Open-Source Information and Repression

Manuel Oechslin*

March 3, 2023

- - - Preliminary - - -

Abstract

As the digital space becomes an ever more attentive "chronicler" of human activity, the chance that acts of government incompetence or wrongdoing leave incriminating traces rises steadily. The present paper proposes a simple game-theoretic model to explore how an incumbent government with reelection concerns responds to this development. In the model, if executive power is not sufficiently checked, the government intensifies (concealed) repression against free speech; as a result, voters receive less and less, rather than more, information about the type of their government—and the chance that incompetence spells electoral defeat erodes. The model's predictions are consistent with recent global trends in freedom of expression. Among other things, the present analysis stresses the increasing importance of fortifying institutions that safeguard free speech and warns that international institutions like the European Union will be subject to growing centrifugal forces.

JEL classification: D72, D82, F53

Keywords: Open-source info, hidden repression, institutional safeguards, manipulated votes

*University of Lucerne, Department of Economics, Frohburgstrasse 3, P.O. Box 4466, CH-6002 Lucerne, Switzerland. Phone: +41 41 229 58 40. Email: manuel.oechslin@unilu.ch.

1 Introduction

With smartphone cameras and microphones present in ever more situations, with an ever larger share of social and economic activities leaving a "digital footprint" (Binswanger and Oechslin 2020), and with the number of (private) satellites orbiting Earth rising fast (Burke et al. 2021), the amount of information available in the public sphere—digital open-source information—has been surging lately. This has nurtured hopes that it will become easier for voters and civil society to hold governments to account. Referring to the surge in digital open-source information, *The Economist* (Aug 7, 2021) wrote that this "is a welcome threat to ... governments with something to hide".[1] Such hopes are not unfounded. A recent *Amnesty International Publication* (May 1, 2020) revealed a large number of human rights violations that were identified on the basis of digital open-source information only. Similarly, various international bodies have recently relied on digital open-source information to monitor whether governments comply with specific international obligations (Bochert 2021). Some investigative journalistic outfits (such as Bellingcat) are heavy users of open-source information too.

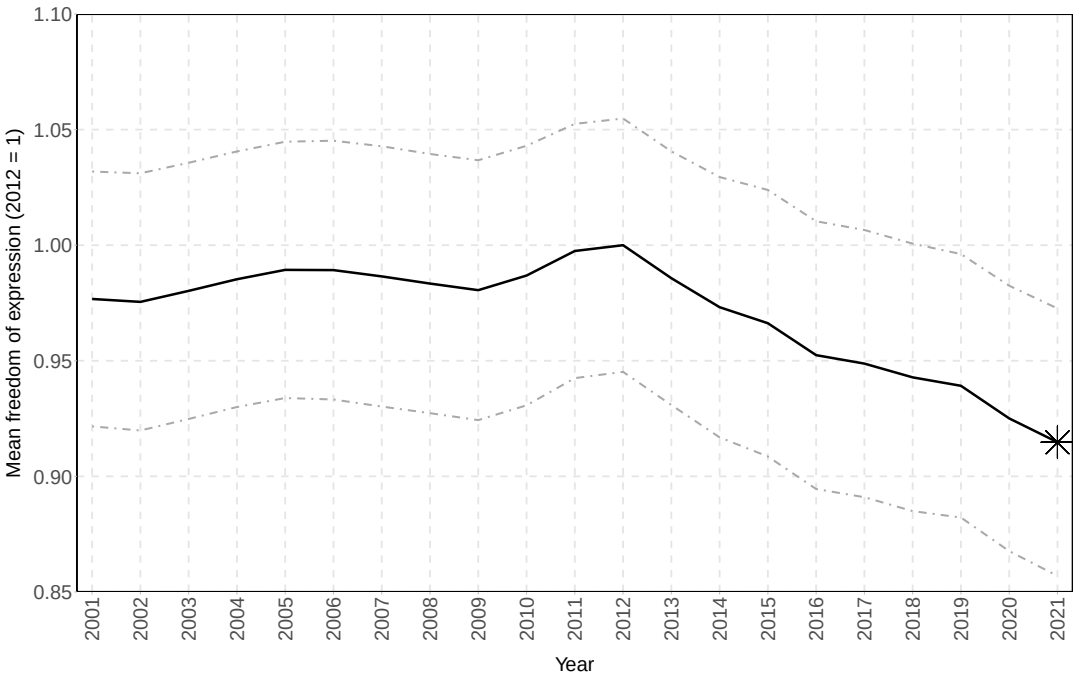
Still, an increase in the chance that acts of government incompetence or wrongdoing leave incriminating evidence somewhere in the public sphere does not necessarily mean that unworthy governments now face a higher chance of losing power. Obviously, such governments may have access to a number of countermeasures. In particular, they may try to prevent incriminating evidence from reaching the general public by silencing (or co-opting) those actors who would filter out such evidence of the large sea of unstructured digital open-source information—opposition groups, journalists, or bloggers.[2] And indeed, in parallel with the rise of digital open-source information, freedom of expression has substantially deteriorated over the past ten years (Boese et al. 2022). Figure 1 illustrates this finding with the help of the freedom of expression index provided by Coppedge et al. (2022), a measure of the extent to which a government respects the freedom of the press and that of ordinary people to discuss political matters in the public sphere. As can be seen from the figure, by 2021, global freedom of expression had suffered a decline of about 8.5% relative to its all-time high in 2012.

This paper offers a simple game-theoretic model to study the relationship between richer open-source information, the use of repression against free speech, and the electoral prospects of governments with something to hide. In the model, an incumbent politician is either competent or not, a quality not directly observed by the electorate. Yet the latter forms a belief in

[1] ChatGPT agrees: open-source intelligence "can serve as a powerful tool for watchdog organizations and civil society groups, who can use it to monitor and hold governments accountable for their actions." (Feb 10, 2023).

[2] Open source does not mean that digital information can be accessed without effort. Gathering and connecting relevant data from the digital space requires resources such as time, data skills, and computing power.

Figure 1: Freedom of expression has been under attack since 2012



Notes. Own calculations based on the variable freedom of expression (variable v2x_freexp from Coppedge et al. 2022). Freedom of expression is measured on a scale from 0 to 1, with higher values indicating greater freedom. The score is based on information provided by country experts whose identity is strongly protected. The figure shows how the mean in a sample of 141 countries evolved over time, with the value for 2012 (all-time high) normalized to 1. Dashed lines: 90% confidence interval. Star: equality with the mean of reference year 2012 rejected at the 10% significance level.

this regard, which then influences the incumbent's reelection chance. Richer digital open-source information is modeled as an increase in the probability that, should the incumbent be incompetent, evidence to that effect is found by an opposition politician. If published, such evidence affects the electorate's belief to the disadvantage of the incumbent, lifting the prospect of an electoral win by the opposition. However, the incumbent can repress the publication of unwelcome evidence, albeit at a cost. The magnitude of the cost is taken to reflect the strength of the safeguards against the misuse of executive power. While the electorate does not directly observe repression by the incumbent, it understands the incentives of the latter to use repression and accounts for them when forming the belief about the incumbent's competence.

This setup captures that nowadays most countries, democracies as well as autocracies, hold elections (Little 2017). However, leaders who lean towards autocracy are more inclined to curb electoral competition. Compared to the previous century, the repressive measures used

more recently tend to be less blatant, more subtle, and often involve manipulating information in order to make voters believe that their leaders are competent and benevolent (Guriev and Treisman 2019, 2022). In practice, repression frequently involves largely concealed efforts to tone down or even silence unwelcome voices (of, e.g., opposition politicians and critical journalist). This is achieved by different means, among them online censorship (Roberts 2020) or the initiation of carefully fabricated court cases (Earl et al. 2022). In the model, the practice of concealed silencing is reflected in the assumption that the electorate does not directly observe whether the non-publication of possible evidence of incompetence is due to repression. However, when forming beliefs, the model electorate is aware of the possibility of hidden repression, something that appears to hold for real-world citizens too (Roberts 2020).

The analysis produces four main results. First, as digital open-source information becomes richer, the incumbent politician, if incompetent, faces stronger incentives to use repression. Second, if this change to incentives does promote repression, the chance that incompetence ends the tenure of the incumbent *falls*. That is, in the "struggle" between better availability of evidence and intensified repression, the latter dominates. In addition, intensified repression when incompetent, by making the signal "absence of evidence" less informative, *reduces* the chance of reelection when competent. Third, the safeguards against the misuse of executive power are a key arbiter of the impact of a secular trend towards richer digital open-source information: with sufficiently strong safeguards, repression remains off the table and the chance that incompetence is punished and competence rewarded steadily improves; with weaker safeguards, by promoting repression, such a trend *erodes* the chance that competence makes a difference in elections. That is, the consequences of richer digital open-source information are not orthogonal to the institutional setting. Sufficiently strong safeguards are a pre-condition for better evidence to be beneficial. A final result is obtained when the safeguards against the misuse of executive power are endogenized. If the incumbent were to set them before learning her type, she would choose *strong* safeguards—anticipating that, if she later turned out to be competent, the option of repression would be useless but would reduce her reelection chance. In the long run, institutions may thus adapt to a changing informational environment.

Connecting the model world with the actual one, we find that the analytical predictions regarding repression precisely identify the subset of countries responsible for the decline in mean freedom of expression shown in Figure 1. In a broader perspective, the present analysis points to the possibility that the secular trends towards richer digital open-source information will amplify cross-country differences along institutional lines. Countries whose institutions prevent a rise in repression will see more competent and honest governments, while in the remaining

countries those qualities will increasingly come under pressure. This pattern may also amplify disparities in economic performance. Another conclusion is that institutions become more important. One might think that, with incompetence and dishonesty more likely to leave traces, the importance of formal safeguards against the misuse of executive power lessens. The present analysis suggests that evidence and institutions are complements, not substitutes.

As many contributions to the more recent political economy literature, we follow Persson and Tabellini (2000) and Alesina and Tabellini (2007, 2008) in assuming that prior to an upcoming election an incumbent politician is concerned about the perception of her competence. We also follow this literature in that the incumbent can take a hidden action to influence the electorate's perception. However, unlike typical career-concern models, the present setup includes an informational asymmetry between the incumbent (who knows her type) and the electorate (who cannot directly observe it). A second deviation is that the incumbent's hidden action does not relate to her effort but the information the electorate will receive.

Both deviations are also part of Guriev and Treisman (2020), a paper that offers a theory of informational autocracy. In that paper, a leader chooses among different political regimes (democracy, informational autocracy, dictatorship) and also selects a tool of manipulation, one of which is elite repression by censorship. The main focus is on how elite size determines regime type and manipulation tool. Here, while keeping the regime fixed, we zoom in on how a steadily improving access of the opposition to information about the incumbent alters the intensity with which repression by censorship is used—and on how the struggle between information and repression affects the chance that incompetence means electoral defeat.

The present analysis, by eventually moving on to an endogenous determination of the institutional provisions against repression, also connects with a literature on how in presence of informational asymmetries institutional constraints may actually work to the benefit of an incumbent leader—for instance, by spurring investment (e.g., Gehlbach and Keefer 2011) or forestalling intra-elite conflict (e.g., Boix and Svolik 2013). Here, stronger institutions work as a commitment not to take concealed actions to sweep possible evidence of incompetence under the carpet. This commitment, while potentially costly when incompetent, is valuable in the opposite case: it means that the electorate takes the non-publication of such evidence as a strong signal of the incumbent's competence—with positive electoral consequence.[3]

At a different level, this paper relates to research on how the digital space, and all the innovations that have come with it, affects the balance of power between leaders, authoritarian

[3]There are many commitment problems unrelated to asymmetric information, of course. See Gehlbach et al. (2016) for a broad review of commitment problems that make autocrats accept some "rules of the game".

or not, and actors in the opposition. A key theme is whether, or when, the digital space and its new communication tools strengthen the opposition by facilitating coordination and mobilization—and when it should be expected to have a weakening effect due to the enhanced possibilities for state propaganda and surveillance (e.g., Edmond 2013; Little 2016; Manacorda and Tesei 2020; Dragu and Lupu 2021).[4] This paper emphasizes another aspect of the digital space, namely its function as a chronicler that increasingly puts facts on the record that in the past would have gone undocumented. Accounting for a leader's response to that, we ask what this chronicler function means for how well voters are informed eventually.

The rest of the paper is organized as follows. The upcoming section introduces the basic setup, which is then discussed in Section 3. Section 4 solves the model and characterizes how information affects the equilibrium. In Section 5, the focus is on how the safeguards against the misuse of executive power moderate the impact of a secular trend towards richer digital open-source information on repression and on the electoral consequences of (in-)competence. That section also returns to the data presented above, but looks at it through the lens of the model. In Section 6, the safeguards are endogenized. Section 7, finally, concludes.

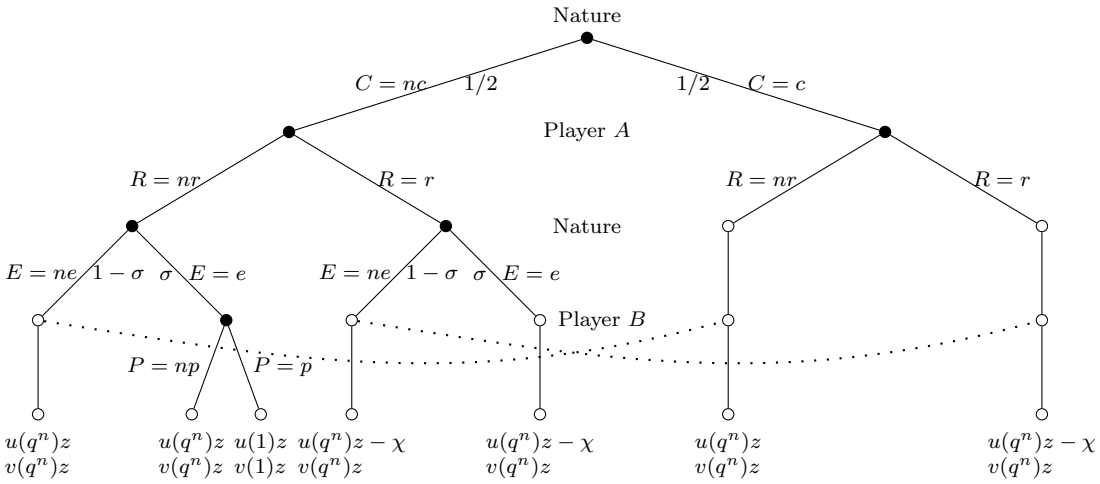
2 Setup

Players and actions. There are two players, player A and player B, and an observer. While player A is always asked to act, it depends on player A's decision, and on chance, whether or not player B is called on to act, too. The observer forms a posterior belief that affects both players' payoffs, as specified in the following two paragraphs. Nature also has a part to play. First, it determines whether or not player A is competent: $C \in \{c, nc\}$, where c stands for competent and nc for not competent. The probability of each option is $1/2$. Second, Nature determines whether in case of $C = nc$ player B obtains evidence that unequivocally documents player A's incompetence: $E \in \{e, ne\}$, where e means that such evidence surfaces and ne means that such evidence does not surface. It is assumed that $\Pr[E = e|C = nc] = \sigma$. Otherwise, if Nature has chosen $C = c$, we have $E = ne$ by default. In the analysis below, richer digital open-source information corresponds to a gain in the parameter σ.

Having learned its competence level C, player A must decide on whether or not to go for repression: $R \in \{r, nr\}$, where r means that player A chooses repression and nr means that it desists from repression. If player A chooses the former option, player B is barred from publishing any evidence of incompetence it might have obtained if $C = nc$. Otherwise, provided

[4] Focusing on the economic rather than the political sphere, the enhanced possibilities for state surveillance, and the associated potential for misuse, are also a key theme in Oechslin and Steiner (2022).

Figure 2: The game in extensive form



Notes. If $C = c$ (player A competent), we have $E = ne$ (no evidence of incompetence) by default. If $R = r$ (repression) and/or $E = ne$, we have $P = np$ (no publication of evidence) by default.

that it has obtained such evidence, player B has to decide on whether or not to publish it: $P \in \{p, np\}$, where p stands for "publish" and np for "not publish". Note that, if player A has chosen $R = r$ and/or $E = ne$, we have $P = np$ by default. Figure 2 offers an overview of the options available to Nature and the two players at any given point.

Beliefs. The observer cannot see Nature's decisions on competence and the surfacing of incriminating evidence.[5] Nor can the observer see player A's repression decision. These are hidden actions. Yet the observer knows the structure of the game, understands the players' incentives, and—if published—heeds incriminating evidence. The observer's posterior belief about the probability of player A being incompetent is denoted by q. If evidence of incompetence is published, $q = 1$ as such evidence is unequivocal; if not, we use q^n for the posterior: $q \in \{q^n, 1\}$. The possibility of repression invites a signal-jamming effect: if used with a strictly positive probability, repression adds another reason (besides $C = c$ and $E = ne$) for why evidence of incompetence may not be published. As a result, the signal "absence of evidence" carries less information and therefore will have less of an impact on the observer's posterior, q^n.

Payoffs. Player A wants to win a fresh term in her current position, while player B seeks to take over that position from the former. The successful player receives a benefit of size $z > 0$,

[5] While not explicitly modeled, incriminating evidence may fail to reach the observer directly for a collective-action problem: if the observer consists of many agents (e.g., voters) and screening digital sources involves a cost that is large for any single agent, only B—who has a chance of replacing A—may want to incur that cost.

the losing player gets nothing. The outcome of this competition is influenced by two factors, the observer's posterior, q, and the incumbency advantage enjoyed by player A, $\alpha \in (0,1)$, where a larger α means a stronger advantage. In particular, the chance that player A wins a fresh term is given by $u(q) = [1-(1-\alpha)q]$; player B takes over with probability $v(q) = (1-\alpha)q$. Thus, given α, the chance of a win by player A is a decreasing function of the observer's posterior belief about the probability that player A is incompetent. On the other hand, given q, the chance of a win by player A is an increasing function of that player's incumbency advantage. In short, incompetence hurts, incumbency helps. The Appendix, relying on an explicit modeling of the observer's objective function, derives u and v from optimizing behavior. The modeling involves a trade-off between competence and closeness in terms of some "value issue".

Repression, while preventing the publication of possible evidence of player A's incompetence, is costly: if player A chooses $R = r$, its payoff is cut by $\chi > 0$. We define the two players' payoffs as the expected benefit (possibly less χ in A's case) after all decisions have been taken and posterior q has been formed (but before chance allocates the position). The payoffs are specified at the bottom of Figure 2. The top (bottom) line refers to player A (B). For the most part of the analysis, χ is treated as an exogenous parameter. However, in Section 6, the parameter is chosen by player A before Nature decides on competence.

3 Discussion

This setup captures in broad terms a standard situation that arises in many different spheres, such as politics, business, and civil society: the fate of an incumbent depends on the perception of her competence; the latter can be influenced by obscuring facts that cast a negative light on the incumbent's record. Avoiding blame for negative outcomes (rather than claiming credit for successes) is indeed a time-tested way of holding on to a position (Hood 2011).

While generic, the concrete modeling is influenced by recent research on the exercise of power in autocracies and weaker democracies. Guriev and Treisman (2019, 2022) argue that, unlike the typical 20th-century autocrat, many modern leaders with a leaning towards autocracy sustain power primarily by manipulating information, by making the broad public "believe—rationally but incorrectly—that they are competent [...]" (Guriev and Treisman 2019, p. 101), in particular when it comes to the stewardship of the economy. In fact, many modern autocrats try to keep a democratic facade by holding elections (Little 2017)—elections that are informationally manipulated. Often, such manipulation involves the hidden silencing of opposition groups, critical journalists, or bloggers. In this regards, Guriev and Treisman (2019,

p. 102) observe that, when using force to silence unwelcome voices, modern autocrats "seek to camouflage the purpose or to conceal the state's role in violent acts." But in many cases, the hidden repression of unwelcome voices does not even involve violence. A frequent non-violent variant is the leveraging of online defamation laws (Earl et al. 2022) or tax laws to initiate fabricated court cases (whose merits more distant observers can hardly discern). Online censorship through friction or flooding are other widespread forms of hidden, non-violent repression (Roberts 2020). The former involves the manipulation of internet search results or a general slowing down of the net (also see Zhuravskaya et al. 2020), the latter the saddling of the net with distracting messages or entertainment (also see King et al. 2017). However, even though 21st-century repression is often hidden and hard to observe directly, a substantial body of empirical work (surveyed by Roberts 2020) points to many instances in which citizens were well aware of the possibility of repression when forming their opinion.

In the present setup, player A can be viewed as an incumbent politician, the holder of a high executive office, who wants to defend her position in an upcoming election. Player B takes the part of opposition politician. The observer is the electorat—with an a priori preference for the incumbent politician (e.g., due to a "value issue"), but a distaste for incompetence (e.g., in economic matters). The election's stochastic element may reflect fluctuations in the salience of the value issue. The incumbent can use repression for the purpose of manipulating the information available to the electorate. For instance, she may fear evidence showing that the incompetent handling of economic policy in the current term will damage the economy in the next term unless the position is given to a competent politician.[6] Any repression is directed at the opposition, serves the purpose of suppressing evidence of incompetence the opposition might have dug up, and is successfully hidden from the electorate. However, the incumbent cannot prevent the electorate from accounting for the possibility of repression. Nor is repression free of charge. Doing so comes at a cost that may reflect the strength of the safeguards against the misuse of executive power. With stronger safeguards, the illegitimate, hidden silencing of unwelcome voices requires more resources in terms of attention, sophistication, and money. What is more, stronger safeguards narrow the scope for deriving private gain from the entrusted powers—and thus may also reduce the benefit form holding the contested position. In practice, such safeguards consist in institutional checks like formal provisions that protect the public administration and judiciary from undue political interference. The formal checks may be complemented by informal ones (Baland et al. 2020), for instance an anticorruption norm that

[6]Suppose that economic incompetence on the part of the incumbent damages the public finances, an effect concealed from the electorate (e.g., by means of shadow households) that impairs future growth if not addressed in time. Then, evidence documenting the hidden damages is what the opposition seeks and the incumbent fears.

strongly disapproves of corruption in the public administration and judiciary.

How does the use of repression change when richer open-source information facilitates the finding of unwelcome evidence in the digital space? What is the net effect on the grip on power by the incumbent? And how do the safeguards against the misuse of executive power moderate how better information affects repression? These are questions to which the analysis turns next.

4 Equilibrium

Pure strategies. From Figure 2, it is clear that there is only one node, a singleton information set, at which player B is called on to act. As $v(1)z \geq v(q^n)z$ for any q^n, player B's optimal action at that node is p. Thus, in the game tree, this subgame can be replaced by a single "leg" with the equilibrium payoffs $u(1)z$ and $v(1)z$ for the players A and B, respectively. In the event of repression and/or absence of evidence, we have $P = np$ by default.

With the help of the reduced game tree, we next determine player As's strategy—given the posterior formed by the observer in absence of published evidence of incompetence. First suppose that player A is competent ($C = c$). Then, independent of q^n, repression, while costly, would not serve any purpose. Thus, player A desists from repression. Now assume that player A is incompetent ($C = nc$). In that case, repression ($R = r$) gives player A an expected payoff of

$$EU^A(nc, r \mid q^n) = [1 - (1 - \alpha)q^n]\, z - \chi, \tag{1}$$

where $[1 - (1 - \alpha)q^n]$ is used for $u(q^n)$. If player A desists from repression ($R = nr$), we obtain

$$EU^A(nc, nr \mid q^n) = (1 - \sigma)\,[1 - (1 - \alpha)q^n]\, z + \sigma\,[1 - (1 - \alpha)]\, z, \tag{2}$$

where $[1 - (1 - \alpha)]$ substitutes for $u(1)$. Given this, repression provides player A with a strictly larger expected payoff if and only if

$$(1 - \alpha)(1 - q^n)\sigma > \chi/z. \tag{3}$$

Condition (3) implies that player A, if incompetent, is more inclined to use repression when the incumbency advantage, α, is smaller; when the observer's absent-evidence posterior about the chance of incompetence, q^n, is smaller; when the chance that, in case of incompetence, evidence to that effect surfaces, σ, is larger; and when the cost of using repression relative to the benefit from holding the contested position is smaller. For a given q^n, condition (3) deter-

mines player A's strategy if $C = nc$:

$$R^*(nc \mid q^n) = \begin{cases} nr & : \quad (1-\alpha)(1-q^n)\sigma \leq \chi/z \\ r & : \quad (1-\alpha)(1-q^n)\sigma > \chi/z \end{cases}. \tag{4}$$

Equilibrium requires player A's strategy to be consistent with the absent-evidence posterior formed by the observer. First assume $R^*(nc \mid q^n) = nr$ (no-repression equilibrium). What is the value of the corresponding q^n? The non-publication of evidence of incompetence could mean that, indeed, player A is competent; or it could mean that the player is incompetent, but evidence to that effect has failed to surface. The probability of the former eventuality is $1/2$, that of the latter one is $(1/2)(1-\sigma)$. As a result, according to Bayes' rule,[7]

$$q^n_{nr} = \frac{1-\sigma}{2-\sigma}. \tag{5}$$

It follows from equation (5) that, as σ rises from 0 (evidence never surfaces) to 1 (evidence always surfaces), q^n_{nr} monotonically falls from $1/2$ to 0. However, given q^n_{nr}, is it indeed optimal for player A to choose $R = nr$? Together, equations (4) and (5) imply that it is if and only if

$$\sigma \leq \bar{\sigma}_{nr} \equiv 2\left(1 + \frac{1-\alpha}{\chi/z}\right)^{-1}. \tag{6}$$

We conclude that, if σ is sufficiently small, the no-repression equilibrium does exist.

Now suppose that $R^*(nc \mid q^n) = r$ (full-repression equilibrium). Then, the non-publication of evidence of incompetence does not contain any information. Hence, the posterior must be equal to the prior: $q^n_r = 1/2$. Given this, is it indeed optimal for player A to choose repression? It follows from equation (4) that it is if and only if σ is sufficiently large. Formally:

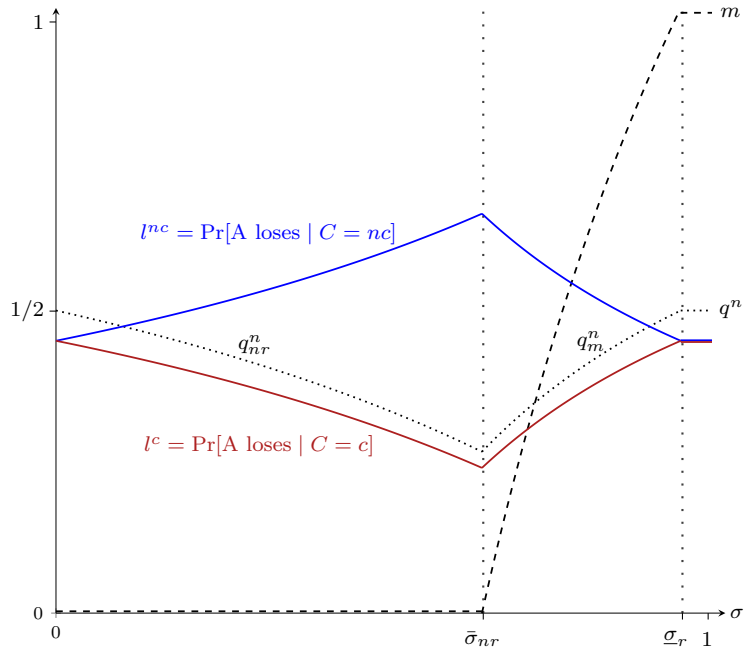
$$\sigma > \underline{\sigma}_r \equiv 2\left(\frac{1-\alpha}{\chi/z}\right)^{-1}. \tag{7}$$

We conclude that, if σ is sufficiently large, the repression equilibrium does exist. Note that, because $(1-\alpha)z/\chi$ is positive, it follows from equations (6) and (7) that $\bar{\sigma}_{nr} < \underline{\sigma}_r$.

Mixed strategy if incompetent. The analysis so far shows that an incompetent player A does not follow a pure strategy if $\bar{\sigma}_{nr} < \sigma \leq \underline{\sigma}_r$. In fact, if incompetent, player A uses a mixed

[7] The rule says that the conditional probability $q^n_{nr} = \Pr[C = nc \mid P = np]_{nr}$ is given by the probability of the second eventuality, divided by the sum of the probabilities of the two possible eventualities.

Figure 3: Key equilibrium outcomes as functions of σ



Notes. The parameter constellation is such that all three equilibria emerge as σ rises from 0 to 1. It is also assumed that the incumbency advantage, α, is strictly positive.

strategy that resorts to repression with probability m. A mixed strategy requires

$$EU^A(nc, r \mid q^n_m) = EU^A(nc, nr \mid q^n_m), \tag{8}$$

where q^n_m refers to the observer's corresponding posterior. To find the posterior, note that the non-publication of evidence of incompetence could mean that, indeed, player A is competent; or it could mean that player A is incompetent, but relies on repression (with evidence of incompetence surfacing or not); or it could mean that the player is incompetent, has desisted from repression, but evidence of incompetence has failed to surface. The probability of the first eventuality is $1/2$, that of the second eventuality is $(1/2)m$, and that of the third one is $(1/2)(1-m)(1-\sigma)$. As a result, according to Bayes' rule,[8]

$$q^n_m = \frac{1 - \sigma(1-m)}{2 - \sigma(1-m)}. \tag{9}$$

Using the explicit expressions for the two expected utilities (equations 1 and 2) and for the

[8]Here the rule says that the conditional probability $q^n_m = \Pr[C = nc \mid P = np]_m$ is the sum of the probabilities of the second and third eventuality, divided by the sum of the probabilities of all three possible eventualities.

posterior (equation 9), equation (8) can be rearranged to obtain an expression for probability m:

$$m = 1 - \left(\frac{2}{\sigma} - \frac{1-\alpha}{\chi/z} \right) \tag{10}$$

Note that $m > 0$ is equivalent to $\sigma > \bar{\sigma}_{nr}$, while $m \leq 1$ is equivalent to $\sigma \leq \underline{\sigma}_r$. Thus, if σ falls in an intermediate range that neither allows for the no-repression nor the repression equilibrium, there is a partial-repression equilibrium in which player A, if incompetent, relies on repression with some positive probability. Using equation (10) to substitute for m in equation (9) yields:

$$q^n_m = 1 - \frac{\chi/z}{1-\alpha} \frac{1}{\sigma}. \tag{11}$$

Provided $\bar{\sigma}_{nr} < 1$, it follows from equations (5) and (11) that $q^n_m = q^n_{nr}$ if σ is at the upper end of the no-repression range, $\bar{\sigma}_{nr}$. Similarly, provided $\underline{\sigma}_r < 1$, we have $q^n_m = q^n_r = 1/2$ if σ is at the lower end of the repression range, $\underline{\sigma}_r$. So, in any case, q^n is a continuous function of σ.

Summary and discussion. The proposition below summarizes the results established so far, assuming that the parameter constellation allows for all three equilibria (proof in the text).

PROPOSITION 1 *Suppose $\chi/z < (1-\alpha)/2$ so that $\underline{\sigma}_r < 1$. Then, as $\sigma \in [0,1]$ rises from 0 to 1, there is a succession of three different types of Bayesian Nash Equilibria (BNE).*

No repression: *If $\sigma \leq \bar{\sigma}_{nr} < \underline{\sigma}_r$, there is a pure-strategy BNE in which A desists from repression no matter its level of competence; B publishes evidence of A's incompetence if such evidence surfaces; the observer's absent-evidence posterior, q^n_{nr}, is given by (5).*

Partial repression: *If $\bar{\sigma}_{nr} < \sigma \leq \underline{\sigma}_r$, there is a BNE in which A desists from repression if competent and else relies on a mixed strategy that uses repression with chance m, where m is given by (10); B publishes evidence of A's incompetence if A desists from repression and such evidence surfaces; the observer's absent-evidence posterior, q^n_m, is given by (11).*

Full repression: *If $\bar{\sigma}_{nr} < \underline{\sigma}_r < \sigma$, there is a pure-strategy BSE in which A desists from repression if competent and uses repression if incompetent; B is never able to publish evidence of A's incompetence; the observer's absent-evidence posterior, q^n_r, is fixed at 1/2.*

Figure 3 visualizes m and q^n. The former rises monotonically from 0 to 1 as σ crosses the partial-repression range. The latter starts and ends at 1/2, the value of the prior probability. If $\sigma = 0$, evidence of incompetence never surfaces; so absence of evidence does not contain any information on player A's competence. At the opposite end, if $\sigma \geq \underline{\sigma}_r$, possible evidence of incompetence is never published—with the result that absence of evidence is again completely

uninformative. In between, q^n follows a V-shape profile. As σ rises from 0, q^n first falls: with evidence of incompetence more likely to surface and make its way to the observer, absence of evidence implies a lower ex post probability that player A is incompetent. Yet, as soon as the threshold to partial repression is crossed, absence of evidence is more and more often due to (hidden) repression. So the signal "absence of evidence" gets increasingly jammed and contains less and less information. That is why q^n eventually returns to 1/2. As will become clear below, for player A, this signal-jamming effect is the downside that comes with the opportunity to use repression. While repression is useful if, in fact, $C = nc$, the opportunity to use it increases the observer's absent-evidence posterior belief about the probability of $C = nc$.

Turnover. Proposition 1 establishes a first main result: as digital open-source information becomes richer, player A, if incompetent, increasingly uses repression. But what is the net effect of repression on player A's tenure? As the formal modeling in the Appendix shows, the present setup is consistent with the idea that, for the observer, holding on to an incompetent incumbent is economically harmful. On the other hand, replacing player A is costly because player B is less close in terms of a value issue. It is thus of interest how σ affects the probability of player A losing if incompetent, $l^{nc} = \Pr[A \text{ loses} \mid C = nc]$, and if competent, $l^c = \Pr[A \text{ loses} \mid C = c]$.

PROPOSITION 2 *Suppose $\chi/z < (1-\alpha)/2$ so that $\underline{\sigma}_r < 1$. Then, as $\sigma \in [0,1]$ rises from 0 to 1, l^{nc} and l^c first diverge, then converge, and eventually take the same value.*

l^{nc}: the equilibrium probability that an incompetent player A loses to B follows a continuous inverted V-shape profile—it first strictly increases (no repression), then strictly decreases (partial repression), and then remains unchanged (full repression).

l^c: the equilibrium probability that a competent player A loses to B follows a continuous V-shape profile—it first strictly decreases (no repression), then strictly increases (partial repression), and then remains unchanged (full repression).

Proof. See Appendix ■

Figure 3 visualizes l^{nc} and l^c. How is the inverted V-shape of l^{nc} explained? As long as repression is not used, an increase in σ lifts the chance that evidence of incompetence surfaces and reaches the observer. So incompetence is more likely to be "punished". With σ in the partial-repression range, the chance of evidence surfacing continues to grow. Yet, at the same time, player A increasingly resorts to repression. In fact, the use of repression rises steeply in σ: the marginal effect of an increase in σ on the use of repression is $2/\sigma^2 > 1$ (from equation 10), while the probability of evidence surfacing rises just 1-for-1 in σ. The net effect is that

evidence of incompetence becomes less and less likely to reach the observer—and the chance that incompetence entails the loss of the position recedes again. Eventually, when σ exceeds the full-repression threshold, evidence of incompetence no longer stands a chance of reaching the observer. So q^n is fixed at 1/2 and l^{nc} is a constant too. Beyond the full-repression threshold, the gap between 1/2 and l^{nc} (see Figure 3) reflects the incumbency advantage, α.

When starting from a low level, an increase in σ not only makes the survival of incompetence less likely, it also lifts the chance that a competent incumbent stays on: at first, as shown in Figure 3, l^c is a decreasing function of σ. Yet, as soon as σ reaches the partial-repression range, l^c changes direction. The reason is that the signal "absence of evidence" gets increasingly jammed and contains less and less information. As a result, the absent-evidence posterior belief about incompetence re-approaches the less favorable prior of 1/2—and the chance of losing despite being competent grows. Eventually, when σ exceeds the full-repression threshold, absence of evidence no longer offers information on competence. Hence, $l^{nc} = l^c$.

For the observer, the paths of l^{nc} and l^c are associated first with rising welfare (no-repression range) and then with falling welfare (partial-repression range). Consider the no-repression range: with σ increasing from 0 to $\bar{\sigma}_{nr}$, the probability that an incompetent incumbent is fired rises ($l^{nc} \uparrow$), while the probability that a competent one has to leave falls ($l^{nc} \downarrow$). With a higher l^{nc}, there is a lower probability that the observer makes the mistake of not replacing an incompetent incumbent when doing so would be optimal. With a lower l^c, there is a lower probability that the observer makes the mistake of replacing a competent incumbent with a player that is less close in terms of a value issue. With those mistakes becoming less likely, welfare improves.[9] But as soon as σ crosses into the partial-repression range, the probabilities of both mistakes start to increase again—with adverse consequences for welfare.

Finally, the result that for all $\sigma < \underline{\sigma}_r$ the conditional probabilities l^{nc} and l^c move in opposite directions suggests that the unconditional probability of player A losing, $l = \Pr[A \text{ loses}]$, could be rather stable as open-source information becomes richer. This is indeed true:

PROPOSITION 3 *Suppose* $\chi/z < (1-\alpha)/2$ *so that* $\underline{\sigma}_r < 1$. *Then, as* $\sigma \in [0,1]$ *rises from* 0 *to* 1, *the unconditional equilibrium probability that player A loses to B is a constant:* $l = (1-\alpha)/2$.

Proof. See Appendix ■

Together, Propositions 2 and 3 imply that richer digital open-source information, as well as any response to it in terms of repression, entails just "redistribution" between the two different

[9] One can frame the two mistakes also in the terminology of hypothesis testing. Suppose the observer's null is that player A is competent. Then, the first mistake is akin to a type II error, the second one to a type I error.

types of player A. Any win by the competent type in terms of a lower chance of losing is mirrored in a loss of exactly equal size by the incompetent type in terms of a higher chance of losing (and vice versa). By contrast, the analysis does not imply that richer digital open-source information leads to systematic shifts in the a priori success chances of players A and B. In other words, the analysis does not deliver any argument that richer digital open-source information would accelerate or slow down the turnover in contested positions.

5 Exogenous Safeguards

Predictions. Proposition 2 establishes a second main result: paradoxically, by intensifying repression, richer digital open-source information may lower the chance of a loss when player A is incompetent and may lift that chance when the player is competent. But the use of repression by an incompetent incumbent is not only influenced by a shifting informational environment. Repression also depends on structural parameters. This section, using the elections interpretation of the setup, studies the role of those parameters.

A key magnitude when it comes to the use of repression is the ratio χ/z, which gives the cost of using repression relative to the benefit from holding the contested position. The larger χ/z, the larger the thresholds that separate the three different equilibria (equations 6 and 7). In the elections interpretation of the setup, χ/z is influenced by the strength of the safeguards against the misuse of executive power: all else equal, stronger safeguards increase the level of resources needed for successful repression and narrow the extent to which the entrusted position can be misused for private gain. In practice, as discussed in Section 3, the safeguards may consist in institutional as well as norms-based checks. In formal terms: $\chi/z = s(\gamma, \nu, ...)$, with $\partial s/\partial\gamma > 0$ and $\partial s/\partial\nu > 0$, where γ captures the institutional checks (e.g., formal provisions that protect the public administration and/or judiciary from undue political interference) and ν captures the norms-based checks (e.g., the general disapproval of corruption on the part of officials in the public administration and/or judiciary). In what follows, while not ruling out that χ/z is affected by factors other than γ and ν, we understand χ/z as an indicator of the strength of the safeguards against the misuse of executive power.

Figure 4 illustrates l^{nc} and l^{c} with stronger (a) and weaker (b) safeguards. Assuming that the start level of σ, σ_0, gives rise to the no-repression equilibrium, the figure shows that with stronger safeguards there is a broad range over which the departure probability of an incompetent incumbent rises and that of a competent one falls. But with weaker safeguards, l^{nc} and l^{c} soon change direction. Accordingly, the safeguards against the misuse of executive

power moderate the relationship between information and the chance of the incumbent winning:

PROPOSITION 4 *Consider an arbitrary starting level of* $\sigma \in [0,1]$*,* σ_0*, that is strictly less than* 1*. Suppose* σ *rises from* σ_0 *to* 1*.*

Strong safeguards: *There exists a minimum level of safeguards against the misuse of executive power,* $\underline{\chi/z}$*, such that for all* $\chi/z \geq \underline{\chi/z}$ *the equilibrium probability* l^{nc} *(*l^{c}*) strictly increases (decreases) as* σ *rises.*

Weak safeguards: *There further exists a maximum level of safeguards against the misuse of executive power,* $\overline{\chi/z}$*, such that for all* $\chi/z < \overline{\chi/z}$ *the equilibrium probability* l^{nc} *(*l^{c}*) decreases (increases)* σ *rises.*

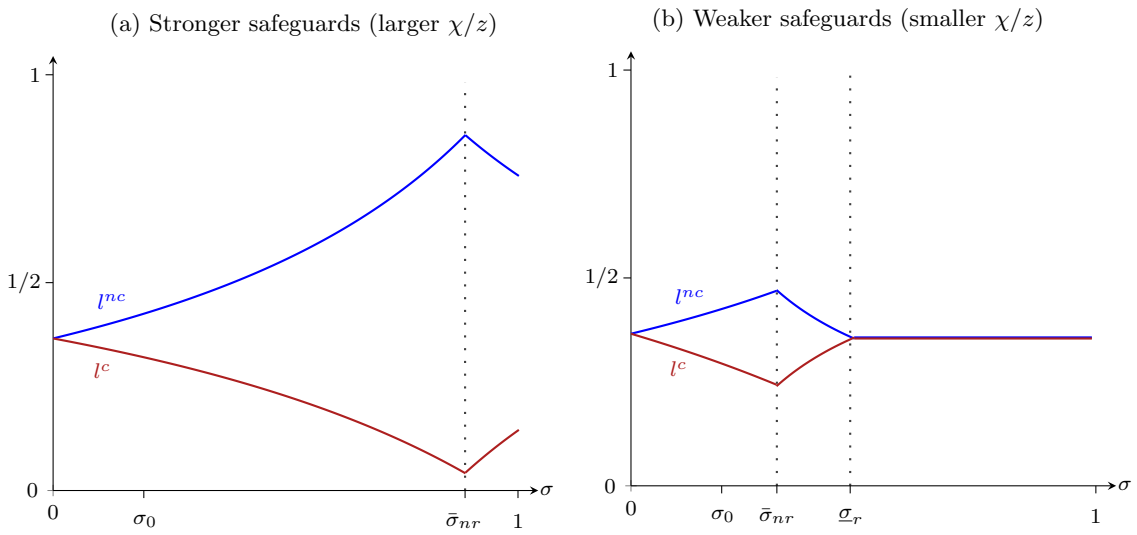
Proof. See Appendix ■

Proposition 4 establishes a third main result: a secular trend towards richer digital open-source information may have exactly opposite consequences, depending on the safeguards against the misuse of executive power. With sufficiently strong safeguards, such a trend unambiguously raises the chance that incompetence is punished and competence is rewarded; with sufficiently weak safeguards, the trend unambiguously reduces these probabilities. In-between, the effect of the trend is non-monotonic (see Figure 4). Viewed from a different angle, the above analysis suggests that the relationship between, on the one hand, institutions and norms and, on the other hand, information is complementary: when it comes to punishing incompetence and rewarding competence, richer information helps if the safeguards against the misuse of executive power are strong; strong safeguards help if the available information is rich. Clearly, our results do not support any notion that one could substitute for the other.

Besides χ and z, the setup knows one more parameter, α. The incumbency advantage is a structural parameter to such a degree as the value issue favoring the incumbent is a persistent topic. The incumbency advantage, too, influences the effect of richer information: all else equal, the minimum level of safeguards for a globally beneficial effect of σ, $\underline{\chi/z}$, is larger when α is smaller. This is intuitive: the more precarious the position of the incumbent, the stronger the safeguards needed to deter her from using repression. It is equally intuitive that also the maximum level of safeguards for a globally harmful effect of σ, $\overline{\chi/z}$, is larger when α is smaller.

Empirical patterns. A key implication of the above analysis is that the effect on repression of a secular trend towards richer digital open-source information is moderated by the safeguards against the misuse of executive power (Proposition 4). This implication can be confronted with the data by identifying the countries responsible for the post-2012 global decline in freedom of

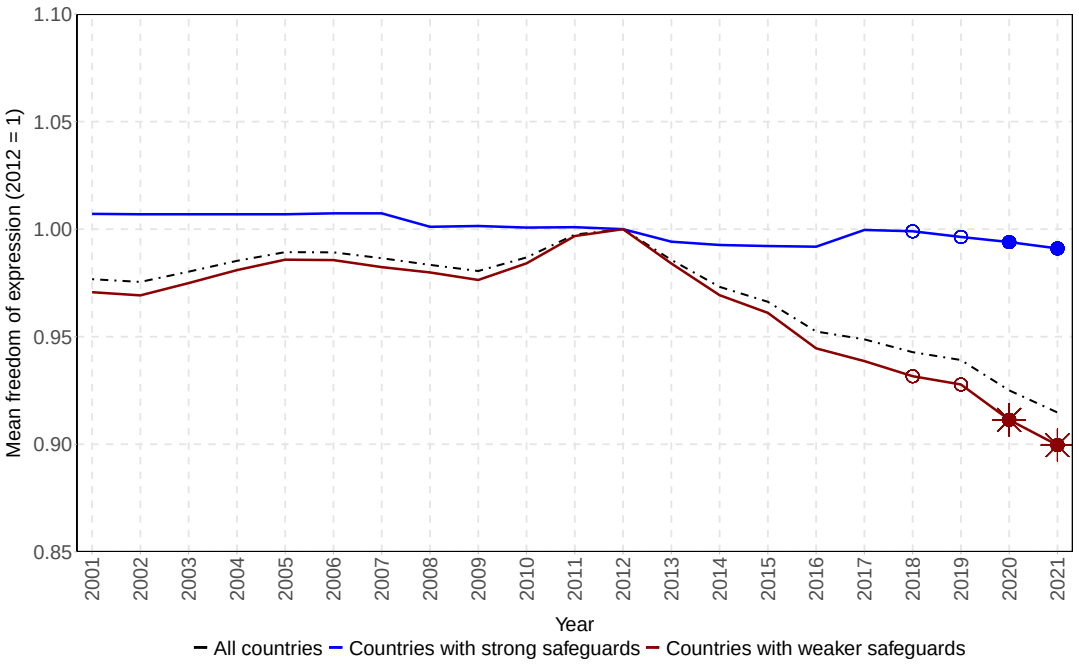
Figure 4: Departure probabilities with stronger and weaker safeguards



expression shown in Figure 1. In the model, the strength of the safeguards becomes manifest in the ratio $\chi/z = s(\gamma, \nu, ...)$, the cost of using repression relative to the benefit from holding the contested position. While this ratio does not have a direct empirical counterpart, two of its supporting determinants—γ and ν, the institutional and norms-based checks on executive power—do. As a result, an empirically verifiable prediction of the theoretical analysis is that the decline shown in Figure 1 is primarily driven by countries whose institutional and/or norms-based checks on executive power are relatively weak.

As an empirical counterpart of the institutional checks, γ, it is natural to use the checks on government index from the World Bank's GovData360 database. The index measures the extent to which the other branches of government (legislature, judiciary) control the executive (rather than the other way round). Below, a country is considered to have strong institutional checks if it is above the 75th percentile in that dimension; and to have weaker institutional checks if it is below that percentile. The classification is based on the mean over the years from 2001 to 2021, the period covered in Figure 1. An empirical proxy for the strength of the norms-based checks, ν, can be found with the help of Fisman and Miguel (2007), a paper that offers information on the strength of the anticorruption norm based on the number of parking violations by foreign diplomats positioned in New York City. Because until November 2002 parking violations by foreign diplomats were not enforced, Fisman and Miguel interpret high numbers in the period prior to November 2002 as an indication of a weak anticorruption norm in the

Figure 5: Free expression suffers most where institutional and/or norms-based checks are weak



Notes. Own calculations based on the variables freedom of expression (see notes to Figure 1), checks on government (variable 41825, World Bank, GovData360 database), and violations per diplomat, pre-enforcement (variable viol_pc–pre, Fisman and Miguel 2007). The figure shows how mean freedom of expression evolved over time, separately for two subsamples. One subsample includes 17 countries with strong safeguards against the misuse of executive power (strong institutional checks and strong norms-based checks), the other subsample 124 countries with weaker safeguards (weaker institutional checks and/or weaker norms-based checks). As in Figure 1, the 2012 means are normalized to 1. Stars: equality with the mean of reference year 2012 rejected at the 10% significance level. Dots (circles): equality with the other (same year) subsample mean rejected at the 5% (10%) significance level.

corresponding foreign country. The paper offers country-level data on the number of parking violations per diplomat and year, averaged over the period from 11/1997 to 11/2002. Consistent with the definition of strong institutional checks, we consider a country to have strong norms-based checks if it is below the 25th percentile in terms of parking violations; and to have a weaker norms-based checks if it is above that percentile.[10]

The sample of countries that underlies Figure 1 can now be split into two subsamples according to the above classifications in terms of institutional and norms-based checks. Because the theory predicts that a secular trend towards richer digital open-source information does not lead to pressure on the freedom of expression if the safeguards against the misuse of executive

[10]The basic empirical pattern, shown in Figure 5 below, does not change when we apply a somewhat stricter (best 20% of countries) or a somewhat more lenient (best 30% of countries) definition of "strong".

power are strong (large $\chi/z = s(\gamma, \nu, ...)$), one subsample pools countries that combine strong institutional checks (γ among the top 25%) with strong norms-based checks (ν among the top 25%). The second subsample thus consists of countries that qualify as weaker in at least one of the two dimensions.[11] Figure 5 shows mean freedom of expression separately for the two subsamples, along with the full sample mean from Figure 1. Again, the 2012 values are normalized to 1. So the lines show the relative deviations from 2012.

It is immediately apparent from Figure 5 that, consistent with the theoretical prediction, the post-2012 global decline in freedom of expression is almost entirely driven by the subsample of countries with weaker safeguards (weaker institutional checks and/or weaker norms-based checks). Countries with strong safeguards contributed close to nothing to the decline. More precisely, in the period from 2012 to 2021, mean freedom of expression fell by 10.04% in the former subsample and by just 0.89% in the latter. Towards the end of the observation period, the divergence in trends becomes statistically significant. To be clear: Figure 5 is not an empirical test of the model. Trends unrelated to information might explain, partially or completely, the decline. But the figure does demonstrate that the model is helpful in identifying the countries responsible for the recent rise of repression in the form of curbs on the freedom of expression.

6 Endogenous Safeguards

In practice, the determinants of the safeguards against the misuse of executive power by the government—institutional and norms-based checks—are deep-rooted and inert factors. They cannot be quickly engineered or adjusted (Acemoglu and Robinson 2020). But it is still legitimate to ask how strong player A would prefer the safeguards to be if she were given a choice. This section thus turns the strength of the safeguards, in the model captured by the ratio χ/z, into a choice variable. We consider χ/z as belonging to the list of factors that set the stage on which the strategic interaction unfolds. As a result, the determination of χ/z is placed at the very beginning. In particular, player A is asked to choose χ/z before Nature determines the player's level of competence. That is, player A has to make a choice behind the "veil of ignorance".[12] When deciding on χ/z, player A aims to maximize the expected payoff, anticipating that the game will continue as stated in Proposition 1.

With χ/z being an endogenous variable, it is appropriate to make two further assumptions:

[11] One might expect that the countries with strong institutions are also those whose diplomats commit few violations (strong anticorruption norm). However, with -0.15, the correlation is only mildly negative. As a result, although the top 25% consist of 35 countries in each dimension, the intersecting set counts just 17 observations.

[12] In the elections interpretation of the setup, this choice would probably concern γ, the strength of the institutional checks that, e.g., protect the administration or judiciary from undue political interference.

ASSUMPTION 1 *When setting χ/z to maximize the expected payoff, player A (i) has to observe a strictly positive lower bound, $(\chi/z)_0$; (ii) prefers the smallest maximizer in $[(\chi/z)_0, \infty)$.*

Assumption 1 (i) rules out two possibilities that would render the analysis simplistic, namely that player B could be silenced at no cost at all and that the benefit from holding the contested position is infinite. Assumption 1 (ii) is introduced because the choice of χ/z is pinned down only to the extent that it leads to a particular equilibrium.

PROPOSITION 5 *Given Assumption 1, player A chooses χ/z so as to just induce the no-repression equilibrium. Therefore, the ratio χ/z is an increasing function of σ, with $\chi/z = (\chi/z)_0$ if $\sigma \leq \bar{\sigma}_{nr}((\chi/z)_0)$ and $\chi/z = (1-\alpha)\sigma(2-\sigma)^{-1}$ otherwise.*

Proof. See Appendix ■

In a situation where player A has just learned that her level of competence is low ($C = nc$), she prefers the relative cost of repression, χ/z, to be small. However, at an earlier stage, where the level of competence is not yet chosen, player A prefers χ/z to be such that it prevents the use of repression should she turn out to be incompetent later on.[13] What at first sight might look like a contradiction simply reflects that the benefits of a large relative cost in case of $C = c$ outweigh the benefits of small relative costs in case of $C = nc$. But how would player A, should she turn out to be competent, benefit from a large relative cost? As a large relative cost rules out repression even if $C = nc$, the observer—who sees neither C nor R—knows that absence of evidence cannot be due to repression. So the observer's absent-evidence posterior belief about the probability of $C = nc$, q^n, is lower in the no-repression equilibrium than under partial or even full repression. That is, in the no-repression equilibrium, player A does not suffer from the signal-jamming effect that works to her disadvantage. At the same time, for a competent player A, eschewing the option of repression has no drawbacks.

Returning to the elections interpretation of the setup, what could Proposition 5 mean for the dynamics of the safeguards against the misuse of executive power in places where such safeguards are weak? As real-world institutional checks are not engineered in response to an optimization exercise, but are deeply rooted in specific historical circumstances, one should not expect the safeguards to evolve smoothly in lockstep with the trend in digital open-source information. As a result, in the short or medium run, the institutional development in places with weak safeguards might not keep up with the informational—and thus fail to prevent a further increase in repression. Proposition 5 does suggest, however, that the political elite

[13] If asked, player B would agree with player A. Thus, if the decision on χ/z required unanimity between the two, the result would still be safeguards that are just sufficiently strong to induces the no-repression equilibrium.

itself, in parallel with the electorate, would increasingly benefit from stronger safeguards. In the long run, this might promote institutional reforms that allow the trend towards richer digital open-source information to yield beneficial consequences.

7 Conclusion

Edmond (2013, p. 1423) observed that "information optimism has a long and somewhat mixed history". As the digital space becomes an ever more attentive chronicler, there is again optimism that it will become easier to hold governments to account. This paper sets up a game-theoretic model to explore whether, or to what extent, the current optimism is warranted. The model highlights the possibility that richer digital open-source information leads to intensified repression against free speech; in turn, more repression means that the electorate is less well informed—with the paradoxical result that incompetent governments have a better, and competent ones a worse, chance of reelection. Such a negative outcome arises when the safeguards against the misuse of executive power are weak such that repression is cheap and public offices are lucrative. We show that the model is helpful in identifying the countries responsible for the steady decline in mean freedom of expression that has started after 2012.

The present analysis leads to two broad conclusions. The first conclusion is that a secular trend towards richer digital open-source information will amplify differences among countries along institutional lines. In places where strong institutional safeguards against the misuse of executive power prevent a slide into repression, "bad"—incompetent or corrupt—governments will increasingly be exposed as such; to the extent that voters value competence and integrity, this will help the quality of government. But in places where the institutional safeguards are weak, the trend will intensify repression against free speech; as a result, incompetence is less often punished and competence less often rewarded. Over time, the increasing quality gap between places with strong and weak institutional safeguards may amplify cross-country differences in economic performance. This is a particular worry for international bodies like the European Union that aim to reduce economic differences via transfers but count among their members countries with widely differing institutions (Alesina et al. 2017).

A second broad conclusion concerns the importance of institutional safeguards against the misuse of executive power. It might be tempting to think that, as acts of government incompetence or wrongdoing become more likely to leave detectable traces that in principle can be shown to voters, such safeguards lose in importance. The results here point in the opposite direction. The role of institutional checks that, for instance, prevent the executive from tram-

pling free speech, like an independent judiciary that stands as a bulwark against fabricated court cases, becomes bigger, not smaller. That is, information and institutions are complements, not substitutes. As the digital space chronicles an ever larger part of what is going on, fortifying institutional safeguards will become increasingly important. While institutions are characterized by inertia in the short and the medium run, it is not implausible that, in fact, they get strengthened over a longer time horizon: the present analysis identifies circumstances in which even the political elite might accede to demands for stronger institutional safeguards.

References

Acemoglu, Daron and James A. Robinson (2020). *The Narrow Corridor. How Nations Struggle for Liberty*. New York: Penguin Books.

Alesina, Alberto and Guido Tabellini (2007). "Bureaucrats or Politicians? Part I: A Single Policy Task". *American Economic Review* 97 (1), pp. 169–179.

— (2008). "Bureaucrats or politicians? Part II: Multiple policy tasks". *Journal of Public Economics* 92, pp. 426–447.

Alesina, Alberto, Guido Tabellini, and Francesco Trebbi (2017). "Is Europe an Optimal Political Area?" *Brookings Papers on Economic Activity* Spring 2017, pp. 169–213.

Baland, Jean-Marie et al. (2020). *The Handbook of Economic Development and Institutions*. Princeton: Princeton University Press. Chap. Economic Development and Institutions, pp. 1–20.

Binswanger, Johannes and Manuel Oechslin (2020). "Better statistics, better economic policies?" *European Economic Review* 130, 103588.

Bochert, Florian (2021). "OSINT – The Untapped Treasure Trove of United Nations Organizations". *Harvard International Review* FALL 2021, 29–32.

Boese, Vanessa A. et al. (2022). *Autocratization Changing Nature? Democracy Report 2022*. Gothenburg: Varieties of Democracy Institute.

Boix, Carles and Milan W. Svolik (2013). "The Foundations of Limited Authoritarian Government: Institutions, Commitment, and Power-Sharing in Dictatorships". *Journal of Politics* 75 (2), pp. 300–316.

Burke, Marshall et al. (2021). "Using satellite imagery to understand and promote sustainable development". *Science* 371 (6535).

Coppedge, Michael et al. (2022). *VDem Dataset v12*. Gothenburg: Varieties of Democracy (V-Dem) Project.

Dragu, Tiberiu and Yonatan Lupu (2021). "Digital Authoritarianism and the Future of Human Rights". *International Organization* 75 (4), pp. 991–1017.

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

Edmond, Chris (2013). "Information Manipulation, Coordination, and Regime Change". *Review of Economic Studies* 80 (4), pp. 1422–1458.

Fisman, Raymond and Edward Miguel (2007). "Corruption, Norms, and Legal Enforcement: Evidence from Diplomatic Parking Tickets". *Journal of Political Economy* 115 (6), pp. 1020–1048.

Gehlbach, Scott and Philip Keefer (2011). "Investment without democracy: Ruling-party institutionalization and credible commitment in autocracies". *Journal of Comparative Economics* 39 (2), pp. 123–139.

Gehlbach, Scott, Konstantin Sonin, and Milan W. Svolik (2016). "Formal Models of Nondemocratic Politics". *Annual Review of Political Science* 19, pp. 565–84.

Guriev, Sergei and Daniel Treisman (2019). "Informational Autocrats". *Journal of Economic Perspectives* 33 (4), 100–127.

— (2020). "A theory of informational autocracy". *Journal of Public Economics* 186, 104158.

— (2022). *Spin Dictators. The Changing Face of Tyranny in the 21st Century*. Princeton: Princeton University Press.

Hood, Christopher (2011). *The Blame Game. Spin, Bureaucracy, and Self-Preservation in Government*. Princeton: Princeton University Press.

King, Gary, Jennifer Pan, and Margaret E. Roberts (2017). "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument". *American Political Science Review* 111 (3), pp. 484–501.

Little, Andrew T. (2016). "Communication Technology and Protest". *Journal of Politics* 78 (1), pp. 152–166.

— (2017). "Are non-competitive elections good for citizens?" *Journal of Theoretical Politics* 29 (2), pp. 214–242.

Manacorda, Marco and Andrea Tesei (2020). "Liberation Technology: Mobile Phones and Political Mobilization in Africa". *Econometrica* 88 (2), pp. 533–567.

Oechslin, Manuel and Elias Steiner (2022). "Statistical Capacity and Corrupt Bureaucracies". *Review of International Organizations* 17, pp. 143–174.

Persson, Torsten and Guido Tabellini (2000). *Political Economics: Explaining Economics Policy*. Cambridge: MIT Press.

Roberts, Margaret E. (2020). "Resilience to Online Censorship". *Annual Review of Political Science* 23, pp. 401–419.

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov (2020). "Political Effects of the Internet and Social Media". *Annual Review of Economics* 12, pp. 415–38.

Appendix

Derivation of u and v

The information on expected utility given at the bottom of Figure 2 is based on an ad-hoc assumption on how the observer's posterior, q, translates into the probability of player A being reelected, $u(q)$, and the probability of player B being elected instead, $v(q)$. But v and u can easily be derived from optimizing behavior by the observer. In practice, the observer could be an electorate that elects a prime minister or a company owner who appoints a CEO.

Suppose that the observer's utility is affected by two factors. On the one hand, keeping player A in her current position generates a benefit. This may reflect that player A is a better match for the observer in terms of a non-economic—or "value"—issue that is of particular importance to the observer. But the salience of this value issue is subject to random forces. For concreteness, denote the size of the benefit (in utility terms) by $a \geq 0$. Moreover, assume that Nature draws a from a uniform distribution with support $[0, \omega]$, where $\omega > 1$.

On the other hand, there is the possibility of an economic loss whose size (in utility terms) is normalized to 1. This loss materializes if and only if player A is incompetent and kept in her position. The loss may result from damages that, for instance, an economy (with A as the incumbent prime minister) or a company (with A as the incumbent CEO) takes if the position is in incompetent hands for two consecutive terms. There is no uncertainty, however, when it comes to player B, who is a "safe pair of hands", known for being competent in positions like the contested one. As a result, opting for player B would avoid any economic loss.

Nature draws a after player B's decision (if any) on the publication of possible evidence. Knowing a, the observer then chooses the player who maximizes expected utility,

$$W(P) = \begin{cases} a - q & : \quad P = A \\ 0 & : \quad P = B \end{cases}, \tag{12}$$

where $P = A$ means a decision in favor of the incumbent and $P = B$ one in favor player B. Assuming that in case of equal expected utility the observer prefers the incumbent, it

follows from equation (12) that the incumbent is reelected if $a \geq q$ and that player B is elected otherwise. Thus, at the point where the observer has formed q but Nature has not yet drawn a from the uniform distribution, the chance of the incumbent being reelected is

$$u(q) = \Pr[A \text{ reelected} \mid q] = (1/\omega)(\omega - q). \tag{13}$$

Finally, using the transformation $\alpha = 1 - (1/\omega)$ yields $u(q) = [1 - (1 - \alpha)q]$. So, the more salient the value issue can become, the larger the incumbency advantage α enjoyed by player A. The expression for $v(q)$, the chance that player B wins, now follows immediately: $(1 - \alpha)q$.

Proofs

Proof of Proposition 2. In general terms, the probability that player A loses its position is $(1 - \alpha)q$, where $q = 1$ if evidence of incompetence is published and $q = q^n$ otherwise. In turn, the functional form of q^n depends on the nature of the equilibrium.

First consider $\Pr[A \text{ loses} \mid C = nc]$, the probability of losing when incompetent. In the no-repression equilibrium (nr), the law of iterated expectations implies

$$\begin{aligned}\Pr[A \text{ loses} \mid C = nc]_{nr} = (1 - \sigma)\Pr[A \text{ loses} \mid C = nc, P = np]_{nr} \\ + \sigma \Pr[A \text{ loses} \mid C = nc, P = p].\end{aligned} \tag{14}$$

Equation (14) uses the result that, whenever evidence of incompetence surfaces $(E = e)$ and player A desists from repression $(R = nr)$, player B publishes that evidence $(P = p)$. Now consider the two conditional probabilities on the right-hand side of equation (14). The first one is given by $(1 - \alpha)q^n_{nr}$ and the second one by $(1 - \alpha)$ 1. Therefore,

$$\Pr[A \text{ loses} \mid C = nc]_{nr} = (1 - \alpha)\left[(1 - \sigma)q^n_{nr} + \sigma\right]. \tag{15}$$

Replacing in equation (15) q^n_{nr} with the explicit expression provided in equation (5), and then rearranging terms, results in $\Pr[A \text{ loses} \mid C = nc]_{nr} = (1 - \alpha)/(2 - \sigma)$.

Let's now turn to the partial-repression equilibrium (m). Again accounting for the fact that, without repression, surfacing evidence get published, the law of iterated expectations produces

$$\begin{aligned}\Pr[A \text{ loses} \mid C = nc]_{m} = (1 - m)\{(1 - \sigma)\Pr[A \text{ loses} \mid C = nc, P = np]_{m} \\ + \sigma \Pr[A \text{ loses} \mid C = nc, P = p]\} + m \Pr[A \text{ loses} \mid C = nc, P = np]_{m}.\end{aligned} \tag{16}$$

The first (and third) conditional probability on the right-hand side of equation (16) is given by $(1-\alpha)q_m^n$. As above, the second one is $(1-\alpha)$ 1. As a result, one gets

$$\Pr[A \text{ loses} \mid C = nc]_m = (1-\alpha)\left\{(1-m)\left[(1-\sigma)q_m^n + \sigma\right] + mq_m^n\right\}. \tag{17}$$

Explicit expressions for m and q_m^n are provided by equations (10) and (11), respectively. Using those in equation (17), and then rearranging terms, yields $\Pr[A \text{ loses} \mid C = nc]_m = \chi/(z\sigma)$.

In the full-repression equilibrium (r), evidence of incompetence never gets published and the absence-evidence posterior equals 1/2. Therefore, $\Pr[A \text{ loses} \mid C = nc]_r = (1-\alpha)(1/2)$.

Combining the results so far, and observing the subranges of σ stated in Proposition 1, one obtains

$$l^{nc} = \Pr[\text{A loses} \mid C = nc] = \begin{cases} (1-\alpha)/(2-\sigma) & : \quad \sigma \leq \bar{\sigma}_{nr} < \underline{\sigma}_r \\ \chi/(z\sigma) & : \quad \bar{\sigma}_{nr} < \sigma \leq \underline{\sigma}_r \\ (1-\alpha)/2 & : \quad \bar{\sigma}_{nr} < \underline{\sigma}_r < \sigma \end{cases}. \tag{18}$$

The inverted V-shape profile mentioned in the proposition arises immediately from equation (18). As for continuity, observe that $(1-\alpha)/(2-\sigma)$ equals $\chi/(z\sigma)$ at the border between the no-repression and the partial-repression equilibrium and that $\chi/(z\sigma)$ equals $(1-\alpha)/2$ at the border between the partial-repression and the full-repression equilibrium.

Now consider $\Pr[A \text{ loses} \mid C = c]$. If $C = c$, evidence of incompetence does not surface and thus is never published. So $\Pr[A \text{ loses} \mid C = c] = (1-\alpha)q^n$. Given this, Proposition 1 implies

$$l^{c} = \Pr[\text{A loses} \mid C = c] = \begin{cases} (1-\alpha)(1-\sigma)/(2-\sigma) & : \quad \sigma \leq \bar{\sigma}_{nr} < \underline{\sigma}_r \\ (1-\alpha) - \chi/(z\sigma) & : \quad \bar{\sigma}_{nr} < \sigma \leq \underline{\sigma}_r \\ (1-\alpha)/2 & : \quad \bar{\sigma}_{nr} < \underline{\sigma}_r < \sigma \end{cases}. \tag{19}$$

The V-shape profile mentioned in the proposition is that of q^n, scaled by the factor $(1-\alpha)$. Finally, continuity follows from the continuity of q^n.

Proof of Proposition 3. The unconditional probability of player A losing to B is given by

$$\Pr[A \text{ loses}] = (1/2)\Pr[A \text{ loses} \mid C = nc] + (1/2)\Pr[A \text{ loses} \mid C = c] = (1/2)(l^{nc} + l^{c}). \tag{20}$$

Replacing in equation (20) l^{nc} and l^{c} with the expressions in equations (18) and (19), respectively, immediately leads to $l = \Pr[A \text{ loses}] = (1-\alpha)/2$ for all $\sigma \in [0,1]$.

Proof of Proposition 4. Regarding the first claim of the proposition, note that $\bar{\sigma}_{nr}$ is a function of $\chi/z \in (0,\infty)$ that under any parameter constellation increases from $\lim_{\chi/z\to 0}\bar{\sigma}_{nr}(\chi/z) = 0$ to $\lim_{\chi/z\to\infty}\bar{\sigma}_{nr}(\chi/z) = 2$ in a strictly monotonic fashion. So there always exists a $\underline{\chi/z} < \infty$ such that $\bar{\sigma}_{nr}(\underline{\chi/z}) = 1$ and $\bar{\sigma}_{nr}(\chi/z) > 1$ for all $\chi/z > \underline{\chi/z}$. It follows from $\sigma \leq 1$ and equations (18) and (19) that with $\chi/z \geq \underline{\chi/z}$ only the strictly increasing arm of l^{nc} is relevant and only the strictly decreasing arm of l^{c} is relevant. As a result, for any $\chi/z \geq \underline{\chi/z}$, equilibrium probability l^{nc} is a strictly increasing function of σ on $[\sigma_0, 1]$ and equilibrium probability l^{c} is a strictly decreasing function of σ on $[\sigma_0, 1]$.

Regarding the second claim, note that the above exposition immediately implies that under any parameter constellation and for any $\sigma_0 > 0$ there must exist a $\overline{\chi/z} > 0$ such that $\bar{\sigma}_{nr}(\overline{\chi/z}) = \sigma_0$ and $\bar{\sigma}_{nr}(\chi/z) < \sigma_0$ for all $\chi/z < \overline{\chi/z}$. It follows from equations (18) and and (19) that with $\chi/z < \overline{\chi/z}$ the strictly increasing arm of l^{nc} is irrelevant and the strictly decreasing arm of l^{c} is irrelevant. As a result, for any $\chi/z < \overline{\chi/z}$, equilibrium probability l^{nc} is a decreasing function of σ on $[\sigma_0, 1]$ and equilibrium probability l^{c} is an increasing function of σ on $[\sigma_0, 1]$.

Proof of Proposition 5. To determine the choice of χ/z prior to the start of the game shown in Figure 2, player A computes the payoffs to be expected under each of the three equilibria that are specified in Proposition 1. In doing so, player A takes into account that the chance of being of high incompetence is 1/2. Below, I use $EU^{A}_{I,X}$ to denote those payoffs, where I stands for ignorance (competence level not yet known) and $X \in \{nr, m, r\}$ means that χ/z is chosen so as to induce equilibrium X (given the other parameters). For the no-repression equilibrium, the law of iterated expectations then implies

$$EU^{A}_{I,nr} \;=\; \frac{1}{2}\left\{(1-\sigma)\left[1-(1-\alpha)q^{n}_{nr}\right]z + \sigma\left[1-(1-\alpha)\;1\right]z\right\} \;+\; \frac{1}{2}\left[1-(1-\alpha)q^{n}_{nr}\right]z. \quad (21)$$

Replacing in equation (21) q^{n}_{nr} with the explicit expression provided in equation (5), and then rearranging terms, results in $EU^{A}_{I,nr} = z(1+\alpha)/2$, an expression independent of σ. On the one hand, if competent, an increase in σ is beneficial for player A as it makes absence of evidence a stronger signal of competence. On the other hand, if incompetent, an increase in σ harms A as it means that evidence is more likely to emerge. The net effect is exactly zero.

Moving to the partial-repression equilibrium, one finds that the analogue to equation (21) is

$$EU^{A}_{I,m} = \frac{1}{2}\left\{\left[1-(1-\alpha)q^{n}_{m}\right]z - \chi\right\} + \frac{1}{2}\left[1-(1-\alpha)q^{n}_{m}\right]z. \quad (22)$$

Equation (22) uses that under partial repression $(1-\sigma)\left[1-(1-\alpha)q^{n}_{m}\right]z + \sigma\left[1-(1-\alpha)\;1\right]z$

must be equal to $[1-(1-\alpha)q_m^n]\,z-\chi$. Replacing in equation (22) q_m^n with the explicit expression provided in equation (9), and then rearranging terms, results in $EU_{I,m}^A = z\alpha+\chi/\sigma-\chi/2$.

If player A sets χ/z to induce full repression, the payoff to be expected without knowing C is

$$EU_{I,r}^A = \frac{1}{2}\left\{[1-(1-\alpha)q_r^n]\,z-\chi\right\}+\frac{1}{2}\,[1-(1-\alpha)q_r^n]\,z. \tag{23}$$

Taking into account that $q_r^n = 1/2$, and then rearranging terms, yields $EU_{I,r}^A = z(1+\alpha)/2-\chi/2$. As under no repression, the expected payoff is independent of σ. This time, the reason is that evidence of incompetence, no matter how often it emerges, gets never published.

Combining the results so far, and observing the subranges of σ stated in Proposition 1, one obtains

$$EU_I^A(\chi) = \begin{cases} z(1+\alpha)/2 & : \quad \sigma \leq \bar{\sigma}_{nr}(\chi/z) < \underline{\sigma}_r(\chi/z) \\ z\alpha+\chi/\sigma-\chi/2 & : \quad \bar{\sigma}_{nr}(\chi/z) < \sigma \leq \underline{\sigma}_r(\chi/z) \\ z(1+\alpha)/2-\chi/2 & : \quad \bar{\sigma}_{nr}(\chi/z) < \underline{\sigma}_r(\chi/z) < \sigma \end{cases}. \tag{24}$$

Assumption 1 (i) ensures that $z(1+\alpha)/2 > z(1+\alpha)/2-\chi/2$, i.e., that player A strictly prefers all values of χ/z that induce the no-repression equilibrium to any $\chi/z \in [(\chi/z)_0,\infty)$ that induces the repression equilibrium. Further note that $\bar{\sigma}_{nr}(\chi/z) < \sigma$ is equivalent to $z(1+\alpha)/2 > z\alpha+\chi/\sigma-\chi/2$. Thus, player A also strictly prefers all values of χ/z that induce the no-repression equilibrium to any χ/z that induces the partial-repression equilibrium. Since there is no upper limit to χ/z, and because $\bar{\sigma}_{nr}(\chi/z)$ is a strictly increasing function of χ/z with $\lim_{\chi/z\to\infty}\bar{\sigma}_{nr}(\chi/z) = 2$ (proof of Proposition 4), player A can always ensure that $\sigma \leq \bar{\sigma}_{nr}(\chi/z)$. That is, player A can always choose χ/z so as to induce the preferred no-repression equilibrium.

Given this, Assumption 1 (ii) implies that χ/z equals $(\chi/z)_0$ if $\sigma \leq \bar{\sigma}_{nr}(\chi^{min})$ and otherwise is implicitly determined by $\sigma = \bar{\sigma}_{nr}((\chi/z)_0)$. The explicit expression for χ/z, stated in the proposition, now follows immediately—as does that χ/z is an increasing function of σ.