Competitive Capture of Public Opinion*

Ricardo Alonso[†] and Gerard Padró i Miquel[‡]

December 26, 2022

Abstract

We propose a general equilibrium model of competitive media capture in which citizens with possibly heterogeneous priors watch one of multiple news channels to gather information on an underlying state of the world. Two opposite lobbyists compete to "capture" each of the news feeds in an effort to push their own agenda (one lobby to persuade citizens towards one state of the world, the other towards the alternative state of the world). We characterize the equilibrium level of capture of media outlets by competing lobbyists as well as the equilibrium level of information transmission. We show that capture polarizes expected media coverage but viewers discount extreme news. As a consequence, opposite capturing efforts do not cancel each other and instead result in a social loss in learning. At the media outlet-level, we show that capture efforts are strategic substitutes: since citizens are skeptical of messages favoring the view of the lobby that is expected to capture that channel, the incentives of the opposite lobby are naturally dampened. This force induces horizontal differentiation in coverage (slant) even when news channels are ex ante identical in quality and viewership.

Keywords: collective action, communication technology

JEL classification: D74, D72

Declarations of interest: none

This research did not receive any specific grant from funding agencies in the public or commercial sectors.

*We thank participants at Princeton, Stanford, Columbia, CEPR Media, Technology, Politics and Society Conference, NBER Political Economy and Queens' Economics of Media Workshop for useful comments. All mistakes are, of course, our own. Funding from the Whitney Betty MacMillan Centre for International and Area Studies is gratefully acknowledged.

[†]Department of Management, London School of Economics. *e-mail*: r.alonso@lse.ac.uk.

[‡]Yale University. *e-mail*: gerard.padroimiquel@yale.edu.

1 Introduction

Since public opinion over issues shapes which policies can be implemented, special interest groups (henceforth SIGs) care about public opinion. One method of influencing such opinion is to lean on media outlets to cover issues of interest in a way favorable to their long-term agendas. For example, Petrova (2008) describes how a SIG successfully lobbied media to spread the use of the term “death tax” to refer to the inheritance tax, Beattie, Durante, Knight, and Sen (2021) shows how newspaper coverage of car recalls varies as a function of car advertisement revenue, and Durante, Fabiani, Laeven, and Peydro (2021) shows that media-bank links colored media coverage of the European debt crisis. These examples of slant in coverage suggest that media, just as government, is subjected to influencing activities by SIGs who would like to secure favorable coverage, and that several opposite interest groups could be plausibly active for every given issue.

However, there are at least two marked differences between lobbying the media and lobbying a government. First, while the government is a monopoly, there are multiple media that a SIG could try to influence. Second, while the exchange between the government and a SIG is about provisions in policy which directly concern the SIG, the SIG only cares about media coverage in-so-far media coverage affects the public's beliefs.[1] This implies that a proper analysis of these relationships must take into account how media consumers update their views – the object of lobby interest – when they suspect the coverage of an issue may be tainted by SIG influence. These strategic interactions at multiple levels pose the following questions. How should SIGs try to influence the “court of public opinion” in the presence of competing SIGs and multiple media outlets? Do SIGs with opposite interests cancel each other in their influencing activities? How does SIG activity affect public learning on specific issues in equilibrium?

To make headway on these questions, we propose a general equilibrium model with two opposed SIGs, multiple (possibly heterogenous) media outlets and media consumers with heterogeneous priors over a binary state of the world. In the model, SIGs who care about the posterior beliefs of media consumers can simultaneously and covertly spend resources to capture each media's coverage

[1] Indeed, the lobbying literature has focused on quid-pro-quo exchanges of funds for policy in which the lobby directly cares about policy. See, among others, Grossman and Helpman (2001).

of an issue. Each outlet receives an informative signal on the state of the world and, if capture by either SIG fails, the media proceeds to honestly publish the signal. However, if capture is successful, the triumphant SIG can induce the media to publish *any* message. Consumers observe the published message in their chosen media outlet and rationally update their beliefs. Messages are not certifiable and there is no commitment to either the resources spent to capture media or the communication strategy of the captured media.

The model yields the following important insights. First, SIG capture leads to polarization in published messages. More specifically, extreme messages, which would be very informative in the absence of capture, increase in frequency. In contrast, centrist messages, which are less informative, are published less often. This is because the optimal manipulation strategy of a SIG is to mix over a set of favorable messages in such a way that the posterior of each consumer is equalized upon observing any message plausibly published by a captured media. For each possible message, the consumer weights two possibilities. On the one hand, how informative that message would be if the media was honest. On the other hand, what are the chances that the media outlet was captured and sent this message. If the SIG sent the message most favorable to its preferred state of the world too often, it would destroy its credibility and the consumer would discard it. The best the SIG can do is therefore to mix across a range of relatively favorable messages. Knowing this, the consumer treats each suspect message with educated skepticism.

Second, despite the fact that capture leads to the publication of more extreme messages –which are more informative if taken at face value– there is less learning in equilibrium. This is because, as noted, the possibility of capture makes rational media consumers skeptical of messages that are too favorable to each of the states: these are indeed the messages that SIGs are sending if they manage to capture the issue. This is very deleterious to learning: the messages that would lead to faster updating are the ones that are being jammed. It follows that the capturing efforts of opposed SIGs do not cancel each other: besides the waste of resources being expended on media capture, the overall informativeness of media is being damaged.

Third, at the media outlet level and due to the rational skepticism of viewers, capturing efforts by the two SIGs are strategic substitutes. The higher the effort exerted by one SIG on a given media outlet, the more skeptical are consumers of that media when they observe messages favorable to that SIG. This dampens the effect of capture and therefore reduces the marginal benefit of capture for the other SIG. Because of this fundamental force, the model supports equilibria with horizontal differentiation: two ex ante identical media outlets with identical viewership may, in equilibrium, not be symmetrically captured. In fact, under some additional conditions we can construct equilibria with full differentiation: each media is captured with positive probability by only one SIG. We thus provide a novel reason for horizontal differentiation in media slant which does not rely on demand effects or on differences across media outlets such as owner ideology.

Fourth, demand for information leads to partial sorting of viewers across media. Under general conditions, viewers that have low priors will sort into media most likely captured by the SIG who desires low posteriors, and the same is true at the other end of the distribution of priors. Centrists, however, may sort non-monotonically.[2] Interestingly, viewer sorting across media does not necessarily mean more horizontal differentiation in media slant. This depends on the relative importance that SIGs place in influencing different demographics of the public, an intuitive feature that emerges in this model with heterogeneous priors. In particular, the curvature of SIG preferences over posteriors determines whether soothing opponents (e.g. reaching viewers with opposite priors) is more important than firing up the base (e.g. reaching viewers who have priors aligned with the SIG). Only when firing up the base is the overriding concern of the SIGs does viewer sorting lead to horizontal differentiation in slant. Finally, we also show that, paradoxically, higher demand for information may lead to a less informative media landscape as greater demand may increase incentives to capture, with deleterious consequences for learning.

In sum, this general equilibrium model supports equilibria with several empirically appealing features. First, coverage of an issue is systematically different across media, and media publishes lies along the equilibrium path. Second, rational viewers largely sort according to priors, but nonetheless are skeptical of the news they consume. In this sense, despite the lies, consumers are not systematically deceived. This is consistent with evidence that while viewers seek outlets with which they are ideologically aligned, they often question the veracity of news and do not update

[2] This contrasts with Suen (2004) where, in the context of advice which coarsens information into two possible recommendations, preferences over advisors are monotonic in the prior and preferences of decision-makers.

according to news' literal meaning. Gentzkow and Shapiro (2010) show robust alignment between a media outlets slant and their viewership. Angelucci and Prat (2021) find that most viewers are able to identify fake political news. Martin and Yurukoglu (2017) find that cable news have progressively polarized in terms of coverage but that ideological polarization in the population is proportionally much smaller, which is in line with existing research in political science (Ansolabehere, Rodden, and Snyder, 2006). In a recent experiment, Brookman and Kalla (2022) show that partisans forced to watch media with opposite slant moderately revise their views but do not fundamentally change their partisan affiliation or presidential vote, and return to their partisan media as soon as the experiment concludes.[3] These findings are all aligned with the equilibrium structure of our model.

The theoretical literature on the political economy of media capture has advanced dramatically in recent decades.[4] The incumbent government is the primordial example of a capturing agent, as shown in most detail in McMillan and Zoido (2004). Models of government capture focus therefore on the case with a single special interest group. Besley and Prat (2006) relies on a disclosure game where printed news are never lies.[5] In Gehlbach and Sonin (2014) commitment to an editorial line means media filter information, but do not distort it.[6] Similarly, Petrova (2008) focuses on capture by a single social group –the rich– and assumes exogenous costs of lying by the media. Corneo (2006), in contrast, offers a model with multiple SIG potentially capturing a single media outlet. These existing models consider viewers with homogeneous priors and limit the message space to a binary signal. We advance on the literature by considering SIGs with opposing interests, which influence multiple media outlets consumed by viewers with heterogeneous priors.[7] In addition, we put no restrictions on the message space and assume no commitment to a publishing rule.[8] These features allow us to have predictions on i. differential capture across media by the different SIG

[3] Gentzkow, Shapiro, and Sinkinson (2011) similarly find that the introduction of partisan newspapers did not affect party vote shares.

[4] For a theoretical survey see Prat (2015)

[5] See Milgrom (1981), Dye (1985), Milgrom and Roberts (1986), and Okuno-Fujiwara et al (1990) for certifiable disclosure of private information.

[6] See Gentzkow and Kamenica (2011) for the analysis of information transmission when the sender can commit to the disclosure rule, and Bergemann and Morris (2019) for a survey of this class of models and their applications. Gitmez and Molavi (2022) also follows this modeling tradition and considers heterogeneous receivers but a single sender.

[7] To our knowledge, Petrova (2012) is the only previously existing model with multiple lobbyists and media outlets. However, it is not a model with information transmission.

[8] We model communication by the SIGs as cheap talk. See Crawford and Sobel (1982) for the seminal contribution and Sobel (2013) for a recent survey.

ii. the polarizing effects of media capture on published news and iii. the resulting compression of viewers' beliefs.

The equilibrium co-existence of media with different slants has been justified in one of three ways in the literature.[9] First, suppliers such as owners or journalists may have different ideologies which they are trying to push on the population (i.a. Baron, 2006; Anderson and McLaren, 2012). Second, rational demand for news by viewers with heterogeneous priors or ideology can lead to a segmented market (i.a. Chan and Suen, 2008; Gentzkow and Shapiro, 2006; Sobbrio, 2014). Finally, demand effects may also be due to cognitive biases or other ideological effects on consumer demand (i.a. Gabszewicz, Laussel, and Sonnac, 2001; Mullainathan and Shleifer, 2005; Bernhardt, Krasa, and Polborn, 2008). In our model, opposite slants result even if media are *ex ante* identical and there are no demand effects. The novel reason is the strategic substitutability of capturing efforts by SIG at the media outlet level. Holding everything else constant, an SIG prefers to direct capturing effort to the outlet that is less likely to be captured by the opposing SIG.

We also contribute to the literature on communication by a sender with uncertain motives.[10] Sobel (1985) shows how a biased sender can maintain a reputation for honesty when an honest sender always tells the truth. In our setup, the honest media also relays the truth to the public, although capturing SIGs do not have an incentive to build a reputation for honesty. Morgan and Stoken (2003) and Li and Madarasz (2008) show that information transmission may be reduced if the sender discloses his preferences. In our model, however, knowing the identity of the sender would always lead to (weakly) more informative media. Thus, in our setup concealment of motives reduces information transmission but incentivizes capture. Finally, Wolinsky (2003) and Dziuda (2011) study models with partial verifiability: the sender may be biased in favor or against a given issue, but can only conceal evidence, not fabricate them. Interestingly, their equilibria also feature receivers' skepticism towards extreme views and a constant informational content of extreme messages.[11] We obtain a similar equilibrium structure despite the fact that in our model

[9] For a theoretical survey see Gentzkow, Shapiro, and Stone (2015)

[10] See also Shin (1993) and Morris (2001).

[11] In Wolinsky (2003), the sender can underreport the state but never overreport it. In equilibrium, any message above a threshold is fully trusted, while messages below that threshold lead to the same posterior. This equilibrium

SIGs are free to fabricate the news.

The rest of the paper is organized as follow. Section 2 sets out the model. Section 3 describes the optimal publishing strategy of SIGs and its effects on news distribution and information transmission. Section 4 studies incentives to capture a media monopoly, shows that capturing efforts are strategic substitutes and explores how the heterogeneous priors of viewers affect capturing incentives. Section 5 introduces multiple media and shows that the model supports full horizontal differentiation. Finally, Section 6 explores the implications of viewership sorting across media. We then offer some conclusions.

2 Model

We propose the following general equilibrium model of the market for news. There are $n \geq 1$ media outlets which may differ in their informativeness about an underlying state of the world. There are two SIGs with opposed preferences over consumers' perceptions as informed by the media's coverage of a newsworthy issue. For example, the underlying state of the world may be the gravity of the climate crisis and the newsworthy issue may be a sequence of weather events. Carbon-dependent energy companies may want to downplay the connection of these events with global warming, while ecologists may want to highlight it. These SIGs can covertly devote resources to capture the coverage of the newsworthy issue, separately for each outlet. Viewers select which media to consume according to some news-independent preference (entertainment value) or because of its instrumental value (to learn about the state of the world), and discount the news they consume according to the anticipated level of capture.

State space and Prior Beliefs: There is a newsworthy issue which is informative of an unknown binary state $\theta \in \Theta = \{-1, 1\}$. Viewers have heterogeneous prior beliefs $p = \Pr[\theta = 1]$ over the state, with a mass $F_p(p)$ of citizens with priors not exceeding p and $M = \int_0^1 dF_p(p)$ the total mass of media consumers.

Special Interest Groups and Media Outlets: There are two strategic SIGs, R and L with opposed

is similar to our model in which only the L-SIG engages in capture. In Dziuda (2011), the sender privately obtains several arguments in favor or against an issue and can conceal arguments. For the case of a single type of bias, equilibria exhibit, as in our model, receiver's skepticism when a small number of arguments either in favor or against are presented.

preferences. Specifically, R wants to induce in viewers the highest posterior belief over θ while L wants to induce the lowest. If μ is the posterior belief of a citizen, then the SIGs utility functions are $v_R(\mu)$ and $v_L(\mu)$ with v_R strictly increasing and v_L strictly decreasing with $|v_i'|$, $i \in \{L, R\}$ bounded away from zero. Thus, if $\mu(m;p)$ is the posterior belief of a citizen with prior p after observing message m, then the indirect utility over messages of an i- SIG, $i \in \{R, L\}$, when faced with a viewership characterized by $F_p(p)$, is

$$V_i(m) \equiv \int_0^1 v_i\left(\mu(m;p)\right) dF_p(p). \tag{1}$$

There are $n \geq 1$ different media outlets, whose coverage of an issue can be captured by a SIG. Media outlets function as a Blackwell-experiment: they observe an informative signal $m \in \mathscr{M} \subset \mathbb{R}$ which is generated according to $\Pr\left[m|\theta = i\right] = p_i^j(m)$, $i \in \{-1, 1\}$, with $j \in \{1, ...n\}$ indexing media outlets. If coverage of media j is *not* captured by either SIG we say that the coverage is *honest* and in this case the media outlet simply publishes –i.e., truthfully conveys to their viewers– the signal it observes. Thus, the posterior belief of a $p-$viewer after observing message m from media j which is known to have an honest coverage of the issue is

$$\mu_H(m;p) = \Pr\left[\theta = 1|m, H, p\right] = \frac{p_1^j(m)p}{p_1^j(m)p + p_{-1}^j(m)(1-p)}. \tag{2}$$

Without loss of generality in this binary-state case, we order messages according to the likelihood ratio $\lambda_H^j(m) = \frac{p_1^j(m)}{p_{-1}^j(m)}$ (so that $\lambda_H^j(m)$ is increasing). Following this convention, we will say that a message m is higher (lower) when viewers would update more towards state 1 (-1) should that message be published in an outlet j whose coverage of the issue is known to be honest. To characterize the informativeness of an honest media, we let $F_{H,\theta}^j(\lambda) \equiv \Pr[\lambda_H^j(m) \leq \lambda|\theta]$.

Competitive Media Capture: For each media j, SIGs simultaneously and covertly decide how much effort to expend in capturing coverage of the newsworthy issue. These efforts determine whether the coverage remains honest or whether it is captured by one of the SIG. We denote these possible states of capture by $S^j \in \{R, L, H\}$. More specifically, if r^j and l^j are the efforts expended by R and L on capturing coverage on media j, then we have $r^j = \Pr\left[S^j = R\right]$, $l^j = \Pr\left[S^j = L\right]$

and $1 - l^j - r^j = \Pr[S^j = H]$. In other words, we assume that the SIGs are locked in a linear contest function to capture the coverage of the issue for each media outlet.[12] Capturing effort takes resources: the total cost of capture for the R-SIG is $C_R(\sum_{j=1}^n \beta_R^j r^j)$ and for the L-sender is $C_L(\sum_{j=1}^n \beta_L^j l^j)$ with C_R and C_L satisfying standard Inada conditions.[13] For most of our results we will look at the case with no differences in the cost of capture across media channels so that we set $\beta_i^j = 1, i \in \{L, R\}, j \in 1, ...n$.

If coverage by media j is captured by either SIG, then the capturing SIG can have the media publish *any* message m.[14] We assume the message space is independent of state of capture or state of the world so there is no restriction on the message m a captured media can publish. We allow SIGs to follow mixed strategies in deciding which messages to publish.

Media Consumption: viewers choose one of the n available media to consume.[15] We assume that each viewer chooses according to a media outlet's exogenous entertainment value with probability $1 - \gamma$. Entertainment value of media j to a viewer of prior p is $\vartheta^j(p)$. With probability γ the viewer chooses which media to consume in order to acquire information about the state of the world, taking into account the anticipated endogenous levels of capturing effort. We micro-found in Section 6 the decision problem that induces the viewer's demand for information.

Timing: Simultaneously, SIGs R and L decide on $r^j, j = 1...n$ and $l^j, j = 1...n$ and viewers choose which media to consume. Then, nature selects $S^j \in \{R, L, H\}$ according to r^j and l^j but S^j is not observed by consumers. For media j such that $S^j = R$ ($S^j = L$) SIG-R (SIG-L) decides which message to publish. Viewers then observe the message published by their chosen media and update their beliefs. After this, payoffs are realized. We look for a Perfect Bayesian Equilibrium of this capture and communication game.

This model displays a few noteworthy features. First, the model focuses on the competition between SIGs and the inference problem it induces on rational consumers of media. To simplify the analysis and highlight new insights, we model media outlets as passive subjects of pressure

[12] The linear contest function greatly simplifies the exposition but it is not essential, as shown in Proposition 2.

[13] In particular, we assume $C_i'(x) > 0, C_i''(x) > 0$, and $\lim_{x \to 0} C'(x) = 0$ and $\lim_{x \to 1} C'(x) = \infty$.

[14] For simplicity, we assume that the choice of message by a successful SIG is independent of media $j's$ realized signal. As we show in the online Appendix, conditioning on the realized signal does not change the distribution of viewers' posterior beliefs, nor the equilibrium capture efforts, but increases the notational burden.

[15] This single homing assumption is widespread in the literature on media bias. See, for example Gentzkow and Shapiro (2006), Chan and Suen (2008) and Duggan and Martinelli (2011) among many others.

from SIGs and sidestep the media owner trade-off between audience and bias which is already well-understood in the literature. Second, we model the capture of the coverage of a specific news item and not the full capture of a media outlet. For this reason, it makes sense to fix the number of existing outlets. However, the model can perfectly accommodate the fact that some outlets more frequently bias coverage towards one end of the ideological spectrum. For example, Fox News can be conceptualized as having low β_R^{FOX} so that it is cheap for the R-SIG to capture coverage at FOX. This is known by viewers, who take this into account when updating their beliefs. These consumers are asking themselves: "is FOX's coverage of this ongoing weather event honestly reporting possible links with global warming or has it (again) been compromised by the R-SIG?"

Third, messages m have an *accepted meaning* in our model, following the terminology of Sobel (2020).[16] In particular, everyone agrees how message m is to be interpreted –that is, how priors are to be updated– if a message m is published in a media outlet which is known to be honest. The shadow of capture, however, drives a wedge between m's accepted meaning and m's interpretation *in equilibrium*. This allows us to separately keep track of published news –i.e. equilibrium m – and the effect of such news –i.e. equilibrium viewers' posteriors. This is important because, empirically, slant is reflected in m, not necessarily on viewer posteriors.

Finally, we impose no restrictions on the message space of captured media. More specifically, messages are not certifiable and there is no ex ante commitment to any communication strategy. In this sense we have a genuine model of lying in which capturing SIGs can have media manufacture fake news at will, completely untethered to the true state of the world.

3 Communication Equilibria of a Captured Media Outlet

We start our analysis by characterizing communication equilibria for a given media outlet when the capturing efforts l and r are common-knowledge. In this section, we drop the subscript j to characterize a given outlet and we also assume that $\gamma = 0$ so we take as given the viewership of this outlet. Let $F_p(p)$ denote the mass of viewers with a prior at most p watching the media outlet. We have three main insights: first, potential media capture leads to a polarization of expected

[16] Sobel (2020) defines lies as statements whose accepted meaning is different from what the sender knows. Media do lie along the equilibrium path in our model.

published news. Second, rational viewers discount the informativeness of such extreme messages. Third, as a consequence, potential capture is very deleterious to social learning.

3.1 Optimal Lying, Optimal Skepticism

Recall that if a media remains honest, it publishes a message m which is informative about the state of the world θ. The likelihood ratio $\lambda_H(m) = p_1(m)/p_{-1}(m)$ captures the informational content of this message from a known honest source and it is sufficient to compute the posterior of a viewer with any prior p according to (2). If there is no media capture, therefore, viewers interpret each message according to its accepted meaning, although viewers with different priors will typically reach different posteriors–i.e, viewers agree on what message m means (i.e. agree on $\lambda_H(m)$) but differ on the conclusions they draw about the underlying state of the world.

The media outlet, however, is only honest with probability $1-l-r$. When it is captured, m is generated according to the interests of the capturing SIG. Consequently, viewers need to modify the way they update their beliefs upon observing m. In a communication equilibria, the lying strategies of SIGs and the updating process of viewers are consistent with each other. To describe communication equilibria, let $\tau_R^*(m)$ and $\tau_L^*(m)$ be the $R-$SIG and $L-$SIG equilibrium (mixed) strategies. These specify the probability of selecting message m if they successfully capture the media outlet. Let $\mu^*(m;p)$ be the posterior belief of a media consumer with prior p after observing m consistent with strategies $\tau_R^*(m)$ and $\tau_L^*(m)$. Then, for $i \in \{L, R\}$ the $i-$SIG selected message maximizes $V_i(m) = \int v_i(\mu^*(m;p))dF_p(p)$.

The following proposition shows that equilibrium behavior takes a simple form: mixing by SIG equalizes the *equilibrium informational content* of the highest (lowest) messages.

Proposition 1. *Consider a single media outlet and fix levels of capture r and l, with $r + l <$ 1. There are unique $\overline{\lambda}$, $\underline{\lambda}$, $\overline{m}^*$, and $\underline{m}^*$, with $\overline{\lambda} = \lambda_H(\overline{m}^*)$ and $\underline{\lambda} = \lambda_H(\underline{m}^*)$, so that for every communication equilibrium, with $\tau_R^*(m)$ and $\tau_L^*(m)$ the SIGs' equilibrium (mixed) strategies, we have*

1. *$m \in supp(\tau_R^*)$ iff $\lambda_H(m) \geq \overline{\lambda}$; $m \in supp(\tau_L^*)$ iff $\lambda_H(m) \leq \underline{\lambda}$.*

2. *The equilibrium likelihood ratio of message* m, $\lambda^*(m) \equiv \frac{\Pr[m|\theta=1]}{\Pr[m|\theta=-1]}$, *satisfies*

$$\lambda^*(m) = \begin{cases} \underline{\lambda} & \text{if } m \leq \underline{m}^* \\ \lambda_H(m) & \text{if } \underline{m}^* < m < \overline{m}^*. \\ \overline{\lambda} & \text{if } m \geq \overline{m}^* \end{cases} \tag{3}$$

3. *The maximum and minimum likelihood ratios* $\overline{\lambda} = \max_{m \in \mathscr{M}} \lambda^*(m)$ *and* $\underline{\lambda} = \min_{m \in \mathscr{M}} \lambda^*(m)$ *satisfy*

$$\int_{\overline{\lambda}}^{\infty} \left(\lambda - \overline{\lambda}\right) dF_{H,-1}(\lambda) = \frac{r}{1-l-r}\left(\overline{\lambda} - 1\right), \tag{4}$$

$$\int_{0}^{\underline{\lambda}} \left(\underline{\lambda} - \lambda\right) dF_{H,-1}(\lambda) = \frac{l}{1-l-r}\left(1 - \underline{\lambda}\right). \tag{5}$$

Part 1 of the proposition states that the $R-$SIG randomizes over a set of messages with $\lambda_H(m)$ above a threshold informativeness $\overline{\lambda}$. These are messages that would be very informative and induce higher posteriors, if sent by a media known to be honest. Part 2 describes how viewers update. For all messages possibly sent by $R-$SIG, instead of updating according to $\lambda_H(m)$, viewers just use $\overline{\lambda}$. This has two implications. First, since $\overline{\lambda} \leq \lambda_H(m)$ for $m \in supp(\tau_R^*)$, the informational content of these messages is downgraded: because the outlet is possibly captured by $R-$SIG, viewers are skeptical of messages that are favorable to $\theta = 1$. Second, all such messages are treated identically since $\lambda^*(m) = \overline{\lambda}$, a constant. This means that for messages that are more favorable to $\theta = 1$, viewers are relatively more skeptical.

Of course, the same is true at the other end of the distribution. The $L-$SIG randomizes over a set of messages favorable to state $\theta = -1$ and viewers, skeptical of such messages, treat them all as $\underline{\lambda} \geq \lambda_H(m)$. Again, they downgrade the informational content of messages below $\underline{\lambda}$ and do so more the more such messages are favorable to $\theta = -1$.

The effect of potential capture is therefore to make viewers skeptical of messages that would otherwise be very informative –i.e. either very high or very low $\lambda_H(m)$. Skepticism is well-founded because very informative messages are potential lies –i.e., $\Pr[S = H|m] < 1$ for such m. Moderate messages $m \in (\underline{m}^*, \overline{m}^*)$ are instead taken at face value. Upon observing them, a viewer can infer that the outlet is not captured and updates to $\mu^*(m;p)= \mu_H(m;p)$. The proposition thus implies that $\mu^*(m;p)$ is a two-sided censored distribution of posterior beliefs for every $p-$viewer.

Part 3 of Proposition 1 characterizes the unique $\overline{\lambda}$ and $\underline{\lambda}$ induced by a $\{l, r\}$ configuration. To build intuition note that, given a fixed level of media capture, Bayesian updating requires that the equilibrium posterior beliefs of a $p-$viewer must average to the prior. Hence

$$(1-l-r)\,\mathbb{E}_H\left[\mu^*(m;p);p\right] + l\mu^*(\underline{m}^*;p) + r\mu^*(\overline{m}^*;p) = p.$$

Given the two-sided censored nature of $\mu^*(m;p)$ and mixing behavior by, say, the $R-$SIG we thus have

$$(1-l-r)\int_{\overline{m}^*}^{1}\left(\mu_H(m;p)-\mu_H(\overline{m}^*;p)\right)dF_H(m;p) = r\left(\mu_H(\overline{m}^*;p)-p\right). \tag{6}$$

where $F_H(m;p) = pF_{H,1}(m) + (1-p)F_{H,-1}(m)$ is the distribution of messages that a $p-$viewer expects from an honest media. The left hand side of (6) represents the expected downward distortion in beliefs from messages of an honest sender when viewers fear that the message may be captured –i.e., any $m \geq \overline{m}^*$ is suspected to come from an $R-$SIG. Conversely, the right hand side of (6) is the upward distortion by the $R-$SIG who systematically sends "high" news. In equilibrium, the two distortions cancel each other, which determines $\overline{m}^*$. Then, expressing (6) in terms of likelihood ratios we obtain (4). A similar reasoning applied to the $L-$SIG leads to (5). Finally, (4) is independent of $\underline{\lambda}$ and its right hand side decreases in $\overline{\lambda}$ while its left hand side strictly increases in $\overline{\lambda}$. This implies that the solution to (4) is unique. The same argument applied to (5) yields a unique $\underline{\lambda}$. In sum, the fact that no rational viewer can be fooled in expectation, uniquely determines $\overline{\lambda}$ and $\underline{\lambda}$.

3.2 Published Lies by Captured Media

An interesting feature of this model is that we have predictions of the effect of capture on the (continuous) distribution of messages published by media. When the media is known to be honest, a viewer with prior p expects each message m according to distribution $F_H(m;p)$. When the media is captured, however, the expected frequency of messages is influenced by $\tau_R^*(m)$ and $\tau_L^*(m)$. To understand how the SIGs send messages in equilibrium, note that the $R-$SIG cannot afford to exclusively send the most favorable message, which would be the highest m available. If it did so, viewers would fully discount that specific message as being the result of manipulation and would update very little. Given this, the $R-$SIG could profitably deviate to sending another slightly lower message $m' = m - \epsilon$, which would induce full updating as viewers would trust that such a message could only be sent by honest media. In equilibrium it must therefore be the case that viewers have the same interpretation (i.e. assign the same informational content) for all messages sent by a given SIG.

To achieve this equalization of informational content the equilibrium mixed strategy of the SIG distributes the probability of lying for each m in such a way that the equilibrium likelihood ratio is equalized. Note that the equilibrium likelihood ratio for a message $m \in supp(\tau_R^*)$ sent with positive probability by the $R-$SIG is

$$\lambda^*(m) = \frac{(1-r)p_1(m) + r\tau_R^*(m)}{(1-r)p_{-1}(m) + r\tau_R^*(m)}$$

and this expression is decreasing in $\tau_R^*(m)$: the more often a message m expected to be sent by the $R-$SIG, the less informational content that message is perceived to have. Equalizing $\lambda(m)$ across the various $m \in supp(\tau_R^*)$ thus implies spreading $\tau_R^*(m)$ across messages in a very specific way that fully characterizes the optimal lying strategy of SIG.

Lemma 1. *In every communication equilibria described in Proposition 1 we have that for every two messages* $m \in supp(\tau_R^*)$ and $m' \in supp(\tau_R^*)$,

$$\begin{aligned}\tau_R^*(m')/\tau_R^*(m) &= \left(\lambda_H(m') - \overline{\lambda}\right) p_{-1}(m') / \left(\lambda_H(m) - \overline{\lambda}\right) p_{-1}(m) \\ &= \left(p_1(m') - \overline{\lambda} p_{-1}(m')\right) / \left(p_1(m) - \overline{\lambda} p_{-1}(m)\right).\end{aligned}$$

Equivalently, we have that for every two messages $m \in supp(\tau_L^*)$ and $m' \in supp(\tau_L^*)$,

$$\tau_L^*(m')/\tau_L^*(m) = \left(\underline{\lambda} - \lambda_H(m')\right) p_{-1}(m') / \left(\underline{\lambda} - \lambda_H(m)\right) p_{-1}(m).$$

For instance, if $p_{-1}(m)$ weakly decreases in m while $p_1(m)$ increases in m, then mixing by

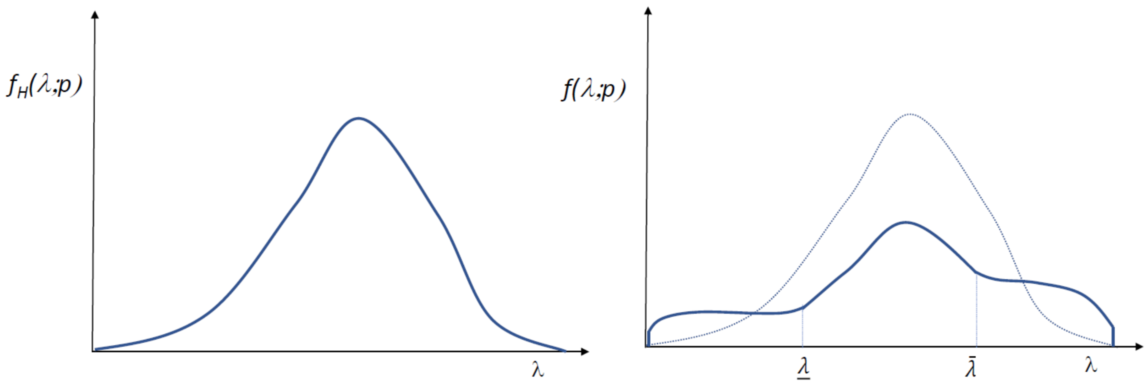

Figure 1: Content of Captured Media

SIGs must put more weight to relatively more informative messages in order to equalize their informational content. In other words, under such conditions, in equilibrium both SIGs send the most extreme messages relatively more often than any other message. Because these messages are more *ex ante* informative, the SIG can assign to them a higher probability of lying.

Figure 1 depicts what happens to the expected distribution of published messages. Compared to the distribution expected from an honest media (drawn in the first panel) captured messages are more polarized as the mass moves towards the extreme messages that the SIGs induce with high probability.

3.3 Informativeness of a Captured Media Outlet

The prior discussion shows that capture affects informativeness by changing the distribution of likelihood ratios of the news published by the outlet. Using (3) in Proposition 1, we have that the distribution of likelihood ratios for a p-viewer is

$$F(\lambda;p) = \begin{cases} 0 & \text{if } \lambda < \underline{\lambda}, \\ l + (1-r-l)F_H(\lambda;p) & \text{if } \underline{\lambda} \leq \lambda < \overline{\lambda}, \\ 1 & \text{if } \lambda \geq \overline{\lambda}. \end{cases} \tag{7}$$

The specter of capture decreases the likelihood that a viewer revises her beliefs to entertain a very high or very low view of the world even when the message is truthful: optimal lying limits the informational content of each message to $\lambda^*(m) \in [\underline{\lambda}, \overline{\lambda}]$. As a consequence, capture reduces the

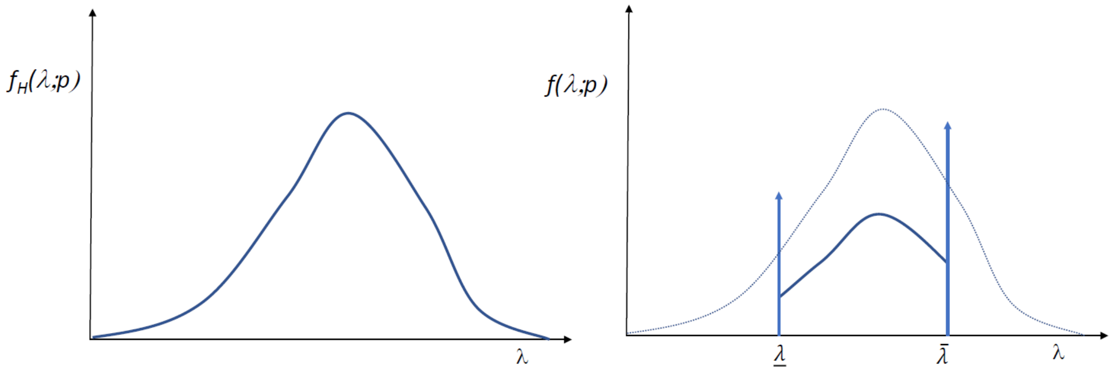

Figure 2: Informational Content with Honest and Captured Media

Blackwell-informativeness of the channel: $F(\lambda; p)$ second-order stochastically dominates $F_H(\lambda; p)$. This reduction in informativeness operates through two channels. First, it limits the informativeness of very informative messages to either $\lambda_H(\underline{m}^*) = \overline{\lambda}$ or $\lambda_H(\overline{m}^*) = \underline{\lambda}$. Second, it reduces the likelihood that a message $m \in (\underline{m}^*, \overline{m}^*)$ is observed. These two effects are depicted in Figure 2. The left hand panel shows the expected distribution of likelihood ratios associated with messages from an honest outlet, while the right hand side shows the expected distribution when there is possible capture by both SIGs. Note the mutually-supporting contrast between Figure 1 and Figure 2. While news become polarized because of SIG interference, beliefs become compressed due to the skepticism generated by this interference. Just as reported in Angelucci and Prat (2021), viewers discount fake news.

We now describe how these bounds on informativeness change with exogenous changes in the parameters of the model. We show that (i) increasing capture by either SIG exacerbates viewers' skepticism over the informative of news at *both* ends of the spectrum; (ii) viewers' prior distribution F_p does not affect the equilibrium informational content of the media outlet; and (iii) viewers entertain more extreme views when the internal news report is Blackwell more informative–i.e., when honest media is more informative, equilibrium lies more successfully influence viewers.

Lemma 2. *Let $\overline{\lambda}$, $\overline{m}^*$, $\underline{\lambda}$ and $\underline{m}^*$ be the equilibrium quantities defined in Proposition 1. Then,*

1. *$\overline{\lambda}$ and $\overline{m}^*$ are decreasing in l and r, while $\underline{\lambda}$ and $\underline{m}^*$ are increasing in l and r.*

2. *$\overline{\lambda}$, $\overline{m}^*$, $\underline{\lambda}$ and $\underline{m}^*$ are invariant in F_p.*

3. *$\overline{\lambda}$ increases and $\underline{\lambda}$ decreases if the honest media is Blackwell more informative.*

Lemma 2.1 shows that a viewer is more skeptical following any increase in media capture, as it lowers the maximum (and raises the minimum) belief that she might entertain and reduces the number of messages that she will trust. Importantly, more intense capture by, say, an $R-$SIG leads viewers to trust less "favorable" reports that the state is high, but also to trust less reports that the state is low. The first effect is clear as more intense capture makes it more likely that high messages are generated by an $R-$SIG. However, an increase in capture of an $R-$SIG also makes it less likely that the outlet is honest. Therefore, upon viewing a low m, rational consumers must place higher probability that the $L-$SIG prevailed.

Lemma 2.2 shows that the equilibrium informativeness of a media outlet is invariant to its viewership given l and r. This is because, as it is clear Lemma 1, in equilibrium mixing equalizes the informational content of each potential message sent. As a consequence the informational content only depends on properties of the honest media, and not on the priors of the viewers. In short, the optimal lies of a SIG are independent of who is watching the media outlet. Viewership, however, will affect incentives to capture, as we show below.

Finally, lemma 2.3 shows that captured media can afford to send more extreme messages if the honest sender is more informative.[17] This result follows readily from a higher dispersion of posterior beliefs generated from a Blackwell-more informative credible sender and its effect on equilibrium conditions (4) and (5). Intuitively, when an honest media is more informative, a given amount of lying has a smaller effect as viewers discounting of messages starts from a higher baseline.

4 Competitive Capture of a Monopoly Media Outlet

Having established the effects of capture on messages, we now turn to the equilibrium level of capture of a monopoly media outlet by competing SIGs. As viewers have no other alternative to become informed, this case will help us understand media capture in the absence of demand-side effects coming from endogenous viewership; that is, equilibrium capture is solely driven by supply-side considerations. We have two main insights. First, each SIG's marginal gain from capture is reduced when viewers expect a higher level of capture by the opposing SIG: capture efforts are

[17] Note that we cannot say how this will change the messages that citizens trust as we impose no structure on the message space of a Blackwell more-informative source.

strategic substitutes. Second, whether capturing is about firing up the base or demobilizing the opposition determines how the distribution of viewers' priors shapes incentives to capture.

4.1 Incentives to Capture Coverage

Our model is one of covert capture: SIGs privately deploy resources to capture the coverage of an issue so as to influence the beliefs of uninformed, but suspecting, viewers. To understand incentives to capture, the equilibrium likelihood ratio $\lambda^*(m)$ suffices to characterize the distribution of viewers' posterior beliefs–see Proposition 1. By expressing each viewer's equilibrium posterior as

$$\mu^*(\lambda; p) = \frac{\lambda p}{\lambda p + 1 - p},$$

we can write the expected value to the $i-$SIG from sending a message m such that $\lambda^*(m) = \lambda$ as

$$V_i(\lambda) \equiv \int_0^1 v_i\left(\mu^*(\lambda; p)\right) dF_p(p) = \int_0^1 v_i\left(\frac{\lambda p}{\lambda p + 1 - p}\right) dF_p(p). \tag{8}$$

Note that this expression varies with the message sent –through its associated λ– and it also depends on the priors of the viewership of the outlet –through $F_p(p)$. What is the benefit of covertly capturing a media outlet? Suppose that viewers suspect a level of capture $\left(\tilde{r}, \tilde{l}\right)$ and believe that the informational content of message m is $\lambda^*(m)$. In our basic model, increasing capture reduces the likelihood that the message originates from an honest media. Then the marginal gain to, say, $R-$SIG from covert capture is

$$\begin{aligned} B_R(r; \tilde{r}, \tilde{l}) = V_R\left(\overline{\lambda}\right) - \mathbb{E}_H\left[V_R\left(\lambda\right); p_R\right] = \int_{\underline{\lambda}}^{\overline{\lambda}} \left(V_R(\overline{\lambda}) - V_R(\lambda)\right) dF_H(\lambda; p_R) \\ + \left(V_R(\overline{\lambda}) - V_R(\underline{\lambda})\right) F_H(\underline{\lambda}; p_R) = \int_{\underline{\lambda}}^{\overline{\lambda}} V_R'(\lambda) F_H(\lambda; p_R) d\lambda, \end{aligned} \tag{9}$$

as she does, upon capture, replace the expected message the honest source generates and sends instead a message inducing the highest credible likelihood $\overline{\lambda}$. Note that the priors of the viewership of the channel matter through $V_R'(\lambda)$, and the $R-$SIG evaluates the distribution of messages from an honest source according to her own prior belief p_R. Therefore, $B_R(r; \tilde{r}, \tilde{l})$ depends both on the prior beliefs of viewers and on the prior belief of the $R-$SIG. We can equivalently compute the marginal gain to the $L-$SIG as

$$B_L(l;\tilde{r},\tilde{l}) = V_L\left(\underline{\lambda}\right) - \mathbb{E}_H\left[V_L\left(\lambda\right);p_L\right] = \int_{\underline{\lambda}}^{\overline{\lambda}} \left(-V_L'(\lambda)\right)\overline{F}_H(\lambda;p_L)d\lambda, \tag{10}$$

where $\overline{F}_H(\lambda;p) = 1 - F_H(\lambda;p)$.

4.1.1 Capturing Efforts are Strategic Substitutes

How do incentives to capture change if the other SIG raises its effort? Our linear probability model implies that the marginal gain is independent of the actual level of capture, and only depends on viewers' anticipated level of capture through its effect on $\overline{\lambda}$ and $\underline{\lambda}$. Then, differentiating (9) and using Lemma 2.1 we have

$$\frac{\partial B_R(r;\tilde{r},\tilde{l})}{\partial \tilde{l}} = V_R'(\overline{\lambda})F_H(\overline{\lambda};p_R)\frac{\partial \overline{\lambda}}{\partial \tilde{l}} - V_R'(\underline{\lambda})F_H(\underline{\lambda};p_R)\frac{\partial \underline{\lambda}}{\partial \tilde{l}} \leq 0.$$

This is one of our key insights: influence efforts are strategic substitutes. The intuition is powerful: increasing $\tilde{l}$ has a double dampening effect on the incentives of the $R-$SIG. On the one hand, viewers of the channel, anticipating higher effort by the $L-$SIG become more skeptical of low messages and discount them more, even when coming from an honest sender, thereby reducing the gains from capture by the $R-$SIG. This is captured by $\partial\underline{\lambda}/\partial\tilde{l} > 0$. On the other hand, higher $\tilde{l}$ makes viewers more skeptical of high messages as well, as the coverage is less likely to be honest. This effect is captured by $\partial\overline{\lambda}/\partial\tilde{l} < 0$. Hence this result is an intuitive and direct corollary of Lemma 2.1. The same argument, of course, applies to the $L-$SIG.

Note that we have assumed a linear contest function in which an increase in covert capture by one SIG does not crowd out influence by the other, i.e., increasing the probability that the $R-$SIG generates the message does not reduce that of an $L-$SIG. This assumption certainly strengthens the second effect which leads to $\frac{\partial\overline{\lambda}}{\partial\tilde{l}} < 0$. However, this effect is not exclusive of this formulation: strategic substitutability of influence efforts holds under more general conditions as noted in the following proposition.

Proposition 2. *Let $\pi_i(r,l)$ be the probability that nature selects $S = i,\ i \in \{R, L\}$, given capture levels r and l, and suppose that increasing i's effort weakly decreases the probability of capture by j and the probability that the media remains honest. If*

$$\frac{\partial^2 \pi_i}{\partial r \partial l} = 0, \tag{11}$$

then $B_R(r; \tilde{r}, \tilde{l})$ decreases in $\tilde{l}$ and $B_L(l; \tilde{r}, \tilde{l})$ decreases in $\tilde{r}$.

Condition (11) allows for capture efforts by a SIG to crowd-out the opposite SIG's influence, but rules out any interaction effect with the level of effort of that SIG.[18] However, it is important to note that this is a sufficient condition. Strategic substitutability also holds in cases where the cross-partial is non-zero, but one needs to keep track of second order effects caused by the contest function.[19]

4.1.2 Firing up the Base versus Demobilizing the Opposition

How are SIGs' capure incentives affected by the priors of the media's viewership? We now explore what happens when SIGs prioritize influencing certain types of viewers over others. In particular, we classify SIGs according to their attitudes towards viewers whose priors are already quite favorable (or unfavorable) to the state they want to push (i.e., low p for $L-$SIG and high p for $R-$SIG). Thus, we say that a SIG wants to "fire up its base" if incentives to capture increase when facing a crowd of convinced partisans, while she wants to "demobilize the opposition" if these incentives are stronger with a crowd of opposite partisanship. Formally, an $R-$SIG ($L-$SIG) wants to fire up its base if $B_R(B_L)$ increases when $F_p(p)$ increases (decreases) in the FOSD sense, with a similar definition for the case in which she wants to "demobilize the opposition". In either case, SIGs that want to fire up the base or demobilize the opposition exhibit a preference for capturing audiences with extreme opinions, a behavior that can exacerbate differentiated capture as we will see in Section 6.

[18] One such contest function would be $\pi_R(r,l) = r - \eta l, \pi_L(r,l) = l - \eta r$, with η a fixed parameter.

[19] See Corchon (2007) and Acemoglu and Jensen (2013) for treatments of the complexity of comparative statics for arbitrary contest functions.

From (9) and (10), viewership affects capture incentives only through

$$V_i'(\lambda) = \int (\partial v_i(\mu(\lambda, p))/\partial \lambda)\, dF_p(p). \tag{12}$$

For $i = R$, $\partial v_R(\mu(\lambda, p))/\partial\lambda$ represents the $R-$SIG's marginal payoff from sending a more favorable message to a viewer with prior p. Then, (12) averages this payoff across all viewers. Clearly, the $R-$SIG wants to fire up its base if $\partial v_R(\mu(\lambda, p))/\partial\lambda$ increases in p, as she gains more from changing the coverage when viewers already hold favorable beliefs, while she wants to demobilize the opposition if $\partial v_R(\mu(\lambda, p))/\partial\lambda$ decreases in p as she prefers to bring towards the centre those viewers with priors opposite to her preferred state. Likewise, the $L-$SIG wants to fire up its base (demobilize the opposition) if $-\partial v_L(\mu(\lambda, p))/\partial\lambda$ decreases (increases) in p. In both cases, an $i-$SIG wants to fire up its base if and only if $\partial v_i^2(\mu(\lambda, p))/\partial\lambda\partial p \geq 0$. The next proposition shows that SIGs have preference for extreme distributions if v_i is sufficiently convex or concave.

Proposition 3. *Let $\mathbb{M} = [\underline{\mu}, \overline{\mu}]$ be the range of posteriors induced on viewers when consuming the news of a known honest media. Let $K(\mu) = \mu/(1-\mu) - (1-\mu)/\mu$ be the difference between the odds of $\theta = 1$ and $\theta = -1$. Then*

i-If $\min_{\mu\in[\underline{\mu},\overline{\mu}]} \frac{v_R''(\mu)}{v_R'(\mu)} > K(\overline{\mu})$, then the $R-$SIG wants to fire up its base, while if $\min_{\mu\in[\underline{\mu},\overline{\mu}]} \frac{v_L''(\mu)}{\left|v_L'(\mu)\right|} > -K(\underline{\mu})$ then the $L-$SIG wants to fire up its base.

ii-If $\max_{\mu\in[\underline{\mu},\overline{\mu}]} \frac{v_R''(\mu)}{v_R'(\mu)} < K(\underline{\mu})$, then the $R-$SIG wants to demobilize the opposition, while if $\max_{\mu\in[\underline{\mu},\overline{\mu}]} \frac{v_L''(\mu)}{\left|v_L'(\mu)\right|} < -K(\overline{\mu})$ then the $L-$SIG wants to demobilize the opposition.

If v_i is sufficiently convex, then influencing public opinion is mostly about firing one's base, while if v_i is sufficiently concave, it is mostly about demobilizing the opposition. This is intuitive for an $R-SIG$ as the gain from raising the beliefs of the public is higher (lower) for those holding very favorable beliefs if v_R is convex (concave). For an $L-SIG$, convexity (concavity) of v_L implies that the gain from lowering the beliefs of the public is higher (lower) for those holding already low priors. The extra conditions in both cases are needed to account for the fact that a higher λ has a smaller (larger) effect on viewers posteriors if viewers hold a higher (lower) prior belief.

Proposition 3 links the SIGs preferences for capturing extreme viewerships to the curvature of their utility function. The conditions depend on the informativeness of the honest media as well as the audience priors (as they both affect the set of possible posteriors $\mathbb{M}$). However, convexity in the odds of a favorable state are sufficient to guarantee that SIGs want to fire up their base. This insight will be used for our analysis of viewer sorting in Section 6.

Lemma 3. *Suppose that $v_R = g_R(\mu/(1-\mu))$ and $v_L = g_L((1-\mu)/\mu)$, with g_i, $i \in \{L, R\}$, increasing and convex. Then both SIGs want to fire up their base.*

4.2 Equilibrium Capture

The next proposition determines the equilibrium level of capture;

Proposition 4. *Suppose that each SIG can invest in media capture at an increasing and convex cost $C_R(r)$ and $C_L(l)$. With $V_i(\lambda)$ defined in (8), every equilibrium level of capture r^*, l^* has unique $\overline{\lambda}$ and $\underline{\lambda}$ satisfying*

$$\int_{\underline{\lambda}}^{\overline{\lambda}} V_R'(\lambda) F_H(\lambda; p_R) d\lambda = C_R'(r^*), \tag{13}$$

$$\int_{\underline{\lambda}}^{\overline{\lambda}} (-V_L'(\lambda))\, \overline{F}_H(\lambda; p_L) d\lambda = C_L'(l^*), \tag{14}$$

$$\int_{\overline{\lambda}}^{\infty} \left(\lambda - \overline{\lambda}\right) dF_{H,-1}(\lambda) = \frac{r^*}{1 - l^* - r^*} \left(\overline{\lambda} - 1\right), \tag{15}$$

$$\int_{0}^{\underline{\lambda}} \left(\underline{\lambda} - \lambda\right) dF_{H,-1}(\lambda) = \frac{l^*}{1 - l^* - r^*} \left(1 - \underline{\lambda}\right) \tag{16}$$

Equations (13-16) encapsulate the main equilibrium tension in our model: (13) and (14) show that each SIG's marginal benefit from capturing coverage increases if the public trusts media more–resulting in a higher $\overline{\lambda}$ and lower $\underline{\lambda}$– but more intense media capture lowers citizens' trust as indicated by (15) and (16). Equations (13) and (14) imply that each SIG has no incentive to increase effort in capture given the anticipated levels of capture while, following Proposition 1, (15) and (16) represent the most R-favorable and L-favorable equilibrium likelihood ratios.

A direct inspection of (13) shows that the marginal benefit from media capture decreases if the prior belief of the $R-$SIG increases. This follows as $F_H(\lambda; p_R)$ increases in a FOSD sense with increases in p_R. Thus, an $R-$sender that is more optimistic of "good news" from the honest media will profit less from capturing that coverage. Strategic substitutability then implies that, if the equilibrium is unique, then the $R-$sender's capture level must decrease. Thus, the level of $R-$capture unambiguously decreases as the $R-$sender becomes more optimistic about the underlying state θ.

The effect of changes in viewers' prior beliefs on equilibrium capture is less immediate: even though $F_p(p)$ does not affect the way SIGs communicate given the anticipated level of capture –see Lemma 2– it does affect the returns from capture through its effect on the marginal gain/loss from a higher message $V_i'(\lambda)$ – see Proposition 3 and Lemma 3. We can be more specific, however, if SIGs have a preference for influencing partisan audiences. For instance, if both SIGs want to fire up their base, a FOSD increase in $F_p(p)$ would increase capture by the $R-SIG$ while simultaneously reducing that of an $L-SIG$. Conversely, $L-SIG$ capture would be exacerbated (and $R-SIG$'s capture reduced) if instead $F_p(p)$ reduces in the FOSD.

5 Competitive Capture of a Media Duopoly

In the previous section we have demonstrated that competitive efforts to capture the coverage of a monopoly media outlet are strategic substitutes. Now consider a duopoly with two otherwise identical media channels, 1 and 2. Strategic substitutability still holds in the duopoly case when considering each individual channel. As a consequence, this force now invites horizontal differentiation. As $R-$SIG increases efforts to capture, say, outlet 1, it dampens the marginal return to $L-$SIG to try to capture 1. This pushes $L-$SIG to divert more effort towards outlet 2. We now explore this logic by studying conditions under which maximum differentiation can be sustained in a general equilibrium with duopolistic media.

5.1 Maximum Differentiation

Media polarization along political parties is a stylized fact of several media markets, with some media heavily favoring a right wing ideology while other favoring a left wing ideology. In the context of our equilibrium model, we can show that systematic capture of an issue by only one of the SIG is possible, even if the two media are identical ex ante. If this is possible, then *a fortiori* full horizontal differentiation can be the norm if we allow for systematic differences of media, such as capturing costs β_j^i, which model extant differences such as ownership or political leaning of editorial board.

To gain some intuition, suppose that viewers believe that media outlet $j = 1$ is only captured by the $R-$SIG with effort $r_1 = r^*$ while media outlet $j = 2$ is only captured by the $L-$SIG with effort $l_2 = l^*$. Assume further that the distribution of priors of the viewership of outlet i is F_p^i (whether exogenous or endogenous through anticipated levels of media capture as we develop in Section 6). A necessary condition for maximum differentiation is that the marginal gain for the $R-$SIG to capture the outlet where $L-$SIG is focusing its effort cannot exceed that of devoting resources to the outlet that viewers anticipate to be potentially captured by $R-$SIG. Using (9), the difference between these two marginal gains when viewers anticipate capture $(\widetilde{r}, \widetilde{l}) = ((r^*, 0), (0, l^*))$ is

$$\Delta_R = B_R(r_1; (\tilde{r}, \tilde{l})) - B_R(r_2; (\tilde{r}, \tilde{l})) = \int_{\lambda_{min}^1}^{\overline{\lambda}_1} V_R'(\lambda; F_p^1) F_H^1(\lambda; p_R) d\lambda - \int_{\underline{\lambda}_2}^{\lambda_{max}^2} V_R'(\lambda; F_p^2) F_H^2(\lambda; p_R) d\lambda$$

$$= - \int_{\overline{\lambda}_1}^{\lambda_{max}^2} V_R'(\lambda; F_p^2) F_H^2(\lambda; p_R) d\lambda + \int_{\lambda_{min}^1}^{\underline{\lambda}_2} V_R'(\lambda; F_p^1) F_H^1(\lambda; p_R) d\lambda \tag{17}$$

$$+ \int_{\underline{\lambda}_2}^{\overline{\lambda}_1} \left(V_R'(\lambda; F_p^1) F_H^1(\lambda; p_R) - V_R'(\lambda; F_p^2) F_H^2(\lambda; p_R) \right) d\lambda. \tag{18}$$

and maximum differentiation requires that $\Delta_R \geq 0$. To understand this expression, consider first the terms in (17). Note that because viewers do not anticipate the $R-$SIG to try to capture outlet 2, there is no cap on the trusted "high" messages in that channel. Thus, by capturing channel 2, the $R-$SIG can credibly send the most informative message in his favor λ_{max}^2, which is a strong incentive. In contrast, the informational content of messages is limited when capturing outlet 1 to $\overline{\lambda}_1$, as viewers of media 1 are skeptical of news favoring $R-$SIG's interest. This explains the first term in (17). However, honest media can potentially avoid more adverse news reaching the public in media 1 than in media 2 as viewers skepticism of $L-$favoring news limits them to $\underline{\lambda}_2$. This explains the second term in (17). Thus, (17) captures the differences in marginal gains to capture by focusing on differences in the range of equilibrium informational content of messages in both media outlets. Finally, the differences in marginal gains can also come from differences in the distribution of F_p^i and F_H^i for the common support of likelihood ratios in both media outlets. This is (18). In summary, differences in marginal gains to capture can come from differences in the range of messages that are credible in each media outlet, but also from differences in viewership and outlet informativeness when honest.

Note that if we allow for differences in viewership across media (i.e. $F_p^1 \neq F_p^2$) it is trivial to generate situations in which $\Delta_R \geq 0$ thanks to (18). However, even with identical F_p^i and F_H^i we can have $\Delta_R \geq 0$. Inspection of (17) shows that this requires extreme concavity of SIG preferences so the overiding concern of the $R-$SIG is to avoid the publishing of very unfavorable and credible λ. In this case, since the activity of the $L-$SIG is already reducing the credibility of unfavorable messages in media 2, the $R-$SIG is better off increasing the capture of media 1 to prevent credible unfavorable messages to filter through.

We can equivalently write the difference in marginal gains to an $L-$sender as

$$\Delta_L = B_L(l_2; (\tilde{r}, \tilde{l})) - B_L(l_1; (\tilde{r}, \tilde{l})) = \int_{\underline{\lambda}_2}^{\lambda_{max}^2} V_L'(\lambda; F_p^2) F_H^2(\lambda; p_L) d\lambda - \int_{\lambda_{min}^1}^{\bar{\lambda}_1} V_L'(\lambda; F_p^1) F_H^1(\lambda; p_R) d\lambda$$

where we can similarly decompose Δ_L to the component due to the difference in equilibrium information content of messages in each outlet and the component due to the difference in viewership and informativeness of those outlets when they are honest.

If we have $\Delta_R \geq 0$ and $\Delta_L \geq 0$ we can support an equilibrium with full differentiation across the two media outlets.

6 Viewer Sorting in a Media Duopoly

We now consider the endogenous choice of viewers across media outlets in response to the anticipated level of capture. To model viewers' value for information, we endow them with the following choice problem. With probability γ, a viewer needs to make a choice between acting ($a = 1$) or not ($a = 0$). For example, acting may be choosing which party to vote, going to a demonstration, or taking some decision influenced by beliefs over the seriousness of climate change. Each viewer is characterized by her prior $p = Pr(\theta = 1)$ and the threshold $\alpha \in (0, 1)$ that her belief needs to cross for her to act. That is, if her posterior $\mu(m, p) > \alpha$, then the viewer chooses $a = 1$.[20] Note

[20] We will use the following viewer's preferences to model this behavior: a viewer obtains $1-\alpha$ if $a = 1$ and $\theta = 1$; α if $a = 0$ and $\theta = -1$; and 0 otherwise.

that we allow viewers to vary in ideology α and in beliefs p.

We can rewrite the characterization of each viewer in terms of (p, λ_{crit}), where

$$\frac{\alpha}{1-\alpha} = \lambda_{crit}\frac{p}{1-p}.$$

This formulation is useful because λ_{crit} is the minimum informational content of a message that will lead a viewer of prior p and threshold α to act. For example, viewers with $\lambda_{crit} > 1$ do not act in the absence of news, and to act they need to see a message that increases their belief that $\theta = 1$. In contrast, viewers with $\lambda_{crit} < 1$ are already convinced of the need to act and they will only change their decision if they see a message that reduces their belief that $\theta = 1$.

We can show that viewers sort across outlets according to ideology.

Proposition 5. *Consider two outlets, 1 and 2 with* $F_H^1 = F_H^2 (= F_H)$. *Consider an equilibrium with outlet 1 mostly captured by R−SIG (so that* $r^1 \geq l^1$*) and outlet 2 by the L−SIG (so that* $l^2 \geq r^2$*) while total capture is not too dissimilar in the sense that*

$$\frac{r^1}{r^2} > \frac{1-(r^1+l^1)}{1-(r^2+l^2)} > \frac{l^1}{l^2} \tag{19}$$

If $\gamma = 1$, *then:*

i-For each p *there are* $\underline{\lambda}(p) \leq \bar{\lambda}(p)$ *such that viewers with* $\lambda_{crit} > \bar{\lambda}(p)$ *consume outlet 2 and viewers with* $\lambda_{crit} < \underline{\lambda}(p)$ *consume outlet 1.*

ii-If $r^1 + l^1 = r^2 + l^2$, *then for each* p *there is* $\tilde{\lambda}(p)$ *such that viewers sort monotonically: viewers consume outlet 2 if* $\lambda_{crit} > \tilde{\lambda}(p)$ *and consume outlet 1 if* $\lambda_{crit} < \tilde{\lambda}(p)$.

Note that viewers with high λ_{crit} are those whose combination of priors and ideology makes them reluctant to act as they believe $\theta = -1$ is likely. These viewers refuse to watch outlet 1, which is the outlet mostly captured by the R−SIG which wants to increase beliefs that $\theta = 1$. These left wing viewers instead endogenously choose to watch outlet 2, which is expected to be captured by the L−SIG. The intuition is that such left wing viewers need a strong credible message that the state is $\theta = 1$ in order to change their decision. However, outlet 1 is often captured by the R−SIG and consequently messages that favor $\theta = 1$ are suspect and not convincing enough. This viewer is better off watching outlet 2: if outlet 2 happens to remain honest, a high message with high λ is possible, and coming from this outlet it would be convincing enough to make the viewer change her choice of action.

This sorting effect is similar to Suen (2004) but we obtain it in a model without filtering in which media can freely transmit information. In fact, in our model, the value of media for *all* consumers diminishes with increased effort of capture. The reason viewers sort in our model is that they cannot trust the messages that would be valuable to them in the media outlet that is captured by the SIG that is ideologically opposed.

However, the fact that capture reduces the value of media does not mean in our model that increased demand for information reduces polarization and makes media more centrist. The following proposition describes a situation in which the opposite is true.

Proposition 6. *Suppose that* $v_R = g(\frac{\mu}{1-\mu})$ *and* $v_L = g(\frac{1-\mu}{\mu})$ *with* g *increasing and convex and two media outlets 1 and 2 with* $F_H^1 = F_H^2\,(= F_H)$ *and* $\vartheta^1(p) = \vartheta^2(p) = \frac{1}{2}$*. Assume also that* $F_p(\frac{1}{2}) = \frac{1}{2}$*. Consider an asymmetric equilibrium with* $\bar{\lambda}_1$ *(*$\underline{\lambda}_2$*) the highest (lowest) likelihood ratio in media 1 (media 2), dominated by R−SIG (L−SIG). If for any* (p, λ_{crit}) *in the support of the distribution of viewers* $F(p, \lambda_{crit})$ *we have:*

$$\text{if } p > \frac{1}{2}, \text{ then } \lambda_{crit} \leq \underline{\lambda}_2; \text{ If } p < \frac{1}{2}, \text{ then } \lambda_{crit} \geq \bar{\lambda}_1 \qquad (20)$$

then any increase in γ *increases the degree of media polarization.*

The assumptions in this proposition ensure symmetry in viewers. $\vartheta^1 = \vartheta^2 = \frac{1}{2}$ means that in terms of entertainment value (which directs consumer choice with probability $1 - \gamma$) the two outlets get half of viewers each, independent of their priors. Similarly $F_p(\frac{1}{2}) = \frac{1}{2}$ ensures that half of viewers have priors that lean towards each state being likely. The proposition states that when there are no centrists, then increasing the value of information γ increases polarization.

To see the intuition note that condition (20) ensures by Proposition 5 that all viewers who search for information (a proportion γ of the population) are sorted ideologically across outlets, while the rest are sorted equally across outlets independent of p. Note also that because g is convex, Lemma 3 indicates that SIGs want to capture coverage in order to fire up their base. Increasing γ

increases sorting: as γ increases, the proportion of viewers of outlet 1 who have high p increases and the proportion of viewers of outlet 2 who have low p also increases. As γ increases therefore the $R-$SIG can reach more of the high p viewers through outlet 1 and less through outlet 2. This increases $R-$SIG's incentives to influence outlet 1 and reduces incentives to influence outlet 2. The fact that capturing efforts are strategic substitutes further reinforces this. As a consequence, increased value of information leads to more polarized media.

7 Naive Viewers

Viewer skepticism towards extreme news published by a possibly captured media led to two of our main insights: news polarization and captured efforts being strategic substitutes. In this section we show that our insights are robust to allowing for some viewer naiveté when interpreting the news they consume. Namely, we allow for a fraction $1-\gamma < 1$ of viewers to be "naive" in that they believe all media coverage to be honest, while a fraction γ of viewers are "sophisticated" as they correctly anticipate the likelihood of capture.

To fix ideas, consider the case of a monopoly media outlet. Naive and sophisticated viewers interpret the same news λ differently: naive viewers take news at face value and interpret λ literally, while sophisticated viewers are wary of capture and interpret them as $\lambda_\gamma(\lambda)$. For instance, Proposition 1 shows that when all viewers are sophisticated (i.e., $\gamma = 1$), $\lambda_\gamma(\lambda) = \overline{\lambda}$ for $\lambda \geq \overline{\lambda}$ while $\lambda_\gamma(\lambda) = \underline{\lambda}$ for $\lambda \leq \underline{\lambda}$. The following proposition summarizes the main features of communication equilibria with naive viewers.

Proposition 7. *Fix levels of capture r and l, with $r+l<1$, and let $V_i(\lambda)$, defined in (8), be the expected utility of the $i-SIG$ if viewers interpret the message as λ. There exists a unique equilibrium interpretation of the news by sophisticated viewers $\lambda_\gamma(\lambda)$, with unique $\overline{\lambda}$ and $\underline{\lambda}$, satisfying*

1. *$\lambda_\gamma(\lambda)$ is given by*

$$\lambda_\gamma(\lambda) = \begin{cases} V_L^{-1}(V_L(\underline{\lambda}) + \frac{1-\gamma}{\gamma}(V_L(\underline{\lambda}) - V_L(\lambda))) & \text{if } \lambda \leq \underline{\lambda}, \\ \lambda & \text{if } \underline{\lambda} < \lambda < \overline{\lambda}, \\ V_R^{-1}(V_R(\overline{\lambda}) + \frac{1-\gamma}{\gamma}(V_R(\overline{\lambda}) - V_R(\lambda))) & \text{if } \lambda \geq \overline{\lambda}. \end{cases} \tag{21}$$

2. *The associated* $\overline{\lambda}$ *and* $\underline{\lambda}$ *satisfy*

$$\int_{\overline{\lambda}}^{\infty}\left(\frac{\lambda-\lambda_\gamma(\lambda)}{\lambda_\gamma(\lambda)-1}\right)dF_{H,-1}(\lambda)=\frac{r}{1-l-r}, \tag{22}$$

$$\int_{0}^{\underline{\lambda}}\left(\frac{\lambda_\gamma(\lambda)-\lambda}{1-\lambda_\gamma(\lambda)}\right)dF_{H,-1}(\lambda)=\frac{l}{1-l-r} \tag{23}$$

3. $\overline{\lambda}$ *decreases in* l,r, *and* γ *while* $\underline{\lambda}$ *is increasing in* l,r, *and* γ. *Fixing* $\overline{\lambda}$ *and* $\underline{\lambda}$, *then* $\lambda_\gamma(\lambda)$ *decreases (increases) in* l,r, *and* γ *for* $\lambda \geq \overline{\lambda}$ ($\lambda \leq \underline{\lambda}$).

The presence of naive viewers does not qualitatively change our insights regarding news polarization and viewer skepticism: the $R-$SIG selects messages with a literal meaning above some $\overline{\lambda}$ while the $L-$SIG chooses messages below $\underline{\lambda}$; this results in an increase incidence of extreme news which, in turn, are not trusted by sophisticated viewers. However, SIGs reporting must now balance the effect of news on each type of viewer: as naive viewers take news at face value, selecting news with more favorable literal meanings must be offset by a less favorable interpretation by sophisticated viewers. This effect is captured in (21) as $\lambda_\gamma(\lambda)$ is decreasing for both$\lambda > \overline{\lambda}$ and for $\lambda < \underline{\lambda}$–see Figure XXX. It follows from (21) that more extreme news are more heavily discarded by sophisticated viewers and lead to a more centrist interpretation when SIGs also target naive viewers.

A key difference between Proposition 1 and 7 is that, in the presence of naive viewers, communication equilibria can vary with the distribution of viewers' priors. The reason is that each SIGs' indifference among all potential lies relies on balancing his utility from naive and sophisticated viewers, but a SIG's utility from each message interpreted as face value does depend on viewers' priors. This also implies that the highest and lowest trusted news, as given by Part 2 of the Proposition, now depend on the priors of viewers.

Finally, increased viewer sophistication (higher γ) makes them trust a smaller set of news–this is in part 3 of Proposition 7. This is intuitive as each SIG gains less from pandering to naive viewers, reducing the likelihood of sending the most extreme messages. This, however, means that more centrist messages are now used by either SIG.

A key feature of Proposition 7, as shown in part 3, is that increasing the capture level of, say, the $L-$SIG, not only reduces $\overline{\lambda}$ and increases $\underline{\lambda}$, but it also affects in a monotonic way the interpretation of the news by sophisticated viewers: increasing l worsens the intrerpretation of $R-$lies by sophisticated viewers–by reducing $\lambda_\gamma(\lambda)$ for $\lambda \geq \overline{\lambda}$–but makes his own lies more favorable to the $R-$SIG–by increasing $\lambda_\gamma(\lambda)$ for $\lambda \leq \underline{\lambda}$. Both effects unambiguously reduce the marginal gain for the $R-$SIG from capture.

Proposition 8. *Suppose that there is a single media outlet and the probability that the $R-$SIG ($L-$SIG) captures the coverage is $r(l)$. Then, for any fraction $\gamma > 0$ of sophisticated viewers, capture efforts are strategic substitutes.*

8 Conclusion

We have developed a model of media capture by special interest groups. In the model, two opposing interest groups covertly devote effort to capture coverage of an issue by multiple media outlets who broadcast to a viewership with heterogeneous priors. Captured media can publish any fake news, untethered to the underlying state of the world, and with no commitment to any editorial line. We characterize the optimal lying strategies of special interest groups and show that capture leads to polarization in the news. However, rational viewers are not deceived by extreme messages and become skeptical. We also show that capturing efforts are strategic substitutes at the media outlet level. This strategic substitution implies that horizontal differentiation is possible when multiple media outlets are present. When we allow viewers to choose which outlet to consume, they sort ideologically, which can reinforce horizontal differentiation.

In focusing on the capturing and manipulating decisions of special interest groups, and on the informational consequences for viewers, we take a simplified view of the media outlets themselves. In particular, media are passive receivers of pressure by special interest groups and, if they remain free of capture, they are honest conveyors of information. The rich existing literature on media capture has emphasized a trade-off between profit/viewership maximizing and yielding to pressure which we do not consider in this model. We leave for further research to study the conditions under which this trade-off reinforces or weakens the novel mechanisms we have uncovered in this paper.

References

Acemoglu, D., and M. K. Jensen (2013): "Aggregate Comparative Statics," *Games and Economic Behavior*, 81, 27–49. 19

Anderson, S., and J. McLaren (2012): "Media Mergers and Media Bias with Rational Consumers," *Journal of the European Economic Association*, 10(4), 831–859. 1

Angelucci, C., and A. Prat (2021): "Is Journalistic Truth Dead? Measuring How Informed Voters Are about Political News," Working Paper 6132, MIT Sloan. 1, 3.3

Ansolabehere, S., J. Rodden, and J. Snyder (2006): "Purple America," *Journal of Economic Perspectives*, 20(2), 97–118. 1

Baron, D. (2006): "Persistent Media Bias," *Journal of Public Economics*, 90(1-2), 1–36. 1

Beattie, G., R. Durante, B. Knight, and A. Sen (2021): "Advertising Spending and Media Bias: Evidence from News Coverage of Car Safety Recalls," *Management Science*, 67(2), 698–719. 1

Bergemann, D., and S. Morris (2019): "Information Design: a Unified Perspective," *Journal of Economic Literature*, 57(1), 44–95. 6

Bernhardt, D., S. Krasa, and M. Polborn (2008): "Political Polarization and the Electoral Effects of Media Bias," *Journal of Public Economics*, 92(5-6), 1092–1104. 1

Besley, T., and A. Prat (2006): "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability," *American Economic Review*, 96(3), 720–736. 1

Brookman, D., and J. Kalla (2022): "The Manifold Effects of Partisan Media on Viewers' Beliefs and Attitudes: A Field Experiment with Fox News Viewers," Working paper, UC Berkeley and Yale University. 1

Chan, J., and W. Suen (2008): "A Spatial Theory of News Consumption and Electoral Competition," *The Review of Economic Studies*, 75(3), 699–814. 1, 15

Corchon, L. (2007): "The Theory of Contests: a Survey," *Review of Economic Design*, 11(2), 69–100. 19

Corneo, G. (2006): "Media Capture in a Democracy: the Role of Wealth Concentration," *Journal of Public Economics*, 90(1-2), 37–58. 1

Crawford, V., and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–1451. 8

Duggan, J., and C. Martinelli (2011): "A Spatial Theory of Media Slant and Voter Choice," *The Review of Economic Studies*, 78(2), 640–666. 15

Durante, R., A. Fabiani, L. Laeven, and J. L. Peydro (2021): "Media Capture by Banks," Discussion Paper 15260, CEPR. 1

Dziuda, W. (2011): "Strategic Argumentation," *Journal of Economic Theory*, 146(1), 1362–1397. 1, 11

Gabszewicz, J., D. Laussel, and N. Sonnac (2001): "Press Adverstising and the Ascent of the 'Pensee Unique'," *European Economic Review*, 45, 641–651. 1

Gehlbach, S., and K. Sonin (2014): "Government control of the media," *Journal of Public Economics*, 118, 163–171. 1

Gentzkow, M., and E. Kamenica (2011): "Bayesian Persuasion," *American Economic Review*, 101(6), 2590–2615. 6

Gentzkow, M., and J. Shapiro (2006): "Media Bias and Reputation," *Journal of Political Economy*, 115(2), 280–316. 1, 15

——— (2010): "What Drives Media Slant? Evidence from US Daily Newspapers," *Econometrica*, 78(1), 35–71. 1

Gentzkow, M., J. Shapiro, and M. Sinkinson (2011): "The Effect of Newspaper Entry and Exit on Electoral Politics," *American Economic Review*, 101(7), 2980–3018. 3

Gentzkow, M., J. Shapiro, and D. Stone (2015): "Media Bias in the Marketplace: Theory," in *Handbook of Media Economics*, ed. by D. S. Simon Anderson, and J. Waldfogel. Elsevier B.V., Amsterdam. 9

Gitmez, A., and P. Molavi (2022): "Polarization and Media Bias," Working paper, Bilkent University and Northwestern University. 6

Grossman, G., and E. Helpman (2001): *Special Interest Politics*. MIT Press, Cambridge, MA and London, UK. 1

Li, M., and K. Madarasz (2008): "When mandatory disclosure hurts: Expert advice and conflicting interests," *Journal of Economic Theory*, 139(1), 47–74. 1

Martin, G., and A. Yurukoglu (2017): "Bias in Cable News: Persuasion and Polarization," *American Economic Review*, 107(9), 2565–2599. 1

McMillan, J., and P. Zoido (2004): "How to Subvert Democracy: Montesinos in Peru," *Journal of Economic Perspectives*, 18(4), 69–92. 1

Morgan, J., and P. Stoken (2003): "When mandatory disclosure hurts: Expert advice and conflicting interests," *Rand Journal of Economics*, 34(1), 183–203. 1

Morris, S. (2001): "Political Correctness," *Journal of Political Economy*, 109(2), 231–265. 10

Mullainathan, S., and A. Shleifer (2005): "The Market for News," *American Economic Review*, 95(4), 1031–1053. 1

Petrova, M. (2008): "Inequality and Media Capture," *Journal of Public Economics*, 92(1-2), 183–212. 1

——— (2012): "Mass Media and Special Interest Groups," *Journal of Economic Behavior and Organization*, 84(7), 17–38. 7

Prat, A. (2015): "Media Capture and Media Power," in *Handbook of Media Economics*, ed. by D. S. Simon Anderson, and J. Waldfogel. Elsevier B.V., Amsterdam. 4

Shin, H. S. (1993): "Measuring the incidence of insider trading in a market for state-contingent claims," *The Economic Journal*, 103, 1141–1153. 10

Sobbrio, F. (2014): "Citizen Editors' Endogenous Information Acquisition and News Accuracy," *Journal of Public Economics*, 113(1), 43–53. 1

Sobel, J. (1985): "A Theory of Credibility," *The Review of Economic Studies*, 52, 557–573. 1

——— (2013): "Giving and Receiving Advice," in *Advances in Economics and Econometrics*, ed. by M. A. Daron Acemoglu, and E. Dekel. Cambridge University Press. 8

——— (2020): "Lying and Deception in Games," *Journal of Political Economy*, 128(3), 907–947. 2, 16

Suen, W. (2004): "The Self-Perpetuation of Biased Beliefs," *The Economic Journal*, 114, 377–396. 2, 6

Wolinsky, A. (2003): "Information transmission when the senderâs preferences are uncertain," *Games and Economic Behavior*, 42(2), 319–326. 1, 11

Appendix

Proof of Proposition 1. Suppose that the $R-$SIG and $L-$SIG's strategies are $\tau_R(m)$ and $\tau_L(m)$ so that $\tau_i(m)$ is the probability that the i-SIG sends m if he captures the coverage. Then, the perceived likelihood ratio $\lambda(m) \equiv \frac{\Pr[m|\theta=1]}{\Pr[m|\theta=0]}$ is

$$\lambda(m) = \frac{(1-l-r)p_1(m) + r\tau_R(m) + l\tau_L(m)}{(1-l-r)p_{-1}(m) + r\tau_R(m) + l\tau_L(m)}. \tag{24}$$

The perceived likelihood ratio is sufficient to compute a p-viewer's posterior

$$\mu(m;p) = \frac{\Pr[\theta=1,m]}{\Pr[m]} = \frac{p\lambda(m)}{1-p+p\lambda(m)},$$

so that the difference in posteriors after observing two different messages m and m' is

$$\mu(m;p) - \mu(m';p) = (\lambda(m) - \lambda(m')) \frac{p(1-p)}{(1-p+p\lambda(m))(1-p+p\lambda(m'))}.$$

Averaging over the posterior of all viewers, the i-SIG's indirect utility from message m when viewers anticipate mixing $\tau_R(m)$ and $\tau_L(m)$ is

$$V_i(m) \equiv \int_0^1 v_i(\mu(m;p))\, dF_p(p) = \int_0^1 v_i\left(\frac{p\lambda(m)}{1-p+p\lambda(m)}\right) dF_p(p). \tag{25}$$

SIGs' optimality requires that if $m, m' \in \mathit{supp}\,\tau_i$ then $V_i(m) = V_i(m')$, $i \in \{L, R\}$. We now show that this implies that $\lambda(m) = \lambda(m')$. Indeed, suppose wlog that $\lambda(m) \geq \lambda(m')$. Then, for $i = R$ we have

$$\begin{aligned}
0 &= \int_0^1 (v_R(\mu(m;p)) - v_R(\mu(m';p)))\, dF_p(p) = \int_0^1 \left(\int_{\mu(m';p)}^{\mu(m;p)} v_R'(s)\, ds\right) dF_p(p) \\
&\geq \inf_{0\leq s\leq 1} (v_R'(s)) \left(\int_0^1 (\mu(m;p)) - \mu(m';p) dF_p(p)\right) \\
&= \inf_{0\leq s\leq 1} (v_R'(s)) (\lambda(m) - \lambda(m')) \int_0^1 \left(\frac{p(1-p)}{(1-p+p\lambda(m))(1-p+p(\lambda(m'))}\right) dF_p(p) \\
&\geq 0
\end{aligned}$$

as the integrand in the last equation is strictly positive. Since v_R' is bounded away from zero, we must then have that $\lambda(m) = \lambda(m')$. A similar argument would establish that $\lambda(m) = \lambda(m')$ if $m, m' \in support\ \tau_L$.

Note that (i) $V_R(m)$ in (25) is strictly increasing in $\lambda(m)$ while $V_L(m)$ in (25) is strictly decreasing in $\lambda(m)$, and (ii) if $\tau_R(m) = \tau_L(m) = 0$ then $\lambda(m) = \lambda_H(m)$. Letting $\lambda^*(m)$ be the equilibrium likelihood ratio of message m with $\overline{\lambda} = \max_{m \in \mathscr{M}} \lambda^*(m)$ and $\underline{\lambda} = \min_{m \in \mathscr{M}} \lambda^*(m)$, then $\lambda^*(m) = \overline{\lambda}$ if $m \in supp(\tau_R^*)$ while (ii) implies that $m \in supp(\tau_R^*)$ only if $\lambda_H(m) \geq \overline{\lambda}$. If $\overline{m}^*$ is defined by $\lambda_H(\overline{m}^*) = \overline{\lambda}$ then $m \in supp(\tau_R^*)$ iff $m \geq \overline{m}^*$. Conversely, if $m \in supp(\tau_L^*)$ then $\lambda^*(m) = \underline{\lambda}$ and $m \in supp(\tau_L^*)$ iff $\lambda_H(m) \leq \underline{\lambda}$. Thus, if $\underline{m}^*$ is defined by $\lambda_H(\underline{m}^*) = \underline{\lambda}$, then $m \in supp(\tau_R^*)$ iff $m \leq \underline{m}^*$.

Note that, generically, the R and L lobbyists will never send the same message with positive probability–this will be always the case if $l + r < 1$ so that $\Pr[S = H] > 0$. That is, we must have in equilibrium that $\tau_R^*(m)\tau_L^*(m) = 0$ for all $m \in \mathscr{M}$.

Using (24) we can write for all m such that $\lambda_H(m) \geq \overline{\lambda}$

$$\frac{r}{1-l-r}\left(\overline{\lambda}\tau_R(m) - \tau_R(m)\right) = \left(\lambda_H(m) - \overline{\lambda}\right) p_{-1}(m), \tag{26}$$

and for all m such that $\lambda_H(m) \leq \underline{\lambda}$ we have

$$\frac{l}{1-l-r}\left(\tau_L(m) - \underline{\lambda}\tau_L(m)\right) = \left(\underline{\lambda} - \lambda_H(m)\right) p_{-1}(m). \tag{27}$$

Integrating (26) over all $\left\{m : \lambda_H(m) \geq \overline{\lambda}\right\}$ gives (4). A similar argument yields (5) from (27). The right hand-side of (4) is increasing, and the left hand side is non-increasing, in $\overline{\lambda}$, thus, guaranteeing a unique solution to (4). The same argument establishes uniqueness of $\underline{\lambda}$ satisfying (5) □

Proof of Lemma 1. We can solve for $\tau_R^*(m)$ and $\tau_L^*(m)$ using (24) to obtain

$$\tau_R^*(m) = \frac{1-l-r}{r}\left(\frac{\lambda_H(m)-\overline{\lambda}}{\overline{\lambda}-1}\right)p_{-1}(m),$$
$$\tau_L^*(m) = \frac{1-l-r}{l}\left(\frac{\underline{\lambda}-\lambda_H(m)}{1-\underline{\lambda}}\right)p_{-1}(m).$$

implying that $\tau_R^*(m')/\tau_R^*(m) = \left(\lambda_H(m')-\overline{\lambda}\right)p_{-1}(m')/\left(\lambda_H(m)-\overline{\lambda}\right)p_{-1}(m)$ and $\tau_L^*(m')/\tau_L^*(m) = \left(\underline{\lambda}-\lambda_H(m')\right)p_{-1}(m')/\left(\underline{\lambda}-\lambda_H(m)\right)p_{-1}(m)$. □

Proof of Lemma 2. (1) Note that both $r/(1-l-r)$ and $l/(1-l-r)$ increase with l and r. This implies that the right hand sides of (4) and (5) increase with l and r. Equilibrium then implies that $\overline{\lambda}$ must decrease (as well as $\overline{m}^*$) with l and r, while $\underline{\lambda}$ must increase (as well as $\underline{m}^*$) with l and r.

(2) Proposition 1 shows that $\overline{\lambda}$, $\overline{m}^*$, $\underline{\lambda}$ and $\underline{m}^*$ do not vary with F_p as the equilibrium conditions (4) and (5) do not depend on viewers' prior distribution.

(3) To prove that $\overline{\lambda}$ increases and λ decreases when the honest sender is Blackwell-more informative, we will exploit the fact that posterior beliefs are more disperse (in the sense of second order stochastic dominance) under the more informative sender (Blackwell Girshwick 1954). To do this, we will express (4) and (5) in terms of posterior beliefs $\mu(m;p)$ for $p\in(0,1)$. First, we can write

$$\begin{aligned}\frac{\lambda_H(m)-\lambda_H(\overline{m}^*)}{\lambda_H(\overline{m}^*)-1}p_{-1}(m) &= \frac{1}{\lambda_H(\overline{m}^*)-1}p_1(m) - \frac{\lambda_H(\overline{m}^*)}{\lambda_H(\overline{m}^*)-1}p_{-1}(m)\\ &= \frac{p(1-\mu_H(\overline{m}^*;p))}{\mu_H(\overline{m}^*;p))-p}p_1(m) - \frac{\mu_H(\overline{m}^*;p)(1-p)}{\mu_H(\overline{m}^*;p))-p}p_{-1}(m)\\ &= \left(\frac{\mu_H(m;p))-\mu_H(\overline{m}^*;p))}{\mu_H(\overline{m}^*;p))-p}\right)\Omega_H(m;p)\end{aligned}$$

with $\Omega_H(m;p) \equiv p_1(m)p + p_{-1}(m)(1-p)$ the $p-$viewer probability of observing m by an honest media. Then, (4) can be expressed as

$$\int_{\{m:\mu_H(m;p)\geq\overline{\mu}(p)\}} (\mu_C(m;p)) - \overline{\mu}(p))\,\Omega_H(m;p)dm = \frac{r}{1-l-r}\left(\overline{\mu}(p)-1\right)$$

where $\overline{\mu}(p) \equiv \mu_H(\overline{m}^*; p))$. Integrating by parts and expressing the result in terms of $\mu_H = \mu_H(m; p)$ we can write

$$\int_{\overline{\mu}(p)}^{1} \overline{F}_H(\mu_H; p) d\mu_H = \frac{r}{1-l-r}\left(\overline{\mu}(p) - p\right) \tag{28}$$

If honest media H' is Blackwell-more informative than honest media H, then Blackwell Girshwick 1954 shows that for every $p \in (0,1)$

$$\int_{\overline{\mu}(p)}^{1} \overline{F}_{H'}(\mu_{H'}; p) d\mu_{H'} \geqq \int_{\overline{\mu}(p)}^{1} \overline{F}_H(\mu_H; p) d\mu_H,$$

so that to satisfy (28), we must have a higher maximum belief in equilibrium under H'. This implies that $\lambda_H(\overline{m}^*)$ must increase. Conversely, from

$$\frac{\lambda_H(\underline{m}^*) - \lambda_H(m)}{1 - \lambda_H(\underline{m}^*)} p_{-1}(m) = \left(\frac{\mu_H(\underline{m}^*; p) - \mu_H(m; p)}{p - \mu_H(\underline{m}^*; p)}\right) \Omega_H(m; p)$$

we have that (5) translates, after integrating by parts, to

$$\int_{0}^{\underline{\mu}(p)} F_H(\mu_H; p) d\mu_H = \frac{l}{1-l-r}\left(p - \underline{\mu}(p)\right) \tag{29}$$

where $\underline{\mu}(p) = \mu_H(\underline{m}^*; p))$. A Blackwell-more informative sender satisfies

$$\int_{0}^{\underline{\mu}(p)} F_{H'}(\mu_{H'}; p) d\mu_{H'} \geq \int_{0}^{\underline{\mu}(p)} F_H(\mu_H; p) d\mu_H$$

so that $\underline{\mu}$ must decrease to satisfy (29), implying a lower $\lambda_H(\underline{m}^*) = \underline{\lambda}$. □

Proof of Proposition 2. Suppose that citizens anticipate a level of capture of $\left(\tilde{r}, \tilde{l}\right)$; a $p-$citizen's posterior belief after observing m is $\mu^*(m; p) = f\left(\frac{p}{1-p}\lambda^*(m)\right)$ with $f(x) = x/(1+x)$ with the equilibrium likelihood $\lambda^*(m)$ satisfying Proposition 1.2; and the $i-$SIG's interim utility from sending message m with likelihood $\lambda = \lambda^*(m)$ is

$$V_i(\lambda) := \int v_i\left(\mu^*(m; p)\right) dF_p(p) = \int v_i\left(f\left(\frac{p}{1-p}\lambda\right)\right) dF_p(p),$$

Therefore, the $R-$SIG's expected utility when investing r in covertly capturing media is

$$\widehat{V}_R\left(r;\left(\widetilde{r},\widetilde{l}\right)\right)=\pi_R(r,\widetilde{l})V_R(\overline{\lambda})+\pi_L(r,\widetilde{l})V_R(\underline{\lambda})+\pi_H(r,\widetilde{l})\mathbb{E}_H\left[V_R(\lambda);p_R\right]-C_R(r).$$

The $R-$SIG's computes the likelihood that an honest media would have sent a message inducing $\lambda=\lambda^*(m)$ according to his prior p_R, so that

$$\mathbb{E}_H\left[V_R(\lambda);p_R\right]=\overline{F}_H(\overline{\lambda};p_R)V_R(\overline{\lambda})+\int_{\underline{\lambda}}^{\overline{\lambda}}V_R(\lambda)dF_H(\lambda;p_R)+F_H(\underline{\lambda};p_R)V_R(\underline{\lambda}). \tag{30}$$

Therefore, the $R-$SIG 's marginal gain from covertly increasing media capture is

$$B_R(r,(\tilde{r},\tilde{l})):=\frac{\partial\widehat{V}_R\left(r;\left(\widetilde{r},\widetilde{l}\right)\right)}{\partial r}=\frac{\partial\pi_R(r,\widetilde{l})}{\partial r}V_R(\overline{\lambda})+\frac{\partial\pi_L(r,\widetilde{l})}{\partial r}V_R(\underline{\lambda})+\frac{\partial\pi_H(r,\widetilde{l})}{\partial r}\mathbb{E}_H\left[V_R(\lambda);p_R\right]$$

as viewers' interpretation of messages only depends on the expected level of capture $\left(\tilde{r},\tilde{l}\right)$ rather than the actual level $\left(r,\tilde{l}\right)$. The change in B_R if viewers expect a higher level of $L-$capture is

$$\begin{aligned}\frac{\partial B_R(r,(\tilde{r},\widetilde{l}))}{\partial\widetilde{l}}&=\frac{\partial^2\pi_R(r,\widetilde{l})}{\partial r\partial l}V_R(\overline{\lambda})+\frac{\partial\pi_L^2(r,\widetilde{l})}{\partial r\partial l}V_R(\underline{\lambda})+\frac{\partial^2\pi_H(r,\widetilde{l})}{\partial r\partial l}\mathbb{E}_H\left[V_R(\lambda);p_R\right]\\ &\text{‘}+\frac{\partial\pi_R(r,\widetilde{l})}{\partial r}V_R'(\overline{\lambda})\frac{\partial\overline{\lambda}}{\partial\widetilde{l}}+\frac{\partial\pi_L(r,\widetilde{l})}{\partial r}V_R'(\underline{\lambda})\frac{\partial\underline{\lambda}}{\partial\widetilde{l}}+\frac{\partial\pi_H(r,\widetilde{l})}{\partial r}\frac{\partial\mathbb{E}_H\left[V_R(\lambda);p_R\right]}{\partial\widetilde{l}}\end{aligned}$$

Differentiating (30) we have

$$\frac{\partial\mathbb{E}_H\left[V_R(\lambda);p_R\right]}{\partial\widetilde{l}}=\overline{F}_H(\overline{\lambda};p_R)V_R'(\overline{\lambda})\frac{\partial\overline{\lambda}}{\partial\widetilde{l}}+F_H(\underline{\lambda};p_R)V_R'(\underline{\lambda})\frac{\partial\underline{\lambda}}{\partial\widetilde{l}}$$

and using the assumption that $\frac{\partial^2\pi_i(r,\widetilde{l})}{\partial r\partial l}=0$ we can write

$$\frac{\partial B_R(r,(\tilde{r},\tilde{l}))}{\partial\widetilde{l}}=\left(\frac{\partial\pi_R(r,\widetilde{l})}{\partial r}+\frac{\partial\pi_H(r,\widetilde{l})}{\partial r}\overline{F}_H(\overline{\lambda};p_R)\right)V_R'(\underline{\lambda})\frac{\partial\overline{\lambda}}{\partial\widetilde{l}}+\left(\frac{\partial\pi_L(r,\widetilde{l})}{\partial r}+\frac{\partial\pi_H(r,\widetilde{l})}{\partial r}F_H(\underline{\lambda};p_R)\right)V_R'(\underline{\lambda})\frac{\partial\underline{\lambda}}{\partial\widetilde{l}} \tag{31}$$

We now show that $\partial B_R(r,(\tilde{r},\tilde{l}))/\partial\widetilde{l}\leq 0$ so the $R-$SIG's incentives to capture decrease with the anticipated level of capture of the $L-$SIG. Since $\sum_{i\in\{H,R,L\}} \pi_i(r,l)=1$, then $\sum_{i\in\{H,R,L\}} \partial\pi_i(r,l)/\partial r = 0$ and by assumption $0 \geq \frac{\partial \pi_H(r,\widetilde{l})}{\partial r}$ and $0 \geq \frac{\partial \pi_L(r,\widetilde{l})}{\partial r}$. Therefore, we must have $\partial\pi_R(r,l)/\partial r = |\partial\pi_H(r,l)/\partial r| + |\partial\pi_L(r,l)/\partial r|$ so that the first term in parenthesis in (31) is positive while the second is negative. From lemma 2.1, the effect of increasing $L-$capture is to decrease $\overline{\lambda}$ and $\underline{\lambda}$ increase. Therefore, (31) must be negative.

A similar analysis applied to capture by the $L-$SIG shows that $\partial B_L(l,(\tilde{r},\tilde{l}))/\partial\widetilde{r} \leq 0$. □

Proof of Proposition 3. With $\mu=\mu(\lambda,p)$ to simplify notation, we show that under (i), $\partial^2 v_i(\mu)/\partial\lambda\partial p > 0$ so that the $i-$SIG wants to fire up its base, while under (ii) we have $\partial^2 v_i(\mu)/\partial\lambda\partial p < 0$ so that the $i-$SIG wants to demobilize its opposition. First, differentiating $v_i(\mu)$ we have

$$\frac{\partial^2 v_i(\mu)}{\partial\lambda\partial p} = v_i''(\mu)\frac{\partial\mu}{\partial\lambda}\frac{\partial\mu}{\partial p} + v_i'(\mu)\frac{\partial^2\mu}{\partial\lambda\partial p}.$$

We have $\frac{\partial\mu}{\partial\lambda} = \frac{p(1-p)}{(\lambda p+1-p)^2}$, $\frac{\partial\mu}{\partial p} = \frac{\lambda}{(\lambda p+1-p)^2}$ and $\frac{\partial^2\mu}{\partial\lambda\partial p} = \frac{1-p-\lambda p}{(\lambda p+1-p)^3}$ so that

$$\begin{aligned}\frac{\partial^2 v_i(\mu)}{\partial\lambda\partial p} &= v_i''(\mu)\frac{\lambda p(1-p)}{(\lambda p+1-p)^4} + v_i'(\mu)\frac{1-p-\lambda p}{(\lambda p+1-p)^3} \\ &= \frac{\lambda p(1-p)}{(\lambda p+1-p)^4}\left(v_i''(\mu) + \frac{(1-p-\lambda p)(\lambda p+1-p)}{\lambda p(1-p)}v_i'(\mu)\right) \\ &= \frac{\lambda p(1-p)}{(\lambda p+1-p)^4}\left(v_i''(\mu) - K(\mu)v_i'(\mu)\right)\end{aligned}$$

with

$$K(\mu) = \frac{\lambda p}{1-p} - \frac{1-p}{\lambda p} = \frac{\mu}{1-\mu} - \frac{1-\mu}{\mu}$$

the difference between the odds of a good state and a bad state. As $K(\mu)$ is increasing in μ, we have $K(\mu)\in\left[K(\underline{\mu}),K(\overline{\mu})\right]$ with $[\underline{\mu},\overline{\mu}]$ the range of posteriors induced on viewers when consuming the news of a media known to be honest.

Consider first the case of the $R-$SIG. As $v_R'(\mu)>0$, then $\partial^2 v_R(\mu)/\partial\lambda\partial p > 0$ if $\min_{\mu\in[\underline{\mu},\overline{\mu}]}\frac{v_R''(\mu)}{v_R'(\mu)} > \max_{\mu\in[\underline{\mu},\overline{\mu}]} K(\mu) = K(\overline{\mu})$ while $\partial^2 v_R(\mu)/\partial\lambda\partial p < 0$ if $\max_{\mu\in[\underline{\mu},\overline{\mu}]}\frac{v_R''(\mu)}{v_R'(\mu)} < \min_{\mu\in[\underline{\mu},\overline{\mu}]} K(\mu) = K(\underline{\mu})$. Turning to the $L-$SIG, we have $v_L'(\mu)<0$ so that $\partial^2 v_L(\mu)/\partial\lambda\partial p > 0$ if $\min_{\mu\in[\underline{\mu},\overline{\mu}]}\frac{v_L''(\mu)}{|v_L'(\mu)|} > \max_{\mu\in[\underline{\mu},\overline{\mu}]} -K(\mu) = -K(\underline{\mu})$ while $\partial^2 v_L(\mu)/\partial\lambda\partial p < 0$ if $\max_{\mu\in[\underline{\mu},\overline{\mu}]}\frac{v_L''(\mu)}{|v_L'(\mu)|} < \min_{\mu\in[\underline{\mu},\overline{\mu}]} -K(\mu) =$

$-K(\overline{\mu})$. $\square$

Proof of Lemma 3. We can express the odds of the good state as $\mu/(1-\mu) = \lambda p/(1-p)$. Then we have

$$\begin{aligned}\frac{\partial^2 v_R(\mu)}{\partial\lambda\partial p} &= \frac{1}{(1-p)^2}\left(g_R''\left(\frac{\lambda p}{1-p}\right)\frac{\lambda p}{1-p} + g_R'\left(\frac{\lambda p}{1-p}\right)\right) \\ &= \frac{1}{(1-p)^2}\left.\frac{d\left(g_R'(x)x\right)}{dx}\right|_{x=\frac{\lambda p}{1-p}}.\end{aligned}$$

Therefore, if $g_R'(x)x$ is increasing then the $R-$SIG wants to fire up its base, while if $g_R'(x)x$ is increasing he wants to demobilize the opposition. A sufficient condition for $g_R'(x)x$ being increasing is that g_R is convex. The same analysis applies to the $L-$SIG once we observe that

$$\begin{aligned}\frac{\partial^2 v_L(\mu)}{\partial\lambda\partial p} &= \frac{1}{\lambda^2 p^2}\left(g_L''\left(\frac{1-p}{\lambda p}\right)\frac{1-p}{\lambda p} + g_L'\left(\frac{1-p}{\lambda p}\right)\right) \geq 0 \\ &= \frac{1}{\lambda^2 p^2}\left.\frac{d\left(g_L'(x)x\right)}{dx}\right|_{x=\frac{1-p}{\lambda p}}.\end{aligned}$$

$\square$

Proof of Proposition 4. Recall that the *SIGs* expected utility when viewers expect capture levels $\left(\widetilde{r},\widetilde{l}\right)$ is

$$\begin{aligned}\widehat{V}_R\left(r;\left(\widetilde{r},\widetilde{l}\right)\right) &= rV_R(\overline{\lambda}) + \widetilde{l}V_R(\underline{\lambda}) + (1-r-\widetilde{l})\mathbb{E}_H\left[V_R(\lambda);p_R\right] - C_R(r), \\ \widehat{V}_L\left(l;\left(\widetilde{r},\widetilde{l}\right)\right) &= \widetilde{r}V_R(\overline{\lambda}) + lV_R(\underline{\lambda}) + (1-\widetilde{r}-l)\mathbb{E}_H\left[V_R(\lambda);p_R\right] - C_R(r),\end{aligned}$$

with $\overline{\lambda}$ and $\underline{\lambda}$ consistent with $\left(\widetilde{r},\widetilde{l}\right)$–i.e., satisfying (4) and (5). In equilibrium, we must have

$$\begin{aligned}r^* &\in \underset{r\in[0,1]}{\arg\max}\widehat{V}_R(r;(r^*,l^*)), \\ l^* &\in \underset{l\in[0,1]}{\arg\max}\widehat{V}_L(l;(r^*,l^*)).\end{aligned}$$

Using (9) and (10) we can express these equilibrium conditions as (13) and (14). As citizens correctly anticipate (r^*, l^*), then (4) and (5) would capture the equilibrium maximum and minimum likelihood ratio. $\square$

Proof of Proposition 5. Recall that for each viewer with prior p and threshold α, $\lambda_{crit} = \frac{\alpha}{1-\alpha}\frac{1-p}{p}$ is the minimum informational content of a message that would lead her to act. We first derive the instrumental value of a (p, λ_{crit}) −viewer from a captured channel, and then study how the difference in instrumental values between outlets 1 and 2 varies with λ_{crit}.

Let $F^i_\lambda(\lambda, p)$, $i \in \{1, 2\}$, be the perceived equilibrium distribution of likelihood ratios of media i by a (p, λ_{crit}) −viewer–see (7)–and $F^i_\mu(\mu, p) \equiv F^i_\lambda(\frac{\mu}{1-\mu}\frac{1-p}{p}, p)$ be the distribution of posterior beliefs after she consumes the news of media i. The instrumental value of that viewer if $\lambda_{crit} < 1$ (so $p > \alpha$) is

$$
\begin{aligned}
W^i_I &\equiv \int_0^\alpha [\alpha(1-\mu) - (1-\alpha)\mu] dF^i_\mu(\mu, p) = \int_0^\alpha [\alpha - \mu] dF^i_\mu(\mu, p) = \int_0^\alpha F^i_\mu(\mu, p) d\mu \\
&= \int_0^{\lambda_{crit}} F^i_\lambda(\lambda, p) \frac{p(1-p)}{(1-p+\lambda p)^2} d\lambda,
\end{aligned}
$$

where we made the change of variables $\lambda = \frac{\mu}{1-\mu}\frac{1-p}{p}$ to obtain the last term. This follows as the viewer will change her decision from $a = 1$ to $a = 0$ only after observing a message that leads her to a posterior belief $\mu \leq \alpha$–i.e., a message with $\lambda \leq \lambda_{crit}$. Equivalently, if $\lambda_{crit} > 1$ (so $p < \alpha$) we have

$$
\begin{aligned}
W^i_I &\equiv \int_\alpha^1 [(1-\alpha)\mu - \alpha(1-\mu)] dF^i_\mu(\mu, p) = \int_\alpha^1 [\mu - \alpha] dF^i_\mu(\mu, p) = \int_\alpha^1 \overline{F}^i_\mu(\mu, p) dp \\
&= \int_{\lambda_{crit}}^1 \overline{F}^i_\lambda(\lambda, p) \frac{p(1-p)}{(1-p+\lambda p)^2} d\lambda.
\end{aligned}
$$

Let $\Delta_F(\lambda, p) = F^1_\lambda(\mu, p) - F^2_\lambda(\mu, p)$ be the difference in the equilibrium distribution of likelihood ratios between media 1 and media 2, and $\Delta_W(\lambda_{crit}, p) \equiv W^1_I(p, \lambda_{crit}) - W^2_I(p, \lambda_{crit})$ be the difference in instrumental value between both media outlets. Then, the (p, λ_{crit}) −viewer with $\lambda_{crit} < 1$ will watch media 1 whenever

$$\Delta_W(\lambda_{crit}, p) = \int_0^{\lambda_{crit}} \Delta_F(\lambda, p) \frac{p(1-p)}{(1-p+\lambda p)^2} d\lambda \geq 0$$

and will watch media 2 otherwise. Similarly, a (p, λ_{crit}) −viewer with $\lambda_{crit} > 1$ will watch media 1 if

$$\Delta_W(\lambda, p) = \int_{\lambda_{crit}}^1 (-\Delta_F(\lambda, p)) \frac{p(1-p)}{(1-p+\lambda p)^2} d\lambda \geq 0.$$

Suppose $r^1 \geq l^1$, $l^2 \geq r^2$, and (19) holds so capture levels are not too dissimilar. We now show that we must have

$$\overline{\lambda}_1 < \overline{\lambda}_2 \; and \; \underline{\lambda}_2 > \underline{\lambda}_1, \tag{32}$$

that is the highest equilibrium likelihood ratio is smaller in the right-dominated media while the lowest one is higher in the left-dominated media. Note first that (19) implies that $\frac{r^1}{1-(r^1+l^1)} > \frac{r^2}{1-(r^1+l^1)}$ and $\frac{l^2}{1-(r^1+l^1)} > \frac{l^1}{1-(r^1+l^1)}$, i.e., the likelihood that the message is sent by the $R-$SIG rather than the honest sender is higher in media 1, while the likelihood that the message is sent by the $L-$SIG rather than the honest sender is higher in media 2. As $F_H^1 = F_H^2 \, (= F_H)$ so that $F_{H,-1}^1(\lambda) = F_{H,-1}^2(\lambda)$, (4) and (5) imply (32).

Given symmetry of the channel and the relation between the maximum and minimum likelihood ratios (32), we can write $\Delta_F(\lambda, p)$ as

$$\Delta_F(\lambda, p) = \begin{cases} 0 & if \quad \lambda < \underline{\lambda}_1 \\ (1-(r^1+l^1)) \, F_H(\lambda, p) + l^1 & if \quad \underline{\lambda}_1 \leq \lambda < \underline{\lambda}_2 \\ ((r^2+l^2)-(r^1+l^1)) \, F_H(\lambda, p) - (l^2 - l^1) & if \quad \underline{\lambda}_2 \leq \lambda < \overline{\lambda}_1 \\ 1-(1-(r^2+l^2)) \, F_H(\lambda, p) - l^2 & if \quad \overline{\lambda}_1 \leq \lambda < \overline{\lambda}_2 \\ 0 & if \quad \lambda \geq \overline{\lambda}_2 \end{cases}$$

Note that $\Delta_F(\lambda, p) \geq 0$ if $\lambda < \underline{\lambda}_2$ or if $\overline{\lambda}_1 \leq \lambda$. Therefore, $\Delta_W(\lambda_{crit}, p) \geq 0$ if $\lambda_{crit} < \underline{\lambda}_2$ but $\Delta_W(\lambda_{crit}, p) \leq 0$ if $\lambda_{crit} > \overline{\lambda}_1$. This proves part i.

Suppose, in addition, that $r^1 + l^1 = r^2 + l^2$. Then $\Delta_F(\lambda, p) = -(l^2 - l^1)$ in $\underline{\lambda}_2 \leq \lambda < \overline{\lambda}_1$ which does not change sign for $\lambda \in [\underline{\lambda}_2 \leq \lambda < \overline{\lambda}_1]$. We now show that this implies that $\Delta_W(\lambda_{crit}, p)$ is strictly single-crossing in λ_{crit},which proves part *ii*. Note that, for $\lambda_{crit} < 1$, $\Delta_W(\lambda_{crit}, p)$ must be single-crossing, from positive to negative, as $\Delta_F(\lambda, p)$ changes sign at most once from positive to negative (i.e., at $\lambda_{crit} = \underline{\lambda}_2$ if $l^2 > l^1$). Likewise, for $\lambda_{crit} > 1$, $\Delta_W(\lambda_{crit}, p)$ must be single-crossing, from positive to negative as $\Delta_F(\lambda, p)$ changes sign at most once, from negative to positive–i.e., at $\lambda_{crit} = \overline{\lambda}_1$ if $l^2 > l^1$. Continuity of $\Delta_W(\lambda_{crit}, p)$ at $\lambda_{crit} = 1$ implies that the sign of $\Delta_W(\lambda_{crit}, p)$ must not change for either $\lambda_{crit} < 1$ or $\lambda_{crit} > 1$, proving that $\Delta_W(\lambda_{crit}, p)$ is single-crossing in λ_{crit}. □

Proof of Proposition 6. The functional forms of v_R and v_L guarantee that both SIGs want to fire up their base–see Lemma 3. From the proof of Proposition 5, the instrumental value of media 1 relative to media 2, $\Delta_W(\lambda_{crit}, p) = W_I^1(p, \lambda_{crit}) - W_I^2(p, \lambda_{crit})$ is positive for $\lambda_{crit} \leq \underline{\lambda}_2$ and negative for $\lambda_{crit} \geq \bar{\lambda}_1$. As the distribution of viewers (p, λ_{crit}) is such that $\lambda_{crit} \leq \underline{\lambda}_2$ if $p > \frac{1}{2}$ and $\lambda_{crit} \geq \bar{\lambda}_1$ if $p < \frac{1}{2}$, this implies that all viewers with $p > 1/2$ ($p < 1/2$) prefer to consume media 1 (media 2) if their preference is driven by instrumental value, at the current equilibrium levels of capture.

Suppose that there is an exogenous increase in γ and let $F_p^{>1/2}(p) = \Pr[p' \leq p | p' \geq 1/2]$ and $F_p^{<1/2}(p) = \Pr[p' \leq p | p' < 1/2]$ be the distribution of priors of viewer with $p \geq \frac{1}{2}$ and $p < \frac{1}{2}$, respectively. At the current capture levels and since $\vartheta^1 = \vartheta^2 = \frac{1}{2}$, the distribution of viewers watching media 1 and media 2 are

$$F_p^1(p) = \gamma F_p^{>1/2}(p) + (1 - \gamma) F_p(p), \tag{33}$$

$$F_p^2(p) = \gamma F_p^{<1/2}(p) + (1 - \gamma) F_p(p). \tag{34}$$

This follows as viewers that sort on entertainment value are equally likely to choose either media, while those that sort according to instrumental value sort according to their prior. Note also that, given $F_p(\frac{1}{2}) = \frac{1}{2}$, the total mass of viewers in both media is the same.

An increase in γ has two effects in equilibrium: it affects viewers sorting and changes the SIGs incentives to capture given this sorting. Suppose that in the new equilibrium we have $\bar{\lambda}'_1 \leq \bar{\lambda}_1$ and $\underline{\lambda}'_2 \geq \underline{\lambda}_2$, that is the maximum likelihood ratio in the R-dominated media did not increase while the minimum likelihood ratio in the L-dominated media did not decrease–we will later confirm that this is indeed the case. Then, viewers incentives to sort on instrumental value did not change as all $p < \frac{1}{2}$ ($p > \frac{1}{2}$) satisfy $\lambda_{crit} \geq \bar{\lambda}_1 \geq \bar{\lambda}'_1$ ($\lambda_{crit} \leq \underline{\lambda}_2 \leq \underline{\lambda}'_2$) and thus derive no value from consuming media 1 (2). As the sign of the relative instrumental value of media 1 relative to media 2 has not changed for any consumer, then increasing γ leads to a FOSD increase in $F^1_p(p)$ in (33) and a FOSD increase in $F^2_p(p)$ in (34). As both SIGs want to fire up its base, this increases the R-SIG incentives to capture media 1, and lowers its incentives to capture media 2, while it decreases the L-SIG incentives to capture media 2, and lowers its incentives to capture media 1. Strategic substitutability then implies that the equilibrium level of $R-$capture increases in media 1, and decreases in media 2, while L-capture increases in media 2 and decreases in media 1. As long as the first effect dominates the second, we then have that $\frac{r^1}{1-r^1-l^1}$ and $\frac{l^2}{1-r^2-l^2}$ leading to $\bar{\lambda}'_1 \leq \bar{\lambda}_1$ and $\underline{\lambda}'_2 \geq \underline{\lambda}_2$–see Proposition 1-3. □

Proof of Proposition 7. Suppose that the $R-$SIG and $L-$SIG's strategies are $\tau_R(\lambda)$ and $\tau_L(\lambda)$. Then, the perceived likelihood ratio by sophisticated viewers, $\lambda_\gamma(\lambda) \equiv \frac{\Pr[\lambda|\theta=1]}{\Pr[\lambda|\theta=0]}$, is

$$\lambda_\gamma(\lambda) = \frac{(1-l-r)p_1(\lambda) + r\tau_R(\lambda) + l\tau_L(\lambda)}{(1-l-r)p_{-1}(\lambda) + r\tau_R(\lambda) + l\tau_L(\lambda)}, \tag{35}$$

while the $i-$SIG expected utility from a message that is interpreted as λ is $V_i(\lambda)$ as given by (8). Then, the expected utility of the $i-$SIG when sending a message with literal meaning λ is

$$\tilde{V}_i(\lambda) \equiv (1-\gamma)\, V_i(\lambda) + \gamma V_i(\lambda_\gamma(\lambda)). \tag{36}$$

SIGs' optimality requires that if $\lambda, \lambda' \in supp\, \tau_i$, then $\tilde{V}_i(\lambda) = \tilde{V}_i(\lambda')$, $i \in \{L, R\}$. We now show that if the distribution $F_H(\lambda)$ is continuous, then (i) $supp\, \tau_i$ is an interval of the form $supp\, \tau_R = [\overline{\lambda}, \lambda_{max}]$ and $supp\, \tau_L = [\lambda_{min}, \underline{\lambda}]$, (ii) $\lambda_\gamma(\overline{\lambda}) = \overline{\lambda}$ and $\lambda_\gamma(\underline{\lambda}) = \underline{\lambda}$, and (iii) λ_γ must satisfy (21) for any level of capture given $\overline{\lambda}$ and $\underline{\lambda}$.

First, suppose that $F_H(\lambda)$ is a continuous distribution with convex support $supp\, F_H$ and let $\overline{\lambda} \equiv \max\{\lambda : \lambda_\gamma(\lambda) = \lambda, \lambda \in supp\, F_H\}$ be the highest news that sophisticated viewers interpret at face value. Since $\lambda_\gamma(\lambda) \neq \lambda$ implies that $\lambda \in supp\, \tau_R \cup \tau_L$, we must have $\min\{\lambda : \lambda \in supp\, \tau_R\} \leq \overline{\lambda}$.

We show that $\min\{\lambda : \lambda \in supp\, \tau_R\} = \overline{\lambda}$. Suppose by contradiction that $\min\{\lambda : \lambda \in supp\, \tau_R\} < \overline{\lambda}$. Then the $R-$SIG obtains utility $\tilde{V}_i(\overline{\lambda}) = V_i(\overline{\lambda})$ from $\overline{\lambda}$, while any $\lambda' \in (\min\{\lambda : \lambda \in supp\ \tau_R\}, \overline{\lambda})$ gives strictly less utility as $\tilde{V}_i(\lambda') \leq V_i(\lambda') < V_i(\overline{\lambda})$. Thus, the $R-$SIG can improve by sending instead $\overline{\lambda}$, thus reaching a contradiction. A similar argument applied to the $L-$SIG implies that $supp\ \tau_L = [\lambda_{min}, \underline{\lambda}]$ and $\lambda_\gamma(\underline{\lambda}) = \underline{\lambda}$. Finaly, we obtain (21) by solving for $\lambda_\gamma(\lambda)$ in

$$\begin{aligned}(1-\gamma)\, V_L(\lambda) + \gamma V_L(\lambda_\gamma(\lambda)) &= V_L(\underline{\lambda}) \quad \text{if } \lambda \leq \underline{\lambda},\\ (1-\gamma)\, V_R(\lambda) + \gamma V_R(\lambda_\gamma(\lambda)) &= V_R(\overline{\lambda}) \quad \text{if } \lambda \geq \overline{\lambda}.\end{aligned}$$

Note that the equilibrium interpretation (21) depends on $\overline{\lambda}$ and $\underline{\lambda}$. These are pinned down in equilibrium by the condition that each SIGs probability of sending each potential lie aggregate to one. Solving for $\tau_R(\lambda)$ and $\tau_L(\lambda)$ in (35)

$$\begin{aligned}\frac{r}{1-l-r}\tau_R(\lambda) &= \frac{\lambda - \lambda_\gamma(\lambda)}{\lambda_\gamma(\lambda) - 1} p_{-1}(\lambda),\\ \frac{l}{1-l-r}\tau_L(\lambda) &= \frac{\lambda_\gamma(\lambda) - \lambda}{1 - \lambda_\gamma(\lambda)} p_{-1}(\lambda),\end{aligned}$$

and integrating these expressions over the respective supports we obtain (22) and (23).

To complete the proof, we write (21) as $\lambda_\gamma(\lambda; \overline{\lambda}, \underline{\lambda})$ to make explicit the dependence on $(\overline{\lambda}, \underline{\lambda})$ and define

$$\overline{w}(\overline{\lambda}) \equiv \int_{\overline{\lambda}}^{\infty} \frac{\lambda - \lambda_\gamma(\lambda; \overline{\lambda}, \underline{\lambda})}{\lambda_\gamma(\lambda; \overline{\lambda}, \underline{\lambda}) - 1} dF_{H,-1}(\lambda), \tag{37}$$

$$\underline{w}(\underline{\lambda}) \equiv \int_{0}^{\underline{\lambda}} \frac{\lambda_\gamma(\lambda; \overline{\lambda}, \underline{\lambda}) - \lambda}{1 - \lambda_\gamma(\lambda; \overline{\lambda}, \underline{\lambda})} dF_{H,-1}(\lambda). \tag{38}$$

First, we show that $\lambda_\gamma(\lambda; \overline{\lambda}, \underline{\lambda})$ is monotonic in $(\overline{\lambda}, \underline{\lambda})$. Indeed, as V_R is strictly increasing (and V_L strictly decreasing), then $V_R(\overline{\lambda}) + \frac{1-\gamma}{\gamma}(V_R(\overline{\lambda}) - V_R(\lambda))$ increases in $\overline{\lambda}$ and decreases in γ for any $\lambda > \underline{\lambda}$; similarly, $V_L(\underline{\lambda}) + \frac{1-\gamma}{\gamma}(V_L(\underline{\lambda}) - V_L(\lambda))$ decreases in $\underline{\lambda}$ and increases in γ for any $\lambda < \underline{\lambda}$. Looking at (21) we conclude that, for a fixed value of λ, $\lambda_\gamma(\lambda; \overline{\lambda}, \underline{\lambda})$ is non-increasing in $\overline{\lambda}$ and non-decreasing in $\underline{\lambda}$.

Second, we will make use of the fact that $\frac{\lambda - x}{x-1}$ is decreasing in x for $1 < x < \lambda$, while $\frac{x-\lambda}{1-x}$ is decreasing in x for $\lambda < x < 1$. This fact and the monotonicity of $\lambda_\gamma(\lambda; \overline{\lambda}, \underline{\lambda})$ in $(\overline{\lambda}, \underline{\lambda})$ imply that $\overline{w}(\overline{\lambda})$ in (37) is an strictly decreasing function of $\overline{\lambda}$ with $\overline{w}(\lambda_{max}) = 0$ while $\underline{w}(\underline{\lambda})$ in (38) is an strictly increasing function of $\underline{\lambda}$ with $\underline{w}(\lambda_{min}) = 0$. Furthermore, conditions (22) and (23) translate to $\overline{w}(\overline{\lambda}) = r/(1-r-l)$ and $\underline{w}(\underline{\lambda}) = l/(1-r-l)$. We can then establish uniqueness: As the left hand side of (22) is an strictly decreasing function of $\overline{\lambda}$ and the left hand side of (23) is strictly increasing function of $\underline{\lambda}$, a unique solution to (22-23) is guaranteed for every r and l.

Finally, increasing r or l raises the right hand side of (22) and (23) leading to a lower $\overline{\lambda}$ and higher $\underline{\lambda}$. Likewise, increasing γ lowers both $\overline{w}(\overline{\lambda})$ and $\underline{w}(\underline{\lambda})$, leading to a lower equilibrium $\overline{\lambda}$ and higher $\underline{\lambda}$. □

Proof of Proposition 8. Suppose that citizens anticipate a level of capture of $\left(\tilde{r}, \tilde{l}\right)$. Therefore, the $R-$SIG's expected utility when investing r in covertly capturing media is

$$\widehat{V}_R\left(r; \left(\widetilde{r}, \widetilde{l}\right)\right) = r\tilde{V}_R(\overline{\lambda}) + \widetilde{l}\mathbb{E}_{\tau_L}\left[\tilde{V}_R(\lambda); p_R\right] + (1-r-l)\mathbb{E}_H\left[\tilde{V}_R(\lambda); p_R\right] - C_R(r).$$

with

$$\mathbb{E}_H\left[\tilde{V}_R(\lambda); p_R\right] = \overline{F}_H(\overline{\lambda}; p_R)V_R(\overline{\lambda}) + \int_{\lambda_{min}}^{\overline{\lambda}} \left((1-\gamma)\, V_R(\lambda) + \gamma V_R(\lambda_\gamma(\lambda))\right) dF_H(\lambda; p_R). \tag{39}$$

Therefore, the $R-$SIG 's marginal gain from covertly increasing media capture $B_R(r, (\tilde{r}, \tilde{l})) := \frac{\partial \widehat{V}_R\left(r; \left(\widetilde{r}, \widetilde{l}\right)\right)}{\partial r}$ is

$$\begin{aligned} B_R(r, (\tilde{r}, \tilde{l})) &= V_R(\overline{\lambda}) - \mathbb{E}_H\left[V_R(\lambda); p_R\right] \\ &= \int_{\lambda_{min}}^{\overline{\lambda}} \left(V_R(\overline{\lambda}) - V_R(\lambda)\right) dF_H(\lambda; p_R) \qquad (40) \\ &\quad - (1-\gamma) \int_{\lambda_{min}}^{\underline{\lambda}} \left(V_R(\lambda_\gamma(\lambda)) - V_R(\lambda)\right) dF_H(\lambda; p_R) \qquad (41) \end{aligned}$$

By increasing capture efforts, the $R-$SIG obtains $V_R(\overline{\lambda})$ instead of the utility derived from an honest coverage $\mathbb{E}_H\left[V_R(\lambda); p_R\right]$.Thus, the $R-$SIG gains $V_R(\overline{\lambda}) - V_R(\lambda)$ whenever $\lambda \leq \overline{\lambda}$ and all viewers (including sophisticated ones) interpret the message at face value (this is (40)) except when $\lambda \leq \underline{\lambda}$ and sophisticated viewers discount the news (this is (41)).

We now show that $\partial B_R(r,(\tilde{r},\tilde{l}))/\partial\widetilde{l} \leq 0$ so the $R-$SIG's incentives to capture decrease with the anticipated level of capture of the $L-$SIG. First, part 3 of Proposition 7 shows that $\overline{\lambda}$ decreases with l, so (40) decreases with $\tilde{l}$. Moreover, part 3 of Proposition 7 also shows that increasing l, (i) increases $\lambda_\gamma(\lambda)$ for $\lambda \leq \underline{\lambda}$, and (ii) increases $\underline{\lambda}$. Both effects raise the value of the integral in (41), thus decreasing (41). Therefore, increasing $\widetilde{l}$ lowers $B_R(r,(\tilde{r},\tilde{l}))$. A similar analysis applied to capture by the $L-$SIG shows that $\partial B_L(l,(\tilde{r},\tilde{l}))/\partial\widetilde{r} \leq 0$. □