Propaganda, Conspiracy Theories, and Accountability in Fragile Democracies

Anqi Li[1], Davin Raiha[2], and Kenneth W. Shotts[3]

October 24, 2019

[1]Department of Economics, Washington University in St. Louis. anqili@wustl.edu.
[2]Kelley School of Business, Indiana University. draiha@iu.edu.
[3]Stanford Graduate School of Business. kshotts@stanford.edu.

Abstract

We develop a model of electoral accountability in the presence of mainstream and alternative media outlets. In addition to standard high and low competence types, the incumbent may be an aspiring autocrat, who controls the mainstream media and will cause substantial harm if not removed from office. Alternative media can help voters identify and remove aspiring autocrats and can enable voters to focus on honest mainstream media assessments of incumbents' competence. But malicious alternative media that peddle false conspiracy theories about the incumbent and the mainstream media can induce voters to mistakenly remove nonautocratic incumbents, which in turn demotivates incumbent effort and undermines accountability. The alternative media is most beneficial when it is honest and known to be honest. It is most dangerous when it is sufficiently credible that voters pay attention to it, but sufficiently likely to be malicious that it undermines accountability.

Many countries inhabit a grey area between democracy and autocracy—their leaders are elected, but may act to eliminate institutional checks on their power. One check that aspiring autocrats seek to remove is the media, which can be induced, by a combination of censorship, ownership, and corruption, to refrain from criticism and act as a propaganda vehicle. Recent examples of rulers who have taken this approach include Turkey's Erdoğan, Hungary's Orban, Venezuela's Chávez and Maduro, and Peru's Fujimori.

Citizens who are unsure about their leaders' intentions and the mainstream media's independence may turn to alternative media: opposition television and newspapers, the internet and social media, or foreign news providers. However, alternative media may themselves be untrustworthy, seeking to undermine a legitimate ruler by maliciously spreading fake news or false theories about conspiracies between the incumbent and the mainstream media.

We develop a model of electoral selection and accountability in which citizens are uncertain about the incumbent's type, the independence of the mainstream media, and the reliability of the alternative media. Citizens want to re-elect competent leaders, but when the mainstream media praises the incumbent they don't know whether this is neutral praise of a skillful policymaker or propaganda for an aspiring autocrat. Voters can use alternative media sources to identify and remove autocratic types, but they also sometimes mistakenly remove incumbents who are falsely accused. The alternative media also undermines trust in the mainstream media.

More subtly, the alternative media affects accountability, by altering whether voters use information from the mainstream media to decide whether to re-elect the incumbent. This in turn can undermine incumbents' incentives to demonstrate competence by exerting effort on policymaking. In other circumstances, the alternative media can improve incumbents' incentives, by allowing voters to focus on their competence.

A key parameter in our analysis is the reliability of the alternative media. If it is unlikely to publish false accusations, its main effect is to help citizens remove aspiring autocrats. In contrast, an alternative media that is known to be malicious is ignored by citizens, who expect it to cry wolf about legitimate leaders. The worst case arises when the alternative media is sufficiently reliable that voters pay attention to its claims, yet sufficiently unreliable that its false accusations disrupt the process of accountability; when this happens, incumbent incentives are undermined and voters lose the ability to select competent leaders.

Our theory sits at the intersection of the literatures on media bias, electoral accountability, and fragile democracies. The media's role in accountability has been examined by many scholars, including several who analyze forms of pro- or anti-incumbent bias (Besley and Prat 2006, Ashworth and Shotts 2010, Warren 2012, Wolton 2019). In those models, voters know the media's bias, whereas a key feature of our model is that voters are uncertain about

whether the mainstream media is captured and whether the alternative media is reliable.[1]

Research on autocracies and fragile democracies includes empirical analyses of government-controlled media (for a review, see Enikolopov and Petrova 2015) as well as theories of propaganda and censorship (e.g., Egorov, Guriev, and Sonin 2009; Gehlbach and Sonin 2014; Lorentzen 2014; Cheah 2016; Gehlbach, Sonin, and Svolik 2016; Little 2017; Horz 2018). Very few models analyze accountability in weak democracies (though see Svolik 2013).

Our focus is closest to Guriev and Treisman's (2018) model of propaganda and censorship by a regime that wishes to appear competent. However, they focus on countries that are already non-democratic, in which citizens are only concerned about whether the ruler is competent, not whether she will subvert democracy and move towards autocracy. Another difference is that criticism of the incumbent in their model is always accurate, whereas a key parameter in our analysis is the reliability of the alternative media.

Model

Consider a model of policymaking and elections, with five actors: an incumbent policymaker, replacement policymaker, mainstream media, alternative media, and voter. There are two equally-likely states of the world, $\omega \in \{0,1\}$. The incumbent may learn ω and chooses a policy $x \in \{0,1\}$. Then the mainstream media announces a message $m \in \{0,1\}$ about the policy it believes is correct. The alternative media simultaneously issues a report $r \in \{C, NC\}$ about whether the mainstream media is engaged in a conspiracy with an incumbent who is an aspiring autocrat. Finally, the voter observes x, m, and r but not the true state ω, and decides between the incumbent and replacement.

The incumbent can be a high, low, or autocratic type, $\theta_I \in \{H, L, A\}$. High and low competence types are standard in the accountability literature. The autocratic type is novel–she controls the mainstream media and if she is re-elected she will consolidate her power and deliver a negative expected payoff to voters. The incumbent is autocratic with probability α. Conditional on not being autocratic, she is a high type with probability π. Low and autocratic types only know the prior $\Pr(\omega = 1) = \frac{1}{2}$ and cannot acquire additional information.[2] Type H can exert effort, at cost k, to learn the state of the world before choosing policy. Politicians get an ego-rent of 1 if they win the election.

The mainstream media is either truthful or propagandist. If truthful, it non-strategically issues a report $m \in \{0,1\}$ that matches the true state ω with probability $q \in \left(\frac{1}{2}, 1\right)$. The

[1] Outside of the accountability literature, advice from experts with uncertain biases has been examined in models of electoral competition (Chan and Suen 2009) and policy advice (Minozzi 2011). Whistleblowing about wrongdoing has been examined in many settings (Chassang and Padró i Miquel 2019), including models tailored to bureaucracies (Ting 2008) and judicial hierarchies (Beim, Hirsch, and Kastellec 2014).

[2] Even if the autocratic type could exert high effort at a cost, she would have no incentive to do so because she is always praised by the propagandist mainstream media.

mainstream media is a propagandist if and only if the incumbent is autocratic, in which case it always reports that that the incumbent's policy choice was correct, $m = x$.

The alternative media is either truthful or malicious. If truthful, it non-strategically reports $r = C$, i.e., that the incumbent is autocratic and there is conspiracy, if and only if this is indeed the case. A malicious alternative media always reports $r = C$. A central parameter in our analysis is the probability that the alternative media is malicious, $\phi \in [0, 1]$.

After observing x, m and r, the voter elects the incumbent or replacement. The voter gets utility 1 (in an unmodeled second period) from re-electing a high-type incumbent, 0 from a low type, and $-a$ from an autocratic type.[3] As is standard in accountability models, the replacement doesn't take any actions unless elected. We set $U_R \in [-a, 1]$ as the voter's exogenous expected utility from the replacement.

We characterize Perfect Bayesian equilibria that are symmetric with respect to policies 0 and 1. Symmetry implies that low and autocratic incumbents choose each policy with probability $\frac{1}{2}$. Other components of an equilibrium are: (1) the high-type incumbent's effort decision and policy choice, and (2) the voter's belief about the incumbent, as well as his election decision. For technical simplicity, we assume that a high-type incumbent who is indifferent exerts effort and a voter who is indifferent removes the incumbent.

Baseline

To establish a baseline, suppose there is no alternative media. We say there is *accountability* if two conditions hold: (1) high-type incumbents exert effort to choose policy in the voter's interest and (2) the incumbent is removed from office if $m \neq x$. We characterize equilibria with accountability wherever it is possible, as well as what happens when accountability is impossible. Note that the two conditions for accountability are mutually reinforcing: a high type exerts effort to choose good policies and earn praise from the mainstream media, and the fact that she exerts effort means the media message m is informative about her type.

Voter inferences are complicated by the possibility that the incumbent is autocratic and the media is propagandist. If the incumbent is very likely to be autocratic, the media message is meaningless. But if the autocratic type isn't too likely or too harmful, the possibility of propaganda doesn't overwhelm the positive information conveyed by the report $m = x$. If high-type incumbents choose $x = \omega$ and low and autocratic types choose each policy with probability $\frac{1}{2}$, the voter's expected utility from re-electing the incumbent is:

[3] A possible objection is that some autocrats turn out to be effective policymakers who are well-liked by their subjects. In our model, $-a$ represents voters' *expected* payoff when an autocrat remains in office. This can include autocrats who turn out to be "good" along with those who turn out to be disastrous, as long as the expected payoff is negative.

$$\overline{U} \coloneqq 1 \cdot \Pr\left(\theta_I = H | m = x\right) - a \cdot \Pr\left(\theta_I = A | m = x\right) = \frac{\pi q - al}{\pi q + (1-\pi)\frac{1}{2} + l},$$

where $l \coloneqq \frac{\alpha}{1-\alpha}$ is the likelihood that the incumbent is autocratic. On the other hand, when $m \neq x$ the voter knows the incumbent is nonautocratic, so his utility from re-electing her is his belief about her being type H:

$$\underline{U} \coloneqq 1 \cdot \Pr\left(\theta_I = H | m \neq x\right) = \frac{\pi\left(1-q\right)}{\pi\left(1-q\right) + (1-\pi)\frac{1}{2}}.$$

If the incumbent and replacement are similarly appealing ex ante, i.e., $U_R \in [\underline{U}, \overline{U})$, the voter bases his decision on the mainstream media message. Accountability also requires that the high-type incumbent exert effort. If she doesn't exert effort, then $m = x$ with probability $\frac{1}{2}$, but with effort this probability increases to q. Thus the effort cost must be less than $q - \frac{1}{2}$. In what follows, we focus on parameters such that it is possible to have accountability for some values of U_R.

Assumption 1. *$k \leq q - \frac{1}{2}$, and α and a are sufficiently small that $\underline{U} < \overline{U}$.*

The following result summarizes the baseline.

Proposition 1. *With only the mainstream media.*

1. *If $U_R \in [\underline{U}, \overline{U})$, accountability is possible. High type incumbents exert effort to choose $x = \omega$ and the voter re-elects the incumbent iff $m = x$.*

2. *Otherwise accountability is impossible and the voter re-elects the incumbent iff she is ex ante better than the replacement, i.e., $(1-\alpha)\pi - a\alpha > U_R$.*

Alternative Media

We now analyze the full model, in which the alternative media reports on whether the incumbent is an aspiring autocrat. To assess the effects of the alternative media, we ask two questions. First, how does it affect the incumbent's policymaking effort? Second, how does it affect electoral selection, both in the sense of re-electing high-type incumbents and in the sense of removing autocratic ones? We provide intuitive answers to these questions and then state equilibria in Proposition 2.

Accountability The probability ϕ that the alternative media is malicious plays a key role in determining whether accountability is possible. To see this, we begin with situations with accountability in the baseline model, i.e., $U_R \in [\underline{U}, \overline{U})$ from Proposition 1.

If $\phi = 0$, the alternative media perfectly reveals whether the incumbent is autocratic, and there is an accountability equilibrium in which re-election requires not only $m = x$ but also $r = NC$, i.e., the alternative media doesn't allege a conspiracy. In contrast, if $\phi = 1$, the alternative media always reports $r = C$, alleging a conspiracy, so there is an accountability equilibrium in which the alternative media is ignored and re-election is based solely on the mainstream media's message, as in the baseline model.

For intermediate values, $\phi \in (0, 1)$, the alternative media's effect is twofold. The first effect concerns the high-type incumbent's effort. If re-election requires $m = x$ and $r = NC$, then increases in ϕ demotivate the incumbent, because effort is only rewarded if the alternative media is truthful. To induce effort requires $k \leq \left(q - \frac{1}{2}\right)(1 - \phi)$ or, equivalently,

$$\phi \leq \phi_e := 1 - \frac{k}{q - \frac{1}{2}}.$$

The second effect concerns whether the voter uses information from the alternative media. The voter removes the incumbent upon hearing $m = x$ and $r = C$ if

$$U_R \geq U_v := \frac{\pi q \phi - al}{\pi q \phi + (1 - \pi)\frac{1}{2}\phi + l}.$$

This requires that the alternative media be sufficiently reliable,

$$\phi \leq \phi_v := \frac{(a + U_R)\, l}{\pi q\,(1 - U_R) - (1 - \pi)\frac{1}{2}U_R}.$$

Otherwise the voter ignores the alternative media.

Combining these effects, we see how ϕ affects accountability. If $\phi_v \leq \phi_e$, then if there is accountability in the baseline, there is also accountability with an alternative media (Proposition 2.2(a)). The voter attends to the alternative media if $\phi \leq \phi_v$ and ignores it otherwise.

Things are more interesting if $\phi_v > \phi_e$ (Proposition 2.2(b)). If $\phi \leq \phi_e$ or $\phi > \phi_v$, there is accountability and the voter attends to the alternative media in the first case and ignores it in the second case. But for intermediate values $\phi \in (\phi_e, \phi_v]$, accountability is impossible.

Under what circumstances does the presence of the alternative media disrupt accountability? First, the replacement must be moderately appealing compared to the incumbent, $U_R \in [\underline{U}, \overline{U})$. Second, the alternative media must be sufficiently trustworthy that the voter attends to it ($\phi \leq \phi_v$) but sufficiently untrustworthy that the incumbent's fear of being falsely criticized demotivates her effort ($\phi > \phi_e$). A necessary condition is $\phi_e < \phi_v$ or, equivalently $k > \left(q - \frac{1}{2}\right)(1 - \phi_v)$. Thus, disruption of accountability due to the presence of an alternative media occurs when policymaking is difficult, in the sense that it is costly for the

incumbent to learn the correct policy that serves the voter's interest.

In other situations, the alternative media can have positive effects on accountability. In the baseline model, when the replacement is highly appealing ($U_R \geq \overline{U}$), accountability is impossible and the voter removes the incumbent even when $m = x$. The alternative media reveals additional information, because $r = NC$ means the incumbent is nonautocratic. If the high-type incumbent exerts effort, the voter's expected utility from re-electing the incumbent after $m = x$ and $r = NC$ is:

$$\overline{\overline{U}} \coloneqq \Pr\left(\theta_I = H | m = x \text{ and } r = NC\right) = \frac{\pi q}{\pi q + (1-\pi)\frac{1}{2}} > \overline{U}.$$

As is shown in Proposition 2.3, the presence of the alternative media makes accountability possible when the replacement is reasonably highly-appealing and false criticism is not so likely as to demotivate effort, i.e., $U_R \in [\overline{U}, \overline{\overline{U}})$ and $\phi \leq \phi_e$.

Selection We analyze how the probability ϕ that the alternative media is malicious affects selection, starting with cases where the replacement is moderately appealing: $U_R \in [\underline{U}, \overline{U})$ from Proposition 1.

The effect is straightforward if $\phi_v \leq \phi_e$. For $\phi \leq \phi_v$, there is accountability and the voter listens to both media outlets. Increasing ϕ worsens selection, as nonautocratic incumbents are more frequently falsely accused and removed from office. For $\phi > \phi_v$, there is accountability but the alternative media is ignored, so local changes in ϕ don't affect selection.

Things are more dramatic if $\phi_v > \phi_e$, because changes in ϕ can affect the incumbent's effort. At $\phi = \phi_e$, the equilibrium transitions from one with accountability to one without accountability. Thus as ϕ crosses the threshold of ϕ_e, the voter completely loses the benefit of selecting based on the mainstream media message. However, as the alternative media becomes more unreliable and ϕ increases above ϕ_v, accountability is restored, and the voter is able to benefit from selecting based solely on the mainstream media message.

We also note how ϕ affects selection based on the mainstream media report when $U_R \notin [\underline{U}, \overline{U})$. With an unappealing replacement ($U_R < \underline{U}$), there is no accountability and the voter never selects based on the mainstream media message. With a highly-appealing replacement ($U_R \in [\overline{U}, \overline{\overline{U}})$), there is accountability if the alternative media is not demotivating, i.e., $\phi \leq \phi_e$, in which case the voter uses information from both media outlets.

Finally, we note that absent accountability, the voter benefits from selecting based on the alternative media report if two conditions hold. First, the alternative media must be sufficiently trustworthy for the voter to remove the incumbent when $r = C$. This requires

$$\phi \leq \phi_s \coloneqq \frac{(a + U_R)\, l}{\pi - U_R} \text{ or, equivalently, } U_R \geq U_s \coloneqq \frac{\pi\phi - al}{\phi + l}.$$

Second, the incumbent must be sufficiently likely to be a high type for the voter to re-elect her when he sees $r = NC$, learns that she is nonautocratic, and doesn't learn anything about her competence. This requires $\pi > U_R$ or, equivalently, $\phi_s > 0$.

Voter welfare Our results on accountability and selection establish non-monotonicity of voter welfare as we vary the probability that the alternative media is malicious. The voter can be better off with a highly unreliable alternative media, which he ignores, than with a somewhat reliable one that cannot be ignored but demotivates the incumbent's policy effort.

Summary The following proposition summarizes equilibria with accountability whenever it is possible as well as what happens in equilibrium when accountability is impossible.

Proposition 2. *With an alternative media.*

1. *For $U_R < \underline{U}$, the incumbent never exerts effort and two situations can happen:*
   (a) *if $\phi \leq \phi_s$, the incumbent is re-elected iff $r = NC$;*
   (b) *if $\phi > \phi_s$, the incumbent is always re-elected.*
2. *For $U_R \in [\underline{U}, \overline{U})$.*
   (a) *If $\phi_v \leq \phi_e$, two situations can happen:*
      i. *if $\phi \leq \phi_v$, the incumbent exerts effort and is re-elected iff $m = x$ and $r = NC$;*
      ii. *if $\phi > \phi_v$, the incumbent exerts effort and is re-elected iff $m = x$.*
   (b) *If $\phi_v > \phi_e$, three situations can happen:*
      i. *if $\phi \leq \phi_e$, the incumbent exerts effort and is re-elected iff $m = x$ and $r = NC$;*
      ii. *if $\phi \in (\phi_e, \phi_v]$, the incumbent exerts no effort, and is either never re-elected (if $\phi_s < 0$), always re-elected (if $\phi > \phi_s > 0$), or re-elected iff $r = NC$ (if $\phi \leq \phi_s$);*
      iii. *if $\phi > \phi_v$ the incumbent exerts effort and is re-elected iff $m = x$.*
3. *For $U_R \in [\overline{U}, \overline{\overline{U}})$,*
   (a) *if $\phi \leq \phi_e$, the incumbent exerts effort and is re-elected iff $m = x$ and $r = NC$;*
   (b) *if $\phi > \phi_e$, the incumbent exerts no effort and two situations can happen:*
      i. *If $U_R < \pi$, the incumbent is re-elected iff $r = NC$;*
      ii. *If $U_R \geq \pi$, the incumbent is never re-elected;*
4. *For $U_R \geq \overline{\overline{U}}$, the incumbent exerts no effort and is never re-elected.*

Discussion

We conclude by briefly discussing additional implications of our model. It is straightforward to show that the voter's belief about the probability that the mainstream media is truthful decreases when the alternative media accuses it of being a propaganda vehicle for the incumbent. This is consistent with the decline in trust in the media at a time when there is a proliferation of alternative media sources that disparage the mainstream media.

At a broader level, our model contributes to the growing literature on democratic backsliding and autocratization in countries such as Brazil, Ecuador, Hungary, India, Poland, Russia, Turkey, and Venezuela. Backsliding takes many forms, as aspiring autocrats undermine multiple institutions, including elections, opposition parties, the bureaucracy, and the judiciary. Scholars have identified behind-the-scenes control of the media as one of the most common forms of backsliding (Bermeo 2016; Mechkova, Lührmann, and Lindberg 2017), and evidence on bribes by Peru's Fujimori shows that he placed an especially high value on the media (McMillan and Zoido 2004). Alternative media and civic organizations can sometimes counteract such actions, serving as independent information sources and undermining support for autocratic rulers (Knight and Tribin 2019). However, as noted by Bermeo (2016), they often face a credibility problem because rulers may accuse them of being special interests, representatives of a discredited old order, or tools of foreign powers.

Another application of our model is to fake news that alleges conspiracies between politicians and the mainstream media. Fake news has been studied by empiricists who analyze effects on voting and electoral outcomes (e.g., Allcott and Gentzkow 2017). Theoretical analyses include Allcott and Gentzkow's verbal theory, Yea's (2018) Bayesian persuasion game, Taylor's (2019) model of message substitution, and Bräuninger and Marinov's (2019) model of conspiracy theories and skepticism, all of which analyze the generation of false claims as well as effects on citizens' beliefs and behavior.

Our model suggests that the effect of fake news goes beyond fooling voters into making errors. Rather, fake news, like other malicious alternative media sources, can undermine accountability, disincentivize incumbent policymaking effort, and make it impossible for voters to assess an incumbent based on her policy record. This implication dovetails with Jee, Lueders, and Myrick's (2019) argument that fake news and propaganda contribute to democratic backsliding and "negatively affect[-] citizens' ability to make informed choices in elections and to hold their elected representatives accountable."

Our analysis makes it possible to determine when alternative media will have a negative effect, when it will have positive effect, and when it will have no effect. The alternative media is most beneficial when it is honest and known to be honest. In contrast, an alternative media that is known to be malicious is not much of a problem, because voters ignore it. What is

most problematic is an alternative media that is sufficiently likely to be honest that voters pay attention to it, but sufficiently likely to be malicious that it interferes with selection, undermines incumbent incentives, and destroys accountability.

References

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31:211-236.

Ashworth, Scott, and Kenneth W. Shotts. 2010. "Does Informative Media Commentary Reduce Politicians' Incentives to Pander?" *Journal of Public Economics* 94:838-847.

Beim, Deborah, Alexander V. Hirsch, and Jonathan P. Kastellec. 2014. "Whistleblowing and Compliance in the Judicial Hierarchy." *American Journal of Political Science* 58:904-918.

Bermeo, Nancy. 2016. "On Democratic Backsliding." *Journal of Democracy* 27:5-19.

Besley, Timothy, and Andrea Prat. 2006. "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability." *American Economic Review* 96:720-736.

Bräuninger, Thomas, and Nikolay Marinov. 2019. "Conspiracies as a Strategy of Political Communication." Unpublished manuscript.

Chan, Jimmy, and Wing Suen. 2009. "Media as Watchdogs: The Role of News Media in Electoral Competition." *European Economic Review* 53:799-814.

Chassang, Sylvain, and Gerard Padró i Miquel. 2019. "Corruption, Intimidation, and Whistleblowing: A Theory of Inference from Unverifiable Reports." Forthcoming in *Review of Economic Studies.*

Cheah, Hon Foong. 2016. "Does Foreign Media Entry Discipline or Provoke Local Media Bias?" *International Journal of Economic Theory* 12:335-359.

Egorov, Georgy, Sergei Guriev, and Konstantin Sonin. 2009. "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *American Political Science Review* 103:645-668.

Enikolopov, Ruben, and Maria Petrova. 2015. "Media Capture: Empirical Evidence." *Handbook of Media Economics, Volume 1B.*

Gehlbach, Scott, and Konstantin Sonin. 2014. "Government Control of the Media." *Journal of Public Economics* 118:163-171.

Gehlbach, Scott, Konstantin Sonin, and Milan W. Svolik. 2016. "Formal Models of Nondemocratic Politics." *Annual Review of Political Science* 19:565-584.

Guriev, Sergei, and Daniel Treisman. 2018. "Informational Autocrats." Unpublished manuscript.

Horz, Carlo M. 2018. "Propaganda and Skepticism." Unpublished manuscript.

Jee, Haemin, Hans Lueders, and Rachel Myrick. 2019. "Towards a Unified Concept of Democratic Backsliding." Unpublished manuscript.

Knight, Brian, and Ana Tribin. 2019. "The Value of Opposition Media: Evidence from Chavez's Venezuela." Unpublished manuscript.

Little, Andrew T. 2017. "Propaganda and Credulity." *Games and Economic Behavior* 102:224-232.

Lorentzen, Peter. 2014. "China's Strategic Censorship." *American Journal of Political Science* 58:402-414.

McMillan, John, and Pablo Zoido. 2004. "How to Subvert Democracy: Montesinos in Peru." *Journal of Economic Perspectives* 18:69-92.

Mechkova, Valeriya, Anna Lührmann, and Staffan I. Lindberg. 2017. "How Much Democratic Backsliding?" *Journal of Democracy* 28:162-169.

Minozzi, William. 2011. "A Jamming Theory of Politics." *Journal of Politics* 73:301-315.

Svolik, Milan W. 2013. "Learning to Love Democracy: Electoral Accountability and the Success of Democracy." *American Journal of Political Science* 57:685-702.

Taylor, Zachary. 2019. "Persuasion with Fake News." Unpublished manuscript.

Ting, Michael M. 2008. "Whistleblowing." *American Political Science Review* 102:249-267.

Warren, Patrick L. 2012. "Independent Auditors, Bias, and Political Agency." *Journal of Public Economics* 96:78-88.

Wolton, Stephane. 2019. "Are Biased Media Bad for Democracy." *American Journal of Political Science* 63:548-562.

Yea, Sangjun. 2018. "Persuasion Under the Influence of Fake News." Unpublished manuscript.

Appendix: Proofs

For any parameter values there exists an equilibrium in which the type H incumbent does not exert effort. We analyze whether it is possible to have an equilibrium with accountability.

In the proofs, we use μ^H and μ^A to denote voter beliefs about the probability that the incumbent is type H or A.

Before we prove the results, three things are worth noting.

First, we have assumed that if the incumbent is autocratic the mainstream media always reports $m = x$, i.e., it says that the incumbent chose the correct policy. A natural question is whether a strategic propagandist who seeks to maximize the incumbent's probability of re-election has an incentive to instead report $m \neq x$. In the equilibria that we characterize, a strategic propagandist can never help the incumbent by doing this.

Second, although we have assumed symmetry, which directly implies that type L and A incumbents choose each policy $x \in \{0, 1\}$ with probability $\frac{1}{2}$, Lemma 4 in the Additional Results section at the end of the appendix shows that our results on the existence of equilibria with accountability do not depend on this assumption.

Third, we note that our assumption that a type H incumbent who is indifferent about effort exerts effort and a voter who is indifferent removes the incumbent only affects our analysis for knife-edge parameter cases, specifically $k \in \left\{q - \frac{1}{2}, \left(q - \frac{1}{2}\right)(1 - \phi)\right\}$, $U_R \in \left\{\underline{U}, \overline{U}, \overline{\overline{U}}, \pi\right\}$, and $\phi \in \{\phi_e, \phi_v, \phi_s\}$. These are the only parameters for which these actors can be indifferent when each of them plays pure strategies.

Proof of Proposition 1

Part 1. We first calculate voter beliefs, then turn to voter and incumbent strategies.

Suppose type H exerts effort and chooses $x = \omega$. Recall that the mainstream media reports $m = \omega$ with probability q if the incumbent is type H or L and always reports $m = x$ if the incumbent is type A.

For voter beliefs when $m = x$: with probability $(1 - \alpha)\pi$ the incumbent is type H in which case $m = x$ with probability q, with probability $(1 - \alpha)(1 - \pi)$ the incumbent is type L in which case $m = x$ with probability $\frac{1}{2}$, and with probability α the incumbent is type A in which case $m = x$ always. Thus, using the notation $l \coloneqq \frac{\alpha}{1-\alpha}$, Bayes' rule implies that

$$\mu^H(m = x) = \frac{(1 - \alpha)\pi q}{(1 - \alpha)\pi q + (1 - \alpha)(1 - \pi)\frac{1}{2} + \alpha} = \frac{\pi q}{\pi q + (1 - \pi)\frac{1}{2} + l},$$

$$\text{and} \qquad \mu^A(m = x) = \frac{\alpha}{(1 - \alpha)\pi q + (1 - \alpha)(1 - \pi)\frac{1}{2} + \alpha} = \frac{l}{\pi q + (1 - \pi)\frac{1}{2} + l}.$$

For beliefs when $m \neq x$: with probability $(1-\alpha)\pi$ the incumbent is type H, in which case $m \neq x$ with probability $(1-q)$; with probability $(1-\alpha)(1-\pi)$ the incumbent is type L, in which case $m \neq x$ with probability $\frac{1}{2}$; and with probability α the incumbent is type A, in which case the media never reports $m \neq x$. Thus

$$\mu^H(m \neq x) = \frac{\pi(1-q)}{\pi(1-q) + (1-\pi)\frac{1}{2}}$$
$$\text{and} \qquad \mu^A(m \neq x) = 0.$$

The voter gets utility 1 from type H, 0 from type L, and $-a$ from type A, so his expected utility from re-electing when $m = x$ is

$$1 \cdot \mu^H(m = x) - a \cdot \mu^A(m = x) = \frac{\pi q - al}{\pi q + (1-\pi)\frac{1}{2} + l} = \overline{U},$$

and his expected utility from re-electing when $m \neq x$ is

$$1 \cdot \mu^H(m \neq x) - a \cdot \mu^A(m \neq x) = \frac{\pi(1-q)}{\pi(1-q) + (1-\pi)\frac{1}{2}} = \underline{U}.$$

Thus if the voter's utility from the replacement is $U_R \in [\underline{U}, \overline{U})$, he re-elects the incumbent iff $m = x$.

For type H, given the voter's strategy, she is re-elected with probability $\frac{1}{2}$ if she doesn't exert effort and with probability q if she exerts effort, so she exerts effort if the cost satisfies $k \leq q - \frac{1}{2}$.

Part 2. To prove this part of the proposition we first assume accountability and derive a contradiction. If type H exerts effort and chooses $x = \omega$, then for $U_R < \underline{U}$ the voter always re-elects the incumbent, so type H will not exert effort. Likewise, if type H exerts effort and chooses $x = \omega$, then for $U_R \geq \overline{U}$ the voter always removes the incumbent so type H will not exert effort.

Given that the incumbent does not exert effort and her choice of x does not reveal anything about her type, the voter's expected utility from the incumbent is the same regardless of whether $m = x$ or $m \neq x$. With probability $(1-\alpha)\pi$ she is type H, with probability $(1-\alpha)(1-\pi)$ she is L, and with probability α she is A. So the voter's expected utility from the incumbent is

$$1 \cdot (1-\alpha)\pi + 0 \cdot (1-\alpha)(1-\pi) - a \cdot \alpha = (1-\alpha)\pi - a\alpha,$$

and the voter re-elects the incumbent iff this is strictly greater than U_R.■

To set up the proof of Proposition 2, we develop three lemmas to characterize voter beliefs, voter best responses, and incumbent best responses.

Lemma 1. *Voter beliefs depend on type H's effort choice as follows.*

1. *If type H exerts effort, then*

   (a) $\mu^H(m = x, r = NC) = \frac{\pi q}{\pi q + (1-\pi)\frac{1}{2}}$ *and* $\mu^A(m = x, r = NC) = 0$;

   (b) $\mu^H(m = x, r = C) = \frac{\pi q \phi}{\pi q \phi + (1-\pi)\frac{1}{2}\phi + l}$ *and* $\mu^A(m = x, r = C) = \frac{l}{\pi q \phi + (1-\pi)\frac{1}{2}\phi + l}$;

   (c) $\mu^H(m \neq x, r = NC) = \mu^H(m \neq x, r = C) = \frac{\pi(1-q)}{\pi(1-q)+(1-\pi)\frac{1}{2}}$; *and*
   $\mu^A(m \neq x, r = NC) = \mu^A(m \neq x, r = C) = 0$.

2. *If type H does not exert effort, then*

   (a) $\mu^H(m = x, r = NC) = \mu^H(m \neq x, r = NC) = \pi$ *and*
   $\mu^A(m = x, r = NC) = \mu^A(m \neq x, r = NC) = 0$;

   (b) $\mu^H(m = x, r = C) = \mu^H(m \neq x, r = C) = \frac{\pi\phi}{\phi + l}$ *and*
   $\mu^A(m = x, r = C) = \mu^A(m \neq x, r = C) = \frac{l}{\phi + l}$.

Proof. All of the results follow directly from Bayes' rule. We note a few details for different parts of the lemma.

1(a). The alternative media reports $r = NC$ only if the voter is not type A. Conditional on $r = NC$, with probability π the incumbent is type H, in which case $m = x = \omega$ with probability q. With probability $(1 - \pi)$ the incumbent is type L, in which case $m = x$ with probability $\frac{1}{2}$.

1(b). Recall that the alternative media reports $r = C$ for types H and L with probability ϕ (the probability that it is malicious) and it always reports $r = C$ when the incumbent is type A. So with probability $(1 - \alpha)\pi q\phi$ the incumbent is type H and the media reports are $m = x$ and $r = C$, with probability $(1 - \alpha)(1 - \pi)\frac{1}{2}\phi$ the incumbent is type L and the reports are $m = x$ and $r = C$, and with probability α the incumbent is type A and the reports are $m = x$ and $r = C$. Applying Bayes' rule and using $l = \frac{\alpha}{1-\alpha}$ yields the result.

1(c). When $m \neq x$ the voter knows the incumbent is not type A, and straightforwardly updates on the competence of the incumbent.

2(a). The alternative media announcement $r = NC$ eliminates the possibility that the incumbent is type A. Because type H does not exert effort, the voter doesn't update about the relative likelihood of type H and L based on m.

2(b). The voter doesn't update based on m, but does update based on r about the probability that the incumbent is nonautocratic (in which case $r = C$ with probability ϕ) or autocratic (in which case $r = C$ with probability 1).■

Lemma 2. *The voter's best responses to type H's effort choice are as follows.*

1. *Suppose type H exerts effort.*
   (a) *When $m = x$ and $r = NC$, the voter reelects the incumbent iff $U_R < \overline{\overline{U}} = \frac{\pi q}{\pi q + (1-\pi)\frac{1}{2}}$.*
   (b) *When $m = x$ and $r = C$, the voter reelects the incumbent iff $U_R < U_v = \frac{\pi q \phi - al}{\pi q \phi + (1-\pi)\frac{1}{2}\phi + l}$ or, equivalently, $\phi > \phi_v = \frac{(a+U_R)l}{\pi q(1-U_R) - (1-\pi)\frac{1}{2}U_R}$.*
   (c) *When $m \neq x$, regardless of r the voter reelects the incumbent iff $U_R < \underline{U} = \frac{\pi(1-q)}{\pi(1-q)+(1-\pi)\frac{1}{2}}$.*
2. *Suppose type H does not exert effort.*
   (a) *When $r = NC$, regardless of m the voter reelects the incumbent iff $U_R < \pi$ or, equivalently, $\phi_s = \frac{(a+U_R)l}{\pi - U_R} > 0$.*
   (b) *When $r = C$, regardless of m the voter reelects the incumbent iff $U_R < U_s = \frac{\pi\phi - al}{\phi + l}$ or, equivalently, $\phi > \phi_s = \frac{(a+U_R)l}{\pi - U_R}$.*

Proof. The voter gets utility 1 from type H, 0 from type L, and $-a$ from type A, so his expected utility from re-electing the incumbent is

$$1 \cdot \mu^H - a\mu^A.$$

The parts of the lemma follow from the corresponding parts of Lemma 1 and algebra.■

Lemma 3. *Type H's best responses to the voter's reelection rule are as follows.*

1. *If the voter re-elects the incumbent iff $m = x$ and $r = NC$, then type H exerts effort iff $k \leq \left(q - \frac{1}{2}\right)(1 - \phi)$ or, equivalently, $\phi \leq \phi_e = 1 - \frac{k}{q - \frac{1}{2}}$.*
2. *If the voter re-elects the incumbent iff $m = x$, then type H exerts effort iff $k \leq q - \frac{1}{2}$.*
3. *If the voter re-elects the incumbent iff $r = NC$, then type H does not exert effort.*
4. *If the voter always re-elects the incumbent or never re-elects the incumbent, then type H does not exert effort.*

Proof. For Part 1: with effort type H's expected utility is $q(1-\phi)-k$, and without effort it is $\frac{1}{2}(1-\phi)$, which yields the condition. For Part 2: with effort type H's expected utility is $q-k$, and without effort it is $\frac{1}{2}$, which yields the condition. For Parts 3 and 4, type H does not exert effort because effort is costly and does not affect the probability of re-election.■

Proof of Proposition 2

For each parameter region, we conjecture that the incumbent *behaves accountably* in the following sense: type H exerts effort and chooses $x=\omega$, whereas L and A choose each $x\in\{0,1\}$ with probability $\frac{1}{2}$. We then characterize the voter's best response. If it's optimal for type H to exert effort given the voter's best response, there is an equilibrium with accountability. Otherwise there isn't, and we characterize what happens in equilibrium when all incumbent types choose each policy with probability $\frac{1}{2}$. Note that $\underline{U}, U_v, U_s < \overline{U} < \overline{\overline{U}}$ and that $\underline{U}, U_s < \pi < \overline{\overline{U}}$.

Proposition 2.1 Suppose $U_R < \underline{U}$ and the incumbent behaves accountably. By Lemma 2.1(c), the voter re-elects the incumbent if $m \neq x$. Thus type H has no incentive to exert effort, so there cannot be an equilibrium with accountability. Now suppose all incumbent types choose each policy with probability $\frac{1}{2}$. Because $U_R < \underline{U} < \pi$, Lemma 2.2(a) implies that when $r=NC$ the incumbent is re-elected. Lemma 2.2(b) shows that when $r=C$ the voter's best response depends on whether $\phi > \phi_s$, thus completing the proof of Proposition 2.1.

Proposition 2.2 Suppose $U_R \in [\underline{U}, \overline{U})$ and the incumbent behaves accountably. The voter's best response depends on whether $\phi > \phi_v$. Consider two cases.

Case 1. If $\phi \leq \phi_v$, then $U_v \leq U_R < \overline{\overline{U}}$, so Lemma 2.1 implies that the incumbent is reelected iff $m=x$ and $r=NC$. By Lemma 3.1, it is optimal for type H to exert effort iff $\phi \leq \phi_e$, and combining yields the existence of equilibria with accountability as in Proposition 2.2(a)i and 2.2(b)i. On the other hand, if $\phi \in (\phi_e, \phi_v]$, then Lemma 3.1 implies that it is not optimal for the incumbent to exert effort given the voter's best response. Without incumbent effort, Lemma 2.2 characterizes voter behavior for Proposition 2.2(b)ii.

Case 2. If $\phi > \phi_v$, then $U_R < U_v < \overline{\overline{U}}$, so Lemma 2.1 implies that the incumbent is reelected iff $m=x$. From Lemma 3.2, it is optimal for type H to exert effort iff $k \leq q - \frac{1}{2}$, which holds true under Assumption 1. Combining yields Proposition 2.2(a)ii and 2.2(b)iii.

Proposition 2.3 Suppose $U_R \in [\overline{U}, \overline{\overline{U}})$ and the incumbent behaves accountably. Since $\overline{U} > U_v$ and hence $U_R > U_v$, Lemma 2.1 implies that the incumbent is reelected iff $m=x$

and $r = NC$. By Lemma 3.1, type H exerts effort iff $\phi \leq \phi_e$, and combining yields the condition for existence of an equilibrium with accountability as in Proposition 2.3(a).

Without effort by type H, first note that $\pi > U_s$. Consider two cases.

Case 1. If $U_R \geq \pi$, then $U_R \geq \pi > U_s$. By Lemma 2, the incumbent is removed regardless of r as in Proposition 2.3(b)ii.

Case 2. If $U_R < \pi$ and $r = NC$, then from Lemma 2.2(a) the incumbent is re-elected. If $U_R < \pi$ and $r = C$ then the incumbent is removed. To see this, note that $U_R \geq \overline{U}$ by assumption and that $\overline{U} > U_s$. Combining yields $U_R > U_s$, so from Lemma 2.2(b) the incumbent is removed when $r = C$. This establishes Proposition 2.3(b)i.

Proposition 2.4 Suppose $U_R \geq \overline{\overline{U}}$ and the incumbent behaves accountably. Since $\overline{\overline{U}} > \max\{U_v, \underline{U}, \pi, U_s\}$, Lemma 2 implies that the incumbent is always removed from office regardless of whether she exerts effort and regardless of r and m. By Lemma 3.4, type H exerts no effort.■

Additional Results

We now show that our restriction to equilibria that are symmetric with respect to the policies does not affect our results on the existence of equilibria with accountability. The next lemma says that if there exists an equilibrium with accountability in which types L and A mix with different probabilities than $\frac{1}{2}$, then we can construct another equilibrium with mixing probability $\frac{1}{2}$:

Lemma 4. *Fix model primitives. If there exists an equilibrium in which type H exerts effort and chooses $x = \omega$, then there exists an equilibrium in which types L and A choose each $x \in \{0, 1\}$ with probability $\frac{1}{2}$, type H exerts effort and chooses $x = \omega$, and voter strategy is unchanged compared to the original equilibrium.*

The intuition is as follows. Symmetry implies that types A and L choose each policy with probability $\frac{1}{2}$. If type A or L were to play asymmetrically, e.g., by choosing $x = 1$ with probability less than $\frac{1}{2}$, this would decrease the voter's expected utility from re-electing the incumbent in information sets where $x = 0$ and increase the voter's expected utility from re-electing in information sets where $x = 1$. If such an asymmetric strategy has any effect on the voter's best response (relative to what happens with types A and L playing symmetrically), it would be to give types A and L an incentive to play $x = 1$ rather than $x = 0$ in order to increase their chances of winning re-election, which means that the asymmetric strategy couldn't be part of an equilibrium. On the other hand, if the asymmetric strategy

doesn't affect the voter's best response relative to what happens with types A and L playing symmetrically, then it has no effect on accountability and the incentive of type H.

Proof. Let p_L and p_A denote the probabilities that types L and A choose $x = 1$ in the original equilibrium, respectively. Lemma 4 claims that changing these probabilities to $\tilde{p}_L = \tilde{p}_A = \frac{1}{2}$ while holding other players' strategies fixed yields a new equilibrium. This requires showing that types L and A are indifferent between $x \in \{0, 1\}$ and thus can mix. It also requires showing that the voter's strategy in the original equilibrium remains a best response after the change, and so the incentives for type H to exert effort and choose $x = \omega$ remain unchanged. The proof proceeds in four steps.

Step 1. Show that in the original equilibrium, the voter must re-elect the incumbent with the same probability in the following two information sets: $m = x = 0, r = C$ and $m = x = 1$, $r = C$.

Consider a type A incumbent. For her, the voter always observes $m = x$ and $r = C$. Thus the above two information sets are the only ones that matter for type A's incentive. Suppose, to the contrary, that the voter behaves asymmetrically in these information sets. This means that either

$$EU_V\left(reelect | m = x = 0, r = C\right) \leq U_R < EU_V\left(reelect | m = x = 1, r = C\right),$$

or

$$EU_V\left(reelect | m = x = 1, r = C\right) \leq U_R < EU_V\left(reelect | m = x = 0, r = C\right),$$

where $EU_V\left(reelect \mid \cdot\right)$ denotes the voter's expected utility from reelecting the incumbent in an information set. Consider the former case. In this case, type A is re-elected iff she chooses $x = 1$, so her strict best response is $p_A = 1$. Using this, along with the conjectured p_L and type H's behavior in the conjectured equilibrium, voter beliefs when $x = 0$ are

$$\mu^H(m = x = 0, r = NC) = \mu^H(m = x = 0, r = C) = \frac{\pi q}{\pi q + (1 - \pi)(1 - p_L)},$$
$$\mu^H(m = 1, x = 0, r = NC) = \mu^H(m = 1, x = 0, r = C) = \frac{\pi (1 - q)}{\pi (1 - q) + (1 - \pi)(1 - p_L)},$$
$$\text{and} \quad \mu^A(m, x = 0, r) = 0, \ \forall m, r.$$

Given these beliefs, the voter never re-elects the incumbent when $x = 0$. This follows because $U_R \geq EU_V\left(reelect | m = x = 0, r = C\right)$ by assumption, and straightforward algebra

establishes that

$$\begin{aligned} &EU_V\left(reelect|m=x=0,r=C\right)=EU_V\left(reelect|m=x=0,r=NC\right)\\ &\quad>EU_V\left(reelect|m=1,x=0,r=C\right)=EU_V\left(reelect|m=1,x=0,r=NC\right). \quad (1)\end{aligned}$$

Meanwhile, the voter re-elects the incumbent when $x=m=1$ and $r=C$ by assumption. Combining shows that all incumbent types strictly prefer to choose $x=1$ rather than to choose $x=0$, so in particular type H does not exert effort, thus reaching a contradiction. The proof of the latter case is similar and is therefore omitted.

Step 2. Show that if $\tilde{p}_L=\tilde{p}_A=\frac{1}{2}$ while type H exerts effort and chooses $x=\omega$, then the voter's best responses in the two information sets considered in Step 1 ($m=x=0$, $r=C$ and $m=x=1$, $r=C$) are the same as in the original equilibrium.

There are two ways that the voter can behave symmetrically: re-electing the incumbent in both information sets, or removing her in both information sets. We focus on the former case, i.e., $U_R<EU_V\left(reelect|m=x=0,r=C\right),EU_V\left(reelect|m=x=1,r=C\right)$, since the proof of the latter case is similar.

In the original equilibrium, by Bayes' rule, voter beliefs in the above information sets are

$$\begin{aligned} \mu^H(m=x=0,r=C)&=\frac{\pi q\phi}{\pi q\phi+(1-\pi)(1-p_L)\phi+2l(1-p_A)},\\ \mu^A(m=x=0,r=C)&=\frac{2l(1-p_A)}{\pi q\phi+(1-\pi)(1-p_L)\phi+2l(1-p_A)},\\ \mu^H(m=x=1,r=C)&=\frac{\pi q\phi}{\pi q\phi+(1-\pi)p_L\phi+2lp_A},\text{ and}\\ \mu^A(m=x=1,r=C)&=\frac{2lp_A}{\pi q\phi+(1-\pi)p_L\phi+2lp_A}. \end{aligned}$$

Moreover, we have

$$\begin{aligned} EU_V\left(reelect|m=x=0,r=C\right)&=\frac{\pi q\phi-a2l(1-p_A)}{\pi q\phi+(1-\pi)(1-p_L)\phi+2l(1-p_A)}>U_R\\ \text{and }EU_V\left(reelect|m=x=1,r=C\right)&=\frac{\pi q\phi-a2lp_A}{\pi q\phi+(1-\pi)p_L\phi+2lp_A}>U_R \end{aligned}$$

by assumption, and rearranging yields

$$\pi q\phi(1-U_R)>(1-p_L)U_R(1-\pi)\phi+(1-p_A)2l(a+U_R), \quad (2)$$

and

$$\pi q\phi\left(1-U_R\right) > p_L U_R\left(1-\pi\right)\phi + p_A 2l\left(a+U_R\right). \tag{3}$$

Equations 2 and 3 imply that for any $z \in [0,1]$,

$$\begin{aligned}\pi q\phi\left(1-U_R\right) \;>\; & z\left[\left(1-p_L\right)U_R\left(1-\pi\right)\phi + \left(1-p_A\right)2l\left(a+U_R\right)\right] \\ & + \left(1-z\right)\left[p_L U_R\left(1-\pi\right)\phi + p_A 2l\left(a+U_R\right)\right].\end{aligned}$$

In particular, for $z = \frac{1}{2}$,

$$\pi q\phi\left(1-U_R\right) > \frac{1}{2}U_R\left(1-\pi\right)\phi + \frac{1}{2}2l\left(a+U_R\right),$$

which, using $\tilde{p}_L = \tilde{p}_A = \frac{1}{2}$, can be written as

$$\frac{\pi q\phi - a2l\tilde{p}_A}{\pi q\phi + \left(1-\pi\right)\tilde{p}_L\phi + 2l\tilde{p}_A} > U_R. \tag{4}$$

Now, notice that after we change p_L and p_A to $\tilde{p}_L = \tilde{p}_A = \frac{1}{2}$ while holding type H's strategy fixed, the voter's expected utilities from reelecting the incumbent, denoted by $\widetilde{EU}_V\left(reelect \mid \cdot\right)$, become

$$\begin{aligned}\widetilde{EU}_V\left(reelect \mid m=x=1, r=C\right) &= \widetilde{EU}_V\left(reelect \mid m=x=0, r=C\right) \\ &= \frac{\pi q\phi - a2l\tilde{p}_A}{\pi q\phi + \left(1-\pi\right)\tilde{p}_L\phi + 2l\tilde{p}_A}\end{aligned}$$

and are strictly greater than U_R by Equation 4. Thus in the above information sets, the voter always reelects the incumbent as in the original equilibrium, thus completing the proof of Step 2.

From Steps 1 and 2, we conclude that type A is indifferent between $x \in \{0,1\}$ given the voter's best response to $\tilde{p}_L = \tilde{p}_A = \frac{1}{2}$ and type H exerting effort and choosing $x = \omega$. That is, type A can mix after the proposed change.

For the rest of the analysis we focus on type L. We begin with three observations:

1. For type L, whether $m = x$ is independent of whether $r = C$, and neither depends on her choice of x. Thus when characterizing L's best responses we can focus exclusively on whether she is re-elected in the various information sets without worrying about the probability that each information set is reached.

2. Because the probabilities of reaching the two information sets $m = x = 0$, $r = C$ and $m = x = 1$, $r = C$ do not depend on x for type L, and Step 2 established that the voter treats the policies symmetrically in these two information sets (either re-electing in both or removing the incumbent in both), we can ignore these two information sets when analyzing type L's incentives.

3. The remainder of the proof focuses on the other six information sets: $m = x = 0, r = NC$; $m = x = 1, r = NC$; $m = 1, x = 0, r = NC$; $m = 0, x = 1, r = NC$; $m = 1, x = 0, r = C$; and $m = 0, x = 1, r = C$. Since these information sets are never reached for a type A incumbent (for whom $m = x$ and $r = C$ always), there the voter knows that the incumbent is not type A, and his beliefs about the incumbent only depend on p_L, not p_A.

With these observations in place, we continue the analysis of type L.

Step 3. Show that in the original equilibrium, the μ^H's in the above six information sets are ordered, with the voter having a higher opinion of an incumbent who chooses the policy that is less likely to be chosen by type L.

These beliefs are

$$\mu^H(m = x = 0, r = NC) = \frac{\pi q}{\pi q + (1-\pi)(1-p_L)},$$
$$\mu^H(m = x = 1, r = NC) = \frac{\pi q}{\pi q + (1-\pi) p_L},$$
$$\mu^H(m = 1, x = 0, r = NC) = \mu^H(m = 1, x = 0, r = C) = \frac{\pi (1-q)}{\pi (1-q) + (1-\pi)(1-p_L)},$$
$$\text{and } \mu^H(m = 0, x = 1, r = NC) = \mu^H(m = 0, x = 1, r = C) = \frac{\pi (1-q)}{\pi (1-q) + (1-\pi) p_L},$$

and straightforward algebra establishes that

$$\begin{aligned} & \mu^H(m = x = 0, r = NC) < \mu^H(m = x = 1, r = NC) \\ \iff \quad & \mu^H(m = 1, x = 0, r = NC) < \mu^H(m = 0, x = 1, r = NC) \\ \iff \quad & \mu^H(m = 1, x = 0, r = C) < \mu^H(m = 0, x = 1, r = C) \\ \iff \quad & p_L < \frac{1}{2}. \end{aligned} \tag{5}$$

Also note that if $p_L > \frac{1}{2}$, then the above inequalities are reversed, and if $p_L = \frac{1}{2}$ then the above inequalities become equalities. Moreover, because the voter knows the incumbent is

not type A the above beliefs directly determine his expected utilities from re-electing the incumbent. That is, $EU_V\left(reelect \mid m, x, r\right) = \mu^H\left(m, x, r\right)$ in the above six information sets.

Step 4. We now complete the proof, analyzing two cases.

Case 1. $p_L = \frac{1}{2}$. After changing p_L to $\tilde{p}_L = \frac{1}{2}$, the voter's best responses remain unchanged in the above six information sets, and because beliefs are equal in each line of Equation 5 he treats the policies symmetrically. Combined with Step 2 (which established that the voter treats the policies symmetrically in the other two information sets where $m = x$ and $r = C$), we obtain that type L is indifferent between the policies and thus can mix after the proposed change. Furthermore, the voter's strategy in the original equilibrium remains a best response, and so type H's incentives remain unchanged, too. Taken together, there exists an equilibrium with accountability in which $\tilde{p}_A = \tilde{p}_L = \frac{1}{2}$.

Case 2. $p_L \neq \frac{1}{2}$. By symmetry, consider the case of $p_L < \frac{1}{2}$, in which the voter's belief given $x = 0$ is lower than its counterpart given $x = 1$ in each line of Equation 5.

We first prove that the voter treats the policies symmetrically. If the contrary were true, then

$$\begin{aligned} U_R &\in [\mu^H(m = x = 0, r = NC), \mu^H(m = x = 1, r = NC)), \text{ or} \\ U_R &\in [\mu^H(m = 1, x = 0, r = NC), \mu^H(m = 0, x = 1, r = NC)), \text{ or} \\ U_R &\in [\mu^H(m = 1, x = 0, r = C), \mu^H(m = 0, x = 1, r = C)), \end{aligned}$$

implying that the incumbent is re-elected for $x = 1$ but not for $x = 0$ in at least one of the following three types of information sets: $m = x$ and $r = NC$, $m \neq x$ and $r = NC$, or $m \neq x$ and $r = C$. However $p_L < \frac{1}{2}$ implies that in the other two information sets, the voter's belief is strictly higher when $x = 1$ than when $x = 0$, so in those information sets he must be weakly more likely to re-elect the incumbent when $x = 1$. Thus type L's best response is to choose $x = 1$ always, which contradicts the assumption that $p_L < \frac{1}{2}$.

Thus the voter must treat the policies symmetrically in the original equilibrium, which requires that

$$\begin{aligned} U_R &\notin [\mu^H(m = x = 0, r = NC), \mu^H(m = x = 1, r = NC)), \\ U_R &\notin [\mu^H(m = 1, x = 0, r = NC), \mu^H(m = 0, x = 1, r = NC)), \text{ and} \\ U_R &\notin [\mu^H(m = 1, x = 0, r = C), \mu^H(m = 0, x = 1, r = C)). \end{aligned}$$

The voter continues to treat the policies symmetrically after we replace p_L with $\tilde{p}_L = \frac{1}{2}$

because, using ˜s to denote the new beliefs, Equation 5 implies that the new beliefs are sandwiched by the old ones:

$$\begin{aligned}
&\mu^H(m = x = 0, r = NC) < \tilde{\mu}^H(m = x = 0, r = NC) \\
&\quad = \tilde{\mu}^H(m = x = 1, r = NC) < \mu^H(m = x = 1, r = NC), \\
&\mu^H(m = 1, x = 0, r = NC) < \tilde{\mu}^H(m = 1, x = 0, r = NC) \\
&\quad = \tilde{\mu}^H(m = 0, x = 1, r = NC) < \mu^H(m = 0, x = 1, r = NC), \text{ and} \\
&\mu^H(m = 1, x = 0, r = C) < \tilde{\mu}^H(m = 1, x = 0, r = C) \\
&\quad = \tilde{\mu}^H(m = 0, x = 1, r = C) < \mu^H(m = 0, x = 1, r = C).
\end{aligned}$$

Combined with Step 2 (which established that the voter treats the policies symmetrically in the two information sets where $m = x$ and $r = C$), we obtain that type L is indifferent between the policies and thus can mix after the proposed change. The remainder of the proof is the same as in Case 1 and is therefore omitted. ■