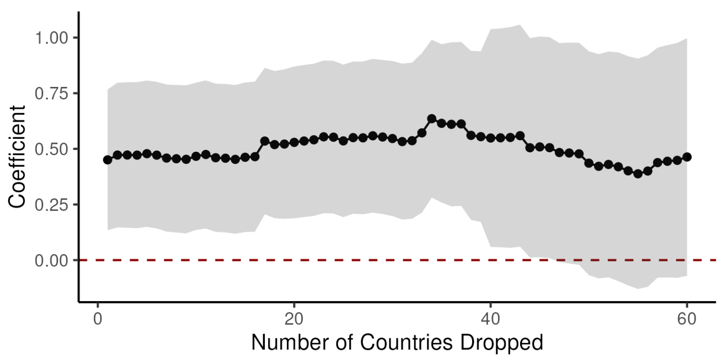
Panel A: Impact of Commodity Export Prices

	Preventiveness
	(1)
log(Commodity export price)	0.451
	(0.159)
	[0.138, 0.771]
Number of country-months	8,748
Number of countries	90

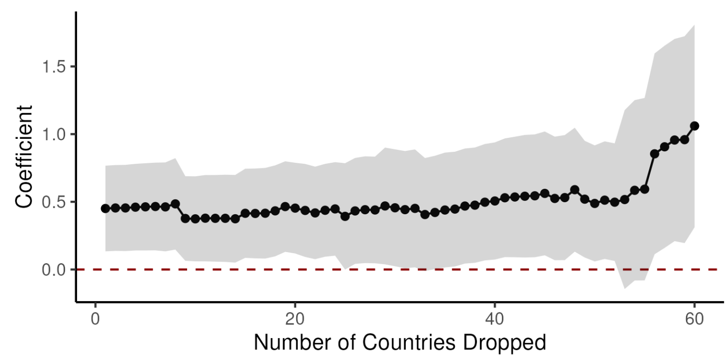
Panel B: Dynamics of Commodity Export Price Effects



Panel C: Sensitivity to Excluding Large Exporters



Panel D: Sensitivity to Excluding Low Protest Incidence



Notes: Panel A presents an estimate of Equation (7) in annual differences. "Commodity export price" refers to a price index aggregating prices of 45 commodities weighted by each commodity's share in a country's exports and the country's export share of its GDP. Below the estimated coefficient, we report an asymptotic standard error in parentheses and a 95% Bayes bootstrap credible interval in brackets, both clustered by country. Panel B plots estimates from $\Delta(1-\overline{p}_{ik}) = \Delta\varphi_k^p + \sum_{l=-2}^{2}\gamma_{-12\ell}^{p,x}\Delta x_{i,k+12\ell} + \Delta\varepsilon_{ik}^p$, where Δ is the annual difference operator, with bars representing 95% pointwise confidence intervals based on asymptotic inference clustered by country. Panel C shows the changes in the point estimate and asymptotic 95% confidence interval for the coefficient in Panel A as we successively exclude countries from the analysis in descending order of their maximum share of world exports across all commodity-years during the sample period. We construct each country's share of global direct exports for each commodity-year using data from the UN Comtrade database (United Nations Statistics Division, 2024). We map each individual good's SITC Revision 1 code to 65 common commodity categories defined in Bazzi and Blattman (2014) using a crosswalk modified based on World Bank and IMO (2022). We then conduct our sensitivity analysis using the 36 commodity groups that are in both Bazzi and Blattman (2014) and Kebhaj and Gruss (2019). Panel D shows the changes in the point estimate and asymptotic 95% confidence interval for the coefficient in Panel A as we successively exclude countries from the analysis in ascending order of the share of months with protest during the sample period.

protest, thus excluding countries in which protest is relatively more rare, and decreasing the mechanical connection between prevention and occurrence. Again, we see that the point estimate is fairly stable, and the statistical significance is preserved even when many countries are dropped, though of course the precision declines as the sample becomes smaller.

The greater precision of the analysis at the monthly level also allows us to study heterogeneity in parameters. Supplemental Appendix S.E shows suggestive evidence that the effects of commodity prices on preventiveness are concentrated in regimes that face a stronger opposition. This finding is consistent with a mechanism whereby greater rents for the regime create a greater incentive to prevent protest, the more so in environments where the opposition is better able to capitalize on protest to threaten the regime.

Across a range of specifications, then, we find clear evidence that greater government resources are associated with greater preventiveness, consistent with a mechanism in which greater resources both enable and encourage regimes to assert more control over dissent. We are able to document this finding because we can separately measure different forms of repression and because we can measure them subannually. That said, the analysis has important limitations, such as the relatively few economic variables measured subannually, and the relatively short panel, which among other things means we rely heavily on period fixed effects to account for aggregate changes. Addressing these and other limitations seems like a fruitful direction for future research.

Appendix: Proofs

Proof of Proposition 1

We will show the existence of a unique cutoff $\hat{\mu}$ for the opposition; the uniqueness of the regime's cutoff is immediate from the arguments in the text. The text addresses the case of $\overline{p}^* = 1$; here we consider the remaining case of $0 < \overline{p}^* < 1$. Because the regime's beliefs depend on the opposition's strategy, write $\hat{p}(x;\hat{\mu})$. Then, $\hat{\mu}$ constitutes an equilibrium if and only if it satisfies $\hat{\mu} = B^* \Pr(\hat{p}(x;\hat{\mu}) \leq \overline{p}^* | m=1)$. By Bayes' rule, $\hat{p}(x;0) = 0$, so that $B^* \Pr(\hat{p}(x;0) \leq \overline{p}^* | m=1) > 0$ and therefore $\hat{\mu} = 0$ is not an equilibrium. Moreover, $B^* \geq B^* \Pr(\hat{p}(x;\hat{\mu}) \leq \overline{p}^* | m=1)$ for any $\hat{\mu}$, so that $\hat{\mu} \geq B^*$ is not an equilibrium. It follows that if $B^* \Pr(\hat{p}(x;\hat{\mu}) \leq \overline{p}^* | m=1)$ is decreasing and continuous in $\hat{\mu}$, then there is a unique $\overline{\mu}^* \in (0, B^*]$ that satisfies $\overline{\mu}^* = B^* \Pr(\hat{p}(x;\overline{\mu}^*) \leq \overline{p}^* | m=1)$, and a unique equilibrium belief of the regime $p^*(x) = \hat{p}(x;\overline{\mu}^*)$.

It remains to show that $B^* \Pr(\hat{p}(x;\hat{\mu}) \leq \overline{p}^* | m=1)$ is decreasing and continuous in $\hat{\mu}$. For x such that $\Pr(\omega = 1 | x) = 0$, we have that $\hat{p}(x;\hat{\mu}) = 0$ for any $\hat{\mu}$. For all other x, $\hat{p}(x;\hat{\mu})$ is interior, weakly increasing, and continuous in $\hat{\mu}$.[36] Thus, $\Pr(\hat{p}(x;\hat{\mu}) \leq \overline{p}^*|m=1)$ is decreasing in $\hat{\mu}$ and, because we rule out mass points in the distribution of $\hat{p}(x;\hat{\mu})$, it is also continuous, which completes the proof.

Proof of Proposition 2

From the arguments in the text, all that remains to show is that $\overline{p}=\overline{p}^*$ when $\chi(\cdot)$ is one-to-one. To see this, note that, for any $x\in\mathscr{X}_\omega$ that arises only when citizens are aggrieved, Bayes' rule implies that

$$p^*(x)=\frac{\Lambda(x|m=1,\omega=1)\Pr(\mu<\overline{\mu}^*)}{\Lambda(x|m=1,\omega=1)\Pr(\mu<\overline{\mu}^*)+\Lambda(x|m=0,\omega=1)(1-\Pr(\mu<\overline{\mu}^*))},$$

where we suppress conditioning on $a=0$ in describing the density $\Lambda(\cdot)$ for notational ease. By the assumption that the image of the support of x in $\lambda(\cdot)$ on $\mathscr{X}_\omega$ is $(0,\infty)$, we have that there is some x^* in the support of $\mathscr{X}$ for which $p^*(x)=\overline{p}^*$, which completes the proof.

Proof of Proposition 3

Part 1. Write $\tilde{L}(y,s;\sigma)=-s\sigma-(1-s)q^*(y)L$ for the regime's expected payoff in the second stage of the game given information $y\in\mathscr{Y}$, and $s^*(y;\sigma)$ for its optimal action given our tiebreaking assumption. Pick some $\sigma'>\sigma$. We know that $\tilde{L}(y,s^*(y;\sigma);\sigma)\geq\tilde{L}(y,s^*(y;\sigma');\sigma)\geq\tilde{L}(y,s^*(y;\sigma');\sigma')$, where the first inequality follows from the optimality of $s^*(y;\sigma)$ and the second from the definition of $\tilde{L}(y,s;\sigma)$. We also know that $s^*(y;\sigma)$ is monotone in σ so that $\tilde{L}(y,s^*(y;\sigma);\sigma)>\tilde{L}(y,s^*(y;\sigma');\sigma')$ whenever $s^*(y;\sigma)=1$. It follows that $L^*(\sigma)=-\mathrm{E}_y\left(\tilde{L}(y,s;\sigma)\right)$ is weakly increasing in σ, with $L^*(\sigma)<L^*(\sigma')$ whenever $s^*(y;\sigma)=1$ with positive probability, which in turn is true whenever $\overline{q}^*<1$. The result then follows because $\overline{p}^*=\min\{\rho/L^*,1\}$.

Part 2. We show that Z^*S^* is increasing in ρ, strictly so when $\overline{p}^*$ is interior. Observe that the outcomes of the second stage, including S^* and B^*, are unaffected by ρ. If $S^*>0$, it therefore suffices to show that Z^* is increasing in ρ, strictly so whenever $\overline{p}^*<1$. To see this, observe that

$$Z^*=\Pr(\omega=1)\Pr(\mu<\overline{\mu}^*\,|\,\omega=1)\Pr(\hat{p}(x;\overline{\mu}^*)\leq\overline{p}^*|m=\omega=1)$$

where $\overline{\mu}^*$ solves

$$\overline{\mu}^*=B^*\Pr(\hat{p}(x;\overline{\mu}^*)\leq\overline{p}^*|m=\omega=1).$$

Recall also that $\overline{p}^*=\min\{\rho/L^*,1\}$ is increasing in ρ, strictly so when it is interior. It follows that $\overline{\mu}^*$, and hence Z^*, are increasing in ρ, strictly so when $\overline{p}^*$ is interior.

[36] If $x\in\mathscr{X}$ is consistent with $\omega=1$, then by Bayes' rule

$$\hat{p}(x;\hat{\mu})=\frac{\Lambda(x|m=1,\omega=1)\Pr(\mu<\hat{\mu})\Pr(\omega=1)}{\Lambda(x|m=1,\omega=1)\Pr(\mu<\hat{\mu})\Pr(\omega=1)+\Lambda(x|m=0,\omega=1)(1-\Pr(\mu<\hat{\mu}))\Pr(\omega=1)+\Lambda(x|m=0,\omega=0)\Pr(\omega=0)},$$

which is strictly greater than zero and strictly less than one because $\lambda(x)\in(0,\infty)$ and $\Pr(\omega=1),\Pr(\mu<\hat{\mu})>0$, and is weakly increasing and continuous in $\hat{\mu}$ because $\Pr(\mu<\hat{\mu})$ has these properties. Notice that we have suppressed conditioning on $a=0$ in describing the density $\Lambda(\cdot)$ for notational ease.

References

Abrahamian, E. (1999). *Tortured Confessions: Prisons and Public Recantations in Modern Iran*. Berkeley: University of California Press.

Acemoglu, D., G. Egorov, and K. Sonin (2015). Political economy in a changing world. *Journal of Political Economy 123*(5), 1038–1086.

Acemoglu, D., T. A. Hassan, and A. Tahoun (2018). The power of the street: Evidence from Egypt's Arab Spring. *Review of Financial Studies 31*(1), 1–42.

Adamovsky, E. (2024). *A History of Argentina: From the Spanish Conquest to the Present*. Durham: Duke University Press. Translated by Rebecca Wolpin.

Aghion, P., G.-M. Angeletos, A. Banerjee, and K. Manova (2010, April). Volatility and growth: Credit constraints and the composition of investment. *Journal of Monetary Economics 57*(3), 246–265.

Aktan, D. (2022a). How does violence deter? Functional and informational effects of preemptive repression. *Working Paper*. Accessed August 2023.

Aktan, D. (2022b). Signaling and screening in contentious politics. *Working Paper*. Accessed August 2023.

Al Jazeera (2013a, July). Mali lifts state of emergency. Accessed October 2024.

Al Jazeera (2013b, November). Mali's Tuareg fighters end ceasefire. Accessed November 2024.

Alimagham, P. (2020). *Contesting the Iranian Revolution: The Green Uprisings*. Cambridge: Cambridge University Press.

Andrews, I. and J. M. Shapiro (2024, January). Bootstrap diagnostics for irregular estimators. Working Paper 32038, National Bureau of Economic Research.

Archibong, B., T. Moerenhout, and E. Osabuohien (2022). Protest matters: The effects of protests on economic redistribution. *Brookings Global Working Paper 172*. Accessed August 2023.

Arnon, D., P. Haschke, and B. Park (2023). The right accounting of wrongs: Examining temporal changes to human rights monitoring and reporting. *British Journal of Political Science 53*(1), 163–182.

Ascher, A. (2004). *The Revolution of 1905: A Short History*. Stanford: Stanford University Press.

Asher, S. and P. Novosad (2023, January). Rent-Seeking and Criminal Politicians: Evidence from Mining Booms. *Review of Economics and Statistics 105*(1), 20–39.

Associated Press in Bangkok (2015, April). Thailand 'still in the same boat' after martial law lifted. Accessed October 2024.

Bagozzi, B. E., S. Chatterjee, and U. K. Mukherjee (2019). Predicting anti-government violence in Mexico with big data on citizen-government interactions. Accessed August 2020.

Bajoria, J. (2023). *"No Internet Means No Work, No Pay, No Food": Internet Shutdowns Deny Access to Basic Rights in "Digital India"*. Human Rights Watch. Accessed November 2024.

Baliga, S. and T. Sjöström (2023). Optimal compellence. Accessed October 2024.

Barigazzi, J. and J. H. Vela (2022). EU's unanimity rules are here for now, despite chatter. *Politico*

Europe 20.

Barrett, P., M. Appendino, K. Nguyen, and J. de Leon Miranda (2022). Measuring social unrest using media reports. *Journal of Development Economics 158*, 102924.

Bazzi, S. and C. Blattman (2014, October). Economic Shocks and Conflict: Evidence from Commodity Prices. *American Economic Journal: Macroeconomics 6*(4).

BBC (2010, September). Thai red-shirts stage defiant protest. Accessed October 2024.

Bearak, M. (2018). At least 3 dead after military opens fire on post-election protests in Zimbabwe. *Washington Post*. Accessed August 2023.

Benkard, C. L., A. Bodoh-Creed, and J. Lazarev (2018). Simulating the dynamic effects of horizontal mergers: US airlines. *Manuscript, Stanford University*. Accessed November 2024.

Beraja, M., A. Kao, D. Y. Yang, and N. Yuchtman (2023, March). AI-tocracy. *Quarterly Journal of Economics 138*(3), 1349–1402.

Bueno de Mesquita, B., A. Smith, R. M. Siverson, and J. D. Morrow (2003). *The Logic of Political Survival*. Cambridge: MIT Press.

Burke, J. (2018). Zimbabwe election unrest turns deadly as army opens fire on protesters. *The Guardian*. Accessed August 2023.

Burke, P. J. and A. Leigh (2010). Do output contractions trigger democratic change? *American Economic Journal: Macroeconomics 2*(4), 124–157.

Campante, F. and D. Yanagizawa-Drott (2018). Long-range growth: Economic development in the global network of air links. *Quarterly Journal of Economics 133*(3), 1395–1458.

Cantoni, D., A. Kao, D. Y. Yang, and N. Yuchtman (2024). Protests. *Annual Review of Economics 16*.

Carey, P. D., C. Bell, E. H. Ritter, and S. Wolford (2022). Oil discoveries, civil war, and preventive state repression. *Journal of Peace Research 59*(5), 648–662.

Carey, S. C. (2006). The dynamic relationship between protest and repression. *Political Research Quarterly 59*(1), 1–11.

Carey, S. C. (2009). *Protest, Repression and Political Regimes: An Empirical Analysis of Latin America and sub-Saharan Africa*. London and New York: Routledge.

Carey, S. C. and A. R. Gohdes (2021). Understanding journalist killings. *Journal of Politics 83*(4), 1216–1228.

Carment, David (principal investigator) (2010). Country Indicators for Foreign Policy Project. Accessed August 2020.

Carter, E. B. and B. L. Carter (2022). When autocrats threaten citizens with violence: Evidence from China. *British Journal of Political Science 52*(2), 671–696.

Caselli, F. and A. Tesei (2016, July). Resource windfalls, political regimes, and political stability. *Review of Economics and Statistics 98*(3), 573–590.

Centre for Social Conflict and Cohesion Studies (2020). *Observatory of Conflicts - Cumulative Dataset*.

Harvard Dataverse. Accessed April 2024.

Chau, T.-H., M. Hassan, and A. T. Little (2023). Communication and coordination in the shadow of repression. Accessed October 2024.

Chilton, A. S. and M. Versteeg (2015). The failure of constitutional torture prohibitions. *Journal of Legal Studies 44*(2), 417–452.

Cingranelli, D. and M. Filippov (2018). Are human rights practices improving? *American Political Science Review 112*(4), 1083–1089.

Cingranelli, D. L., D. L. Richards, and K. C. Clay (2014b). The CIRI Human Rights Dataset, Version 2014.04.14. Accessed August 2019.

CNN Wire Staff (2011). Zimbabwe sanctions to remain in place, says U.S. official. *CNN World*. Accessed August 2023.

Cole, P. and B. McQuinn (Eds.) (2015). *The Libyan Revolution and its Aftermath*. New York: Oxford University Press.

Collins, W. J. and R. A. Margo (2007). The economic aftermath of the 1960s riots in American cities: Evidence from property values. *Journal of Economic History 67*(4), 849–883.

Conrad, C. R. and W. H. Moore (2010). What stops the torture? *American Journal of Political Science 54*(2), 459–476.

Crisis24 (2022). Alerts database 2009-2022. Transmitted to authors in September 2020, October 2020, November 2020, and November 2022.

Crowley, P. (2011). Detention and allegations of torture of political activists in Zimbabwe. *U.S. Department of State*. Accessed August 2023.

Danneman, N. and E. H. Ritter (2014). Contagious rebellion and preemptive repression. *Journal of Conflict Resolution 58*(2), 254–279.

Davenport, C. (2007). State repression and political order. *Annual Review of Political Science 10*(1), 1–23.

Davenport, C., H. Mokleiv Nygård, H. Fjelde, and D. Armstrong (2019). The consequences of contention: Understanding the aftereffects of political conflict and violence. *Annual Review of Political Science 22*(1), 361–377.

De Jaegher, K. and B. Hoyer (2019). Preemptive repression: Deterrence, backfiring, iron fists, and velvet gloves. *Journal of Conflict Resolution 63*(2), 502–527.

Deng, S., H. Rangwala, and Y. Ning (2019). Learning dynamic context graphs for predicting social events. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage AK USA, pp. 1007–1016. ACM.

Donno, D. (2013). *Defending Democratic Norms: International Actors and the Politics of Electoral Misconduct*. New York: Oxford University Press.

Dragu, T. and A. Przeworski (2019). Preventive repression: Two types of moral hazard. *American Political Science Review 113*(1), 77–87.

Drum Cussac (2018). RiskMonitor Pro. Accessed November 2024.

Dube, O. and J. F. Vargas (2013, October). Commodity price shocks and civil conflict: Evidence from Colombia. *Review of Economic Studies 80*(4), 1384–1421.

Earl, J. (2011). Political repression: Iron fists, velvet gloves, and diffuse control. *Annual Review of Sociology 37*(1), 261–284.

Earl, J., T. V. Maher, and J. Pan (2022). The digital repression of social movements, protest, and activism: A synthetic review. *Science Advances 8*(10), eabl8198.

Egorov, G., S. Guriev, and K. Sonin (2009). Why resource-poor dictators allow freer media: A theory and evidence from panel data. *American Political Science Review 103*(4), 645–668.

Enikolopov, R., A. Makarin, and M. Petrova (2020). Social media and protest participation: Evidence from Russia. *Econometrica 88*(4), 1479–1514.

Esberg, J. (2021). Anticipating dissent: The repression of politicians in Pinochet's Chile. *Journal of Politics 83*(2), 689–705.

Fariss, C. J. (2014). Respect for human rights has improved over time: Modeling the changing standard of accountability. *American Political Science Review 108*(2), 297–318.

Fariss, C. J. (2019). Yes, human rights practices are improving over time. *American Political Science Review 113*(3), 868–881.

Felbermayr, G., A. Kirilakha, C. Syropoulos, E. Yalcin, and Y. V. Yotov (2020). The Global Sanctions Data Base. *European Economic Review 129*, 103561.

Fitzpatrick, S. (2017). *The Russian Revolution*. New York: Oxford University Press.

Fording, R. C. (2001). The political response to black insurgency: A critical test of competing theories of the state. *American Political Science Review 95*(1), 115–130.

Franklin, J. C. (2008). Shame on you: The impact of human rights criticism on political repression in Latin America. *International Studies Quarterly 52*(1), 187–211.

Freedom House (2010-2019a). *Freedom in the World: The Annual Survey of Political Rights and Civil Liberties*. Lanham, Maryland: Rowman & Littlefield. Accessed April 2024.

Freedom House (2019b). Freedom on the net 2019: The crisis of social media. Accessed November 2024.

Freedom House (2021c). Freedom in the world 2021. Accessed August 2023.

Freedom House (2023). Freedom on the net 2023: The repressive power of artificial intelligence. Accessed November 2024.

Freyaldenhoven, S., C. Hansen, and J. M. Shapiro (2019, September). Pre-event trends in the panel event-study design. *American Economic Review 109*(9), 3307–3338.

GardaWorld (2024). The DNA of a dynasty: Our success story. Accessed November 2024.

GDELT Project (2020). GDELT 1.0 Event Database. Accessed June 2020.

GDELT Project (2023). GDELT 2.0 Event Database. Accessed April 2023.

GDELT Project (2024). GDELT Global Knowledge Graph. Accessed November 2024.

Gehlbach, S., K. Sonin, and M. W. Svolik (2016). Formal models of nondemocratic politics. *Annual Review of Political Science 19*, 565–584.

Gibilisco, M. (2021). Decentralization, repression, and gambling for unity. *Journal of Politics 83*(4), 1353–1368.

Global Disaster Alert and Coordination System (2024). GDACS - Global Disaster Alert and Coordination System. Accessed January 2024.

Głowacka, D., R. Youngs, A. Pintea, and E. Wołosik (2021). Digital technologies as a means of repression and social control. Accessed October 2024.

Greitens, S. C. (2016). *Dictators and Their Secret Police: Coercive Institutions and State Violence*. Cambridge; New York: Cambridge University Press.

Greitens, S. C. (2022, December). After a wave of protests, China's silent crackdown. *Journal of Democracy*. Accessed November 2024.

GSDB (2023). The Global Sanctions Data Base. Accessed August 2023.

Gueorguiev, D. D. (2017). In public or in private: Repressing dissidents in China. *Working Paper*. Accessed August 2023.

Gunitsky, S. (2015). Corrupting the cyber-commons: Social media as a tool of autocratic stability. *Perspectives on Politics 13*(1), 42–54.

Guriev, S. and D. Treisman (2022). *Spin Dictators: The Changing Face of Tyranny in the 21st Century*. Princeton: Princeton University Press.

Harlan, C. and L. Samuels (2014, May). Thai military declares a coup, detains key political leaders. *Washington Post*. Accessed November 2024.

Hendrix, C. S. and I. Salehyan (2015). No news is good news: Mark and recapture for event data when reporting probabilities are less than one. *International Interactions 41*(2), 392–406.

Hirano, K. and J. R. Porter (2012). Impossibility results for nondifferentiable functionals. *Econometrica 80*(4), 1769–1790.

Hoegh, A. (2019). Model fusion with spatial partitioning for forecasting civil unrest. Presentation at the Joint Statistical Meeting, Activity Number 320 (Statistical Approaches for Modeling Social Unrest).

Hoegh, A., M. A. Ferreira, and S. Leman (2016). Spatiotemporal model fusion: Multiscale modelling of civil unrest. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 65*(4), 529–545.

Hoegh, A., S. Leman, P. Saraf, and N. Ramakrishnan (2015). Bayesian model fusion for forecasting civil unrest. *Technometrics 57*(3), 332–340.

Human Rights Watch (2011). Universal periodic review of Zimbabwe. Accessed August 2023.

Human Rights Watch (2018). Zimbabwe: At least 6 dead in post-election violence. Accessed August 2023.

Human Rights Watch (2024). Human Rights Watch News. Accessed October 2024.

International Crisis Group (2014, January). Mali: Reform or relapse. Accessed November 2024.

International Monetary Fund (2024, July). Commodity Terms of Trade. Accessed July 2024.

International Network of Civil Liberties Organizations (2013). Take back the streets: Repression and criminalization of protest around the world. Accessed October 2024.

International Organization for Standardization (2021). ISO 3166-1 alpha-2. datahub.io [distributor].

Jenkins, J. C. and T. V. Maher (2016). What should we do about source selection in event data? Challenges, progress, and possible solutions. *International Journal of Sociology 46*(1), pp. 42–57.

Jones, M. O. (2015). Social media, surveillance and cyber-politics in the Bahrain uprising. In A. Shehabi and M. O. Jones (Eds.), *Bahrain's Uprising: Resistance and Repression in the Gulf*, pp. 239–62. London: Zed Books.

Jones, M. O. (2020). *Political Repression in Bahrain*. Cambridge: Cambridge University Press.

Jordan, M. (2006, August). Going mobile: Text messages guide Filipino protesters. *Washington Post*. Accessed October 2024.

Kadivar, J. (2015, May). A comparative study of government surveillance of social media and mobile phone communications during Iran's green movement (2009) and the UK riots (2011). *tripleC: Communication, Capitalism & Critique 13*(1), 169–191.

Kadivar, M. A., N. Ketchley, A. Sotoudeh-Sherbaf, and C. Barrie (2024, April). Online calls for protest and offline mobilization in autocracies: evidence from the 2017 Dey Protests in Iran. *European Sociological Review*.

Kadivar, M. A., S. Khani, V. Abedini, and S. Barzin (2023). Fueling the flames of dissent: Oil, grievances, and protest in iran 2019. Unpublished manuscript.

Kebhaj, S. and B. Gruss (2019). Commodity Terms of Trade: A New Database.

Keith, L. C. and S. C. Poe (2004). Are constitutional state of emergency clauses effective? An empirical exploration. *Human Rights Quarterly 26*(4), 1071–1097.

Keith, L. C., C. N. Tate, and S. C. Poe (2009). Is the law a mere parchment barrier to human rights abuse? *Journal of Politics 71*(2), 644–660.

Kelley, J. G. and B. A. Simmons (2015). Politics by number: Indicators as social pressure in international relations. *American Journal of Political Science 59*(1), 55–70.

Ketchley, N. (2017). *Egypt in a Time of Revolution: Contentious Politics and the Arab Spring*. Cambridge: Cambridge University Press.

Kirilakha, A., G. Felbermayr, C. Syropoulos, E. Yalcin, and Y. V. Yotov (2021). The Global Sanctions Data Base (GSDB): An update that includes the years of the Trump presidency. In P. A. G. van Bergeijk (Ed.), *Research handbook on economic sanctions*, pp. 62–106. Cheltenham, UK: Edward Elgar Publishing.

Kosolpradit, S. (2010, October). Thai "red shirts" rally to demand leaders be freed. *Reuters*. Accessed November 2024.

Kryzhanouski, Y. (2022). Managing dissent in post-Soviet authoritarianism. New censorship of protest music in Belarus and Russia, 2000–2018. *Europe-Asia Studies 74*(5), 760–788.

Kuran, T. (1989). Sparks and prairie fires: A theory of unanticipated political revolution. *Public Choice 61*(1), 41–74.

Kuran, T. (1991). Now out of never: The element of surprise in the East European revolution of 1989. *World Politics 44*(1), 7–48.

Landman, T. (2004). Measuring human rights: Principle, practice, and policy. *Human Rights Quarterly 26*(4), 906–931.

Langørgen, E. (2016). The revolution will not be organized? Organized and spontaneous onsets of nonviolent campaigns. Master's thesis, University of Oslo.

Lei, L. (2023). Distribution-free inference on the extremum of conditional expectations via classification. *Working Paper*. Accessed August 2023.

Lei, L., A. D'Amour, P. Ding, A. Feller, and J. Sekhon (2021). Distribution-free assessment of population overlap in observational studies. *Working Paper*. Accessed August 2023.

Leung, H. (2019, December). Hong Kong's famed New Year's fireworks show is the latest big event canceled over protest fears. *Time*. Accessed November 2024.

Lewbel, A. (2019). The identification zoo: Meanings of identification in econometrics. *Journal of Economic Literature 57*(4), 835–903.

Lewis, D. K. (2003). *The History of Argentina*. New York: St. Martin's Griffin.

Li, Y. (2018). *Playing by the Informal Rules: Why the Chinese Regime Remains Stable Despite Rising Protests*. Cambridge: Cambridge University Press.

Licht, A. A. and S. H. Allen (2018). Repressing for reputation: Leadership transitions, uncertainty, and the repression of domestic populations. *Journal of Peace Research 55*(5), 582–595.

Logan, C. and R. Mattes (2012). Democratising the measurement of democratic quality: Public attitude data and the evaluation of African political regimes. *European Political Science 11*(4), 469–491.

Lyall, J. (2009). Does indiscriminate violence incite insurgent attacks? Evidence from chechnya. *Journal of Conflict Resolution 53*(3), 331–362.

Manacorda, M. and A. Tesei (2020). Liberation technology: Mobile phones and political mobilization in Africa. *Econometrica 88*(2), 533–567.

Mavhunga, C. (2011). Silent streets after apparent protest plans in Zimbabwe. *CNN World*. Accessed October 2024.

McAdam, D., S. Tarrow, and C. Tilly (2001). *Dynamics of Contention*. Cambridge: Cambridge University Press.

Merkel, Wolfgang and Bochsler, Daniel (project leaders), K. Bousbah, M. Bühlmann, H. Giebler, M. Hänni, L. Heyne, A. Juon, L. Müller, S. Ruth, and B. Wessels (2020). Democracy Barometer: Codebook Version 7. Accessed August 2020.

Milani, A. (2011). *The Shah*. New York: Macmillan.

Millennium Challenge Corporation (2020). MCC Selection Indicators. Accessed August 2020.

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

Mueller, H. and C. Rauh (2018). Reading between the lines: Prediction of political violence using newspaper text. *American Political Science Review 112*(2), 358–375.

Muradiniya, M. J. (1998). *Khumein dar Inqilab*. Tehran, Iran: ‘Uruj.

National Intelligence Council (2022). Digital repression growing globally, threatening freedoms. Accessed October 2024.

National Statistics, Republic of China (2010). Number of Resident Population and Population Density, Year End of 2010. Accessed February 2021.

Nikolayenko, O. (2017). *Youth Movements and Elections in Eastern Europe*. Cambridge University Press.

Nugent, E. R. (2020). *After Repression: How Polarization Derails Democratic Transition*. Princeton University Press.

OECD (2023, July). List of OECD member countries - ratification of the convention on the OECD. Accessed July 2023.

Open Government Partnership (2019). Open government partnership global report. Accessed August 2020.

O’Toole, B. and S. Sultoon (2019). Sanctions explained: How a foreign policy problem becomes a sanctions program. *Atlantic Council Commentary*. Accessed November 2024.

Overseas Security Advisory Council (2024). US Department of State travel advisories and US embassy and consulate alerts. Accessed January 2024.

Pakes, A., M. Ostrovsky, and S. Berry (2007). Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *RAND Journal of Economics 38*(2), 373–399.

Pemstein, D., K. L. Marquardt, E. Tzelgov, Y. Wang, J. Medzihorsky, J. Krusell, F. Miri, and J. von Römer (2023). The V-Dem measurement model: Latent variable analysis for cross-national and cross-temporal expert-coded data. *V-Dem Working Paper 21*(8).

Pickel, S., W. Breustedt, and T. Smolka (2016). Measuring the quality of democracy: Why include the citizens’ perspective? *International Political Science Review 37*(5), 645–655.

Qiao, F., P. Li, X. Zhang, Z. Ding, J. Cheng, and H. Wang (2017). Predicting social unrest events with hidden Markov models using GDELT. *Discrete Dynamics in Nature and Society 2017*.

Ramakrishnan, N., P. Butler, S. Muthiah, N. Self, R. Khandpur, P. Saraf, W. Wang, J. Cadena, A. Vullikanti, G. Korkmaz, C. Kuhlman, A. Marathe, L. Zhao, T. Hua, F. Chen, C.-T. Lu, B. Huang, A. Srinivasan, K. Trinh, L. Getoor, G. Katz, A. Doyle, C. Ackermann, I. Zavorin, J. Ford, K. Summers, Y. Fayed, J. Arredondo, D. Gupta, and D. Mares (2014). “Beating the news” with EMBERS: Forecasting civil unrest using open source indicators. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’14)*, New York, NY, USA, pp. 1799–1808. Association for Computing Machinery.

Ritter, E. H. and C. R. Conrad (2016). Preventing and responding to dissent: The observational challenges of explaining strategic repression. *American Political Science Review 110*(1), 85–99.

Robles, F. (2021, November). As Cuba crushes dissent, a nationwide protest fizzles. *New York Times*. Accessed November 2024.

Ross, C. W., A. Daughton, G. Fairchild, and S. Del Valle (2019). Forecasting political instability using heterogeneous data streams. Presentation at the Joint Statistical Meeting. Activity Number 320 (Statistical Approaches for Modeling Social Unrest).

Rozenas, A., S. Schutte, and Y. Zhukov (2017). The political legacy of violence: The long-term impact of Stalin's repression in Ukraine. *Journal of Politics 79*(4), 1147–1161.

ScaleUp Capital (2019, November). Drum Cussac acquired by GardaWorld. Accessed November 2024.

Scharpf, A., C. Gläßel, and P. Edwards (2023). International sports events and repression in autocracies: Evidence from the 1978 FIFA World Cup. *American Political Science Review 117*(3), 909–926.

Shehabi, A. and M. O. Jones (Eds.) (2015). *Bahrain's Uprising: Resistance and Repression in the Gulf*. New York: Bloomsbury Publishing.

Sonin, K. and A. L. Wright (2023). Rebel capacity, intelligence gathering, and combat tactics. *American Journal of Political Science*.

Sullivan, C. M. (2016). Political repression and the destruction of dissident organizations: Evidence from the archives of the Guatemalan National Police. *World Politics 68*(4), 645–676.

The Economist (2022a, June). Costly food and energy are fostering global unrest. *The Economist*.

The Economist (2022b, June). A wave of unrest is coming. Here's how to avert some of it. *The Economist*.

Tilly, C. (Ed.) (1975). *The Formation of National States in Western Europe.* Princeton: Princeton University Press.

Tilly, C. (1978). *From Mobilization to Revolution*. Reading: Addison-Wesley Publishing.

Tilly, C. (1995). *Popular Contention in Great Britain, 1758-1834*. Cambridge: Harvard University Press.

Tilly, C. (2006). *Regimes and Repertoires*. Chicago: University of Chicago Press.

Truex, R. (2019). Focal points, dissident calendars, and preemptive repression. *Journal of Conflict Resolution 63*(4), 1032–1052.

United Nations Statistics Division (2024). UN Comtrade Database. Accessed: May 2024.

U.S. Department of State (2021). Sanctioning Cuban police in response to violent repression of peaceful protests. Accessed November 2024.

U.S. Department of the Treasury (2020). Treasury sanctions two individuals for human rights abuse in Zimbabwe. Accessed August 2023.

U.S. Department of the Treasury (2021a). Treasury sanctions Cuban Ministry of Interior officials and military unit in response to violence against peaceful demonstrators. Accessed November 2024.

U.S. Department of the Treasury (2021b). Treasury sanctions senior Cuban officials in response to violence against peaceful demonstrators. Accessed November 2024.

Voice of America News (2010, October). Thailand's 'red shirts' rally in defiance of emergency decree. Accessed November 2024.

Wehrey, F. M. (2014). *Sectarian Politics in the Gulf: From the Iraq War to the Arab Uprisings*. New York: Columbia University Press.

Weidmann, N. B. and E. G. Rød (2019). Coding protest events in autocracies. In N. B. Weidmann and E. G. Rød (Eds.), *The Internet and Political Protest in Autocracies*. New York: Oxford University Press.

World Bank (2024, September). World Development Indicators DataBank.

World Bank and IMO (2022). *Economic Shocks and Human Trafficking Risks: Evidence from IOM's Victims of Human Trafficking Database*.

Wu, C. and M. S. Gerber (2018). Forecasting civil unrest using social media and protest participation theory. *IEEE Transactions on Computational Social Systems 5*(1), 82–94.

Zhao, L. (2022). Event prediction in the big data era: A systematic survey. *ACM Computing Surveys 54*(5), 1–37.

Supplemental Appendix for

Surveillance of Repression

Veli M. Andirin, *Yale University**
Yusuf Neggers, *University of Michigan*
Mehdi Shadmehr, *University of North Carolina at Chapel Hill*
Jesse M. Shapiro, *Harvard University and NBER*

S.A Additional Theoretical Results

S.A.1 Dynamic Model with Time-Varying Primitives

Here we embed the game in Section 2 in a discrete-time dynamic model with time-varying primitives, and relax many of the restrictions on primitives that we introduce for concreteness in the main text. In the dynamic model the primitives in a given period t are determined by a countable state variable $k_t \in \mathbb{N}$ that we refer to as the *environment.* We will think of $k_t = 0$ as an absorbing state denoting that revolution has occurred. If revolution has not occurred after any given period t, then the environment k_t transitions to another one, k_{t+1}, according to a commonly known probability that depends only on k_t. Let $\mathbf{K}$ be the corresponding transition probability matrix. To close the model, we may suppose that in an initial period an environment is chosen at random. We assume that all random variables are i.i.d. across periods within a given environment, so that the environment fully captures the way in which the game structure evolves over time. Except where stated, random variables are independent of one another within a given period.

In any period t in which revolution has not yet occurred (so that $k_t > 0$), the following game is played.

1. Nature determines grievances $\omega_t \in \Omega(k_t)$ according to a distribution that may depend on the environment k_t.

2. The opposition observes grievances and chooses a public action $a_t \in \{0,1\}$. If $a_t = 0$, the opposition may also choose a private action $m_t \in 0 \cup \mathscr{M}(k_t)$. The private action is costless if $m_t = 0$ but otherwise entails a cost $\mu_t : \mathscr{M}(k_t) \to \mathbb{R}_{>0}$, where the realization of the random function μ_t is privately known, and is drawn according to a commonly known distribution that may depend on the environment k_t.

3. The regime observes information $x_t \in \mathscr{X}$ drawn according to a distribution that depends on ω_t, m_t, a_t, and k_t. The regime then chooses whether to prevent protest at a cost of $\rho(k_t) > 0$ to the regime that may depend on the environment.

4. If the regime has prevented protest, protest occurs with probability $\underline{p}(k_t) \in [0,1)$. Otherwise, protest occurs with probability $p(\omega_t, m_t, a_t, k_t)$,where $p(\omega', m_t, 1, k_t) = 1$ for some $\omega' \in \Omega(k_t)$.

*E-mail: velimurat.andirin@yale.edu, yneggers@umich.edu, mshadmehr@unc.edu, jesse_shapiro@fas.harvard.edu.

5. If protest does not occur, the period ends without revolution.

6. If protest occurs, the period continues.

   (a) The regime observes information $y_t \in \mathscr{Y}$ drawn according to a distribution that may depend on k_t.

   (b) The regime chooses whether to suppress protest, $s_t \in \{0,1\}$ at a cost of $\sigma(k_t) > 0$ to the regime and $\alpha(k_t)$ to the opposition.

   (c) If the regime has suppressed protest, revolution occurs with probability $\underline{q}(k_t) \in [0,1)$. Otherwise, revolution occurs with probability $q^*(y_t,k_t)$ with respect to the regime's information.

   (d) If revolution does not occur, the period ends.

   (e) If revolution occurs, the regime obtains payoff $-\ell(k_t) < 0$ forever, and the opposition obtains payoff $b(k_t) > 0$ forever.

We focus on equilibria in stationary pure strategies, resolving indifferences as in Section 2, and assuming that the regime and opposition have discount factors $\delta, \beta \in (0,1)$, respectively. We assume sufficient conditions on $\ell(\cdot)$ and $b(\cdot)$ so that the regime wishes to avoid revolution and the opposition wishes to bring it about. Online Appendix O.C provides example sufficient conditions for this property, and for existence of equilibrium in stationary pure strategies in the dynamic game that we study here. Note that, because we allow the distributions of the regime's information to depend on the environment, the assumption that the information spaces $(\mathscr{X},\mathscr{Y})$ do not depend on the environment is without loss.

The following result captures the essential aspects of equilibria for our purposes.

Proposition 4. *In any equilibrium in stationary pure strategies, there is a belief function* $p^* : \mathscr{X} \times \mathbb{N} \to [0,1]$*, and a cutoff function* $\overline{p}^* : \mathbb{N} \to (0,1]$*, such that if the regime does not prevent protest in a given period t in environment* k_t*, protest occurs with probability* $p^*(x_t,k_t)$ *with respect to the regime's information set, and the regime prevents protest if and only if* $p^*(x_t,k_t) > \overline{p}^*(k_t) \in \left(\underline{p}(k_t),1\right]$.

Proof. In any equilibrium in stationary strategies, the regime's expected discounted payoff at the start of any period t in environment $k_t > 0$ can be written as a function $V^*(k_t)$ of the environment. For any $k_t > 0$, let $\overline{V}^*(k_t) = \sum_{l \in \mathbb{N}} \mathbf{K}_{k_t,l} V^*(l)$ denote the expectation over $V^*(l)$ with respect to the transition probabilities starting from environment k_t. In the case where revolution has occurred, i.e., where $k_t = 0$, the regime's expected discounted payoffs, $V(k_t) = -\ell(k_t)/(1-\delta)$, does not depend on the equilibrium being played. With prevention the expected discounted payoff is

$$\delta\left(\underline{p}(k_t)Q^*(k_t)V(k_t) + \left(1-\underline{p}(k_t)Q^*(k_t)\right)\overline{V}^*(k_t)\right)$$
$$-\rho(k_t) - \underline{p}(k_t)\sigma(k_t)S^*(k_t),$$

where $S^*(\cdot)$ is the equilibrium probability of suppression and $Q^*(\cdot)$ is the equilibrium probability of revolution, given protest. With no prevention the expected discounted payoff is

$$\delta\left(p^*(x_t,k_t)Q^*(k_t)V(k_t)+(1-p^*(x_t,k_t)Q^*(k_t))\overline{V}^*(k_t)\right) \\ -p^*(x_t,k_t)\sigma(k_t)S^*(k_t),$$

where $p^*(x_t,k_t)$ is the regime's equilibrium belief given x_t in environment k_t. Prevention occurs if and only if the former exceeds the latter, i.e., if and only if

$$\left(p^*(x_t,k_t)-\underline{p}(k_t)\right)\left[\delta Q^*(k_t)\left(\overline{V}^*(k_t)-V(k_t)\right)+\sigma(k_t)S^*(k_t)\right] \;>\rho(k_t).$$

Let

$$\overline{p}^*(k_t)=\min\left\{1,\underline{p}(k_t)+\frac{\rho(k_t)}{\delta Q^*(k_t)\left(\overline{V}^*(k_t)-V(k_t)\right)+\sigma(k_t)S^*(k_t)}\right\}\in\left(\underline{p}(k_t),1\right],$$

with the convention that $\frac{\rho(k_t)}{\delta Q^*(k_t)\left(\overline{V}^*(k_t)-V(k_t)\right)+\sigma(k_t)S^*(k_t)}$ is infinite if its denominator is 0. Then, the regime prevents protest if and only if $p^*(x_t,k_t)>\overline{p}^*(k_t)$. □

Intuitively, any stationary equilibrium in pure strategies induces some continuation payoffs $-L^*(k)<0$ and $B^*(k)>0$ in the first stage of the game that depend, implicitly, on the possibility of revolution and the probabilities of transitioning to other states. Given these continuation payoffs, the essential aspects of the regime's incentives in Proposition 1 are preserved.

To extend the argument for identification, now suppose that the econometrician observes a (possibly environment-specific) coarsening $\chi(x_t,k_t)$ of the regime's information x_t. As we have put little structure on the primitives of the game, we replace the earlier assumptions on the likelihood of the regime's information x with a high-level assumption on the regime's assessed conditional probability $p^*(x_t,k_t)$ of protest absent prevention. With these modifications, the ideas in Proposition 2 extend directly.

Proposition 5. *In any stationary equilibrium in pure strategies, the preventiveness $1-\overline{p}^*(k)$ in some given environment $k\in\mathbb{N}$ is partially identified from the marginal distribution $\Phi^*(\cdot,k)$ of $\Pr(z|\chi(x,k))$ for $\chi(\cdot,k)$ a coarsening, and is point identified if $\chi(\cdot,k)$ is one-to-one and the support of $p^*(x_t,k_t)$ contains a left-open interval that includes $\overline{p}^*(k_t)$. More specifically, $1-\overline{p}^*(k)\leq 1-\overline{p}(k)$, for $\overline{p}(k)=\inf\{p\in[0,1]:\Phi^*(p,k)=1\}$, with equality when $\chi(\cdot,k)$ is one-to-one and there is $\tilde{p}$ such that $(\tilde{p},\overline{p}^*(k)]$ is contained in the support of $p^*(x_t,k_t)$.*

S.A.2 Microfoundation in a Coordination Game

In the model of Section 2 we assume that protest occurs if and only if it is coordinated by the opposition. Here we microfound that assumption in a game of coordination among citizens. To focus on the case where coordination is costly to the opposition, we consider situations where $\overline{p}^*=\min\{\rho/L^*,1\}<1$, so that planning must be in secret.

There is a continuum of citizens $h\in[0,1]$ of unit mass, and a continuum of locations $\ell\in[0,1]$, also of unit mass. The timing of the game is as follows:

- Grievances $\omega \in \{0,1\}$ are realized, and citizens observe ω.
- The opposition learns grievances and either chooses not to plan, or pays the cost $\mu > 0$ to secretly plan a protest at a location $\ell^* \in [0,1]$ of its choosing.
- The regime learns information $x \in \mathscr{X}$ and chooses a Lebesgue-measurable set of locations $\mathscr{L} \subseteq [0,1]$ to block, paying a blocking cost $\rho|\mathscr{L}|$ where $|\mathscr{L}|$ is the (Lebesgue) measure of $\mathscr{L}$.
- If the opposition's choice of ℓ^* is not blocked, each citizen $h \in [0,1]$ observes ℓ^*. Otherwise, citizens do not observe the choices of the opposition or the regime or any other public signal.
- Each citizen $h \in [0,1]$ chooses either not to protest (which has no cost and no benefit) or to protest at some location $\ell_h \in [0,1]$. If the citizen is aggrieved, the location ℓ_h is not blocked, and measure at least $s \in (0,1)$ of citizens protest at the same location, the citizen obtains a net benefit $\beta > 0$. Otherwise, the citizen pays a cost $\kappa > 0$. For simplicity we suppose that citizens do not protest when indifferent.

Protest occurs if a strictly positive measure of citizens protest at a location. Other features of the model are the same as in the static model of preventive repression. In particular, the regime pays a cost $L^* > 0$, and the opposition receives a benefit $B^* > 0$, when protest occurs. When citizens are aggrieved and there is no public announcement, the support of x does not depend on m, and we write $\lambda(x) \in (0,\infty)$ as the likelihood ratio of the information on the opposition's plans, viz., its density if the opposition has made a secret plan, divided by its density if the opposition has not made a secret plan. Among the values of x that arise only when citizens are aggrieved, the image of the support of x in $\lambda(\cdot)$ is $(0,\infty)$, so that some realizations of x strongly indicate the presence of a secret plan, and some strongly indicate its absence.[1] The cost of secret planning μ is privately known to the opposition and distributed according to a commonly known continuous distribution with full support on $(0,\infty)$. As before, for simplicity, we assume that when indifferent the opposition (regime) does not plan (block) protest. Our equilibrium concept is subgame perfect Nash equilibrium.

When $\omega = 0$, so that citizens are not aggrieved, no citizen has an incentive to protest. When $\omega = 1$, the opposition is the only source of a public signal, so that citizens cannot condition their decisions on anything except possibly on l^*. First, we show that secret planning by the opposition is necessary for protest to occur.

Lemma 1. *If there is no secret plan, i.e., if $m = 0$, then the equilibrium probability of protest is zero.*

Proof. Fix an equilibrium and suppose that the opposition does not plan but that protest happens with strictly positive probability. The regime anticipates that, if $\omega = 1$, then with probability at least $p' \geq \frac{\kappa}{\beta+\kappa} \in (0,1)$, some measure $s' \geq s \in (0,1)$ of citizens will protest at some location ℓ'. There can be at most finitely many such locations $\mathscr{L}'$. But because $\mathscr{L}'$ is finite, we have that $|\mathscr{L}'| = 0$ so that the regime

[1] We continue to assume that the distribution of $\Pr(m = 1, \omega = 1 | x, a = 0)$ has no mass points given any non-degenerate prior on m.

can block all protest locations $\mathscr{L}'$ at no cost. The proposed equilibrium therefore entails a profitable deviation by the regime. $\square$

We now characterize an equilibrium of the game.

Lemma 2. *There exists an equilibrium in which the opposition chooses $\ell^* \sim U[0,1]$ when $\mu < \overline{\mu}^*$, and chooses not to plan otherwise; all citizens protest at the location ℓ^* if it is planned and not blocked, and otherwise they do not protest; and the regime blocks all locations if $p^*(x) > \overline{p}^*$, and blocks no location otherwise.*

Proof. Given the behavior of the opposition and citizens, if the regime blocks a set $\mathscr{L} \subseteq [0,1]$ of locations, the probability of protest is $(1-|\mathscr{L}|)$, and so the regime's payoff is $-p^*(x)(1-|\mathscr{L}|)L^* - |\mathscr{L}|\rho$. That the regime blocks all locations or none then follows from the linearity of this payoff and our tiebreaking rule; the threshold for blocking follows from the analysis in the main text. Next consider the citizens. Under the proposed equilibrium, when the opposition plans at location ℓ^*, and ℓ^* is not blocked, any citizen obtains benefit $\beta > 0$ from protesting at ℓ^*, and obtains payoff 0 from deviating. By contrast, when the opposition does not plan, any citizen obtains payoff 0 from the equilibrium strategy, and obtains payoff $-\kappa < 0$ from deviating. Therefore citizens have no profitable deviations. Finally, consider the opposition. Planning according to any distribution other than $U[0,1]$ results in weakly lower payoff than playing $U[0,1]$. And the analysis in the main text directly implies that planning when $\mu \geq \overline{\mu}^*$, or not planning when $\mu < \overline{\mu}^*$, are not profitable deviations. $\square$

We now show a sense in which the equilibrium in Lemma 2 achieves the maximum *ex ante* probability of protest. From Lemma 1, there is no protest absent planning by the opposition, so we can restrict to equilibria in which the opposition plans. We further restrict attention to *direct equilibria* in which the opposition's announced locations follow a non-singular distribution G,[2] and in which citizens only consider protesting at the announced location ℓ^*.

Lemma 3. *No other direct equilibrium features a greater probability of protest.*

Proof. Fix a direct equilibrium with associated distribution G of announced locations. The distribution G does not have mass points, because mass points have measure zero on $[0,1]$, and hence the regime will block the mass points at zero cost. Therefore G has a density; call this density g. Because locations are labels, without loss, we can rearrange them so that g is (weakly) increasing.

We proceed by showing that if g is not constant almost everywhere on $[0,1]$, then there is a profitable deviation for the opposition. Towards contradiction, suppose g is not constant almost everywhere on $[0,1]$. Then there is some $\overline{\ell} \in (0,1)$ and two intervals $I_1, I_2 \subset [0,1]$, with $\sup I_1 < \overline{\ell} < \inf I_2$ such that $g(x) < g(y)$ for all $x \in I_1$ and $y \in I_2$. Let $g_1 = \inf\{\alpha \mid \alpha \geq g(x), x \in I_1\}$ and $g_2 = \sup\{\alpha \mid \alpha \leq g(x), x \in I_2\}$, and note

[2]This rules out pathological cases such as the Cantor distribution.

that $0 \leq g_1 < g_2$. For $i = 1,2$, if the regime blocks a subset I_i' of the interval I_i with measure $\left|I_i'\right|$, then the regime's objective function depends on I_i' through the additive term $-\hat{p}(x)\left(\int_{I_i/I_i'} g(y)dy\right)L^* - \rho\left|I_i'\right|$. Next, consider an alternative g which is uniform on interval I_i with density g_i, defined above. We have $-\hat{p}(x)\left(\int_{I_2/I_2'} g(y)dy\right)L^* - \rho\left|I_2'\right| \leq -\hat{p}(x)g_2\left(|I_2| - \left|I_2'\right|\right)L^* - \rho\left|I_2'\right|$ and the marginal gain of increasing $\left|I_2'\right|$ is higher on the left hand side. Thus, in any equilibrium, the regime blocks I_2 if $\hat{p}(x) > p_2 \equiv \frac{\rho}{g_2 L^*}$. Similarly, $-\hat{p}(x)\left(\int_{I_1/I_1'} g(y)dy\right)L^* - \rho\left|I_1'\right| \geq -\hat{p}(x)g_1\left(|I_1| - \left|I_1'\right|\right)L^* - \rho\left|I_1'\right|$. Thus, in any equilibrium, the regime does not block I_1 if $\hat{p}(x) \leq p_1 \equiv \frac{\rho}{g_1 L^*}$, where we set $p_1 = \infty$ if $g_1 = 0$. Moreover, because $0 \leq g_1 < g_2$, we have $0 < p_2 < p_1$, so that the probability of blocking locations in I_2 is strictly greater than the probability of blocking locations in I_1. Thus, the opposition can profitably deviate by modifying G to move measure from I_1 to I_2. □

The following Proposition summarizes these results.

Proposition 6. *In the protest-maximizing direct equilibrium of the coordination game, citizens protest if and only if they are aggrieved, $\omega = 1$, the opposition has made a secret plan, $m = 1$, and the regime has not blocked the location of the secret plan. Moreover, if it is not blocked, protest occurs with probability $p^*(x)$ with respect to the regime's information set, and the regime prevents protest (by blocking all locations) if and only if $p^*(x) > \overline{p}^*$.*

Proposition 6 thus establishes the behavior we assume in Section 2 as the protest-maximizing outcome of the game we describe here.

S.B Additional Details on Parsing of Alerts

S.B.1 Parsing of Crisis24 Security Alerts

S.B.1.1 Elements Used in Main Analysis

We identify protest-related alerts using information provided in alert descriptions.[3] We classify an alert as protest-related if its description includes the words "protest," "demonstration," or "demonstrator," or words with those keywords at the root. We classify an alert as including evidence of use of force if it contains indicative phrases such as "protestors clash with the police" or "hundreds arrested during protests."

We extract, for each alert, any dates mentioned in the alert's title provided that the title contains a protest-related structure such as "protest" or "demonstration." We consider all such dates to be protest dates.[4] We also extract, for each alert, any dates mentioned in the alert's description in the same sentence

[3] We exclude from our search of the description text some generic warnings such as "We advise our clients to stay away from protests."

[4] Date information typically includes the month and day of month, e.g., "May 20." We treat these dates as referring to the year in which the alert is published, unless the alert is published in December (January) and the date is in January (December), in which case we treat the date as referring to the year following (preceding) the alert's publication.

as a protest-related structure. If we do not extract any protest dates from a protest-related alert's title, then we consider the latest date extracted from the alert's description to be a protest date.

Supplemental Appendix Figure 1 illustrates the main elements of our approach using example alerts drawn from our data.

S.B.1.2 Audit of Parser Quality

To assess the quality of the rules we use to parse text fields, we compare the performance of our parser to the performance of human data entry operators on a random sample of alerts. The data entry operators were trained to enter data on protest occurrence based on the text fields, using a web form that we helped to develop. Each alert was keyed by two independent data entry operators, with a reconciliation process for discrepancies. The firm providing data entry services was given a financial incentive for accurate entry.

We used our own judgment to manually date all protests described in 100 randomly chosen alerts, excluding 4 alerts due to ambiguities. Across the 96 remaining alerts, our parsing rules identified 71 dates with protest, of which we classify 57 as correct and 14 as incorrect. Across these same 96 alerts, the human data entry operators identified 119 dates with protest, of which we classify 81 as correct and 38 as incorrect.

When we modify the simulation in Supplemental Appendix S.C.3 to incorporate a fraction of falsely classified protest dates that matches the fraction we measure in the audit, the correlation between the true and estimated target parameters remains large at 0.9991.

S.B.1.3 Elements Used in Applications

We construct an indicator variable for each alert to denote whether it mentions human rights concerns for the analysis in Section 4.2.1. Specifically, we classify an alert as mentioning human rights concerns if it explicitly mentions human rights violations (e.g., terms such as "human rights abuses" or "torture") or references human rights organizations, either generically (e.g., "human rights groups") or specifically (e.g., "Amnesty International").

S.B.1.4 Elements Used in Sensitivity Analysis and Validation

Section S.B.1.1 explains how we obtain information on protest dates for our main analysis. For sensitivity analysis we additionally classify some alerts into temporal categories based on syntactic and semantic information in alert titles. We classify an alert as pertaining to a future protest if its title contains a future-related word such as "announce" along with a protest-related keyword, or if its title contains a protest-related verb preceded by "to," as in "to protest." We classify an alert as pertaining to a present protest if its title contains present-related words such as "continued," "ongoing," or "underway" appearing in the same sentence with a protest-related keyword. We classify a protest alert as pertaining to a past protest if its title contains protest-related verbs in the past tense, such as "protested" or "demonstrated," or if its title does not contain word structures indicative of being about the present or future.[5] We use

[5] We also search for specific clauses such as "hundreds of people gather" that, according to our reading of alerts, typically indicate that the event was in the past even though they are written in the present tense.

Supplemental Appendix Figure 1: Illustrations of Security Alert Parsing

protest-related term
date
generic text

Sample Alert A

Country: Nigeria

Published at: 2017-03-22

Title: Nationwide NULGE Protests Begin

Content: According to local media sources on Wednesday, *22 Mar*ch, the Nigerian Union of Local Government Employees (NULGE) has commenced a nationwide protest calling for a constitutional amendment granting Local Government Areas autonomy. Protests began in Nasarawa state and are predicted to continue across the country until President Muhammadu Buhari intervenes. Travellers are advised to avoid political protests and exercise caution in the vicinity of gathering crowds in order to minimise the risk of exposure to potential crowd disturbances. Monitor local media sources for more information about how the protests will affect specific regions and the effect these may have on overland travel.

Sample Alert B

Country: Peru

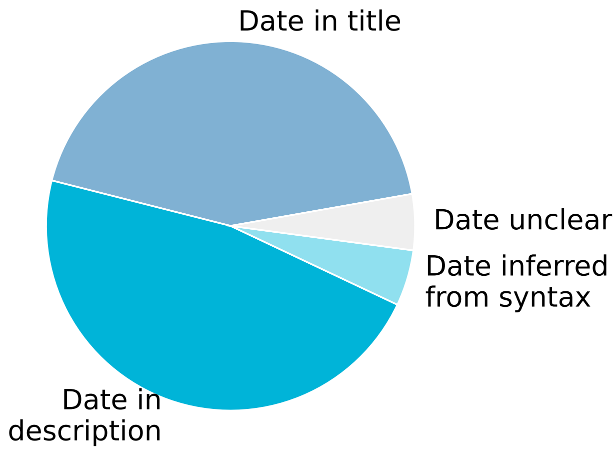
Published at: 2018-07-06

Title: Protest to be Held in Arequipa on *10 Jul*y over Cost of Public Transport

Content: According to local media reports, a protest march is set to be held in Arequipa city over the high cost of public transportation. The action is being organised by a number of civil society organisations and is set to occur on Tuesday, 10 July. Protesters are also angry about poor service and regular accidents. Members in Arequipa are advised to avoid the demonstrations due to the associated risks of exposure to opportunistic crime, unruly crowd behaviour, and police crowd control measures. Monitor local media sources to remain aware of current tensions and for any updates related to planned or ongoing unrest in your area of operation.

Notes: For each sample alert we illustrate (i) terms used to identify protest-related alerts (highlighted in gray), (ii) dates used to identify the date of the protest (italicized), and (iii) generic text that our parser ignores (grayed out).

Supplemental Appendix Figure 2: Sources of Date Information in Protest Alerts



Notes: The pie chart shows the breakdown of protest alerts according to how we obtain information on protest dates. "Date in title" shows the share of protest alerts for which we are able to obtain a date from the alert title. "Date in description" shows the share of protest alerts for which we are not able to obtain a date from the alert title and instead obtain a date from the alert description. "Date inferred from syntax" shows the share of protest alerts for which we are not able to obtain a date from the alert title or description and instead infer the alert date based on grammatical cues. "Date unclear" shows the share of protest alerts for which we were not able to obtain any date information from the alert text.

these classifications to include additional protests described in alerts from whose text fields we were not otherwise able to infer protest dates.[6] Supplemental Appendix Table 1 shows the sensitivity of our main results to including these additional protests. Supplemental Appendix Figure 2 shows a breakdown of protest alerts according to how we obtain information on protest dates.

We extract, for each alert, information about the number of participants in the protests mentioned in the text fields. This information is often in the form of a broad quantitative statement (e.g., "hundreds of protesters gathered" or "around 500 protesters gathered"). We classify an alert as pertaining to small protests if the alert contains information about the number of participants for at least one protest, and if there is no information indicating a protest with one thousand or more participants.

We also identify alerts pertaining to natural disasters. We classify an alert as disaster-related if it is not identified as protest-related, and contains keywords related to natural disasters such as "earthquake" or "cyclone" in its title. We consider a date to have a disaster event if a disaster-related alert was published on that date.

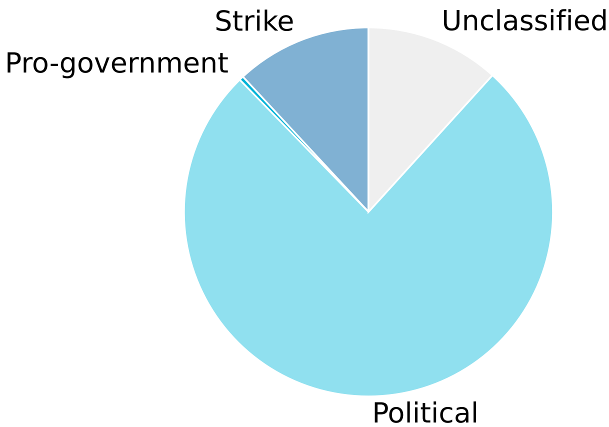
S.B.1.5 Sensitivity to Alternative Definitions of Protest

Our main analysis takes the indicator of protest z_{it} to be intentionally broad. Some protests are not a threat to the regime and are therefore unlikely to be prevented. We can accommodate such events in our model by supposing that there are protests that will not lead to revolution. We can accommodate such events in our empirical analysis by excluding from the definition of protest those events that may not pose a threat to the regime.

Supplemental Appendix Figure 3 reports the share of protests that are political, pro-government, or strikes. Supplemental Appendix Table 1 shows that our results are similar when we exclude, respectively, pro-government protests, strikes, and small protests from the definition of protest.

[6] Specifically, we consider a country-date to have a protest if there was a future protest-related alert published on the preceding day, if there was a continued protest-related alert published on the same day, or if there was a past protest-related alert published the day after.

Supplemental Appendix Figure 3: Categories of Protest Alerts



Notes: The pie chart shows the share of protest alerts that fall into each of four categories. We classify an alert as a "Strike" if its description contains one of the words "worker," "union," "labor," or "labour" and the root "strike" in the same sentence, except when "strike" is part of the word "airstrike." We classify an alert as "Pro-government" if it is not a strike, includes a statement indicating protest in support of the regime (e.g., "government supporters took to the streets"), and does not include a statement indicating protest against the regime (e.g., "thousands protested against the government"). We classify an alert as "Political" if it is not a strike or pro-government protest, and it includes statements related to political actors (e.g., prime minister, far-left groups) or to citizens' demands, rights, or grievances (e.g., indications of citizens denouncing austerity, human-rights related protests, or protests against police violence). We classify all other protests, including those for which the alert text does not indicate a cause, as "Unclassified."

S.B.2 Parsing of Other Alerts

S.B.2.1 OSAC alerts

For each OSAC alert, if there is a country named in the title, we assign the alert to the named country. Otherwise, we assign the alert to the country whose capital city is referenced in the title (United Nations 2018b). For determining whether an OSAC alert is related to a protest, and to infer protest dates, we follow the same procedures outlined in Section S.B.1.1 that we use for Crisis24 alerts.

S.B.2.2 Human Rights Watch Alerts

For each Human Rights Watch alert, we extract the mentions of country names from the alert titles and subtitles. If the country field in the data is not missing and is unique, we use that country as the target of the alert. If the country field is missing or the alert involves multiple countries, we assign the alert to the set of countries inferred from the title fields.

We also create indicators for the content of the alerts. We classify an alert as related to general human rights concerns if it explicitly mentions human rights. We classify whether each alert is related to police deployments (based on inclusion of terms such as "deployed police" or "patrols"), restrictions on press freedom (e.g., "censorship" or "freedom of the press"), restrictions on the internet (e.g., "internet shutdown"), or torture (e.g., "torture" or "ill-treatment"). We classify whether each alert is related to use of force against protesters (based on inclusion of terms such as "tear gas," "water cannon," or "excessive use of force"), dispersal of protests (e.g., "protest" or "demonstration" in proximity with terms such as "disperse" or "disband"), or arrest of protesters (e.g., "protest" or "protester" in proximity with terms such as "arrest," "detain," or "detention").

Supplemental Appendix Table 1: Sensitivity Analysis for Main Index of Preventiveness

	Correlation with baseline	Commodity price effect $\gamma^{p,x}$(Monthly)	HRW alerts
Baseline	1.000	0.451	0.337
		(0.159)	(0.134)
Modifying sample			
Increase population threshold	1.000	0.575	0.359
		(0.194)	(0.140)
Increase alerts threshold	1.000	0.506	0.337
		(0.169)	(0.137)
Modifying regressor			
Use terms of trade	—	0.452	—
		(0.143)	—
Modifying protest definition			
Exclude pro-government protests	0.997	0.445	0.346
		(0.159)	(0.131)
Exclude strikes	0.967	0.478	0.323
		(0.156)	(0.134)
Exclude small protests	0.944	0.456	0.306
		(0.147)	(0.137)
Exclude same-day anticipated protests	0.775	0.417	0.150
		(0.120)	(0.124)
Include protests with uncertain dates	0.962	0.488	0.333
		(0.162)	(0.128)

Notes: "Correlation with baseline" reports the Pearson correlation (in monthly first differences) with respect to the baseline measure. "Commodity price effect" reports the estimated coefficient $\gamma^{p,x}$ from Equation (7), as in Table 6 Panel A. "HRW alerts" reports the estimated coefficient $\gamma_0^{h,p}$ from Equation (1), as in Table 3 column (2). Asymptotic standard errors in parentheses are clustered by country. **Baseline** corresponds to our main index of preventiveness described in Section 3.3. **Modifying sample** corresponds to changing the sample of countries as described in Section 3.1. "Increase population threshold" corresponds to restricting the sample to countries with a population of at least 2,000,000 in 2010, keeping the alerts threshold as in baseline. "Increase alerts threshold" corresponds to restricting the sample to countries in which there is at least one year in the sample period with at least 20 alerts in the Crisis24 database, keeping the population threshold as in baseline. **Modifying regressor** corresponds to modifying regressors in analyses. "Use terms of trade" corresponds to using commodity price indices based on terms of trade instead of just export prices as described in Section 5.4. **Modifying protest definition** corresponds to modifying our definition and parsing of protests from security alerts; see Section 3.3, Supplemental Appendix S.B.1.4, and Supplemental Appendix S.B.1.5 for details.

S.C Additional Details on Predictive Estimation

S.C.1 Data on Predictors of Protest

To construct the covariate vector $\mathbf{x}_{it}$ we obtain the variables described below. For each variable, we standardize its daily value by subtracting its mean and dividing by its standard deviation, both calculated over the preceding 90 days, excluding protest days. We standardize in this way because the distributions of variables may evolve over time; Supplemental Appendix Table 2 shows results where we standardize using the entire sample (e.g., Hastie et al., 2017) or not at all.

If the rolling-window standard deviation is zero, or if data are missing for the given variable, country, and date, we set the standardized value equal to the standardized rolling window mean. Supplemental Appendix Table 2 shows results where we add an indicator for these imputations to the predictive model.[7]

S.C.1.1 Search Query Volume: Google Trends

From Google Health Trends (Google, 2021, 2023), we obtain daily data on searches about political demonstrations in each country.[8] The search data are reported as a (known) multiple of the probability that a given user session includes a search for the given topic, with a value of zero when the query does not meet reporting standards (Google 2019; see also Zepecki et al. 2020).[9]

S.C.1.2 News Mentions: The GDELT Event Databases

From the GDELT Event Database (GDELT Project 2020) we obtain daily data on the number of mentions of protest, as a share of all mentions of a given country, for each country.[10] These data are in turn sourced from various international news sources and wire services (Leetaru and Schrodt 2013).

S.C.1.3 Social Media Mentions: Twitter Data

Using Twitter's advanced search functionality (Twitter, 2020, 2023), for each country we obtain daily data on the number of non-withheld English-language tweets containing both the name of the country and either the keyword "protest" or the keyword "demonstration."[11]

S.C.2 Implementation

S.C.2.1 Predictor Variables

For our main predictive exercise, we take the predictors $\mathbf{x}_{it}$ to consist of the indicator a_{it} of anticipated protest described in Section 3, indicators for days of the week, and seven lags each of the standardized values of demonstration search query volume, protest news mentions, and protest Twitter mentions in English described in Supplemental Appendix S.C.1. Supplemental Appendix Table 2 shows results where we add a number of additional predictor variables.[12]

---

[7] At least one covariate is missing on 5.3 percent of country-dates. Because the standardized rolling-window mean is zero, adding indicators for missing values subsumes a specification (discussed in Gelman and Hill 2006, Section 25.3) where these indicators are interacted with the covariate.

[8] The searches use entity code `/m/0gnwz4`. For use in sensitivity analysis, for each country we also obtain daily data on searches conducted in the US and UK about the country.

[9] These probabilities are calculated on a random sample of searches. The random sample is redrawn daily. We compute an average of valid returned values across runs executed on at least 15 different days, and therefore corresponding to at least 15 different random samples.

[10] The event code for protest is 14.

[11] We obtain analogous data on tweets in the country's official language for use in sensitivity analysis. Our list of countries and official languages is based on UNGEGN's list of country names (United Nations Group of Experts on Geographical Names 2019). In some cases we use short versions of country names rather than official ones (e.g., "Venezuela" rather than "Bolivarian Republic of Venezuela").

[12] These are, seven lags of standardized values of search query volume about the country in the US and UK (Supplemental Appendix S.C.1); protest Twitter mentions in the country's official languages (Supplemental Appendix S.C.1); a set of financial indices (Online Appendix O.D); an indicator for whether an election occurred in the given country within seven days of the given date (Online Appendix O.D); and an indicator for whether a protest occurred in the given country on

S.C.2.2 Predictive Function

For our main predictive exercise, we take the predictive function $\hat{p}(\cdot)$ to be logistic, so that

$$\ln\left(\frac{\hat{p}(\mathbf{x})}{1-\hat{p}(\mathbf{x})}\right)=\mathbf{x}\theta \tag{8}$$

for coefficients θ. The logistic function is a choice made for convenience; Supplemental Appendix Table 2 shows that results are similar when $\hat{p}(\cdot)$ takes the probit form. Assuming the same coefficients θ apply throughout the sample allows precise inferences on θ but risks ignoring important heterogeneity. Supplemental Appendix Table 2 shows that the results are similar when we instead allow the coefficients θ to differ by country, θ_i.[13]

S.C.2.3 Sample Splitting

For our main predictive exercise, we estimate $\hat{p}(\cdot)$ separately using data from even years and data from odd years, and use the estimate from the opposite-parity years to form the prediction $\hat{p}(\mathbf{x}_{it})$ for each country i and date t. This sample-splitting scheme ensures balance in the two samples across countries and broad time periods. Supplemental Appendix Table 2 shows results where we instead use a random split of years rather than one based on year parity, and where we use a leave-one-out estimator that forms the prediction $\hat{p}(\mathbf{x}_{it})$ by estimating the predictive function using all country-months except the one containing country i and date t.

S.C.2.4 Estimator of Predictive Function

For our main predictive exercise, we estimate $\hat{p}(\cdot)$ via maximum likelihood; Supplemental Appendix Table 3 presents the parameter estimates. The maximum likelihood estimator has desirable efficiency properties but can overfit the data when the number of predictors is large. Supplemental Appendix Table 2 shows that results are similar when we instead estimate via penalized maximum likelihood (lasso), a procedure that is designed to control overfitting.[14]

S.C.2.5 Estimator of Preventiveness

Our main estimator of preventiveness for country i and month k is one minus the largest estimated probability of protest in the given country-month, $1-\max_{\{t\in k\}}\hat{p}(\mathbf{x}_{it})$. Supplemental Appendix Table 2 shows results from two alternative estimators described in Section 3.4. One is an unbiased estimate of an upper bound of preventiveness, defined as an indicator for whether protest does not occur on the day with the largest estimated protest probability, $1-z_{it^*(k)}$ for $t^*(k)$ such that $\hat{p}\left(\mathbf{x}_{it^*(k)}\right)\geq\hat{p}(\mathbf{x}_{it})$ for all $t\in k$. The other is an indicator for whether an 80 percent confidence bound, constructed following Lei (2023, Equation 6), contains nearly full preventiveness, $1-\overline{p}^*_{i,k}\geq 0.9$.

the same date in the previous year.

[13]To avoid overfitting, we estimate θ_i via L1-penalized maximum likelihood, excluding countries for which this is infeasible, with the penalty parameter chosen via 10-fold cross-validation where the folds are approximately equal-sized groups of consecutive months.

[14]We choose the lasso penalty parameter via 10-fold cross-validation where the folds are equal-sized groups of country-years.

Supplemental Appendix Table 2: Sensitivity Analysis for Predictive Estimation

	Correlation with baseline	Commodity price effect $\gamma^{p,x}$(Monthly)	HRW alerts $\gamma_0^{h,p}$
Baseline	0.996	0.437	0.352
		(0.151)	(0.139)
Modifying covariates			
Standardize using full sample	0.996	0.443	0.369
		(0.152)	(0.140)
Do not standardize	0.973	0.461	0.326
		(0.151)	(0.141)
Indicators for missing data	0.997	0.436	0.353
		(0.151)	(0.139)
Include additional predictors	0.995	0.417	0.361
		(0.152)	(0.142)
Modifying predictive function and estimator			
Use probit	0.997	0.438	0.352
		(0.151)	(0.139)
Predict separately by country	0.911	0.317	0.442
		(0.101)	(0.158)
Use penalization	0.997	0.437	0.352
		(0.151)	(0.139)
Modifying sample split			
Split sample randomly	0.997	0.438	0.354
		(0.151)	(0.139)
Leave out country-month	0.997	0.438	0.352
		(0.151)	(0.139)
Modifying estimator for preventiveness			
Unbiased estimate of upper bound	0.955	0.475	0.263
		(0.156)	(0.136)
Confidence bound for preventiveness	0.951	0.456	0.257
		(0.158)	(0.137)

Notes: "Correlation with baseline" reports the Pearson correlation (in monthly first differences) with respect to the baseline measure. "Commodity price effect" reports the estimated coefficient $\gamma^{p,x}$ from Equation (7), as in Table 6 Panel A. "HRW alerts" reports the estimated coefficient $\gamma_0^{h,p}$ from Equation (1), as in Table 3 column (2). Asymptotic standard errors in parentheses are clustered by country. **Baseline** corresponds to our main index of preventiveness described in Section 3.3. **Modifying covariates** corresponds to a set of alternative processing steps described in Section S.C.1. **Modifying predictive function and estimator** corresponds to modifying the prediction function $\hat{p}(\cdot)$, as described in Supplemental Appendix S.C.2.2; and to modifying how we estimate the predictive function, as described in Supplemental Appendix S.C.2.4. **Modifying sample split** corresponds to modifying our sample splitting scheme as described in Supplemental Appendix S.C.2.3. **Modifying estimator for preventiveness** corresponds to modifying our estimator of preventiveness as described in Section 3.4 and Supplemental Appendix S.C.2.5.

Supplemental Appendix Table 3: Parameter Estimates for Main Predictive Exercise

		Odd Years		**Even Years**	
		Coefficient	SE	Coefficient	SE
Intercept		-3.654	0.117	-3.875	0.116
Anticipated protest		21.193	0.115	21.330	0.118
Day of week					
	Mo	0.060	0.074	0.278	0.068
	Tu	0.025	0.053	0.190	0.059
	We	0.095	0.065	0.165	0.063
	Th	0.030	0.070	0.143	0.067
	Sa	-0.271	0.070	-0.220	0.073
	Su	-0.279	0.081	-0.160	0.082
Demonstration search query volume (Standardized)					
Lag of:	1 day	0.055	0.007	0.058	0.008
	2 days	0.015	0.009	0.028	0.006
	3 days	0.014	0.007	0.012	0.010
	4 days	0.016	0.010	0.010	0.009
	5 days	0.015	0.008	0.022	0.006
	6 days	0.016	0.008	0.025	0.007
	7 days	0.025	0.009	0.019	0.006
Protest news mentions (Standardized)					
Lag of:	1 day	0.012	0.004	0.006	0.002
	2 days	0.004	0.002	-0.002	0.001
	3 days	0.004	0.001	0.001	0.004
	4 days	0.004	0.001	0.004	0.002
	5 days	0.003	0.001	-0.001	0.002
	6 days	0.003	0.001	0.002	0.002
	7 days	0.007	0.003	0.002	0.001
Protest social media mentions (Standardized)					
Lag of:	1 day	0.008	0.003	0.003	0.001
	2 days	0.002	0.002	0.004	0.002
	3 days	0.002	0.001	0.002	0.001
	4 days	0.002	0.001	0.002	0.001
	5 days	0.000	0.001	0.002	0.001
	6 days	0.004	0.002	0.001	0.001
	7 days	0.002	0.002	0.002	0.001

Notes: The table presents estimates of the coefficients from the logistic regression defined in Equation (8), where the dependent variable is an indicator for the occurrence of protest and the independent variables are listed in the rows of the table. Standard errors are clustered by country. Estimates are presented separately for odd years and even years following the sample-splitting scheme described in Supplemental Appendix S.C.

Supplemental Appendix Table 4: Simulation Evidence on the Performance of the Estimator

	Baseline	Falsification (Reverse time)
Within-country correlation	0.9991	0.0951
	(0.0002)	(0.0005)
Share with abs. err. ≤ 0.01	0.9581	0.4309
	(0.0052)	(0.0177)

Notes: The table shows average statistics across 50 replicates of a simulation exercise described in Section S.C.3, with standard deviations in parentheses. In the "Baseline" column we apply our baseline multivariate estimator. In the "Falsification (Reverse time)" column we reverse time in estimating the predictive model, predicting protest based on future rather than past values of the predictor variables. The "Within-country correlation" is the Pearson correlation in pooled monthly first differences between the estimated preventiveness and one minus the largest probability of protest in the simulated data-generating process. The "Share with abs. err. ≤ 0.01" is the share of country-months in which the difference between the estimated preventiveness and one minus the largest probability of protest is no more than 0.01 in absolute value.

S.C.3 Monte Carlo Evidence

The first column of Supplemental Appendix Table 4 shows simulation evidence on the performance of our baseline multivariate estimator. In each of a set of simulation replications, we randomly generate an indicator of protest occurrence as a sequence of independent Bernoulli draws, with success probabilities given by the estimated equilibrium protest probabilities $\hat{p}(\mathbf{x}_{it})$. We then re-implement the estimator, taking the observed values of $\mathbf{x}_{it}$ as given and using the simulated protest occurrence indicator in place of the observed indicator.[15]

We compare estimated preventiveness in each country-month to (one minus) the maximum true probability of protest in each country-month, which is known from the simulated data-generating process. The first column of Supplemental Appendix Table 4 shows that the estimated preventiveness is very correlated with, and very close numerically to, this target value.

The second column of Supplemental Appendix Table 4 shows evidence from a falsification exercise, in which we use the same simulated data, but reverse time in estimating the predictive model, predicting protest based on future rather than past values of the predictor variables. If the variation in the estimator were driven by variation in ancillary factors such as the quality of the measurement of the predictors $\mathbf{x}_{it}$, instead of by variation in the *ex ante* predictability of protest, the performance of the estimator might be similar in both columns of Supplemental Appendix Table 4. In fact the table shows that the performance of the estimator is much worse in the falsification exercise than in the baseline exercise.

[15] Appendix S.B.1.2 discusses findings from an additional simulation exercise in which we intentionally introduce misclassification in our assignment of protest dates.

S.D Additional Details on Salient Episodes of Repression

S.D.1 Dynamics of Protest and Repression During Interrupted States of Emergency

Supplemental Appendix Figure 4 plots three additional instances of interrupted states of emergency that occurred over multiple years and phases. The changes in enforcement over these episodes make it difficult to isolate a ground truth, but we can still understand them through the lens of our data. Egypt was under state of emergency during large portions of our sample period, and had been so continuously between 1981 and 2012.[16] Measured preventiveness tends to be higher during than outside of emergencies, but it also tracks other salient changes, such as the ousting of Mubarak.

Tunisia has been in a declared state of emergency since July 2015, at which time there was an armed attack at a resort. At the beginning of the emergency and since, reports indicated restrictions on protest and acts of suppression (Human Rights Watch 2015; Freedom House, 2016, pp. 709-714), as well as evidence of authorities permitting protest or showing restraint (United States Department of State 2015; Freedom House, 2017, pp. 529-535). Our data show sporadic periods of preventiveness before and especially during the emergency.

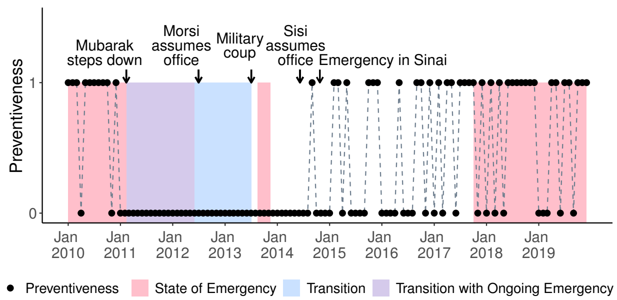
Ethiopia has entered and exited emergencies multiple times during our sample period, notably during 2016-2018 following large-scale protests against the government. Multiple accounts indicate both prevention and suppression of protest during these episodes (Freedom House, 2017, pp. 168-175; Horne 2017; Wikipedia 2023), both of which are corroborated by our data.

[16] These events prompted Auf (2018) to ask whether emergency is "an exception or rule" in Egypt.

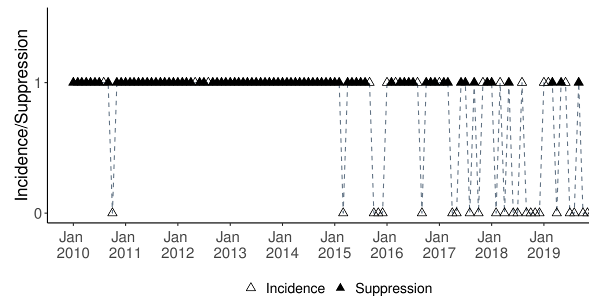
Supplemental Appendix Figure 4: Dynamics of Protest and Repression During Interrupted States of Emergency

Panel A: Egypt 1981-2021
(2010-2019 depicted)

Preventiveness

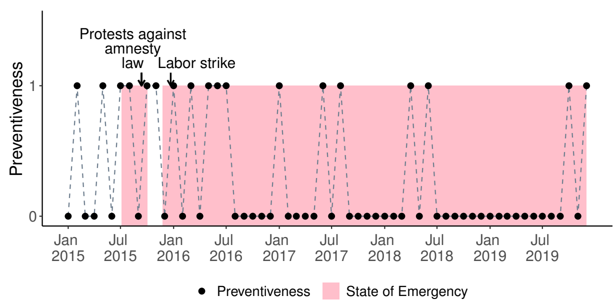


Incidence and Suppression of Protest

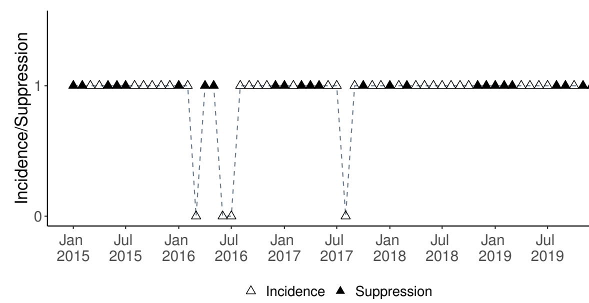


Panel B: Tunisia 2015 - Present
(2015-2019 depicted)

Preventiveness

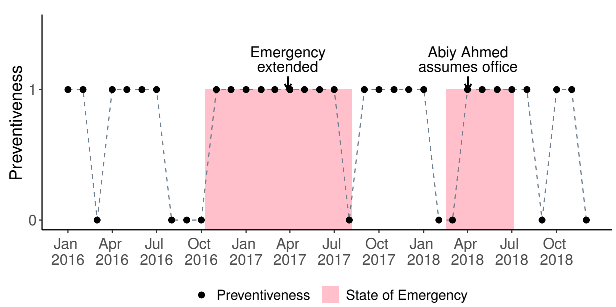


Incidence and Suppression of Protest

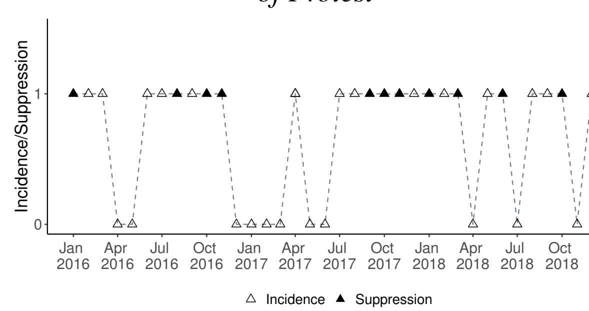


Panel C: Ethiopia 2016-2018

Preventiveness



Incidence and Suppression of Protest



Notes: The plots show the evolution of protest preventiveness, incidence, and suppression during interrupted states of emergency, with annotations for important dates. Shaded areas correspond to the exact dates of the labeled period within the plotted window.

S.D.2 Full Set of Episodes

Supplemental Appendix Table 5: Episodes of Martial Law and States of Emergency

Panel A: Main Episodes

Country	Year	Type	Annotated Quotations
Bahrain	2011	Martial law	In March, after hundreds of thousands of Bahrainis demonstrated in various parts of Manama, the government declared **martial law** and summoned military and security forces from regional allies, including Saudi Arabia and the United Arab Emirates, **to backstop a prolonged crackdown that aimed to clear the streets** and collectively punish the Shiite community.
Thailand	2014-2015	Martial law	Under both the January–March 2014 state of emergency and the **martial law** regime declared in May, **any gathering of more than five people could be banned**. In practice, demonstrations continued unabated until the military takeover. While multiple small and a few larger protests against the coup were held initially, they soon dissipated after the military significantly increased its **presence on the streets** of major cities and began making arrests.
Thailand	2010	State of emergency	In response, the government declared a **state of emergency** in Bangkok and 23 other provinces on April 7, **banned demonstrations**, and attempted to regain control of the occupied area on April 10.
Mali	2013	State of emergency	Under the state of emergency that was in effect in 2013, **gatherings of more than 50 people were banned.**

Notes: The table shows selected episodes of martial law and state of emergency. To select these episodes, we built an automated parser that captures mentions of martial law and state of emergency in Freedom House country reports (Freedom House, 2010-2019). We retained unique episodes by excluding duplicate occurrences of martial law and state of emergency mentions, and dropped mentions that pertain to historical episodes, or aim to provide legal context for current events. We further excluded mentions that refer to episodes that are limited in their geographic scope and those that last less than a month or longer than a year. Panel A shows episodes plotted in the main text. In each panel, the table shows the country, time period, and annotated quotations from Freedom House country reports that provide useful context.

Supplemental Appendix Table 5: Episodes of Martial Law and States of Emergency (continued)

Panel B: Interrupted Episodes

Country	Year	Type	Annotated Quotations
Egypt	1981-2021	State of emergency	**Egypt's Emergency Law**, in effect since 1981, was again **renewed in 2010**, despite Mubarak's 2005 promise that it would be replaced with specific antiterrorism legislation. A state of emergency and a related curfew that lasted from August to November [2014] gave police broad discretion to **break up demonstrations and detain participants** without regard to due process. The **state of emergency** declared by President Sisi in April [2017] grants security forces **additional powers of arrest and detention**, increasing opportunities for physical abuse.
Tunisia	2015-2024	State of emergency	In September [2015], the government began enforcing a **ban on all public demonstrations** under the state of emergency imposed **in response to the shooting in Sousse**. On at least three occasions that month, police used excessive force to disperse protests against the proposed reconciliation law.
Ethiopia	2016-2018	State of emergency	In October [2016], the government admitted that more than 500 people had been killed in connection with the protests since November 2015, though some rights organizations reported that the true figure is at least 800. **In early October [2016], the government announced a nationwide six-month state of emergency**, enacting sweeping powers to deploy the military, restrict speech and the media, impose **curfews** and **movement restrictions**, and monitor communications.

Notes: The table shows selected episodes of martial law and state of emergency. To select these episodes, we built an automated parser that captures mentions of martial law and state of emergency in Freedom House country reports (Freedom House, 2010-2019). We retained unique episodes by excluding duplicate occurrences of martial law and state of emergency mentions, and dropped mentions that pertain to historical episodes, or aim to provide legal context for current events. We further excluded mentions that refer to episodes that are limited in their geographic scope and those that last less than a month or longer than a year. Panel B shows interrupted episodes plotted in Supplemental Appendix Figure 4. In each panel, the table shows the country, time period, and annotated quotations from Freedom House country reports that provide useful context.

Supplemental Appendix Table 5: Episodes of Martial Law and States of Emergency (continued)

Panel C: Excluded Episodes

Country	Year	Type	Annotated Quotations
Liberia	2014	State of emergency	Government **responses to the Ebola epidemic threatened freedom of assembly**. West Point residents responded violently when barricades were erected to enforce the quarantine there, and were dispersed by police officers firing live ammunition and tear gas. In November, the Disciplinary Board of the Armed Forces of Liberia found four soldiers guilty of excessive force during the West Point incident, and in December the UN Security Council cited "mis use of firearms" by state security forces in West Point as justification for renewal of an arms embargo on Liberia. In December, **the government banned rallies and other public gatherings** during the two weeks prior to the senatorial elections, and for 30 days thereafter.
Zambia	2017	State of emergency	**Following a series of arson attacks** in the capital, ruling-party lawmakers voted to approve a restrictive 90-day **state of emergency**. The 90-day state of emergency imposed in July 2017 gave the government the authority to impose **broad media restrictions**, and was seen as a threat to **press freedom**.
Sudan	2019	State of emergency	President Omar al-Bashir, who came to power in a coup d'état in 1989, was **overthrown by the military** in April, after a protest movement beginning in December 2018 placed growing pressure on the government... This took place in October, when the transitional government extended a nationwide **state of emergency imposed by al-Bashir** in February.

Notes: The table shows selected episodes of martial law and state of emergency. To select these episodes, we built an automated parser that captures mentions of martial law and state of emergency in Freedom House country reports (Freedom House, 2010-2019). We retained unique episodes by excluding duplicate occurrences of martial law and state of emergency mentions, and dropped mentions that pertain to historical episodes, or aim to provide legal context for current events. We further excluded mentions that refer to episodes that are limited in their geographic scope and those that last less than a month or longer than a year. Panel C shows episodes with more ambiguous descriptions that we do not plot. In each panel, the table shows the country, time period, and annotated quotations from Freedom House country reports that provide useful context.

S.E Additional Evidence on Correlates of Repression

Supplemental Appendix Table 6: Total Number of HRW Alerts and Repression

	Number of HRW alerts
Preventiveness:	
Current month	-0.116
	(0.053)
	[-0.224, -0.027]
Previous month	-0.121
	(0.069)
	[-0.236, 0.006]
Suppression:	
Current month	0.224
	(0.047)
	[0.130, 0.338]
Previous month	0.070
	(0.057)
	[-0.034, 0.192]
Baseline mean	1.077
Only months with protest	X
Control for protest incidence	
Number of countries	119
Number of country-months	6,512

Notes: The table presents results from a regression of the total number of HRW alerts on our measures of repression and their first lags. The unit of analysis is a country-month. Coefficients are reported as a fraction of the baseline mean of the dependent variable, defined as the sample mean in country-months with protest but without preventiveness or suppression. For each coefficient we report an asymptotic standard error in parentheses and a 95% Bayes bootstrap credible interval in brackets, both clustered by country. The model includes country and month fixed effects. We include only country-months with a protest.

Supplemental Appendix Table 7: Economic Determinants of Suppression

	(1)	(2)	(3)	(4)	(5)
log(GDP)	-0.006		0.081		
	(0.198)		(0.219)		
	[-0.345, 0.354]		[-0.277, 0.467]		
Commodity export price component					-1.040
					(1.161)
					[-4.591, 1.153]
Remainder component					0.122
					(0.225)
					[-0.244, 0.523]
log(Government revenue)		-0.069	-0.076		
		(0.057)	(0.064)		
		[-0.168, 0.054]	[-0.187, 0.055]		
Commodity export price component				0.108	0.106
				(0.290)	(0.267)
				[-0.497, 0.811]	[-0.456, 0.679]
Remainder component				-0.079	-0.089
				(0.054)	(0.060)
				[-0.173, 0.037]	[-0.197, 0.030]
Number of country-years	729	729	729	729	729
Number of countries	90	90	90	90	90

Notes: The table presents an analogue of the estimates in Table 4, replacing preventiveness with the incidence of suppression. All models are estimated in annual differences. Economic variables are in 2015 US dollars. Annual measures of preventiveness, the incidence of suppression, and the incidence of protest are calculated as averages across months. Below each coefficient we report an asymptotic standard error in parentheses, and a 95% Bayes bootstrap credible interval in brackets, both clustered by country.

Supplemental Appendix Table 8: Preventiveness and Opposition Characteristics

	Opposition Size		Ethnic Power Relations	
	Small	Moderate/ Large	Monopoly/ Irrelevant	Dominant/ Share power
	(1)	(2)	(3)	(4)
log(Commodity export price)	0.377	0.500	0.125	0.482
	(0.178)	(0.219)	(0.291)	(0.166)
	[0.031, 0.683]	[0.108, 0.910]	[-0.515, 0.837]	[0.165, 0.781]
Difference in coefficient	0.124		0.357	
	(0.282)		(0.334)	
	[-0.414, 0.651]		[-0.399, 1.056]	
Number of country-months	3,564	5,184	1,920	6,720
Number of countries	35	55	20	69

Notes: Each column presents results from a single regression as specified in Equation (7), with the dependent variable and log(Commodity export price) in annual differences. Columns (1) and (2) compare results from countries with different sizes of opposition to the regime. Columns (3) and (4) compare results from countries with different ethnic power relations. Opposition size is constructed from the V-Dem Dataset v13 (Coppedge et al. 2023). Ethnic power relations variables are constructed from the Ethnic Power Relations (EPR) Core Dataset 2021 (Vogt et al. 2015). "Commodity export price" refers to a price index aggregating prices of 45 commodities weighted by each commodity's share in a country's exports, and the country's export share of its GDP. Below each coefficient estimate we report an asymptotic standard error in parentheses, and a 95% Bayes bootstrap credible interval in brackets, both clustered by country.

References

Auf, Y. (2018, February). The state of emergency in Egypt: An exception or rule? Accessed October 2024.

Coppedge, M., J. Gerring, C. H. Knutsen, S. I. Lindberg, J. Teorell, D. Altman, M. Bernhard, A. Cornell, M. S. Fish, L. Gastaldi, et al. (2023). V-Dem Dataset v13. Accessed March 2023.

Freedom House (2010-2019a). *Freedom in the World: The Annual Survey of Political Rights and Civil Liberties*. Lanham, Maryland: Rowman & Littlefield. Accessed April 2024.

GDELT Project (2020). GDELT 1.0 Event Database. Accessed June 2020.

Gelman, A. and J. Hill (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge, UK: Cambridge University Press.

Google (2019). Google Health Trends API - Getting Started Guide. Transmitted to authors in April 2019.

Google (2021, 2023b). Google Health Trends API. Accessed June-August 2021 and April-May 2023.

Hastie, T., R. Tibshirani, and J. Friedman (2017). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.

Horne, F. (2017, August). State of emergency ends in Ethiopia. Accessed October 2024.

Human Rights Watch (2015, September). Tunisia: Crackdown on peaceful protests. Accessed October 2024.

Leetaru, K. and P. A. Schrodt (2013). GDELT: Global Data on Events, Location and Tone. International Studies Association Annual Conference 2013. International Studies Association, San Diego, CA, 1–49.

Lei, L. (2023). Distribution-free inference on the extremum of conditional expectations via classification. *Working Paper*. Accessed August 2023.

Twitter (2020, 2023b). Advanced search tool. Accessed February-April 2023.

United Nations (2018b). World Urbanization Prospects: The 2018 Revision, Online Edition. Department of Economic and Social Affairs, Population Division. Accessed in February 2024.

United Nations Group of Experts on Geographical Names (2019). Database of Country Names and Official Languages. Transmitted to authors in May 2020.

United States Department of State (2015). Tunisia 2015 human rights report. Accessed October 2024.

Vogt, M., N.-C. Bormann, S. Rüegger, L.-E. Cederman, P. Hunziker, and L. Girardin (2015). Integrating data on ethnicity, geography, and conflict: The ethnic power relations data set family. *Journal of Conflict Resolution 59*(7), 1327–1342.

Wikipedia (2023, November). 2016-2018 Ethiopian state of emergency. Accessed October 2024.

Zepecki, A., S. Guendelman, J. DeNero, N. Prata, et al. (2020). Using application programming interfaces to access Google data for health research: Protocol for a methodological framework. *JMIR Research Protocols 9*(7), e16543.

Online Appendix for

Surveillance of Repression

Veli M. Andirin, *Yale University*[*]
Yusuf Neggers, *University of Michigan*
Mehdi Shadmehr, *University of North Carolina at Chapel Hill*
Jesse M. Shapiro, *Harvard University and NBER*

[*]E-mail: velimurat.andirin@yale.edu, yneggers@umich.edu, mshadmehr@unc.edu, jesse_shapiro@fas.harvard.edu.

O.A Additional Validation of Security Alerts

O.A.1 Comparison to GDACS Natural Disaster Alerts

Online Appendix Table 1: Summary of Crisis24, OSAC and GDELT Coverage of GDACS Disaster Alerts (Monthly) Over Time

		Proportion with:			Average # of:			
		C24	OSAC	GDELT	GDACS	C24	OSAC	GDELT
Alert Type	Obs.	alerts	alerts	alerts	alerts	alerts	alerts	alerts
Panel A: Between 2010 and 2014								
Green	2,573	0.49	0.07	-	7.28	1.29	0.10	-
Orange	217	0.84	0.25	-	18.22	3.34	0.37	-
Red	48	0.94	0.35	-	34.17	5.27	0.54	-
None	6,427	0.12	0.01	-	-	0.19	0.02	-
Panel B: Between 2015 and 2019								
Green	2,570	0.55	0.08	0.65	6.77	1.81	0.12	474.59
Orange	267	0.87	0.27	0.55	13.03	4.43	0.50	1099.52
Red	84	0.87	0.39	0.51	16.20	4.80	0.75	1100.04
None	6,430	0.17	0.01	0.59	-	0.29	0.01	34.80

Notes: The table shows statistics relating to GDACS, Crisis24, OSAC, and GDELT coverage of country-months with natural disasters. "Alert Type" is the classification given by GDACS to a natural disaster, originally coded as green (least severe), orange (medium), or red (most severe). "Obs" displays the number of country-months with at least one alert of the type given by the column "Alert Type" or a higher severity alert type. These country-months are the sample used for calculating the statistics displayed in other columns. "Proportion with C24 alerts" refer to the proportions of country-months with a GDACS alert which are also flagged as having a disaster by Crisis24. "Proportion with OSAC alerts" and "Proportion with GDELT alerts" are analogously defined. "Average # of GDACS alerts" is the average number of total GDACS alerts in country-months within the row's sample. "Average # of C24 alerts", "Average # of OSAC alerts" and "Average # of GDELT alerts" are analogously defined. Panel A and B show results for the subsets of alerts between 2010 and 2014, and between 2015 and 2019, respectively. GDELT alerts for natural disasters are sourced from the Global Knowledge Graph 2.0 (available starting February 2015) and are identified by filtering for alerts that list natural disaster as the first theme. We classify natural disasters as having occurred in a country if that country constitutes at least 30% of all mentions of countries in the GDELT alert.

Online Appendix Table 2: Summary of Crisis24 and OSAC Coverage of GDACS Disaster Alerts, Daily

		Proportion with:		Average # of:		
Alert Type	Obs.	C24 alerts	OSAC alerts	GDACS alerts	C24 alerts	OSAC alerts
Panel A: All countries						
Green	22,917	0.39	0.04	1.58	0.76	0.06
Orange	598	0.80	0.23	2.88	2.19	0.38
Red	143	0.88	0.34	3.78	3.00	0.66
None	524,883	0.09	0.01	-	0.14	0.01
Panel B: OECD countries						
Green	8,485	0.45	0.06	1.76	0.97	0.10
Orange	95	0.97	0.41	7.25	3.36	0.83
Red	17	1.00	0.53	11.59	4.24	1.59
None	104,727	0.16	0.01	-	0.27	0.02
Panel C: Non-OECD countries						
Green	14,432	0.36	0.03	1.47	0.64	0.04
Orange	503	0.77	0.19	2.06	1.97	0.30
Red	126	0.87	0.32	2.73	2.83	0.54
None	420,156	0.07	0.01	-	0.10	0.01

Notes: The table shows statistics relating to GDACS, Crisis24 and OSAC coverage of country-days with natural disasters. “Alert Type” is the classification given by GDACS to a natural disaster, originally coded as green (least severe), orange (medium), or red (most severe). “Obs” displays the number of country-days with at least one alert of the type given by the column “Alert Type” or a higher severity alert type. These country-days are the sample used for calculating the statistics displayed in other columns. “Proportion with C24 alerts” refer to the proportions of country-days with a GDACS alert which are also flagged as having a disaster by Crisis24 within 7 days. “Proportion with OSAC alerts” is analogously defined. “Average # of GDACS alerts” is the average number of total GDACS alerts in country-days within the row’s sample. “Average # of C24 alerts” is the average number of Crisis24 disaster alerts within 7 days of a GDACS alert of a given level. “Average # of OSAC alerts” is analogously defined. Panel A shows results for all countries in our sample. In Panel B, the sample includes OECD countries. In Panel C, the sample includes non-OECD countries.

O.A.2 Comparison to DONs Disease Outbreak Alerts

Online Appendix Table 3: Summary of Crisis24 and OSAC Coverage of DONs Health Alerts, Monthly

		Proportion with:		Average # of:		
		C24	OSAC	DONS	C24	OSAC
Alert Type	Obs.	alerts	alerts	alerts	alerts	alerts
Panel A: All countries						
Alerts with no cases	269	0.47	0.07	1.50	10.20	0.94
Alerts with cases	110	0.63	0.29	1.52	19.38	5.11
No alerts	17,621	0.12	0.03	-	1.95	0.35
Panel B: OECD countries						
Alerts with no cases	29	0.34	0.07	1.28	7.83	0.69
Alerts with cases	24	0.71	0.17	1.54	28.38	3.00
No alerts	3,667	0.10	0.01	-	1.45	0.16
Panel C: Non-OECD countries						
Alerts with no cases	240	0.49	0.08	1.53	10.48	0.97
Alerts with cases	86	0.60	0.33	1.51	16.87	5.70
No alerts	13,954	0.13	0.03	-	2.08	0.40

Notes: The table shows statistics relating to Crisis 24 and OSAC coverage of country-months with health emergencies as reported on by the World Health Organization's Disease Outbreak News (DONs). We sample DONs alerts from data collected by Carlson et al. (2023) and exclude alerts explicitly marked as updates in their headlines, alerts which lack a description, and alerts where the affected country is not mentioned in the alert headline. "Alert Type" is the severity of the DONS alert based on the number of cases. "Alerts with cases" are all alerts with at least one case of a disease counted. "Alerts with no cases" are all alerts with no cases. "Obs" displays the number of country-months with at least one alert of the type given by the column "Alert Type". These country-months are the sample used for calculating the statistics displayed in other columns. "Proportion with C24 alerts" refers to the proportion of country-months with a DONS alert which are also flagged as having a health emergency by Crisis 24. "Proportion with OSAC alerts" is defined analogously. "Average # of DONS alerts" is the average number of total DONS alerts in country-months within the row's sample. "Average # of C24 alerts" is the average number of Crisis 24 health alerts in country-months within the row's sample. "Average # of OSAC alerts" is defined analogously. Panel A shows results for all countries in our sample. In Panel B, the sample includes OECD countries. In Panel C, the sample includes non-OECD countries.

Online Appendix Table 4: Summary of Crisis24 and OSAC Coverage of DONs Health Alerts, Daily

		Proportion with:		Average # of:		
Alert Type	Obs.	C24 alerts	OSAC alerts	DONs alerts	C24 alerts	OSAC alerts
Panel A: All countries						
Alerts with no cases	416	0.37	0.06	1.04	0.77	0.08
Alerts with cases	126	0.52	0.20	1.10	1.01	0.33
No alerts	547,108	0.06	0.01	-	0.07	0.01
Panel B: OECD countries						
Alerts with no cases	46	0.48	0.17	1.04	2.17	0.37
Alerts with cases	25	0.64	0.16	1.04	1.40	0.24
No alerts	113,110	0.04	0.01	-	0.05	0.01
Panel C: Non-OECD countries						
Alerts with no cases	370	0.36	0.05	1.04	0.59	0.05
Alerts with cases	101	0.49	0.21	1.11	0.91	0.35
No alerts	433,998	0.06	0.01	-	0.08	0.01

Notes: The table shows statistics relating to Crisis 24 and OSAC coverage of country-days with health emergencies as reported on by the World Health Organization's Disease Outbreak News (DONs). We sample DONs alerts from data collected by Carlson et al. (2023) and exclude alerts explicitly marked as updates in their headlines, alerts which lack a description, and alerts where the affected country is not mentioned in the alert headline. "Alert Type" is the severity of the DONS alert based on the number of cases. "Alerts with cases" are all alerts with at least one case of a disease counted, "Alerts with no cases" are all alerts with no cases. "Obs" displays the number of country-days with at least one alert of the type given by the column "Alert Type". These country-days are the sample used for calculating the statistics displayed in other columns. "Proportion with C24 alerts" refers to the proportion of country-days with a DONS alert which are also flagged as having a health emergency by Crisis 24 within 7 days. "Proportion with OSAC alerts" is defined analogously. "Average # of DONS alerts" is the average number of total DONS alerts in country-days within the row's sample. "Average # of C24 alerts" is the average number of Crisis 24 health alerts within 7 days of a DONS alert of a given level. "Average # of OSAC alerts" is defined analogously. Panel A shows results for all countries in our sample. In Panel B, the sample includes OECD countries. In Panel C, the sample includes non-OECD countries.

O.A.3 Comparison to Academic Databases of Protests

Online Appendix Table 5: Comparison of Crisis24, OSAC, and GDELT to Kadivar et al. (2023, 2024) Data on Protests in Iran

Alert Locations	# in Kadivar	# in Crisis24	Proportion of Crisis24 in Kadivar	Proportion of Kadivar in Crisis24
All	460	48	0.81	0.08
Capital	12	6	1.00	0.50
Large Cities	66	18	0.89	0.24
Other Cities	382	24	0.71	0.04
Alert Locations	# in Kadivar	# in OSAC	Proportion of OSAC in Kadivar	Proportion of Kadivar in OSAC
All	460	0	—	0.00
Capital	12	0	—	0.00
Large Cities	66	0	—	0.00
Other Cities	382	0	—	0.00
Alert Locations	# in Kadivar	# in GDELT	Proportion of GDELT in Kadivar	Proportion of Kadivar in GDELT
All	460	628	0.19	0.26
Capital	12	17	0.71	1.00
Large Cities	66	102	0.49	0.76
Other Cities	382	509	0.12	0.15

Notes: The table compares the coverage of the 2017–2018 and 2019 Iranian protest episodes in Crisis24, OSAC, GDELT, and data compiled by Kadivar et al. (2023, 2024) and shared with us via e-mail in February 2024. Kadivar's data include protests in Iran during the periods from December 28, 2017 to January 6, 2018 and from November 15 to 21, 2019. Kadivar's data record protests at the district-date level. Each row considers either all protests, protests in the district containing the capital (Tehran), protests in districts containing large cities (ranked 2-10 by population), and protests in other districts (outside the top 10 by population). We exclude from our GDELT sample the 28% of GDELT protests in Iran during these periods that do not indicate the city/district of the protest. Each column labeled "Proportion of SOURCE in TARGET" presents the ratio of the number of protests in both SOURCE and TARGET matching the labeled characteristics, to the number of protests in SOURCE matching the labeled characteristics.

Online Appendix Table 6: Comparison of Crisis24, OSAC, and GDELT to Enikolopov et al. (2020) Data on Protests in Russia

Alert Locations	# in Enikolopov	# in Crisis24	Match Type	Proportion of Crisis24 in Enikolopov	Proportion of Enikolopov in Crisis24
All	84	3	Exact Day Match	0.00	0.00
			+/- 1 Day Match	1.00	0.04
Alert Locations	# in Enikolopov	# in OSAC	Match Type	Proportion of OSAC in Enikolopov	Proportion of Enikolopov in OSAC
All	84	0	Exact Day Match	—	0.00
			+/- 1 Day Match	—	0.00
Alert Locations	# in Enikolopov	# in GDELT		Proportion of GDELT in Enikolopov	Proportion of Enikolopov in GDELT
All	84	0	Exact Day Match	—	0.00
			+/- 1 Day Match	—	0.00

Notes: The table compares coverage of protest episodes from December 10-16, 2011 in Crisis24, OSAC and GDELT with data compiled by Enikolopov et al. (2020). The data from Enikolopov et al. (2020) include protests in Russia (excluding Moscow and St. Petersburg) from December 10, 2011 to December 16, 2011. The data from Enikolopov et al. (2020) record protests at the city-date level. Each row in the table considers all protests. We use two types of matches to compare protest coverage. The first type, the exact day match, requires the same city-date protest pair to appear in the data from Enikolopov et al. (2020) and the comparison dataset. The second type, the +/- 1 day match, allows for the date part of the city-date protest pair for a protest to vary by at most one day. GDELT and OSAC data do not record protests in Russia during the period covered by Enikolopov et al. (2020). Each column labeled "Proportion of SOURCE in TARGET" presents the ratio of the number of protests in both SOURCE and TARGET matching the labeled characteristics, to the number of protests in SOURCE matching the labeled characteristics.

Online Appendix Table 7: Comparison of Crisis24, OSAC, and GDELT to Li (2018) Data on Protests in China

Alert Locations	# in Li	# in Crisis24	Proportion of Crisis24 in Li	Proportion of Li in Crisis24: All	Small	Medium	Large	Very Large
All	621	253	0.055	0.019	0.017	0.007	0.018	0.067
Large Cities	258	77	0.065	0.016	0.029	0.000	0.013	0.000
Other Cities	240	70	0.071	0.013	0.000	0.009	0.000	0.080
No City	123	106	0.038	0.041	0.000	0.017	0.053	0.095
Alert Locations	**# in Li**	**# in OSAC**	**Proportion of OSAC in Li**	**Proportion of Li in OSAC: All**	**Small**	**Medium**	**Large**	**Very Large**
All	621	79	0.000	0.000	0.000	0.000	0.000	0.000
Large Cities	258	18	0.000	0.000	0.000	0.000	0.000	0.000
Other Cities	240	14	0.000	0.000	0.000	0.000	0.000	0.000
No City	123	47	0.000	0.000	0.000	0.000	0.000	0.000
Alert Locations	**# in Li**	**# in GDELT**	**Proportion of GDELT in Li**	**Proportion of Li in GDELT: All**	**Small**	**Medium**	**Large**	**Very Large**
All	621	15	0.000	0.000	0.000	0.000	0.000	0.000
Large Cities	258	5	0.000	0.000	0.000	0.000	0.000	0.000
Other Cities	240	4	0.000	0.000	0.000	0.000	0.000	0.000
No City	123	6	0.000	0.000	0.000	0.000	0.000	0.000

Notes: The table compares coverage of protest episodes in China from Crisis24, OSAC, GDELT with data compiled by Li (2018). During our sample period, the data from Li (2018) include protests in China from January 2010 through December 2012. We exclude from our GDELT sample the 43% of GDELT protests in China during these periods that do not indicate the city and province of the protest. Each protest in the data from Li (2018) is a location-date range pair and each protest in our comparison datasets is a location-date pair. A match requires a location-date protest from the comparison dataset to fall in the location-date range for a protest from Li (2018). The data from Li (2018) provide the province, and sometimes, the city where a protest occurred. To match the location part for protests, we first consider whether both protests occurred in the same city. For all protests without city of occurrence, we then try to match the location part based on whether both protests occurred in the same province. We exclude protests for which we are unable to obtain date information, and protests that occurred online. Each row considers all protests, protests in large cities (15 of China's largest cities by estimated population in 2021 according to United Nations 2018a), protests in other cities (outside the top 15 by population in 2021), or protests that do not list a city. In the rightmost columns, we subset the protest data from Li (2018) by protester count, using the following categories: small (1-50), medium (51-500), large (501-5000), and very large (>5000). The data from Li (2018) contain 181 small protests, 246 medium protests, 160 large protests, and 34 very large protests. Each column labeled "Proportion of SOURCE in TARGET" presents the ratio of the number of protests in both SOURCE and TARGET matching the labeled characteristics, to the number of protests in SOURCE matching the labeled characteristics.

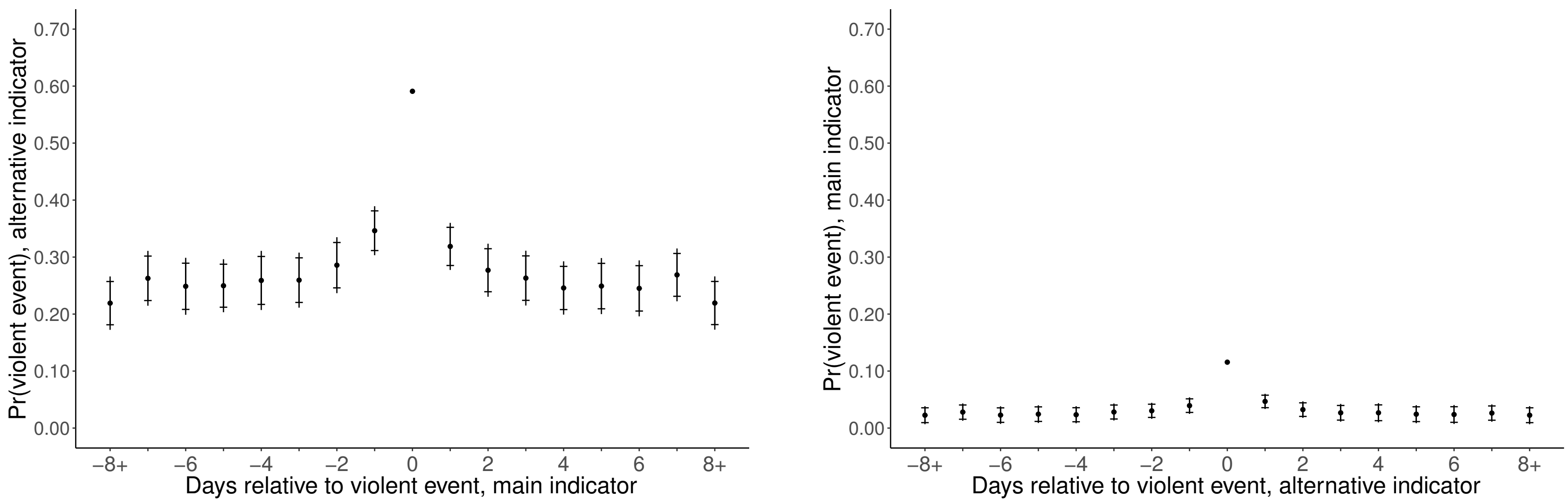
Online Appendix Table 8: Comparison of Crisis24, OSAC, and GDELT to Centre for Social Conflict and Cohesion Studies (COES, 2020) Data on Protests in Chile

	# in	# in	Proportion of	Proportion of COES in C24					
Alert Locations	COES	C24	C24 in COES	All	Size Unknown	Small	Medium	Large	Very Large
All	31,754	711	0.60	0.01	0.01	0.02	0.02	0.04	0.22
Capital	1,969	327	0.73	0.12	0.11	0.12	0.15	0.15	0.45
Large Cities	5,418	118	0.58	0.01	0.01	0.01	0.02	0.04	0.06
Other Cities	24,367	266	0.45	0.00	0.01	0.00	0.00	0.01	0.06
	# in	# in	Proportion of	Proportion of COES in OSAC					
Alert Locations	COES	OSAC	OSAC in COES	All	Size Unknown	Small	Medium	Large	Very Large
All	31,754	47	0.62	0.00	0.00	0.00	0.00	0.00	0.02
Capital	1,969	28	0.71	0.01	0.01	0.00	0.01	0.01	0.04
Large Cities	5,418	9	0.67	0.00	0.00	0.00	0.00	0.00	0.01
Other Cities	24,367	10	0.30	0.00	0.00	0.00	0.00	0.00	0.00
	# in	# in	Proportion of	Proportion of COES in GDELT					
Alert Locations	COES	GDELT	GDELT in COES	All	Size Unknown	Small	Medium	Large	Very Large
All	31,754	2,208	0.48	0.03	0.04	0.04	0.04	0.07	0.18
Capital	1,969	1,657	0.59	0.49	0.56	0.39	0.42	0.41	0.44
Large Cities	5,418	78	0.22	0.00	0.00	0.00	0.00	0.00	0.00
Other Cities	24,367	473	0.13	0.00	0.00	0.00	0.00	0.00	0.01

Notes: The table compares the coverage of protests in Crisis24, OSAC, GDELT, and a dataset compiled by the Centre for Social Conflict and Cohesion Studies (COES; 2020). COES data include protests in Chile from 2010 to 2019 at the commune-date level. Each row considers either all protests, protests in the capital (the commune of Santiago), protests in large communes (9 of Chile's largest communes by population, other than the commune of Santiago), and protests in other communes (outside the top 10 by population). We exclude from our GDELT sample the 50% of GDELT protests in Chile during this period that do not indicate the city of the protest. In the rightmost columns, we restrict to subsets of COES protests of the following size, categorized by the number of protesters: size unknown, small (1-50), medium (51-500), large (501-5000), and very large (>5000). COES samples contain 21,652 commune-date pairs with protests of unknown size, 6,127 with small protests, 6,174 with medium protests, 2,490 with large protests, and 491 with very large protests. Each column labeled "Proportion of SOURCE in TARGET" presents the ratio of the number of protests in both SOURCE and TARGET matching the labeled characteristics, to the number of protests in SOURCE matching the labeled characteristics.

O.A.4 Comparison to Alternative Indicator of Suppression

Online Appendix Figure 1: Relationship Among Alternative Indicators of Suppression



Notes: The plots show the relationship between our main indicator of suppression (s_{it}), and an alternative indicator constructed from the Armed Conflict Location & Event Data Project (ACLED 2022). We perform the analysis on the subset of our sample countries and period that is covered by ACLED. This results in an unbalanced panel of 122 countries over the period from 2010 through 2019. To construct the alternative indicator of suppression, we create an indicator for whether ACLED includes any events categorized as “Protest” or “Riot” with “military versus rioters” or “military versus protesters” interactions on the given country-date. Each plot is constructed from a regression in which the unit of analysis is the country-date and the model includes country and date fixed effects. The independent variables of interest are seven leads (relative days -7 through -1) and seven lags (relative days 1 through 7) of our main indicator in the left plots and seven leads and lags of the alternative indicator in the right plots. Two variables reflecting the cumulative number of events more than seven days in the future and more than seven days in the past are included in both columns as well. The contemporaneous event indicator is excluded as a normalization. The dependent variable is the alternative indicator in the left plots and our main indicator in the right plots. We recenter the y-axis in each plot by adding a constant equal to the sample mean of the dependent variable on dates on which there is an event according to the indicator used in the x-axis. The inner bars depict 95% pointwise confidence intervals and the outer lines depict 95% uniform sup-t bands, both based on inference clustered by country.

O.A.5 Comparison to OSAC alerts

Online Appendix Table 9: Comparison of Crisis24 and OSAC Coverage

Crisis24 Coverage of OSAC Alerts				
Alert Type	Obs.	Proportion with C24 alerts	Average # of OSAC alerts	Average # of C24 alerts
Panel A: All countries				
Occurred Protest	2,188	0.850	1.862	3.448
Anticipated Protest	1,431	0.804	1.456	2.320
No Protest	15,812	0.445	-	1.214
Panel B: OECD countries				
Occurred Protest	710	0.870	2.472	3.672
Anticipated Protest	535	0.843	1.707	2.761
No Protest	3,010	0.577	-	1.905
Panel C: Non-OECD countries				
Occurred Protest	1,478	0.840	1.569	3.340
Anticipated Protest	896	0.781	1.306	2.057
No Protest	12,802	0.414	-	1.051
OSAC Coverage of Crisis24 Alerts				
Alert Type	Obs.	Proportion with OSAC alerts	Average # of C24 alerts	Average # of OSAC alerts
Panel A: All countries				
Occurred Protest	8,897	0.209	3.005	0.411
Anticipated Protest	5,782	0.199	2.274	0.304
No Protest	9,103	0.036	-	0.046
Panel B: OECD countries				
Occurred Protest	2,355	0.262	3.541	0.688
Anticipated Protest	1,808	0.249	2.765	0.442
No Protest	1,365	0.067	-	0.099
Panel C: Non-OECD countries				
Occurred Protest	6,542	0.190	2.811	0.311
Anticipated Protest	3,974	0.176	2.050	0.240
No Protest	7,738	0.031	-	0.037

Notes: The table shows statistics relating to OSAC and Crisis 24 coverage of country-months. For "Crisis24 coverage of OSAC alerts", "Alert Type" refers to whether statistics pertain to months with OSAC alerts related to occurred protests, anticipated protests, or to months with no protests according to OSAC. "Obs" displays the number of country-months with at least one alert of the type given by the column "Alert Type." These country-months are the sample used for calculating the statistics displayed in other columns. "Proportion with C24 alerts" refers to the proportion of country-months with an OSAC alert which are also flagged as having an alert of the same type in Crisis24. "Average # of OSAC alerts" is the average number of total OSAC alerts of the same type in country-months within the row's sample. "Average # of C24 alerts" is the average number of Crisis24 alerts of the same type in country-months within the row's sample. Panel A shows results for all countries in our sample. In Panel B, the sample includes OECD countries. In Panel C, the sample includes non-OECD countries. "OSAC coverage of Crisis24 alerts" follows the same structure, reversing the roles of Crisis24 and OSAC.

Online Appendix Table 10: Comparison of Crisis24 and OSAC Coverage Over Time

Crisis24 Coverage of OSAC Alerts				
Alert Type	Obs.	Proportion with C24 alerts	Average # of OSAC alerts	Average # of C24 alerts
Panel A: Between 2010 and 2014				
Occurred Protest	919	0.789	1.496	3.011
Anticipated Protest	551	0.719	1.285	2.024
No Protest	8,081	0.413	-	1.139
Panel B: Between 2015 and 2019				
Occurred Protest	1,269	0.894	2.127	3.764
Anticipated Protest	880	0.858	1.563	2.506
No Protest	7,731	0.478	-	1.292
OSAC Coverage of Crisis24 Alerts				
Alert Type	Obs.	Proportion with OSAC alerts	Average # of C24 alerts	Average # of OSAC alerts
Panel A: Between 2010 and 2014				
Occurred Protest	4,065	0.178	2.944	0.277
Anticipated Protest	2,371	0.167	2.257	0.224
No Protest	4,935	0.039	-	0.050
Panel B: Between 2015 and 2019				
Occurred Protest	4,832	0.235	3.055	0.523
Anticipated Protest	3,411	0.221	2.285	0.359
No Protest	4,168	0.032	-	0.042

Notes: The table shows statistics relating to OSAC and Crisis 24 coverage of country-months before and after 2015. For "Crisis24 coverage of OSAC alerts", "Alert Type" refers to whether statistics pertain to months with OSAC alerts related to occurred protests, anticipated protests, or to months with no protests according to OSAC. "Obs" displays the number of country-months with at least one alert of the type given by the column "Alert Type." These country-months are the sample used for calculating the statistics displayed in other columns. "Proportion with C24 alerts" refers to the proportion of country-months with an OSAC alert which are also flagged as having an alert of the same type in Crisis24. "Average # of OSAC alerts" is the average number of total OSAC alerts of the same type in country-months within the row's sample. "Average # of C24 alerts" is the average number of Crisis24 alerts of the same type in country-months within the row's sample. Panel A shows results for all countries in our sample. In Panel A, the sample includes years between 2010 and 2014. In Panel B, the sample includes years between 2015 and 2019. "OSAC coverage of Crisis24 alerts" follows the same structure, reversing the roles of Crisis24 and OSAC.

O.B Additional Summary Statistics on Protests

Online Appendix Table 11: Statistics on Runs of Protests Within Months, Across Countries

Mean difference in share of days with protest following given number of consecutive days with vs. without protest:					
1 Day	2 Days	3 Days	4 Days	5 Days	6 Days
0.072	0.133	0.200	0.238	0.232	0.371
(0.037, 0.044)	(0.063, 0.097)	(0.095, 0.166)	(0.111, 0.254)	(0.121, 0.332)	(0.092, 0.413)

Notes: The table reports the mean, across countries, of the difference between the share of days with protest following the given number of consecutive days of protest in the same calendar month and the frequency following the same number of consecutive days of non-protest (Ritzwoller and Romano 2021). In parentheses below each mean is the range, from the 2.5th to the 97.5th percentiles, of the same statistic across 200 random permutations of the protest indicator across dates within each country.

O.C Existence of an Equilibrium in the Dynamic Model

Here we give example sufficient conditions for the existence of an equilibrium in stationary pure strategies in the dynamic extension with a finite number of environments. Following the static model, we assume that, for all environments $k>0$, we have that $\underline{p}(k)=\underline{q}(k)=0$, $\mathscr{M}(k)$ is a singleton, $\Omega(k)=\{0,1\}$, $p(0,m,a,k)=0$, $p(1,m,1,k)=p(1,1,a,k)=1$. We continue the notation from the proof of Proposition 4. Similarly to the proof of Proposition 4, the following Lemma is immediate.

Lemma 4. *In any equilibrium, there is a cutoff $\overline{q}^*(k_t)\in(0,1]$ such that the regime suppresses a protest if and only if $q^*(y_t,k_t)>\overline{q}^*(k_t)$, where $\overline{q}^*(k_t)=\min\left\{1,\frac{\sigma(k_t)}{\delta\max\left\{\left(\overline{V}^*(k_t)-V(k_t)\right),0\right\}}\right\}$, so that $S^*(k)=\Pr(q^*(y,k)>\overline{q}^*(k)\,|\,k)$ and $Q^*(k)=\Pr(q^*(y,k)\leq\overline{q}^*(k)\,|\,k)\mathrm{E}(q^*(y,k)|q^*(y,k)\leq\overline{q}^*(k),k)$.*

Under a regularity condition on the regime's information, Lemma 4 implies a uniform lower bound on the equilibrium probability of revolution.

Assumption 1. *In every environment k, the distribution of $q^*(k,y)$ has full support on* $[0,1]$.

Lemma 5. *Under Assumption 1, there exists a constant $0<\underline{Q}\leq 1$ such that in any equilibrium in stationary pure strategies, $Q^*(k)\geq\underline{Q}$ for all k.*

Proof. Because the number of environments is finite, the regime's stage-game payoff has an upper bound, $\overline{v}$, and a lower bound $\underline{v}$, which implies that the difference in continuation payoff between any two states can differ by no more than $\Delta=(\overline{v}-\underline{v})/(1-\delta)$. Therefore take

$$\underline{Q}=\min_j\left\{\Pr\left(q^*(j,y)\leq\min\left\{1,\frac{\sigma(j)}{\delta\Delta}\right\}|\,j\right)\mathrm{E}\left(q^*(j,y)|q^*(j,y)\leq\min\left\{1,\frac{\sigma(j)}{\delta\Delta}\right\}|\,j\right)\right\}$$

where we take $\min\left\{1,\frac{\sigma(j)}{\delta\Delta}\right\}=1$ if $\Delta=0$. By Assumption 1, $\underline{Q}>0$, which completes the proof. □

Assumption 2. *Recalling that $V(k)=-\frac{l(k)}{1-\delta}<0$, we assume that $-\frac{l(k)-\delta\max_j\{l(j)\}}{1-\delta}<-\max_j\{\sigma(j)+\rho(j)\}$ for all k.*

Assumption 2 implies that, in any environment, the regime strictly prefers the continuation payoff to the payoff under revolution.

Lemma 6. *Under Assumption 2, in any equilibrium in stationary pure strategies, $\overline{V}^*(k)-V(k)\geq\frac{l(k)-\delta\max_j\{l(j)\}}{1-\delta}-\max_j\{\sigma(j)+\rho(j)\}>0$ for all k.*

Proof. It is immediate that $\overline{V}^*(k)-V(k)\geq-\max_j\{\sigma(j)+\rho(j)\}+\delta\min_j\{V(j)\}-V(k)$. The result then follows from Assumption 2 noticing that

$$\delta\min_j\{V(j)\}-V(k)=\frac{\ell(k)-\delta\max_j\{\ell(j)\}}{1-\delta}.$$

□

We now turn to the opposition's strategies. In any equilibrium in stationary pure strategies, the opposition's expected discounted payoff at the start of any period t in environment $k_t > 0$ can be written as a function $W^*(k_t)$ of the environment. For any $k_t > 0$, let $\overline{W}^*(k_t) = \sum_l \mathbf{K}_{k_t,l} W^*(l)$ denote the expectation over $W^*(l)$ with respect to the transition probabilities $\mathbf{K}_{k,\ell} = \Pr(k_t = j \,|\, k_{t-1} = k)$ starting from environment k_t. In the case where revolution has occurred, i.e., where $k_t = 0$, the opposition's expected discounted payoff, $W(k_t) = b(k_t)/(1-\beta)$, does not depend on the equilibrium being played.

By Lemma 4, in any equilibrium in stationary pure strategies, in any environment k, the opposition's continuation payoff if protest occurs is

$$Q^*(k)\beta W(k) - S^*(k)\alpha(k) + \beta(1 - Q^*(k))\overline{W}^*(k)$$

whereas the opposition's continuation payoff if protest does not occur is $\beta \overline{W}^*(k)$. The difference between the opposition's continuation payoff with and without protest is then given by

$$Q^*(k)\beta\left(W(k) - \overline{W}^*(k)\right) - S^*(k)\alpha(k).$$

We impose sufficient conditions so that the opposition strictly prefers that protest occurs.

Assumption 3. *For $\underline{Q}$ defined as in Lemma 5, we assume that*

$$\underline{Q}\beta\left(-\max\left\{0, \max_j\{-\alpha(j)\}\right\} + W(k) - \beta\max_j\{W(j)\}\right) > \max\{0, \alpha(k)\}$$

for all k.

A sufficient condition for Assumption 3 is that $W(k) - \beta\max_j\{W(j)\} > 0$ and $|\alpha(k)|$ has sufficiently small magnitude for all k.[1] Under Assumptions 1 and 3, the opposition strictly prefers for protest to occur in any environment.

Lemma 7. *Under Assumptions 1 and 3, in any equilibrium in stationary pure strategies,*

$$Q^*(k)\beta\left(W(k) - \overline{W}^*(k)\right) - S^*(k)\alpha(k) > 0$$

in any environment k.

Proof. By Assumption 3 we know that, for all k,

$$\max_j\{W(j)\} - \beta\max_j\{W(j)\} \geq W(k) - \beta\max_j\{W(j)\} > \max\left\{0, \max_j\{-\alpha(j)\}\right\}$$

where the first inequality follows because $\max_j\{W(j)\} \geq W(k)$ and the second from Assumption 3. From the above we know that

$$\max_j\{W(j)\} > \frac{\max\left\{0, \max_j\{-\alpha(j)\}\right\}}{1-\beta}$$

[1]In particular, it suffices that $|\alpha(k)| < \frac{\underline{Q}\beta}{\underline{Q}\beta+1}\left(\min_j\{W(j)\} - \beta\max_j\{W(j)\}\right) = \frac{\underline{Q}\beta}{\underline{Q}\beta+1}\frac{\min_j\{b(j)\} - \beta\max_j\{b(j)\}}{1-\beta}$ for all k.

and therefore that the opposition prefers revolution in the best possible environment to any feasible stage game payoffs absent revolution. It follows that $\overline{W}^*(k) \leq \max\left\{0, \max_j\{-\alpha(j)\}\right\} + \beta \max_j\{W(j)\}$, from which we know that

$$W(k) - \overline{W}^*(k) \geq -\max\left\{0, \max_j\{-\alpha(j)\}\right\} + W(k) - \beta \max_j\{W(j)\}.$$

Notice next that

$$S^*(k)\alpha(k) \leq \max_{S\in[0,1]} S\alpha(k) = \max\{0, \alpha(k)\}.$$

The result is then immediate from Assumption 3 and the definition of $\underline{Q}$. □

Lemma 8. *Under Assumptions 1 and 3, in any equilibrium in stationary pure strategies:*

1. $a^*(k_t, \omega_t) = 1$ if and only if $\omega_t = 1$ and $\overline{p}^*(k_t) = 1$, and
2. $m^*(k_t, \omega_t) = 1$ if and only if $\omega_t = 1$, $\overline{p}^*(k_t) < 1$, and $\mu_t < \overline{\mu}^*(k_t)$, where
$$\overline{\mu}^*(k_t) = \Pr(p^*(x_t, k_t) \leq \overline{p}^*(k_t) \mid k_t, \omega_t = 1, m_t = 1)\left(Q^*(k_t)\beta\left(W(k_t) - \overline{W}^*(k_t)\right) - \alpha(k_t)S^*(k_t)\right) \geq 0,$$
and is uniformly bounded from above by some $M > 0$ across all equilibria and environments.

Proof. Part 1. The "if" part follows from Lemma 7. The "only if" part follows from our tiebreaking assumption.

Part 2. From part 1, because $\mu_t > 0$, we only need to consider the case where $\overline{p}^*(k_t) < 1$. Because protest does not occur when $\omega_t = 0$, we only need to consider the case where $\omega_t = 1$. If the opposition takes $m_t = 1$ in environment k, its expected discounted payoff is

$$-\mu(k_t) + \Pr(p^*(x_t, k_t) > \overline{p}^*(k_t) \mid k_t)\beta\overline{W}^*(k_t)$$
$$+ \Pr(p^*(x_t, k_t) \leq \overline{p}^*(k_t) \mid k_t)\left(Q^*(k_t)\beta W(k_t) + (1 - Q^*(k_t))\beta\overline{W}^*(k_t) - \alpha(k_t)S^*(k_t)\right),$$

where $p^*(x_t, k_t)$ is the regime's belief that protest happens given x_t. If the opposition chooses $m_t = 0$, the expected discounted payoff is $\beta\overline{W}^*(k_t)$. The characterization of $\overline{\mu}^*(k_t)$ then follows from comparing these payoffs. That $\overline{\mu}^*(k_t)$ is uniformly bounded is immediate given that all stage game payoffs are bounded. That $\overline{\mu}^*(k_t) \geq 0$ follows from Lemma 7. □

Assumption 4. *The environments $k > 0$ are partitioned into two sets K_1 and K_2 such that $\rho(k) < \underline{\rho}$ for all $k \in K_1$ and $\rho(k) > \overline{\rho}$ for all $k \in K_2$, where*

$$\underline{\rho} = \min_k\left\{\inf_{\overline{q}\in[0,1]}\left\{\delta Q(k,\overline{q})\left(-\max_j\{\sigma(j) + \rho(j)\} + \frac{l(k) - \delta\min_j\{l(j)\}}{1-\delta}\right) + \sigma(k)S(k,\overline{q})\right\}\right\},$$

for $Q(k,\overline{q}) = \Pr(q^*(y,k) \leq \overline{q} \mid k)\mathrm{E}(q^*(y,k) \mid q^*(y,k) \leq \overline{q}, k)$ and $S(k,\overline{q}) = \Pr(q^*(y,k) > \overline{q} \mid k)$, and where

$$\overline{\rho} = \max_k\left\{\frac{\delta}{1-\delta}l(k) + \sigma(k)\right\}.$$

Notice that by Assumption 2, $-\max_j\{\sigma(j)+\rho(j)\}+\frac{l(k)-\delta\min_j\{l(j)\}}{1-\delta}>0$. Notice also that by Assumption 1, if $S(k,\overline{q})\to 0$, then $Q(k,\overline{q})\to \mathrm{E}(q^*(y,k)|k)>0$, so that the infimum is strictly positive for all k. Thus, $\underline{\rho}>0$. It is also immediate that $\overline{\rho}>0$.

Remark 1. There is an open set of parameters in which both Assumptions 2 and 4 hold.

Proof. For all $k\in K_2$,from Assumptions 2 and 4,

$$\max_j\left\{\frac{\delta}{1-\delta}l(j)+\sigma(j)\right\}<\rho(k)\leq\max_j\rho(j)<-\max_j\{\sigma(j)\}+\frac{l(k)-\delta\min_j\{l(j)\}}{1-\delta}.$$

Such $\rho(k)$ exists, if and only if $\max_j\left\{\frac{\delta}{1-\delta}l(j)+\sigma(j)\right\}<-\max_j\{\sigma(j)\}+\frac{l(k)-\delta\min_j\{l(j)\}}{1-\delta}$; that is, if and only if $2\max_j\{\sigma(j)\}<\frac{l(k)-\delta\left(\min_j\{l(j)\}+\max_j\{l(j)\}\right)}{1-\delta}$. Moreover, $\frac{l(k)-\delta\left(\min_j\{l(j)\}+\max_j\{l(j)\}\right)}{1-\delta}\geq\frac{l(k)-2\delta\max_j\{l(j)\}}{1-\delta}$. Thus, it suffices to have $2\max_j\{\sigma(j)\}<\frac{l(k)-2\delta\max_j\{l(j)\}}{1-\delta}$. □

Lemma 9. *Under Assumptions 2 and 4, in any equilibrium in stationary pure strategies, $\overline{p}^*(k)<1$ for all $k\in K_1$ and $\overline{p}^*(k)=1$ for all $k\in K_2$.*

Proof. From Lemma 6,

$$\delta Q^*(k)\left(-\max_j\{\sigma(j)+\rho(j)\}+\frac{l(k)-\delta\min_j\{l(j)\}}{1-\delta}\right)+\sigma(k)S^*(k)$$
$$\leq\delta Q^*(k)\big(\overline{V}^*(k)-V(k)\big)+\sigma(k)S^*(k).$$

Thus, $\overline{p}^*(k)<1$ if

$$\rho(k)<\delta Q^*(k)\left(-\max_j\{\sigma(j)+\rho(j)\}+\frac{l(k)-\delta\min_j\{l(j)\}}{1-\delta}\right)+\sigma(k)S^*(k),$$

which holds by Assumption 4 for all $k\in K_1$. Moreover,

$$\delta Q^*(k)\big(\overline{V}^*(k)-V(k)\big)+\sigma(k)S^*(k)\leq-\delta V(k)+\sigma(k)$$
$$=\frac{\delta}{1-\delta}l(k)+\sigma(k)\leq\max_k\left\{\frac{\delta}{1-\delta}l(k)+\sigma(k)\right\}=\overline{\rho},$$

where the last equality is from the definition of $\overline{\rho}$ in Assumption 4. For all $k\in K_2$, by Assumption 4, $\rho(k)\geq\overline{\rho}$, implying that $\overline{p}^*(k)=1$. □

We now make the following assumptions to ensure the continuity of value functions and the functions defined by $\overline{\mu}^*(\cdot)$, $\overline{p}^*(\cdot)$, $\overline{q}^*(\cdot)$ above, which will guarantee a fixed point and hence the existence of an equilibrium.

From Lemmas 8 and 9, in any equilibrium, $a_t = m_t = 0$ if $\omega_t = 0$. For all $k_t \in K_2$, we have $\overline{p}^*(k_t) = 1$ and $a^*(k_t) = 1$ if $\omega_t = 1$. Now, let $\xi(\cdot) = (\overline{\mu}^*(\cdot), \overline{p}^*(\cdot), \overline{q}^*(\cdot)) \in [0,M]^{K_1} \times [0,1]^{K_1} \times [0,1]^K$, be a cutoff strategy profile. Now define the mapping $\Xi(\cdot) : [0,M]^{K_1} \times [0,1]^{K_1} \times [0,1]^K \to [0,M]^{K_1} \times [0,1]^{K_1} \times [0,1]^K$ such that $\Xi(\xi) = \big(\Xi_M(\xi), \Xi_p(\xi), \Xi_q(\xi)\big)$, with

$$
\begin{aligned}
\Xi_M(\xi) &= P_m(\xi,\cdot)\big(Q^*(\xi,\cdot)\beta\big(W(\cdot) - \overline{W}^*(\xi,\cdot)\big) - \alpha(\cdot)S^*(\xi,\cdot)\big),\\
\Xi_p(\xi) &= \min\left\{1, \frac{\rho(\cdot)}{\delta Q^*(\xi,\cdot)\big(\overline{V}^*(\xi,\cdot) - V(\cdot)\big) + \sigma(\cdot)S^*(\xi,\cdot)}\right\},\\
\Xi_q(\xi) &= \min\left\{1, \frac{\sigma(\cdot)}{\delta\big(\overline{V}^*(\xi,\cdot) - V(\cdot)\big)}\right\},
\end{aligned}
$$

where $P_m(\xi,\cdot) = \Pr(p^*(\xi,x,\cdot) \le \overline{p}^*(\cdot) \mid \cdot, \omega = m = 1)$, with $p^*(\xi,x,\cdot)$ determined according to the Bayes rule when possible, and $\overline{V}^*(\xi,\cdot)$ and $\overline{W}^*(\xi,\cdot)$ are defined as follows. First, to define $\overline{V}^*(\xi,\cdot)$, we take

$$
\begin{aligned}
\overline{V}^*(\xi,k) = \sum_j \mathbf{K}_{kj} \Bigg[&\delta\overline{V}^*(\xi,j) - \mathbf{1}\{j \in K_1\}\Pr(p^*(x,j) > \overline{p}^*(j) \mid j)\rho(j)\\
&+ \Pr(\omega = 1 \mid j) \times \big[Q^*(\overline{q}^*(j),j)\delta\big(V(j) - \overline{V}^*(\xi,j)\big) - \sigma(j)S^*(\overline{q}^*(j),j)\big]\\
&\times [\mathbf{1}\{j \in K_2\} + \mathbf{1}\{j \in K_1\}\Pr(\mu < \overline{\mu}^*(j) \mid j)P_m(\xi,j)]\Bigg].
\end{aligned}
$$

Next, to define $\overline{W}^*(\xi,\cdot)$, we take

$$
\begin{aligned}
\overline{W}^*(\xi,k) = \sum_j \mathbf{K}_{kj} \Bigg[&\beta\overline{W}^*(\xi,j) - \mathbf{1}\{j \in K_1\}\Pr(\omega = 1 \mid j)\Pr(\mu < \overline{\mu}^*(j) \mid j)\mathrm{E}[\mu \mid \mu < \overline{\mu}^*(j), j]\\
&+ \Pr(\omega = 1 \mid j) \times \big[Q^*(\overline{q}^*(j),j)\beta\big(W(j) - \overline{W}^*(\xi,j)\big) - \alpha(j)S^*(\overline{q}^*(j),j)\big]\\
&\times [\mathbf{1}\{j \in K_2\} + \mathbf{1}\{j \in K_1\}\Pr(\mu < \overline{\mu}^*(j) \mid j)P_m(\xi,j)]\Bigg].
\end{aligned}
$$

Recalling Lemmas 4, 6, 8, and the proof of Proposition 4, we recognize that a fixed point of $\Xi(\xi)$ corresponds to the cutoffs of an equilibrium in stationary pure strategies.

Assumption 5. *For all k:*

1. $\Pr(q^*(k,y) \le \eta \mid k)$ is continuous in η,
2. $\Pr(\mu \le \eta \mid k)$ is continuous in η,
3. $\Pr\left(\frac{\Lambda(x|k,\omega=1,m=1)}{\Lambda(x|k)} \le \eta \;\mid\; k, \omega, m\right)$ is continuous in η,
4. μ has full support on $\mathbb{R}_{>0}$.

Lemma 10. *Under Assumptions 1 to 5, $S^*(\xi,\cdot)$, $Q^*(\xi,\cdot)$, $P_m(\xi,\cdot)$, $\overline{V}^*(\xi,\cdot)$, and $\overline{W}^*(\xi,\cdot)$ are continuous in ξ.*

Proof. The continuity of $S^*(\xi,\cdot)$ and $Q^*(\xi,\cdot)$ is immediate from Part (1) of Assumption 5. Next, note that, from Parts (3) and (4) of Assumption 5, and Bayes rule, for $\overline{\mu}^*(\cdot)>0$, we have

$$P_m(\xi,\cdot)=\Pr(p^*(\xi,x,\cdot)\leq\overline{p}^*(\cdot)\,|\,\cdot,\omega=m=1)$$
$$=\Pr\left(\frac{\Lambda(x\,|\,\cdot,\omega=1,m=1)}{\Lambda(x\,|\,\cdot)}\leq\frac{\overline{p}^*(\cdot)}{\Pr(\omega=1\,|\,\cdot)\Pr(\mu<\overline{\mu}^*(\cdot)\,|\,\cdot)}\;|\,\cdot,\omega=m=1\right).$$

The term $\frac{\overline{p}^*(\cdot)}{\Pr(\omega=1|\cdot)\Pr(\mu<\overline{\mu}^*(\cdot)|\cdot)}$ is continuous in ξ from Part (2) of Assumption 5. Moreover, if $\overline{\mu}^*(\cdot)=0$, then $p^*(\xi,x,\cdot)=0$. Thus, from Part (3) of Assumption 5 $P_m(\xi,\cdot)$ is continuous. Finally, from the definitions of $\overline{V}^*(\xi,\cdot)$ and $\overline{W}^*(\xi,\cdot)$ above, the earlier parts of this Lemma, and Part (3) of Assumption 5, $\overline{V}^*(\xi,\cdot)$ and $\overline{W}^*(\xi,\cdot)$ are continuous in ξ. □

Proposition 7. *Under Assumptions 1 to 5, a cutoff equilibrium exists.*

Proof. The mapping $\Xi(\xi)$ is non-empty and convex-valued, from a non-empty, compact, convex subset of a Euclidean space to itself. Moreover, it is continuous by Lemma 10. Therefore, it has a fixed point by the Kakutani Fixed Point Theorem. □

O.D Additional Variables Used in Sensitivity Analysis

We obtain data on the dates of national elections from ElectionGuide (International Foundation for Electoral Systems 2021). We also obtain daily data on a range of financial indices from the sources we describe below. When we include these financial indices in our predictive model, we apply a 99.5% winsorization to their standardized values. Online Appendix Figure 2 gives the dates of coverage of each financial index for each sample country.[2] We now discuss each financial index in more detail.

O.D.0.1 Exchange Rates

We obtain daily data on exchange rates with respect to the US Dollar for a set of currencies from Bloomberg L.P. (2021a). We match currencies to countries using the ISO 4217 (International Organization for Standardization 2018). For countries that adopted the Euro during our sample period, we use the exchange rate of the official currency prior to the transition.

O.D.0.2 Sovereign Bond Indices

We obtain from Bloomberg L.P. (2018) and J.P. Morgan (2021) daily data on the J.P. Morgan Emerging Market Bond Index Global (EMBIG) yield index. We use data quoted to maturity whenever available (usually until 2018) and data quoted to worst otherwise. For 59 countries, we observe data for both forms of quotes for at least 158 dates. Using these periods of overlap we calculate the correlation between the

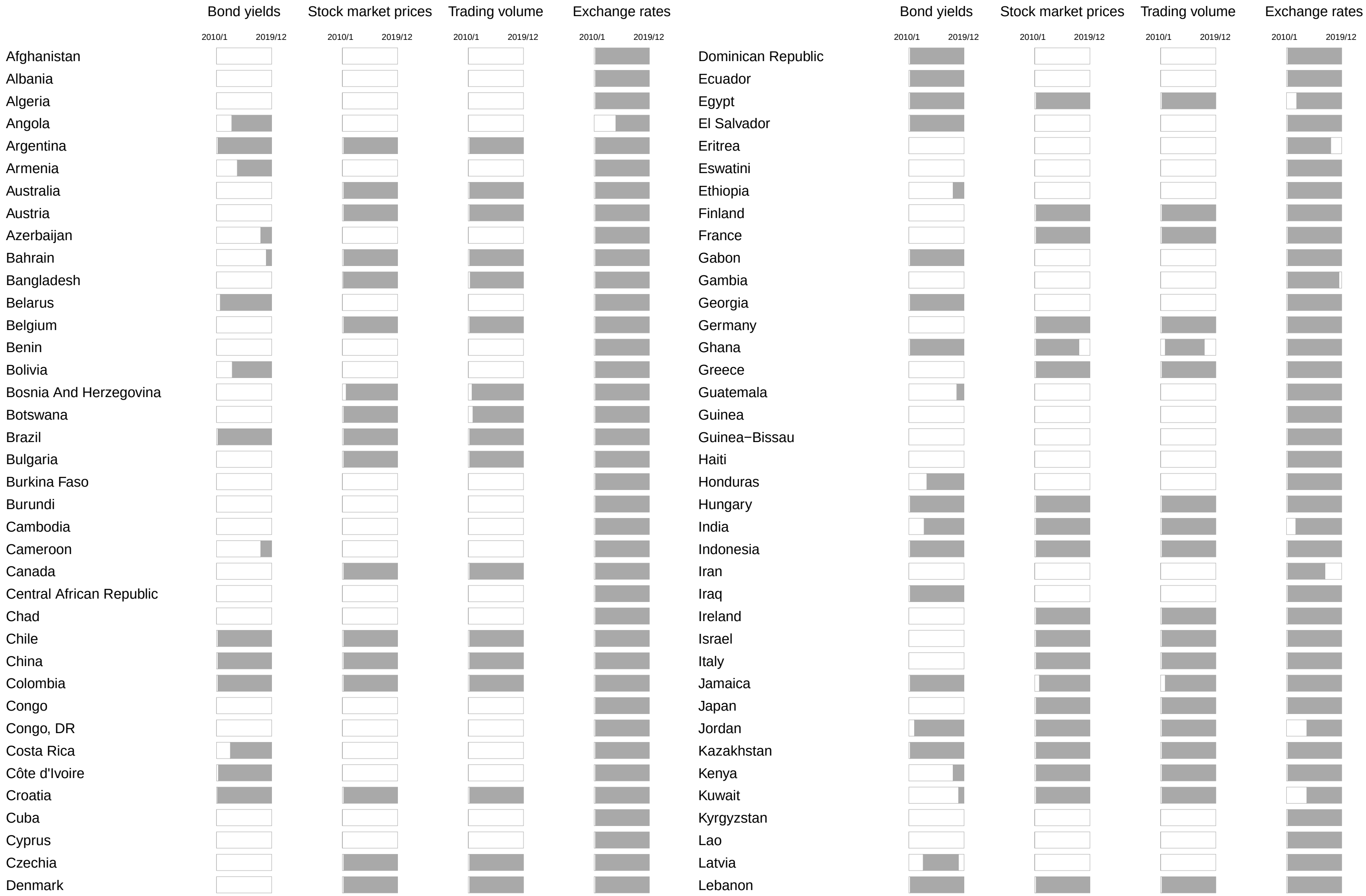
[2]To accommodate weekends and holidays, we impute the value of each indicator to its last recorded value when there is a gap of one or two days between consecutive recorded values.

daily change in the index quoted to maturity and the daily change in the index quoted to worst, separately by country. We find that the correlation is above 0.9 for all 59 countries, and above 0.99 for all except 5 of these countries.

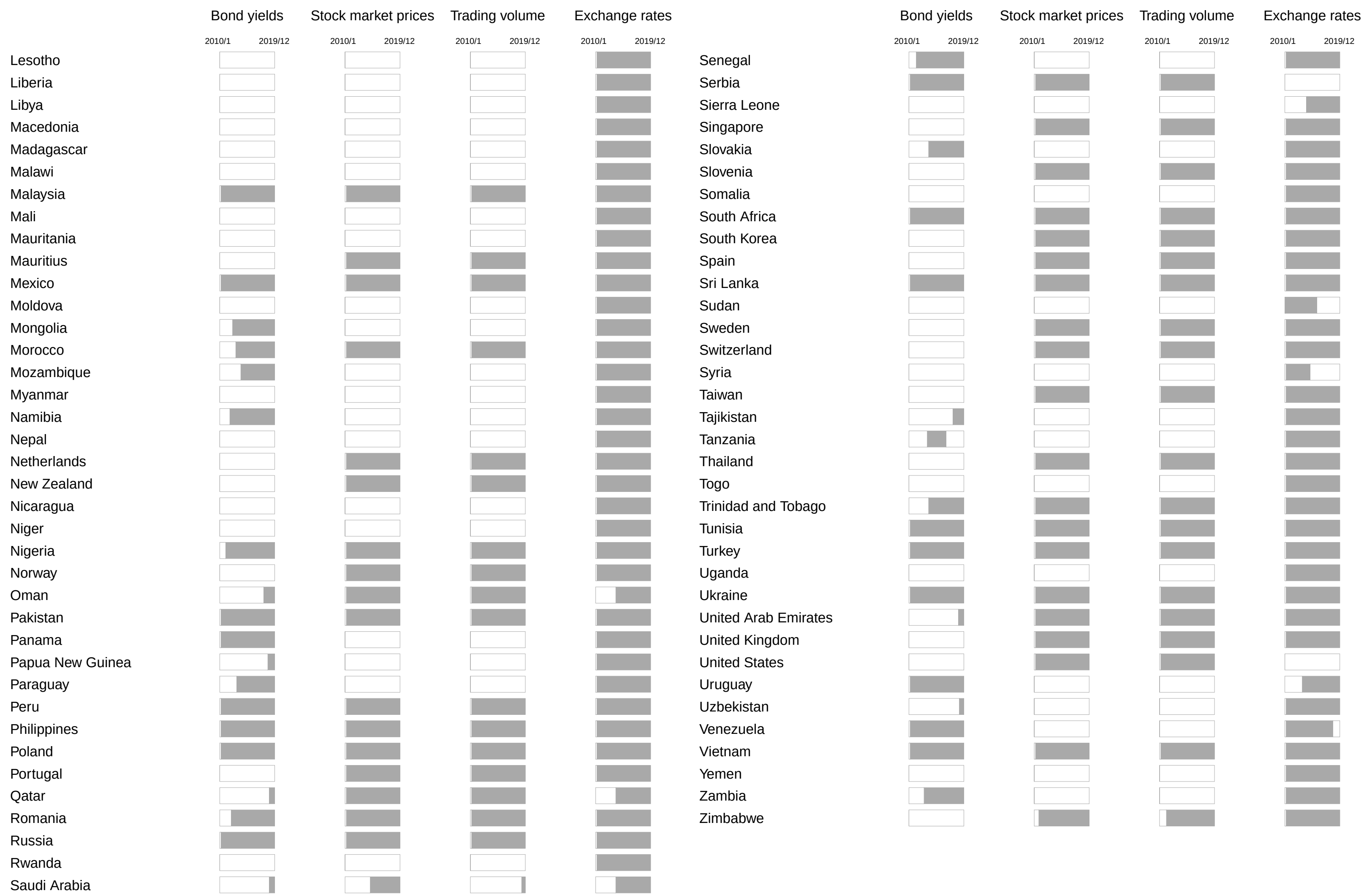
O.D.0.3 Stock Market Indices

We obtain from Bloomberg L.P. (2021b) data on the MSCI Inc. (formerly Morgan Stanley Capital International) indices. For each country, we observe a daily sub-index for the MSCI price, measured in USD whenever possible and the local currency otherwise, and a daily sub-index for the MSCI trading volume.

Online Appendix Figure 2: Coverage of Financial Indices



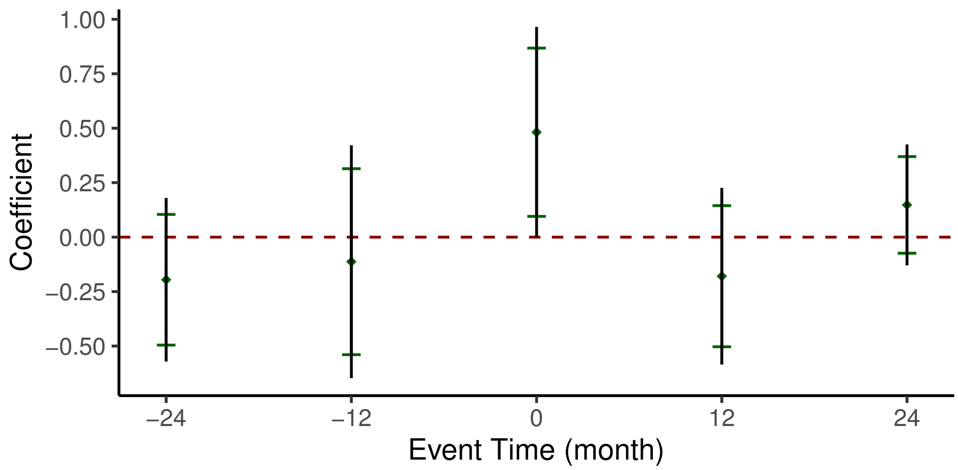
Online Appendix Figure 2: Coverage of Financial Indices (Continued)



Notes: For each sample country (in rows) and each financial indicator (in columns), the filled region of the timeline depicts the dates during the sample period for which data are available.

O.E Additional Evidence on Correlates of Repression

Online Appendix Figure 3: Dynamic Impacts of Commodity Export Prices, Uniform Confidence Bands



Notes: The figure is an analogue of Panel B of Table 6. It plots estimates from $\Delta(1-\overline{p}_{ik})=\Delta\varphi_k^p+\sum_{l=-2}^{2}\gamma_{-12\ell}^{p,x}\Delta x_{i,k+12\ell}+\Delta\varepsilon_{ik}^p$, where Δ is the annual difference operator. The inner bars depict 95% pointwise confidence intervals and the outer lines depict 95% uniform sup-t bands, both based on asymptotic inference clustered by country.

References

ACLED (2022). Armed Conflict Location & Event Data Project. Data accessed December 2022.

Bloomberg L.P. (2018). Daily bond index data. Accessed via Bloomberg Terminal in September 2018.

Bloomberg L.P. (2021a). Daily exchange rate data. Accessed via Bloomberg Terminal in May 2021.

Bloomberg L.P. (2021b). Daily MSCI data. Accessed via Bloomberg Terminal in April 2021.

Carlson, C. J., M. R. Boyce, M. Dunne, E. Graeden, J. Lin, Y. O. Abdellatif, M. A. Palys, M. Pavez, A. L. Phelan, and R. Katz (2023). The World Health Organization's Disease Outbreak News: A retrospective database. *PLOS global public health 3*(1), e0001083.

Centre for Social Conflict and Cohesion Studies (2020). *Observatory of Conflicts - Cumulative Dataset*. Harvard Dataverse. Accessed April 2024.

Enikolopov, R., A. Makarin, and M. Petrova (2020). Social media and protest participation: Evidence from Russia. *Econometrica 88*(4), 1479–1514.

International Foundation for Electoral Systems (2021). ElectionGuide. Accessed May 2020, June 2020, and June 2021.

International Organization for Standardization (2018). ISO 4217:2015. Swiss Association for Standardization [distributor]. Accessed in July 2018.

J.P. Morgan (2021). Daily bond index data. Transmitted to authors in October 2021.

Kadivar, M. A., N. Ketchley, A. Sotoudeh-Sherbaf, and C. Barrie (2024, April). Online calls for protest and offline mobilization in autocracies: evidence from the 2017 Dey Protests in Iran. *European Sociological Review*.

Kadivar, M. A., S. Khani, V. Abedini, and S. Barzin (2023). Fueling the flames of dissent: Oil, grievances, and protest in iran 2019. Unpublished manuscript.

Li, Y. (2018). *Playing by the Informal Rules: Why the Chinese Regime Remains Stable Despite Rising Protests*. Cambridge: Cambridge University Press.

Ritzwoller, D. M. and J. P. Romano (2021, 07). Uncertainty in the Hot Hand Fallacy: Detecting Streaky Alternatives to Random Bernoulli Sequences. *Review of Economic Studies 89*(2), 976–1007.

United Nations (2018a, May). Estimated population of first and new first tier cities in China in 2021, by city (in million inhabitants). Department of Economic and Social Affairs, Population Division. Accessed via Harvard Statista in August 2024.