# Málaga Economic Theory Research Center Working Papers

## A Theory of Media Self-Silence

Ascensión Andina-Díaz and José A. García-Martínez

WP 2015-5
November 2015

Departamento de Teoría e Historia Económica
Facultad de Ciencias Económicas y Empresariales
Universidad de Málaga
ISSN 1989-6908

# A theory of media self-silence*

Ascensión Andina-Díaz† and José A. García-Martínez‡

November 24, 2015

**Abstract**

This paper proposes a theory of media self-silence. The argument is that news organizations have the power to raise public concern and so affect the probability that there is ex-post verification of the true state of the world. Built on the literature of career concerns, we consider a newspaper that seeks to maximize its reputation for high quality. Our results predict more media silence, the higher the prior expectations on the quality of the firm, the greater the probability of ex-post verification, and the higher the power of the newspaper to lead public opinion. We also obtain that the greater the social influence of a news organization, the stricter the firm's vetting process for stories is. Last, competition reduces media silence.

**Keywords:** Feedback power; reputation; quality; competition; media silence

**JEL:** D72; D82

## 1 Introduction

In the last decade, scholars have devoted much attention to the issue of media bias.[1] Not that much to the question of media silence. Though more difficult to measure than other classes of media bias, anecdotal evidence suggests that whereas some news organizations are extremely careful about printing scoops, others do not hesitate a second and run almost every piece of news that arrives to the newsroom. The Lewinsky scandal and the story of bin Laden's death present two good examples to show these differences in media behavior.

*We thank Miguel A. Meléndez-Jiménez for valuable suggestions. We also thank Emilio Calvano, Daniel Cardona, Kenan Huremovic, Antonio Morales, Bernardo Moreno, Martin Peitz, Socorro Puy, Javier Rodero and seminar participants at the University of Málaga, University Milan-Bicocca, University of Illes Balears, JEI 2015 Alicante and ASSET 2015 Granada for useful comments. We gratefully acknowledge the financial support from the Ministerio de Educación y Ciencia through projects ECO2010-21624 and ECO2011-26996, the Ministerio de Economía y Competitividad through project MTM2014-54199-P and the Junta de Andalucía through project SEJ2011-8065. Andina-Díaz also acknowledges the hospitality of Royal Holloway, University of London, where part of this research was carried out. The usual disclaimer applies. This paper previously circulated under the title "Media silence, feedback power and reputation".

†Dpto. Teoría e Historia Económica, Universidad de Málaga, Málaga, Spain. E-mail: aandina@uma.es

‡Dpto. Estudios Económicos y Financieros, Universidad Miguel Hernández, Elche, Spain. E-mail: jose.garciam@umh.es

[1]See surveys by Prat and Strömberg (2011), Gentzkow et al. (2014) and Andina-Díaz (2011).

The first story goes back to January 1998, when Mark Whitaker, the *Newsweek*'s editor at that time, decided not to run the Lewinsky Story that his reporter, Michael Isikoff, had been pursuing for nearly a year. In reference to why he did not publish the story, Mark Whitaker admitted in an interview to *CNN* in November, 2011: "*We didn't fell that we were on firm enough ground to report a story that would be about accusing the president [...]. If we had gotten that wrong could have been [...] a mortal blow to Newsweek's reputation.*" The story that belonged to *Newsweek* was finally published in the Internet by *Drudge Report*, a far less influential outlet than the prestigious magazine *Newsweek*. Despite it, the news hit Internet new groups and the *Drudge Report* web site had thousands of visits. Three days after, *The Washington Post* broke it.[2]

The second story is about the investigative reporter Seymour M. Hersh and his article "The Killing of Osama bin Laden". Hersh, who in the seventies won the Pulitzer Prize for exposing the My Lai Massacre during the Vietnam War and has written other several influential articles, started to investigate the official story of bin Landen's death just a couple of months after the US operation, in May, 2011. More than three years later, he sent a draft of his report to *The New Yorker*. Despite Hersh's strong ties to the magazine, where he is a regular contributor, the *The New Yorker*'s editor, David Remnick, told Hersh that he didn't think he had "the story nailed down" and suggested him to continue his investigation. Instead, Hersh gave the story to *The London Review of Books*, where it was published in May 2015. According to Jonathan Mahler: "*The bin Laden report wasn't the first one by Hersh that Remnick rejected because he considered the sourcing too thin [...] In 2013 and 2014, he passed on two Hersh articles [...] Those articles also landed in The London Review of Books.*"[3]

At the light of these stories, we can tell that news organizations differ in the treatment of scoops. How can we explain such differences in media behavior? Why did neither *Newsweek* nor *The New Yorker* decide to run the articles in their pages, whereas other media outlets found no impediment to publish them? Is it the case that there is something in common between *Newsweek* and *The New Yorker* on the one hand, and *Drudge Report* and *The London Review of Books* on the other hand? If so, what is that?

In this paper we propose a theory of media silence that explains the observed behavior. The interest of this theory is that it only requires newspapers to be concerned with reputation. This is new in the literature, as previous research explains media silence by means of institutional features. Two arguments have been used so far to explain the decision of a newspaper to hold information: Media captured, either by the government or by advertisers (see the theoretical works by Besley and Prat (2006) and Ellman and Germano (2009)); and the existence of defamation lawsuits and/or physical threats to journalists (see Garoupa (1999*a*,*b*) and Stanig (2014)). Beyond these arguments, whose importance is entirely justified, we propose a new reason to explain

[2]See "Scandalous scoop breaks online", *BBC News* January 30, 1998; and "Former Newsweek Editor on Why He didn't Run Lewinsky Story: 'We Didn't Feel We Were on Firm Enough Ground"', *NewsBuster*, November 6, 2011.

[3]See "What Do We Really Know About Osama bin Laden's Death?", written by Jonathan Mahler, in *The New York Times*, October 18, 2015.

media silence. The novelty of our approach is that we argue at the firm level, and this allows us to explain variations in media silence in news organizations that compete under the same rules. In this sense, we talk of media *self-silence*.

At the core of our model is the idea that news organizations have the power to raise public concern and so affect the probability that there is ex-post verification of the true state of the world. It means that news organizations will have in our model, as in real world, the capacity to affect *feedback*, that is, the power to influence the probability that consumers learn the true state. This is new in this literature. Indirectly, it means that a newspaper can affect its reputation. The argument is that a news organization that turns the spotlight on, let us say, a possible corruption scandal, may raise public concern about the consequences of the fraud, may eventually induce a citizen or institution to denounce the facts and take the case to court, which may result in the judge passing sentence and thus, indirectly, determining whether the media's story was true or just another example of a "Jimmy's World" fabrication.[4] On the contrary, a country in which news organizations give no room to scoops on their front pages, but rather print news items on the usual events of a society (economy, politics, sports, etc.), silences citizens and precludes learning.

Interestingly, this theory helps us explain why both *Newsweek* and *The New Yorker* decided to hold the Lewinsky and the bin Laden's death stories, respectively, whereas *Drudge Report* and *The London Review of Books* ran them. More generally, it provides an argument to explain why in competitive environments renown newspapers and magazines, such as *The New York Times*, *The Washington Post* or *The Guardian*, have a very stringent vetting process for stories; whereas smaller newspapers, lacking the power to influence public opinion, are less strict with the quality of their sources, which makes them more prone to print scandals.

The model is as follows. We consider a newspaper that seeks to build a reputation for quality. The newspaper receives an informative signal on the existence (or not) of a corruption scandal in the economy. The media firm is risk neutral and its only concern is to appear competent. The newspaper can take two actions: (i) to publish the scandal and (ii) not to publish the scandal and print instead easy-to-cover stories. Each action corresponds to one of the two possible states of the world: the malpractice does/does not exist. The key assumption is that actions are different in terms of consequences. That is, whereas one action (letting the scandal go in the paper) activates the feedback with positive probability, taking the other action (holding the story and printing instead easy-to-cover stories) guarantees the newspaper a smaller probability that

[4] In reference to a false story written by Janet Cooke, that was front-page in the *Washington Post* on September 29, 1980. Cooke, who was even given the Pulitzer Prize for this article, subsequently confessed the story was false. The confession was printed in the *Post* on April 16, 1981. This malpractice obliged the *Washington Post* to offer numerous explanations and apologies, as well as to publicly return the Pulitzer, to make personnel changes in the newspaper and, naturally, to fire Cooke. More recently, *The New York Magazine* printed on the December 15, 2014, the story of Mohammed Islam, who claimed had won $72 million trading on the stock market. This story turned into a major international news item. However, just one day after, *The New York Observer* published an interview with Islam, who admitted he had previously lied. *The New York Magazine* retracted the story and apologized, concluding: "We were duped. Our fact-checking process was obviously inadequate; we take full responsibility and we should have known better. New York apologizes to our readers."

consumers learn the truth.[5] Our model thus partially endogeneizes the probability of feedback.[6]

We start the analysis with a benchmark case consisting of a monopoly newspaper that operates in an economy in which two states of the world are equiprobable. Our results show that, in the unique equilibrium, the newspaper does not always act on its information, but it chooses a mixed strategy that silences corruption signals with positive probability and thus, it prints, too often, easy-to-cover stories. Interestingly, we obtain that the greater the (prior) expectations on the newspaper being high quality, the more corruption signals it chooses to omit. Similarly, the higher the probability of feedback, the greater the media silence will be. These effects are so strong that they also hold for newspapers with high quality signals. At first sight, it can be a bit puzzling. In fact, these are newspapers that if choosing to print a scandal, they will quite likely bring people down and consequently, will boot their reputation. Then, why do news organizations decide to remain silent? To see it, note that remaining silent ensures for the newspaper that consumers will never know about the omitted story. In contrast, covering a corruption scandal can result in feedback, in which case the state of the world will be realized and citizens will count on hard evidence to build the newspaper's reputation. This means that printing a corruption scandal is the only way for a firm to improve its reputation, but since the scandal can turn out to be false, it also means that the firm may end up being proven wrong. Our results show that the expected payoff to a newspaper from printing a scandal is not enough to completely offset the payoff from reporting easy-to-cover stories. Thus, in equilibrium, risk-neutral news organizations hold scandals. Now, note that the benefit of being silent increases in the consumers' (prior) expectations on the quality of the firm. In fact, with high expectations, a newspaper can guarantee a high reputation even if publishing easy-to-cover stories. Accordingly, the firm reacts choosing to silence more stories. Similarly, an increase in the probability of feedback also reduces the incentive for a newspaper to cover a scandal. Again, the result is more media silence.

Our second main result is that we can relax the assumption that the two states of the world are equiprobable and be certain that previous conclusions hold. There is only one exception: when the prior distribution that the state is corrupt is too strong. The reason is that in this case, consumers are so biased in their prior beliefs that they rate any newspaper that contradicts their priors as low quality. This is the classical "herding on the prior" effect. An argument that, because of the counterbalanced forced that the endogenous feedback introduces in our model, is here only strong enough to drive media behavior when priors are sufficiently high.

Finally, we extend the model to account for competition. Our results show that competition disciplines news organizations with higher quality signals, but not those with lower quality, that even in the presence of

[5] This probability is zero in the basic model with a monopoly newspaper (Section 2), and positive when we introduce competition (Section 3).

[6] The literature on experts and effort choice has also considered situations in which the probability of ex-post verification of the true state may depend on the action chosen by the agent. See Milbourn et al. (2001) or Suurmond et al. (2004). The idea behind these papers is the implementation of a *de novo* project, where success or failure can only be observed if the project is implemented (in which case, ex-post verification of the state always occurs with probability one).

more media firms, in equilibrium use a mixed strategy that silence corruption signals with positive probability. The analysis also serves us to analyze the effect of a change in the marginal impact of a newspaper on the probability that consumers learn the state of the world. We refer to this measure of social power as the *feedback power* of a firm. To this respect, our model predicts more revelation of information by newspapers that, because of an increase in the competitiveness of the media industry, see their feedback power reduced. In the limit, tough competition disciplines all news organizations. This prediction gives support to the empirical evidence that countries with a competitive media industry have more active newspapers and better informed consumers (see Djankov et al. (2003), Prat and Stromberg (2005) and Gentzkow et al. (2006)). The interest of the result is that in our model it is exclusively driven by a pure reputational concern, originated at the firm level. Additionally, we obtain that ceteris paribus, the greater the social influence of a news organization, the more stringent the firm's vetting process for stories is, then the higher its silence. Thus, according to our theory, greater social influence drives newspapers to be stricter with fact-checking, which results in more corruption signals being silenced, but also in a higher accuracy of printed news articles. As already mentioned, this is is line with the casual observation that big powerful newspapers tend to be very cautious about publishing scoops, whereas smaller newspapers with little feedback power, are usually more prone to print scandals.

The closest paper to ours is Gentzkow and Shapiro (2006). They propose a model in which a newspaper seeks to build a reputation for quality and the consumers' prior expectations are in favor of one state of the world. This drives media bias which, in their model, originates in the incentive of the newspaper to slant its reports towards the consumers' popular beliefs. Formally, our paper is also related to Prat (2005), who first showed that an increase in the transparency (on actions) can have detrimental effects. In his model, however, increasing the transparency of consequences (the kind of transparency we talk about in our paper) can only be beneficial. The present work challenges this view. Also related is the work by Fox and Weelden (2012), who obtain that when the prior on the state is too unbalanced, transparency of consequences increases the incentive for the expert to stick more often to the prior. Interestingly, if costs of mistakes are asymmetric, this can decrease the principal's expected welfare.[7]

Topically, our paper belongs to the blooming literature on media economics, and more particularly, it contributes to the analysis of the origins of media bias. Much has been said in this respect. The numerous explanations to date have been grouped into two categories. On the one hand, the supply-side arguments, that account for reasons such as media ownership (Bovitz et al. (2002) and Djankov et al. (2003)), cost structure (see Stromberg (2004)), advertisers and interest groups (Corneo (2006) and Ellman and Germano (2009)), career concerns of journalists (Baron (2006)), government capture (Besley and Prat (2006), Egorov et al. (2009) and Durante and Knight (2012)) or government advertising (Tella and Franceschelli (2011)). On the other hand, there is the demand driven forces, that consider reasons that originate in the consumers' preferences for certain

[7] That reputation can have perverse effects has also been shown in other contexts. See Levy (1997), Morris (2001), Ottaviani and Sorensen (2001) or Hörner (2002).

stories (Mullainathan and Shleifer (2005) and Latham (2015)), or the consumers' prior beliefs (Gentzkow and Shapiro (2006)). The present paper contributes to the latter class of literature, by pointing out that the media's ability to determine what consumers get to know can also result in media bias and, more precisely, in media self-silence.

The rest of the paper is organized as follows. In Section 2, we present the basic model with a monopoly newspaper and show the results. We also analyze the case of an unbalanced prior. In Section 3, we extend the model to allow for competition. Finally, Section 4 concludes. All the proofs are relegated to Appendix A. Appendix B analyzes an alternative model of competition.

## 2 Media silence in a monopoly

In this section we present a simple model with one risk-neutral newspaper and a mass of consumers. The state of the world is $w \in \{N, C\}$, where $C$ corresponds to a situation in which there is a *corruption scandal* in the economy and $N$ to one in which *no corruption scandal* exists. Let $P(w = C) = \theta$ denote the prior probability that the state is *corrupted*. We start considering the case where each state occurs with equal probability.[8]

The newspaper receives a signal $s \in \{n, c\}$ on the state of the world, whose distribution depends on the newspaper's quality. With probability $\alpha$, the newspaper is high quality and has a signal that perfectly reveals the state of the world. With probability $1 - \alpha$, the newspaper is normal and receives an imperfect but informative signal, with $P(n \mid N) = P(c \mid C) = \gamma > \frac{1}{2}$.

Upon receiving the signal, the newspaper publishes a report $r \in \{\hat{n}, \hat{c}\}$, where $\hat{n}$ denotes the action of printing easy-to-cover stories, and $\hat{c}$ the action of printing a scandal. We denote by $\sigma_s(r) \in [0, 1]$ the probability that, conditioned on its signal $s$, a newspaper takes action $r$. We assume that the high type media firm always reports its signal honestly.[9] As for the normal type, we consider it has discretion to report either $\hat{n}$ or $\hat{c}$. This freedom to lie captures two types of media bias: A newspaper that having observed factual (though inconclusive) evidence of a corruption scandal chooses to silence it, i.e., $\sigma_c(\hat{n}) > 0$; and a news organization that without any evidence chooses to let the corruption scandal go in the paper, i.e., $\sigma_n(\hat{c}) > 0$. We refer to the former class of bias as media self-silence.[10]

We assume that a newspaper that reports $\hat{c}$ activates the feedback with probability $\mu \in [0, 1]$, in which case the state of the world will be revealed to consumers. In contrast, publishing $\hat{n}$ ensures the media firm that consumers will never know the true state (or, at least, not before they assign a reputation to the newspaper).[11]

[8] The assumption that both states are equiprobable means that media firms have no incentives to go for the consumers' prior beliefs. This differentiates our analysis from Gentzkow and Shapiro (2006) and ensures that herding effects play no role in generating our conclusions. This will become clearer in Section 2.3, where we relax this assumption and consider the more general case of $P(w = C) = \theta \in (0, 1)$.

[9] This assumption is for expositional purposes. In Section A.3 in Appendix A, we show that playing truthfully is the unique equilibrium strategy of the high type.

[10] In the paper we will refer to media silence and media self-silence indistinctively.

[11] This assumption will be relaxed in Section 3, where we analyze the effects of competition. In that case, publishing $\hat{n}$ will no

Note that this assumption partially endogeneizes the existence of feedback, giving the media firm the power to affect, with its report, whether or not consumers will learn if there is truly a corruption scandal in the economy.[12] This is a quite natural assumption in the media industry. Nonetheless, the power of the media to ignite cascades of accusations and responses and to stimulate coverage by other social spheres may lead to depuration of responsibilities, and thus learning. At the same time, it is difficult to think of a situation in which consumers manage to learn the truth of a story that never received the attention of the media industry, possibly because in this case consumers even ignore that such a story could have ever occurred. We denote by $X \in \{N, C, 0\}$ the feedback received, with $X = 0$ indicating that there is no feedback.

The consumers observe the newspaper's report $r$ and feedback $X$ and, based on this information, update their belief on the newspaper's type. Let $\lambda(r, X)$ denote the consumers' posterior probability that the newspaper is high quality, given $r \in \{\hat{n}, \hat{c}\}$ and $X \in \{N, C, 0\}$.

The newspaper is career-concerned and seeks to maximize reputation. As most papers in the literature, we assume that reputation is captured by the probability that consumers place on the media firm being of high expertise $\lambda(r, X)$.[13] This assumption should be taken as a reduced form of a more complex game, in which the newspaper seeks to appear high quality because future circulation (and thus profits) is increasing in reputation.[14]

Consumers are assumed to value information. To make this point, we can think of our consumers as citizens who, upon observing the report of the newspaper and before the state may be realized, take a decision on a private investment that yields a positive payoff $\pi$ when the newspaper correctly informs on the state of the world, and zero otherwise. Because the newspaper's signal is informative, $\gamma > 1/2$, the expected payoff to a consumer from the investment,

$$\frac{1}{2}\left(\alpha + (1-\alpha)(\gamma\sigma_n(\hat{n}) + (1-\gamma)\sigma_c(\hat{n}))\right)\pi + \frac{1}{2}\left(\alpha + (1-\alpha)(\gamma\sigma_c(\hat{c}) + (1-\gamma)\sigma_n(\hat{c}))\right)\pi,$$

is increasing in the accuracy of news and maximized when the news organization follows its signal, $\sigma_n(\hat{n}) = \sigma_c(\hat{c}) = 1$. Thus, if we define media bias as any deviation of the information the newspaper transmits from the signal it receives, the conclusion is straightforward: Media bias has detrimental effects for consumers.

We begin the analysis by considering the belief that the consumers place on the newspaper being the high

longer guarantee the newspaper that the truth will never come out.

[12] In the paper we use the terms *feedback* (Gentzkow and Shapiro (2006)) and *transparency of consequences* (Prat (2005)) indistinctively, with $\mu = 1$ meaning feedback occurring with probability one, and so there is full transparency of consequences.

[13] See Ottaviani and Sorensen (2001), Prat (2005), Gentzkow and Shapiro (2006) and Fox and Weelden (2012). Note that this modeling assumption implies that printing true stories has a prize - that of being assigned a high type. For an alternative way to model this idea, see Andina-Díaz (2009).

[14] This is in line with empirical evidence. Logan and Sutter (2004), using a cross-section of US newspapers, find that papers that have recently won Pulitzer Prizes have higher circulations, and Kovach and Rosenstiel (2001) observe that media firms with high standards have higher audiences. Also related, Anderson (2004) obtains that market forces penalize media firms whose quality of journalism falls.

type $\lambda(r, X)$.

$$\lambda(\hat{n}, 0) = \frac{\alpha}{\alpha + (1-\alpha)(\sigma_c(\hat{n}) + \sigma_n(\hat{n}))} \tag{1}$$

$$\lambda(\hat{c}, N) = 0 \tag{2}$$

$$\lambda(\hat{c}, C) = \frac{\alpha}{\alpha + (1-\alpha)(\gamma\sigma_c(\hat{c}) + (1-\gamma)\sigma_n(\hat{c}))} \tag{3}$$

$$\lambda(\hat{c}, 0) = \frac{\alpha}{\alpha + (1-\alpha)(\sigma_c(\hat{c}) + \sigma_n(\hat{c}))}, \tag{4}$$

with $\lambda(\hat{c}, C) > \lambda(\hat{c}, 0) > \lambda(\hat{c}, N)$.

Note that a media firm that chooses to print easy-to-cover events, $\hat{n}$, cannot activate the feedback. As a result, only $\lambda(\hat{n}, 0)$ follows a report of $\hat{n}$. This introduces an asymmetry in the consequences of reports, as $\hat{n}$ ensures a certain payoff of $\lambda(\hat{n}, 0)$, whereas printing $\hat{c}$ means playing a lottery with outcomes $\lambda(\hat{c}, 0)$, $\lambda(\hat{c}, N)$ and $\lambda(\hat{c}, C)$.

Let $E\{\lambda(r, X) \mid s\}$ denote the expected payoff to the (normal) newspaper when it observes signal $s \in \{n, c\}$ and publishes $r \in \{\hat{n}, \hat{c}\}$, over the possible realizations of $X \in \{N, C, 0\}$ .[15]

$$\begin{aligned}
E\{\lambda(\hat{n}, X) \mid s\} &= \lambda(\hat{n}, 0) \ \forall s \in \{n, c\} \\
E\{\lambda(\hat{c}, X) \mid n\} &= (1-\mu)\lambda(\hat{c}, 0) + \mu[P(N \mid n)\lambda(\hat{c}, N) + P(C \mid n)\lambda(\hat{c}, C)] \\
E\{\lambda(\hat{c}, X) \mid c\} &= (1-\mu)\lambda(\hat{c}, 0) + \mu[P(C \mid c)\lambda(\hat{c}, C) + P(N \mid c)\lambda(\hat{c}, N)],
\end{aligned}$$

where, given $\lambda(\hat{c}, N) = 0$, the last two expressions reduce to:

$$\begin{aligned}
E\{\lambda(\hat{c}, X) \mid n\} &= (1-\mu)\lambda(\hat{c}, 0) + \mu P(C \mid n)\lambda(\hat{c}, C) \\
E\{\lambda(\hat{c}, X) \mid c\} &= (1-\mu)\lambda(\hat{c}, 0) + \mu P(C \mid c)\lambda(\hat{c}, C).
\end{aligned}$$

Now, we can define the expected gain to reporting $\hat{n}$ rather than $\hat{c}$, after observing signal $s$, as:

$$\Delta_s[\sigma_n(\hat{n}), \sigma_c(\hat{c})] = E\{\lambda(\hat{n}, X) \mid s\} - E\{\lambda(\hat{c}, X) \mid s\}.$$

Substituting, we obtain:

$$\Delta_n[\sigma_n(\hat{n}), \sigma_c(\hat{c})] = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu P(C \mid n)\lambda(\hat{c}, C)) \tag{5}$$

$$\Delta_c[\sigma_n(\hat{n}), \sigma_c(\hat{c})] = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu P(C \mid c)\lambda(\hat{c}, C)). \tag{6}$$

with $P(C \mid n) = \frac{(1-\gamma)\theta}{(1-\gamma)\theta + \gamma(1-\theta)}$ and $P(C \mid c) = \frac{\gamma\theta}{\gamma\theta + (1-\gamma)(1-\theta)}$. In the present case with $\theta = \frac{1}{2}$, $P(C \mid n) = P(n \mid C) = 1 - \gamma$ and $P(C \mid c) = P(c \mid C) = \gamma$.

In the following, we will say that $(\sigma_n(\hat{n}), \sigma_c(\hat{c}))$ is an *equilibrium strategy* if $\sigma_n(\hat{n})$ maximizes the expected payoff to the newspaper after observing signal $n$, and $\sigma_c(\hat{c})$ does it after signal $c$. We will denote an equilibrium strategy by $(\sigma_n(\hat{n})^*, \sigma_c(\hat{c})^*)$.

[15] Note we consider a risk-neutral media firm. Assuming risk aversion would magnify the media bias that our results predict.

**Remark 1.** *An equilibrium strategy* $(\sigma_n(\hat{n})^*, \sigma_c(\hat{c})^*)$ *is of the form:*

1. *If* $\Delta_n\left[\overline{\sigma}_n(\hat{n}), \overline{\sigma}_c(\hat{c})\right] = \Delta_c\left[\overline{\sigma}_n(\hat{n}), \overline{\sigma}_c(\hat{c})\right] = 0$*, then* $(\overline{\sigma}_n(\hat{n}), \overline{\sigma}_c(\hat{c}))$ *is an equilibrium strategy.*

2. *If* $\Delta_n\left[\sigma_n(\hat{n}), \overline{\sigma}_c(\hat{c})\right] > 0$ $(< 0)$ *for all* $\sigma_n(\hat{n}) \in [0,1]$*, then* $(\sigma_n(\hat{n}), \overline{\sigma}_c(\hat{c}))$ *is an equilibrium strategy only if* $\sigma_n(\hat{n}) = 1$ $(0)$*.*

3. *If* $\Delta_c\left[\overline{\sigma}_n(\hat{n}), \sigma_c(\hat{c})\right] > 0$ $(< 0)$ *for all* $\sigma_c(\hat{c}) \in [0,1]$*, then* $(\overline{\sigma}_n(\hat{n}), \sigma_c(\hat{c}))$ *is an equilibrium strategy only if* $\sigma_c(\hat{c}) = 0$ $(1)$*.*

## 2.1 No feedback

Let us start the analysis briefly discussing the case of no feedback, $\mu = 0$. Here, consumers form beliefs on the quality of the news organization with the sole information of the newspaper's report, i.e., $\lambda(r, 0)$. This could be the case in countries with slow institutions or an inefficient judicial system, where processes drag on in time, hence postponing learning.

The analysis of this case yields a clear cut prediction. In the equilibrium without feedback, thus no fear of being proven wrong, the (normal) newspaper simply replicates the frequency of reports of the high type news organization. In a world where each state occurs with equal probability, this means sending $\hat{n}$ and $\hat{c}$ with probability 1/2 each. Next proposition characterizes the equilibria.

**Proposition 1.** *Suppose* $\mu = 0$*. Then,* $(\sigma_n(\hat{c}), \sigma_c(\hat{n}))$ *is an equilibrium strategy if and only if* $\sigma_n(\hat{c}) = \sigma_c(\hat{n})$*, with* $\sigma_n(\hat{c}) \in [0,1]$*.*

## 2.2 Feedback

Let us now consider the more interesting scenario in which consumers can learn the state of the world before assessing a belief. Remember that for this to be the case the newspaper must have previously taken action $\hat{c}$, as otherwise the state is never revealed.

Prior to presenting the result, let $x_0 = 1 - \frac{2\gamma(1-\alpha) - 2\alpha + \alpha\mu(1-\gamma) + \sqrt{4\gamma^2 + \alpha^2(1-\gamma)^2(2-\mu)^2 + 4\alpha(1-\gamma)\gamma - (2-3\mu)}}{4(1-\alpha)\gamma}$.[16] We next characterize the unique equilibrium of the game.

**Theorem 1.** *Suppose* $\mu > 0$*. There is a unique equilibrium. In the equilibrium,* $\sigma_n(\hat{n})^* = 1$ *and* $\sigma_c(\hat{n})^* = min\{1, x_0\} > 0$*.*

First, Theorem 1 shows that a newspaper that receives signal $n$ always reports $\hat{n}$. To have an intuition for this result, note that conditioned on having observed signal $n$, the probability of the occurrence of a scandal is $P(C \mid n) = 1 - \gamma$. Because $\gamma > \frac{1}{2}$, the firm realizes that it is very unlikely that the scandal exists, thus very costly to print it. As a result, the newspaper optimally chooses to stick to the signal. Second, it shows that

[16] The value of $x_0$ is derived in the proof of Theorem 1, in Appendix A.

a firm that receives signal $c$ uses a mixed strategy that silences scandals with probability $\sigma_c(\hat{n})^* > 0$. Thus, a sole media firm with the only concern of maximizing reputation generates media silence.[17] Note, additionally, that the result in Theorem 1 is independent of the value of parameters $\alpha$, $\gamma$ and $\mu$.[18] This is rather surprising, as although we may expect newspapers with low quality signals to misreport facts, it was not so clear a priori that firms with reliable signals (high $\gamma$) would find it optimal to resign from printing a story that would quite likely bring public recognize, and choose instead to silence the facts.[19]

More interesting is the comparative static analysis with respect to parameters $\alpha$ and $\mu$. Regarding $\alpha$, we obtain that the higher the prior probability that the news organization is high type, the more the corruption scandals the firm will silence. In other words, great expectations on the quality of a newspaper generates media silence.

**Corollary 1.** *Media self-silence is increasing in $\alpha$, i.e.* $\frac{\partial \sigma_c(\hat{n})^*}{\partial \alpha} > 0$.

To gain an intuition for this result, note that the reputation of a news organization that takes action $\hat{n}$ is increasing in $\alpha$ and, in the limit as $\alpha$ tends to 1, reputation $\lambda(\hat{n}, 0)$ approaches 1. Thus, when the average quality of the media industry in a country is excellent, consumers without information on the type of a particular newspaper will (most likely) perceive this firm as being of high type too. Because in this case the firm has no need to print a scandal to prove its quality, it optimally chooses to hold scandals with positive probability. A silence that we show can be complete for $\alpha$ sufficiently high.[20] On the contrary, when $\alpha$ is low, consumers without information on the type of a newspaper will (most likely) perceive this firm as low type too. Here, making an error does not imply such a large relative loss as before, whereas being proven right has more beneficial effects (than before). The consequence of all this are firms which, because of their mediocre environment, are forced to follow their signals and publish scandals more often, hoping to run a big story that brings people down. Unexpectedly, this reduces media bias.

Last, we perform the comparative static analysis with respect to parameter $\mu$. We obtain that the incentive of a media firm to stick with action $\hat{n}$ out of fear of being proven wrong increases with $\mu$. That is, increasing transparency has detrimental effects because it induces the news organization to silence more scandals.

**Corollary 2.** *Media self-silence is increasing in $\mu$, i.e.* $\frac{\partial \sigma_c(\hat{n})^*}{\partial \mu} > 0$.

[17] We want to highlight the fact that our results are obtained in a set-up in which the news organization is only concerned about reputation. That is, it is not the fact that a government or an interest group (advertisers) can buy the silence of the media (see Vaidya (2005), Besley and Prat (2006) and Ellman and Germano (2009)); or that libels can be punished (see Garoupa (1999*a*,*b*), Stanig (2014) and Gratton (2015)), what drives our results. In our model, the self-censorship that the newspaper practices exclusively originates in the media's power to preclude citizens learning, which indirectly affects the firm's reputation.

[18] Though the magnitude of the media bias does depend on these values. Substituting for different cases shows that the bias is important and, sometimes, even extreme, making the newspaper's report completely uninformative. For example, when $\alpha = 0.75$, $\gamma = 0.6$ and $\mu = 1$, we have $\sigma_c(\hat{n})^* = 1$.

[19] Here, the static comparative analysis reveals $\frac{\partial \sigma_c(\hat{n})^*}{\partial \gamma} < 0$, as shown in Lemma 4 in Appendix A. This is as expected, that is, the lower the quality of a newspaper's signal, the higher the incentive to be silent.

[20] Lemma 3 in Appendix A shows that there exists $\bar{\alpha} \in (0, 1)$ such that $\forall \alpha > \bar{\alpha}$, $\sigma_c(\hat{n})^* = 1$.

The intuition for this result is as follows. First, note that the payoff to a newspaper for sending $\hat{n}$ does not depend on $\mu$. However, its payoff for sending $\hat{c}$ does. In fact, it turns out that when $\sigma_n^*(\hat{c}) = 0$, $\gamma\lambda(\hat{c}, C) < \lambda(\hat{c}, 0)$. That is, the newspaper' expected payoff for sending $\hat{c}$ with feedback is smaller than its payoff for publishing the scandal without feedback. Now, because increasing $\mu$ increases the probability of receiving the expected payoff $\gamma\lambda(\hat{c}, C)$, we obtain that the higher the probability of feedback, the higher the probability that the news organization holds a scandal. Once more, the argument hinges upon the endogeneity of feedback, which is crucial to explain why, in our model, and in the terminology of Prat (2005), *transparency of consequences* has detrimental effects. This is in contrast to previous contributions in the career concern literature, where increasing transparency of consequences (learning the state) always induces the (low type) expert to send the report that is most likely to match the state of the world. If signals are informative, as in Prat (2005) or Gentzkow and Shapiro (2006), this leads to truthful reporting. If, however, the prior on the state is sufficiently strong,[21] as in Fox and Weelden (2012), this same argument explains why increasing transparency makes the expert more reticent to act on his private information.

## 2.3 Unbalanced prior

We next relax the assumption that the two states of the world are equiprobable. Let $P(w = C) = \theta$, with $\theta \in (0, 1)$. In this case, after report $r \in \{\hat{n}, \hat{c}\}$ and feedback $X \in \{N, C, 0\}$, the posterior probability $\lambda(r, X)$ that the consumers assign to the news organization as being of high type is:

$$\lambda(\hat{n}, 0) = \frac{\alpha(1-\theta)}{\alpha(1-\theta) + (1-\alpha)((1-\theta)(\gamma\sigma_n(\hat{n}) + (1-\gamma)\sigma_c(\hat{n})) + \theta(\gamma\sigma_c(\hat{n}) + (1-\gamma)\sigma_n(\hat{n})))} \tag{7}$$

$$\lambda(\hat{c}, N) = 0 \tag{8}$$

$$\lambda(\hat{c}, C) = \frac{\alpha}{\alpha + (1-\alpha)(\gamma\sigma_c(\hat{c}) + (1-\gamma)\sigma_n(\hat{c}))} \tag{9}$$

$$\lambda(\hat{c}, 0) = \frac{\alpha\theta}{\alpha\theta + (1-\alpha)(\theta(\gamma\sigma_c(\hat{c}) + (1-\gamma)\sigma_n(\hat{c})) + (1-\theta)(\gamma\sigma_n(\hat{c}) + (1-\gamma)\sigma_c(\hat{c})))}. \tag{10}$$

It is interesting to distinguish two cases here: $\theta < \frac{1}{2}$ and $\theta > \frac{1}{2}$. To see the reason for this, note that with an unbalanced prior there are two important force on stage. On the one hand, the endogenous feedback, that drives the media firm towards silencing corruption scandals. On the other hand, the "herding on the prior" effect, that induces the newspaper to send the report that is most likely to confirm the prior belief.[22] When $\theta < \frac{1}{2}$, it is clear that the two effects go in the same direction, whereas when $\theta > 1/2$ they push towards different actions.

Let us first comment the case $\theta < \frac{1}{2}$.[23] Here, there are clear reasons to print easy-to-cover stories. Based on this, we should expect the newspaper to omit signals of corruption. Our results show this to be the case.

[21] Namely, the prior is higher than the quality of the signal of the low type expert.

[22] See Gentzkow and Shapiro (2006) for an explanation of the "herding on the prior" argument and its consequences in terms of media bias. See also Heidhues and Lagerlöf (2003) and Cummins and Nyman (2005) for models of herding applied to other contexts.

[23] The analysis and results that follow are relegated to Section A.2 in Appendix A.

Thus, in line with Theorem 1, we obtain $\sigma_n(\hat{n})^* = 1$ and $\sigma_c(\hat{n})^* > 0$. Additionally, and also as expected, we observe that the higher the prior probability that the state is $N$ (the lower $\theta$), the higher the media silence. That is, a stronger prior (towards $N$) drives a greater bias. Last, we obtain that there exists $\bar{\alpha} \in (0,1)$ such that $\forall \alpha > \bar{\alpha}$, $\sigma_c(\hat{n})^* = 1$. Or, to say it differently, if $\alpha$ is sufficiently high, media silence is complete.[24] This result, which we also derived in the previous section, highlights the perverse effects that great expectations on the quality of the media can have on the number of corruption scandals reported by a news organization. Indeed, it raises a concern about the silent role of the media in countries with high standards of the press (high $\alpha$) and low levels of perceived corruption (low $\theta$). To these cases, our result suggests that media silence might be more the consequence of a career concerned industry than the real image of the country's level of corruption.

Let us now consider $\theta > \frac{1}{2}$. Here, the two aforementioned driving forces push towards opposite directions. This creates a richer and more complex scenario. Prior to presenting the results for this case, let us define $x_1$ such that $\Delta_c\left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 1 - x_1; \theta\right] = 0$ and $x_2$ such that $\Delta_n\left[\sigma_n(\hat{n}) = 1 - x_2, \sigma_c(\hat{c}) = 1; \theta\right] = 0$.[25]

**Theorem 2.** *Let $\theta \in (1/2, 1)$. There exist $\bar{\theta}_1$, $\bar{\theta}_2$ and $\bar{\theta}_3$, with $\frac{1}{2} < \bar{\theta}_1 < \bar{\theta}_2 < \bar{\theta}_3 < 1$. For each $\theta \in (1/2, 1)$ there is a unique equilibrium. In the equilibrium:*

1. *If $\theta \in (1/2, \bar{\theta}_1)$, $\sigma_n(\hat{n})^* = 1$ and $\sigma_c(\hat{n})^* = min\{1, x_1\} > 0$,*
2. *If $\theta \in (\bar{\theta}_1, \bar{\theta}_2)$, $\sigma_n(\hat{n})^* = 1$ and $\sigma_c(\hat{c})^* = 1$,*
3. *If $\theta \in (\bar{\theta}_2, \bar{\theta}_3)$, $\sigma_n(\hat{c})^* = x_2 > 0$ and $\sigma_c(\hat{c})^* = 1$,*
4. *If $\theta \in (\bar{\theta}_3, 1)$, $\sigma_n(\hat{c})^* = 1$ and $\sigma_c(\hat{c})^* = 1$,*

Theorem 2 illustrates how the equilibrium strategy of the media firm changes as $\theta$ increases. Thus, we first observe that when $1/2 < \theta < \bar{\theta}_1$, the "herding on the prior" effect is not strong enough to completely offset the incentive of the newspaper to take the conservative action. The conclusion is that even if consumers believe that the *corrupted* state is the most likely, if this prior is not too strong, we can have news organizations silencing evidence of corruption.[26] Additionally, we also obtain that for $\alpha$ sufficiently high, media silence can here be complete.[27]

Next, when $\theta \in (\bar{\theta}_1, \bar{\theta}_2)$, the two aforementioned arguments cancel out and we obtain that the newspaper truthfully follows its signals. Now, pushing $\theta$ beyond $\bar{\theta}_2$ represents a situation in which the prior on the state is sufficiently strong to offset any counterbalancing force which, for $\theta > \bar{\theta}_3$, results in the newspaper always reporting $\hat{c}$, independently of its signal. This creates a different class of media bias, that talks about newspapers

[24] This is an extreme result that nonetheless does not require of a too strong parameter configuration. For example, if $\theta = 0.2$, $\gamma = 0.7$ and $\mu = 0.8$, we obtain $\bar{\alpha} = 0.8$.

[25] These values are defined in the proof of Theorem 2, in Appendix A.

[26] This is the same result than in cases $\theta < \frac{1}{2}$ and $\theta = \frac{1}{2}$ (Theorem 1). It thus proves the robustness of this conclusion.

[27] See Lemma 15 in Appendix A.

printing too many stories on corruption, in the hope for catering to the people and possibly, bringing them down.

# 3 Competition and feedback power

Thus far we have assumed that there is only one newspaper with the power to activate feedback. We next relax this assumption and consider the case with more firms. To this aim, we follow Gentzkow and Shapiro (2006) and consider one scoop-firm and $J$ feedback-firms. An alternative approach can be found in Appendix B, where we consider two (strategic) scoop-firms.[28]

Consider a game with $J+1$ newspapers and two stages. Each of the $J+1$ firms seeks to maximize its reputation for high quality. In stage 1, the scoop-firm receives a signal $s \in \{n, c\}$ on the state of the world $w \in \{N, C\}$. Each state is equally likely. This firm can be either high type or normal type. A high scoop-firm receives a perfectly accurate signal and reports it honestly; a normal scoop-firm receives an imperfect signal of quality $\gamma > \frac{1}{2}$. After receiving the scoop, the newspaper chooses to print a report $r \in \{\hat{n}, \hat{c}\}$. As in Section 2, a newspaper publishing $\hat{c}$ activates the feedback with positive probability, whereas printing $\hat{n}$ does not affect the feedback. All consumers read the scoop-firm's report. In stage 2, the remaining $J$ newspapers, the feedback-firms, observe the scoop-firm's report and learn the true state. Then, they choose what to report $r \in \{\hat{n}, \hat{c}\}$.

Because feedback-firms observe the true state and have career concerns, we know that in the equilibrium of the second stage, they will report their signals truthfully. This result is in Gentzkow and Shapiro (2006). It means that any consumer reading a feedback-firm's report will learn the true state. Let $\phi(J)$ be the probability that each consumer reads a feedback report, and assume $\frac{d\phi}{dJ} > 0$. Thus, we can define $\mu_J = \phi(J)$ as the probability that consumers learn the state of the world after the report of the $J$ feedback-firms. The idea is that the higher the number of feedback-firms in the industry, the higher the probability that a consumer reads a feedback report, thus the higher the probability that consumers learn the true state.

Next, we come back to stage 1. As already mentioned, in line with Section 2 we consider that the scoop-firm has the ability to affect the probability of feedback with its report. In particular, we assume that if the scoop-firm publishes $\hat{n}$, the probability that consumers learn the state is given by $\mu_J$. However, if the scoop-firm publishes $\hat{c}$, the probability of feedback is $\mu_{J+1}$, with $0 < \mu_J \leq \mu_{J+1} < 1$.[29] Accordingly, we can define $\mu_{J+1} - \mu_J$ as the marginal impact of the scoop-firm on the probability that consumers learn the state.

[28] In Appendix B we obtain that the results in this more strategic (though more complex) scenario are qualitatively the same than those obtained in the main analysis.

[29] The idea is that when a scoop-firm prints a scandal, public awareness of the malpractice increases, which makes also increase the probability that consumers learn the true state. In the present scenario, there are at least two reasons for this: (i) The probability that a consumer reads a feedback-firm's report is higher when the scoop-firm first publish a malpractice and/or (ii) the scoop-firm's report activates other institutional mechanisms (for example the judicial system) that increases the probability that a consumer learns the true state.

We refer to this marginal impact as the *feedback power* of the scoop-firm.

To clarify the concept of feedback power, consider the following exercise. Suppose a competitive media industry, in which all the newspapers are small enough as to affect the probability of feedback. This means the scoop-firm takes the probability of feedback as something exogenous, something it cannot affect. Note that this is the case when $\mu_{J+1} - \mu_J$ is small enough and, in the limit, equal to zero. Indeed, when $\mu_{J+1} - \mu_J = 0$, it happens that there is only one probability of feedback, which does not depend on the action taken by the scoop-firm.[30] Suppose now the other extreme scenario, in which the scoop-firm is a powerful newspaper, with a large influence on the society. Because a corruption scandal printed in this newspaper will surely have a great social and judicial impact on the society, $\mu_{J+1} - \mu_J$ cannot be zero for this firm. Moreover, the greater the power of the scoop-firm, the higher this difference should be.[31]

We now go into the analysis of the scoop-firm game. Given a report $r \in \{\hat{n}, \hat{c}\}$ and feedback $X \in \{N, C, 0\}$, the posterior probability $\lambda(r, X)$ that consumers assign to the scoop-newspaper as being of high type is:

$$\lambda(\hat{n}, N) = \frac{\alpha}{\alpha + (1-\alpha)(\gamma\sigma_n(\hat{n}) + (1-\gamma)\sigma_c(\hat{n}))} \tag{11}$$

$$\lambda(\hat{n}, C) = 0 \tag{12}$$

$$\lambda(\hat{n}, 0) = \frac{\alpha}{\alpha + (1-\alpha)(\sigma_c(\hat{n}) + \sigma_n(\hat{n}))} \tag{13}$$

$$\lambda(\hat{c}, N) = 0 \tag{14}$$

$$\lambda(\hat{c}, C) = \frac{\alpha}{\alpha + (1-\alpha)(\gamma\sigma_c(\hat{c}) + (1-\gamma)\sigma_n(\hat{c}))} \tag{15}$$

$$\lambda(\hat{c}, 0) = \frac{\alpha}{\alpha + (1-\alpha)(\sigma_c(\hat{c}) + \sigma_n(\hat{c}))}. \tag{16}$$

Proceeding as previously, we can obtain the expected payoff to the (normal) scoop-newspaper when it observes signal $s \in \{n, c\}$ and publishes $r \in \{\hat{n}, \hat{c}\}$:

$$\begin{aligned}
E\{\lambda(\hat{n}, X) \mid n\} &= (1-\mu_J)\lambda(\hat{n}, 0) + \mu_J\gamma\lambda(\hat{n}, N) \\
E\{\lambda(\hat{n}, X) \mid c\} &= (1-\mu_J)\lambda(\hat{n}, 0) + \mu_J(1-\gamma)\lambda(\hat{n}, N) \\
E\{\lambda(\hat{c}, X) \mid n\} &= (1-\mu_{J+1})\lambda(\hat{c}, 0) + \mu_{J+1}(1-\gamma)\lambda(\hat{c}, C) \\
E\{\lambda(\hat{c}, X) \mid c\} &= (1-\mu_{J+1})\lambda(\hat{c}, 0) + \mu_{J+1}\gamma\lambda(\hat{c}, C),
\end{aligned}$$

and the expected gain to reporting $\hat{n}$ rather than $\hat{c}$, after observing signal $n$ and $c$, respectively:

$$\Delta_n[\sigma_n(\hat{n}), \sigma_c(\hat{c})] = (1-\mu_J)\lambda(\hat{n}, 0) + \mu_J\gamma\lambda(\hat{n}, N) - ((1-\mu_{J+1})\lambda(\hat{c}, 0) + \mu_{J+1}(1-\gamma)\lambda(\hat{c}, C)) \tag{17}$$

$$\Delta_c[\sigma_n(\hat{n}), \sigma_c(\hat{c})] = (1-\mu_J)\lambda(\hat{n}, 0) + \mu_J(1-\gamma)\lambda(\hat{n}, N) - ((1-\mu_{J+1})\lambda(\hat{c}, 0) + \mu_{J+1}\gamma\lambda(\hat{c}, C)). \tag{18}$$

[30] This is the standard assumption in the literature, and the case in Gentzkow and Shapiro (2006).

[31] Entman (2012) presents extensive evidence supporting the idea that feedback power differs across news organizations. For example, he classifies *New York Times* and *Washington Post* as highly influential; *Time* and *Newsweek* as influential; and *Boston Globe*, *Chicago Tribune* and other major regional papers as occasionally influential.

Prior to presenting the result of the scoop-firm game, let us define $x_3$ such that $\Delta_c[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 1 - x_3] = 0$.[32]

**Theorem 3.** *Suppose* $0 < \mu_J < \mu_{J+1} < 1$. *Let* $\hat{\gamma} = \frac{\mu_J + \alpha(\mu_{J+1} - \mu_J)}{2\mu_J + \alpha(\mu_{J+1} - \mu_J)} \in (0, 1)$. *There is a unique equilibrium in the scoop-firm game. In the equilibrium:*

1. *If* $\gamma < \hat{\gamma} \in (0, 1)$, $\sigma_n(\hat{n})^* = 1$ *and* $\sigma_c(\hat{n})^* = min\{1, x_3\} > 0$,
2. *If* $\gamma > \hat{\gamma} \in (0, 1)$, $\sigma_n(\hat{n})^* = 1$ *and* $\sigma_c(\hat{c})^* = 1$,

From Theorem 3, we observe that introducing competition in the media industry ensures that newspapers with higher quality sources ($\gamma > \hat{\gamma}$) stick to their signals and thus, reduce their bias. However, this watchdog role is not at work on newspapers of lower quality sources, which continue silencing evidence of corruption. The fear of revealing their type and being proven wrong again prevents scoop-firms with lower quality signals to reveal their information.

The comparative static analysis with respect to parameters $\alpha$, $\gamma$, $\mu_{J+1}$ and $\mu_J$, also yields interesting insights. First, that an increase in the probability that the newspaper is perceived as high type, $\alpha$, strengthens the requirement on the quality of a firm to ensure revelation.[33] Thus, in line with the result in Corollary 1, we obtain that also with competition, great expectations on the type of a newspaper drives media silence. Second, that an increase in the quality of a signal, $\gamma$, reduces the probability that the newspaper holds information. This as expected, and also in line with the result in the monopoly scenario.

**Corollary 3.** *Media self-silence is decreasing in* $\gamma$, *i.e.* $\frac{\partial \sigma_c(\hat{n})^*}{\partial \gamma} < 0$.

Third, that an increase in $\mu_J$ reduces media silence, whereas an increase in $\mu_{J+1}$ has the opposite effect and magnifies the bias. This is formally stated next.

**Corollary 4.** *Media self-silence is decreasing in* $\mu_J$ *and increasing in* $\mu_{J+1}$, *i.e.* $\frac{\partial \sigma_c(\hat{n})^*}{\partial \mu_J} < 0$ *and* $\frac{\partial \sigma_c(\hat{n})^*}{\partial \mu_{J+1}} > 0$.

It is useful to interpret the last result in terms of feedback power. To this, let us consider the following exercise. For a given $\mu_{J+1}$, the smaller $\mu_J$, thus the greater the feedback power of the scoop-firm, $\mu_{J+1} - \mu_J$, the higher the incentive of this firm to silence evidence of corruption. Indeed, from Theorem 3, we observe that in the limit, when $\mu_J$ tends to zero, $\hat{\gamma} \longrightarrow 1$. That is, in this case, for any signal's quality, the scoop-firm will optimally choose to silence information with positive probability. This is the same result than in the monopoly scenario. On the contrary, the higher $\mu_J$, thus the smaller the social influence of the scoop-firm, the more likely the firm will stick to its signal. This insight is also observed in Theorem 3, when $\mu_J \longrightarrow \mu_{J+1}$. Here, in the limit, we obtain $\hat{\gamma} = 1/2$. In other words, tough competition serves to discipline all scoop-firms, independently of their quality. A similar exercise can be done for a fixed $\mu_J$. Thus, given $\mu_J$, the higher the

[32] This value is defined in the proof of Theorem 3, in Appendix A.

[33] Note that when $\alpha \longrightarrow 0$, then $\hat{\gamma} \longrightarrow 1/2$. Thus, for any $\gamma \in (0, 1)$, in equilibrium, a scoop-firm always follows its signal. Since, $\frac{\partial \hat{\gamma}}{\partial \alpha} > 0$, the result follows.

political influence and social power of the scoop-firm, as measured by $\mu_{J+1}$, the more often it omits a scandal; and the lower its feedback power, the higher its incentives to reveal its information.

### 3.1 Discussion

Our analysis of competition can be used to explain some interesting phenomena. We discuss them next.

■ We start with the Lewinsky and the bin Laden's death stories discussed in Section 1. The reader will remember the different treatment that different news organizations gave to the same scoop. In fact, both *Newsweek* and *Drugde Report* shared the same source on the Lewinsky scandal, as well as *The New Yorker* and *The London Review of Books* had access to the same report on bin Laden's death. However, neither *Newsweek* nor *The New Yorker* published the stories, whereas *Drugde Report* and *The London Review of Books* found no impediment to run them. To these cases, what can our theory say?

Our model suggests that if we compare two scenarios, each with one scoop-firm with different feedback power, $\mu_{J+1} - \mu_J$, but with the same signals' quality, $\gamma$, the newspaper with the greater social influence will hold scandals with a higher probability than the less influential news organization. Interestingly, these predictions are in line with observed behavior. In fact, neither *Newsweek* ran the Lewinsky scandal, nor *The New Yorker* published Hersh's report on bin Laden's death. Despite both stories first belonged to important magazines, they were finally broken by smaller media outlets, with a more limited feedback power. In the former case, it was an Internet web site. In the latter, it was a magazine that, despite having a thorough review process, it is far less influential than *The New Yorker*.[34]

Does it mean that newspapers with greater feedback power impose stricter conditions on the quality of their sources? Could it be that these more stringent conditions drive influential newspapers to hold stories that less powerful firms run instead? In the light of our examples, we could say that this is the case. In fact, Mark Whitaker, the *Newsweek*'s editor at the time of the Lewinsky scandal, in reference to why he decided not to publish the scoop, argued: "*We didn't feel [...] that we were on firm enough ground*".[35] Similarly, in reference to David Remnick, editor of *The New Yorker*, Jonathan Mahler writes: "*But the bin Laden story wasn't the first one by Hersh that Remnick rejected because he considered the sourcing too thin*".[36]

Interestingly, our theory also explains this behavior. To see it, note that from Theorem 3 we know that

[34] According to Entman (2012), who classifies US news organizations according to their influence, *Newsweek* is ranked as "influential". Similarly, in reference to *The New Yorker*, he writes: *"Books or magazine articles written by journalists affiliated with major national media who are members of the close-knit intermedia networks in Washington have the best chace of influencing scandal production. Examples would include [...] the muckraking investigations of Seymour Hersh in The New Yorker."* As for *Drugde Report*, the author does not explicitly consider it, but he classifies specialized blogs and websites as "rarely influential". See pages 66-70.

[35] See "Former Newsweek Editor on Why He didn't Run Lewinsky Story: 'We Didn't Feel We Were on Firm Enough Ground"', *NewsBuster*, November 6, 2011.

[36] See "What Do We Really Know About Osama bin Laden's Death?", written by Jonathan Mahler, in *The New York Times*, October 18, 2015.

$\hat{\gamma} = \frac{\mu_J + \alpha(\mu_{J+1} - \mu_J)}{2\mu_J + \alpha(\mu_{J+1} - \mu_J)} \in (0, 1)$ is the minimum quality of a corruption signal for the associated scandal to be published. In fact, Theorem 3 shows that for all $\gamma > \hat{\gamma}$ the scoop-firm reveals all its information, whereas for all $\gamma < \hat{\gamma}$ the firm silences corruption scandals with positive probability. Now, if we take the derivative of $\hat{\gamma}$ with respect to $\mu_{J+1}$ and $\mu_J$, we obtain the following result.

**Corollary 5.** *The higher the feedback power of a scoop-firm, $\mu_{J+1} - \mu_J$, the greater its requirement on the quality of a signal to be published, i.e. $\frac{\partial \hat{\gamma}}{\partial \mu_{J+1}} > 0$ and $\frac{\partial \hat{\gamma}}{\partial \mu_J} < 0$.*

That is, the greater the power of a news organization to lead public opinion, $\mu_{J+1} - \mu_J$, the stronger the newspaper's requirement on the quality of a signal to be published. As already mentioned, this is in line with observed behavior. In this sense, our theory helps explain why big newspapers and magazines, such as *The New York Times*, *The Washington Post* or *The Guardian*, have stricter vetting processes for stories and fact-checkers in their staff; whereas smaller newspapers, lacking the power to influence public opinion, have less thorough review processes, which makes them more prone to print scandals. In terms of our example, it helps explain why neither *Newsweek* nor *The New Yorker* chose to run the Lewinsky and the bin Laden's death stories, respectively, whereas *Drudge Report* and *The London Review of Books* found no repair to print them and break the news.

■ Note that the analysis above compares two firms that have access to the same signal. To this case our model has a clear cut prediction: the more influential newspaper will hold scandals with a higher probability than the less influential one. However, a look at the *Pulitzer Prizes* and the list of winners show that although in recent years many of the big *Pulitzer Prizes* have gone to first-time winners and small news organizations, renown newspapers and magazines such as *The New York Times* or *The Washington Post* are far ahead of the rest of media outlets.[37] Because many of these honors are in the categories of Breaking News Reporting and Investigative Reporting, it presents clear evidence that big newspapers do also cover scandals. In the light of this evidence, the question is then whether our theory can accommodate this empirical observation. The answer is yes.

In our view, there are at least two arguments to explain why more influential newspapers, that according to our theory have more incentives to hold information, can end up breaking more news.

- The first argument says that bigger newspapers have easier access to financial resources. In fact, if any, they are those that can afford having a couple of journalists working for month on a story before exposing the facts. Just as in the case of Woodward and Bernstein. In terms of our theory, it means that newspapers may differ not only in their feedback power, but also in the quality of their signals.

  If this is the case, and if we consider that bigger newspapers with higher social influence have access to better sources, we can explain why renown news organizations may cover more scandals. To see it,

[37] Since 1918, *The New York Times* has been awarded 117 *Pulitzer Prizes*, more than any other news organization. *The Washington Post* has won 47.

suppose we compare two scenarios, each with one scoop-firm with different feedback power, $\mu_{J+1} - \mu_J$, and different signal's quality, $\gamma$, such that the firm with the highest social influence receives a better signal. Now, note that our analysis of competition has two clear predictions. On the one hand, we obtain that an increase in the quality of a signal, $\gamma$, reduces media silence. This is Corollary 3. On the other hand, we know that an increase in the feedback power of a firm, $\mu_{J+1} - \mu_J$, increases media silence. This is Corollary 4. Now, because both effects go in opposite directions, the model has not a clear prediction, and the answer will depend on which effect dominates. Despite it, we can say the following. If we have a situation in which the big firm has some feedback power and the small one almost no social influence, then we know that the big newspaper will silence its corruption signals with a higher probability than the small one. To see it, note that from Theorem 3 we know that for a given $\mu_{J+1}$, if $\mu_J \to 0$, then $\hat{\gamma} \to 1$; and if $\mu_J \to \mu_{J+1}$, then $\hat{\gamma} \to \frac{1}{2}$. That is, when the feedback power of the scoop-firm is sufficiently high (in the limit, when $\mu_J \to 0$), the newspaper holds corruption signals with positive probability, independently of their quality. On the contrary, when the feedback power of the scoop-firm is sufficiently low (in the limit, when $\mu_J \to \mu_{J+1}$), the newspaper publishes all its signals of corruption.

According to this, our model suggests that we can have a situation in which the bigger newspaper publish more scandals than the less influential one. However, it also says that for this situation to occur, we need the big firm to have a signal of a significantly higher quality, compared to that of the smaller newspaper. Otherwise, that is if signals are nor qualitatively much different, the smaller newspaper will, more likely, reveal the scoop.

- The second argument says that bigger newspapers may simply receive more scoops, because their name and/or influence makes them more attractive to whistle-blowers. Again, if this is the case, it may perfectly be that renown news organizations publish more scoops, even if they have a more stringent vetting process for stories.

# 4 Conclusion

People receive much of the information from the media. Even in the area of new technologies, a recent study conducted by *Gallup*,[38] shows that 70% of Americans rate traditional media (tv, print newspapers and radio) as *"the main source of news about current events in the US and around the world"*. Internet and social media (Facebook, Twitter, etc) is mentioned by 21% of the population, and only a small 5% talk about other sources such as word of mouth. A tendency that the *American Press Institute* confirms to hold across generations.[39] These numbers reflect the importance of the media industry as to set what citizens get to know, to learn, and how much we lose from a silent media.

[38] "TV Is Americans' Main Source of News", *Politics*, July 8, 2013.

[39] "The Personal News Cycle: How Americans choose to get their news", March 17, 2014.

Based on these facts we build a model in which a news organization, through its printing strategy, has the power to determine how much citizens can ever learn about an issue. In other words, we endogeneize the feedback. Our results put forth an important reputational incentive for a newspaper to self-silence information, showing that silence increases in the prior expectations on the quality of the firm, the probability of feedback, and the political and social influence of the newspaper. We also obtain that silence decreases with competition and that the greater the power of a news organization to lead public opinion, the stricter its vetting process for stories is. These results are much in line with empirical evidence. The stories of Lewinsky and bin Laden's death support this claim.

It is interesting to note that the media silence we obtain in this paper does not depend on the existence or not of defamation lawsuits or physical threats to journalists (see Garoupa (1999*a*,*b*) and Stanig (2014)), which we agree are real phenomena and important sources of media silence. It is neither explained by media captured, either by the government or by advertisers (see the theoretical works by Besley and Prat (2006) and Ellman and Germano (2009)). Introducing these kind of considerations would just reinforce our results. Similarly, our analysis considers risk-neutral newspapers. Again, extending the analysis to account for risk-averse news organizations or journalists would just magnify the result of media bias. In this sense, our contribution is to point out to a more subtle source of media silence, that exclusively originates in the power of the media to raise public concern and so affect the probability that there is ex-post verification of the true state of the world.

Beyond the media industry, where our assumption that actions affect feedback may seem quite natural, we consider that there are other real-world situations that can also fit into our model. Think for example in a judge, court, or any authority with the power to accuse and prosecute somebody for a harmful act. Suppose this authority receives factual (though inconclusive) evidence of a wrongdoing by a powerful personality or firm. In this case, its decision on whether to go further with the inquiry quite closely resembles that of the news organization. Indeed, formally prosecuting means embarking on a process whose details will be argued by citizens beyond the court, thus with a verdict trespassing on public opinion. In contrast, keeping silent on the misconduct will probably preclude citizens' learning about the factual (though inconclusive) evidence of wrongdoing. Similarly, think in a competition authority opening an inspection in an important firm, or a doctor prescribing a new treatment or medicine. The essence to all these examples is clear, and is to do with a really simple question: Do all my actions provoke feedback with the same probability? Or they are instead different in terms of attracting public attention? As the answer is simple as well. If actions differ in their influence magnitude, then we can presume we have another example of endogenous feedback. To these situations, the model in this paper presents new insights into the unexpected effects of quality, transparency and feedback power.

# A Appendix

Appendix A contains the proofs. It is divided into four subsections: A.1) Monopoly with balanced prior; A.2) Monopoly with unbalanced prior; A.3) Strategic high type; and A.4) Competition and feedback power.

Prior to the analysis, note that functions $\Delta_n$ and $\Delta_c$ depend on two variables, $\sigma_c(\hat{c})$ and $\sigma_n(\hat{n})$, and three parameters, $\alpha$, $\gamma$ and $\mu$. In the case of an unbalanced prior, there will be a fourth parameter, $\theta$. We use notation $\Delta_s[\cdot]$, with $s \in \{n, c\}$, when we need to make explicit this dependence.

## A.1 Monopoly with balanced prior

***Proof of Proposition 1***

First, note that

$\lambda(\hat{n}, 0) = \frac{\alpha}{\alpha+(1-\alpha)(\sigma_c(\hat{n})+\sigma_n(\hat{n}))} > \lambda(\hat{c}, 0) = \frac{\alpha}{\alpha+(1-\alpha)(\sigma_c(\hat{c})+\sigma_n(\hat{c}))} \iff$

$\sigma_c(\hat{c}) + \sigma_n(\hat{c}) > \sigma_c(\hat{n}) + \sigma_n(\hat{n}) \iff 1 - \sigma_c(\hat{n}) + \sigma_n(\hat{c}) > \sigma_c(\hat{n}) + 1 - \sigma_n(\hat{c}) \iff$

$\sigma_n(\hat{c}) > \sigma_c(\hat{n})$.

Now, suppose $\mu = 0$. Clearly, $\Delta_n = \Delta_c = \lambda(\hat{n}, 0) - \lambda(\hat{c}, 0)$.

We first prove that, in equilibrium, only $\sigma_c(\hat{n})^* = \sigma_n(\hat{c})^*$ can occur. Let us argue by contradiction:

$\sigma_n(\hat{c})^* > \sigma_c(\hat{n})^* \iff \lambda^*(\hat{n}, 0) > \lambda^*(\hat{c}, 0) \iff \Delta_c = \Delta_n > 0 \iff \sigma_n(\hat{c})^* = 0$ and $\sigma_c(\hat{n})^* = 1$, a contradiction.

$\sigma_n(\hat{c})^* < \sigma_c(\hat{n})^* \iff \lambda^*(\hat{n}, 0) < \lambda^*(\hat{c}, 0) \iff \Delta_c = \Delta_n < 0 \iff \sigma_n(\hat{c})^* = 1$ and $\sigma_c(\hat{n})^* = 0$, a contradiction.

We now prove that $\forall \sigma_c(\hat{n})^* \in [0, 1]$, if $\sigma_c(\hat{n})^* = \sigma_n(\hat{c})^*$, the strategy profile is an equilibrium. Since $\sigma_c(\hat{n})^* = \sigma_n(\hat{c})^*$, it follows that $\lambda^*(\hat{n}, 0) = \lambda^*(\hat{c}, 0)$ and $\Delta_n = \Delta_c = 0$. This completes the proof. ■

***Proof of Theorem 1***

To prove this Theorem we first need the following two Lemmas.

**Lemma 1.** *If $\mu > 0$, then $\Delta_n > \Delta_c$.*

***Proof.*** $\Delta_n > \Delta_c$

$\iff$

$\lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu\,(1-\gamma)\lambda(\hat{c}, C)) > \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu\,\gamma\lambda(\hat{c}, C))$

$\iff$

$-\mu(1-\gamma)\lambda(\hat{c}, C) > -\mu\gamma\lambda(\hat{c}, C)$

$\iff$

$(1-\gamma) < \gamma$.

Since $\gamma > \frac{1}{2}$, the proof follows. ♦

**Lemma 2.** *If $\mu > 0$ and $\sigma_c(\hat{c}) = 1$, then $\Delta_n > 0$*

***Proof.***

$$\Delta_n = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu\,(1-\gamma)\lambda(\hat{c}, C)) =$$
$$= \frac{\alpha}{\alpha+(1-\alpha)(\sigma_c(\hat{n})+\sigma_n(\hat{n}))} - \left(\frac{(1-\mu)\alpha}{\alpha+(1-\alpha)(\sigma_c(\hat{c})+\sigma_n(\hat{c}))} + \frac{\mu(1-\gamma)\alpha}{\alpha+(1-\alpha)(\gamma\sigma_c(\hat{c})+(1-\gamma)\sigma_n(\hat{c}))}\right).$$

Now, $\Delta_n[\sigma_c(\hat{c}) = 1] =$

$$= \frac{\alpha}{\alpha+(1-\alpha)\sigma_n(\hat{n})} - \left(\frac{(1-\mu)\alpha}{1+(1-\alpha)\sigma_n(\hat{c})} + \frac{\mu(1-\gamma)\alpha}{\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_n(\hat{c}))}\right) > 0$$
$$\Longleftrightarrow$$
$$(1+(1-\alpha)\sigma_n(\hat{c}))(\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_n(\hat{c})))$$
$$-(1-\mu)((\alpha+(1-\alpha)\sigma_n(\hat{n}))(\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_n(\hat{c})))$$
$$-\mu(1-\gamma)((\alpha+(1-\alpha)\sigma_n(\hat{n}))(1+(1-\alpha)\sigma_n(\hat{c})) > 0$$
$$\Longleftrightarrow$$
$$(\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_n(\hat{c})))(1+(1-\alpha)\sigma_n(\hat{c}) - \alpha - (1-\alpha)\sigma_n(\hat{n}))$$
$$+\mu\gamma(\alpha+(1-\alpha)\sigma_n(\hat{n}))(1+(1-\alpha)\sigma_n(\hat{c}))$$
$$+\mu(\alpha+(1-\alpha)\sigma_n(\hat{n}))(\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_n(\hat{c})) - 1 - (1-\alpha)\sigma_n(\hat{c})) > 0$$
$$\Longleftrightarrow$$
$$(\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_n(\hat{c})))(1-\alpha)2\sigma_n(\hat{c})$$
$$+\mu(\alpha+(1-\alpha)\sigma_n(\hat{n}))(\gamma+\gamma(1-\alpha)\sigma_n(\hat{c}) + \alpha + \gamma - \alpha\gamma + (1-\alpha)(1-\gamma)\sigma_n(\hat{c})) - 1 - (1-\alpha)\sigma_n(\hat{c})) > 0$$
$$\Longleftrightarrow$$
$$(\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_n(\hat{c})))(1-\alpha)2\sigma_n(\hat{c}) + \mu(\alpha+(1-\alpha)\sigma_n(\hat{n}))(2\gamma - 1 + \alpha(1-\gamma)) > 0$$
$$\Longleftrightarrow$$

$(2\gamma - 1 + \alpha(1-\gamma)) > 0$ which, since $\frac{1}{2} < \gamma < 1$, always holds. ♦

Now, we can prove that if $\mu > 0$, then $\sigma_n(\hat{n})^* = 1$ and $\sigma_c(\hat{n})^* > 0$.
There are nine equilibrium configuration to analyze.

| | | | | | |
|---|---|---|---|---|---|
| 1. | $\sigma_c(\hat{c})^* = 1$ | $\sigma_n(\hat{n})^* = 1$ | $\Longleftrightarrow$ | $\Delta_c \leq 0$ | $\Delta_n \geq 0.$ |
| 2. | $0 < \sigma_c(\hat{c})^* < 1$ | $\sigma_n(\hat{n})^* = 1$ | $\Longleftrightarrow$ | $\Delta_c = 0$ | $\Delta_n \geq 0.$ |
| 3. | $\sigma_c(\hat{c})^* = 0$ | $\sigma_n(\hat{n})^* = 1$ | $\Longleftrightarrow$ | $\Delta_c \geq 0$ | $\Delta_n \geq 0.$ |
| 4. | $\sigma_c(\hat{c})^* = 1$ | $\sigma_n(\hat{n})^* = 0$ | $\Longleftrightarrow$ | $\Delta_c \leq 0$ | $\Delta_n \leq 0.$ |
| 5. | $0 < \sigma_c(\hat{c})^* < 1$ | $\sigma_n(\hat{n})^* = 0$ | $\Longleftrightarrow$ | $\Delta_c = 0$ | $\Delta_n \leq 0.$ |
| 6. | $\sigma_c(\hat{c})^* = 0$ | $\sigma_n(\hat{n})^* = 0$ | $\Longleftrightarrow$ | $\Delta_c \geq 0$ | $\Delta_n \leq 0.$ |
| 7. | $\sigma_c(\hat{c})^* = 1$ | $0 < \sigma_n(\hat{n})^* < 1$ | $\Longleftrightarrow$ | $\Delta_c \leq 0$ | $\Delta_n = 0.$ |
| 8. | $0 < \sigma_c(\hat{c})^* < 1$ | $0 < \sigma_n(\hat{n})^* < 1$ | $\Longleftrightarrow$ | $\Delta_c = 0$ | $\Delta_n = 0.$ |
| 9. | $\sigma_c(\hat{c})^* = 0$ | $0 < \sigma_n(\hat{n})^* < 1$ | $\Longleftrightarrow$ | $\Delta_c \geq 0$ | $\Delta_n = 0.$ |

Note that from Lemma 1, configurations 5, 6, 8 and 9 cannot be. Similarly, from Lemma 2, configurations 4 and 7 can neither be. Consequently, $\sigma_n(\hat{n})^* = 1$. This means only configurations 1-3 are left which, taking into account Lemmas 1 and 2, can be rewritten as:

1. $\sigma_c(\hat{c})^* = 1 \quad \sigma_n(\hat{n})^* = 1 \iff \Delta_c \leq 0 \quad \Delta_n \geq 0.$
2. $0 < \sigma_c(\hat{c})^* < 1 \quad \sigma_n(\hat{n})^* = 1 \iff \Delta_c = 0 \quad \Delta_n > 0.$
3. $\sigma_c(\hat{c})^* = 0 \quad \sigma_n(\hat{n})^* = 1 \iff \Delta_c \geq 0 \quad \Delta_n > 0.$

Let us now consider $\sigma_n(\hat{n})^* = 1$ and analyze how the normal newspaper proceeds when it observes signal $c$.

$$\Delta_c\left[\sigma_n(\hat{n})^* = 1\right] = \frac{\alpha}{\alpha+(1-\alpha)(1+\sigma_c(\hat{n}))} - \left(\frac{(1-\mu)\alpha}{\alpha+(1-\alpha)\sigma_c(\hat{c})} + \frac{\mu\gamma\alpha}{\alpha+\gamma(1-\alpha)\sigma_c(\hat{c})}\right). \tag{19}$$

Now, let us suppose $\sigma_c(\hat{n})^* = 0$. In this case,

$\Delta_c\left[\sigma_n(\hat{n})^* = 1, \sigma_c(\hat{n})^* = 0\right] = \alpha - \left((1-\mu)\alpha + \frac{\mu\gamma\alpha}{\alpha+\gamma(1-\alpha)}\right) = \frac{\mu\alpha^2(1-\gamma)}{\alpha+\gamma(1-\alpha)} > 0$. Hence, in equilibrium, $\sigma_c(\hat{n})^* > 0$,which implies that $\sigma_c(\hat{c})^* < 1$.

Finally, we prove that the equilibrium is unique and characterize it. Note that $\Delta_c\left[\sigma_n(\hat{n})^* = 1, \sigma_c(\hat{c}) = 1\right] > 0$ and $\Delta_c\left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c})\right]$ is strictly increasing in $\sigma_c(\hat{c})$, as:

$\frac{\partial\Delta_c[\sigma_n(\hat{n})=1,\sigma_c(\hat{c})]}{\partial\sigma_c(\hat{c})} = \alpha\left(\frac{1-\alpha}{(\alpha+(1-\alpha)\sigma_c(\hat{c})-2)^2} + \frac{(1-\alpha)(1-\mu)}{(\alpha+(1-\alpha)\sigma_c(\hat{c}))^2} + \frac{\gamma^2\mu(1-\alpha)}{(\alpha+(1-\alpha)\sigma_c(\hat{c})\gamma)^2}\right) > 0.$

Therefore, $\sigma_c(\hat{c})^*$ has to be either 0 or the only root that equation $\Delta_c\left[\sigma_c(\hat{c}); \sigma_n(\hat{n}) = 1\right] = 0$ has in the interval (0,1), which is $\frac{2\gamma(1-\alpha)-2\alpha+\alpha\mu(1-\gamma)+\sqrt{4\gamma^2+\alpha^2(1-\gamma)^2(2-\mu)^2+4\alpha(1-\gamma)\gamma-(2-3\mu)}}{4(1-\alpha)\gamma}$. Let $\tilde{x}_0$ be that root. Then $\sigma_c(\hat{c})^* = \max\{0, \tilde{x}_0\}$ and consequently $\sigma_c(\hat{n})^* = \min\{1, x_0\}$, with $x_0 = 1 - \tilde{x}_0$. ■

***Proof of Corollary 1***

From (19),

$\Delta_c\left[\sigma_n(\hat{n}) = 1\right] = \alpha\left(\frac{1}{\alpha+(1-\alpha)(1+\sigma_c(\hat{n}))} - \left(\frac{(1-\mu)}{\alpha+(1-\alpha)(1-\sigma_c(\hat{n}))} + \frac{\mu\gamma}{\alpha+\gamma(1-\alpha)(1-\sigma_c(\hat{n}))}\right)\right).$

Let us denote

$$F(\sigma_c(\hat{n}), \alpha) = \frac{1}{\alpha+(1-\alpha)(1+\sigma_c(\hat{n}))} - \left(\frac{(1-\mu)}{\alpha+(1-\alpha)(1-\sigma_c(\hat{n}))} + \frac{\mu\gamma}{\alpha+\gamma(1-\alpha)(1-\sigma_c(\hat{n}))}\right). \tag{20}$$

In equilibrium, $\Delta_c\left[\sigma_n(\hat{n})^* = 1, \sigma_c(\hat{n})^*\right] = 0 \iff F(\sigma_c(\hat{n})^*, \alpha) = 0$.

Now, by the implicit function theorem,

$$\frac{\partial\sigma_c(\hat{n})^*}{\partial\alpha} = -\frac{\frac{\partial F(\sigma_c(\hat{n})^*,\alpha)}{\partial\alpha}}{\frac{\partial F(\sigma_c(\hat{n})^*,\alpha)}{\partial\sigma_c(\hat{n})^*}},$$

where,

$\frac{\partial F(\sigma_c(\hat{n})^*,\alpha)}{\partial\alpha} = \frac{\sigma_c(\hat{n})^*}{(\alpha+(1-\alpha)(1+\sigma_c(\hat{n})^*))^2} + (1-\mu)\frac{\sigma_c(\hat{n})^*}{(\alpha+(1-\alpha)(1-\sigma_c(\hat{n})^*))^2} + \gamma\mu\frac{\gamma\sigma_c(\hat{n})^*+1-\gamma}{(\alpha+\gamma(1-\alpha)(1-\sigma_c(\hat{n})^*))^2} > 0,$

$\frac{\partial F(\sigma_c(\hat{n})^*,\alpha)}{\partial \sigma_c(\hat{n})^*} = -\frac{1-\alpha}{(\alpha+(1-\alpha)(1+\sigma_c(\hat{n})^*))^2} - \left((1-\alpha)\frac{1-\mu}{(\alpha+(1-\alpha)(1-\sigma_c(\hat{n})^*))^2} + \gamma^2\mu\frac{1-\alpha}{(\alpha+\gamma(1-\alpha)(1-\sigma_c(\hat{n})^*))^2}\right) < 0.$

Consequently,

$\frac{\partial \sigma_c(\hat{n})^*}{\partial \alpha} > 0.$ ■

***Proof of Corollary 2***

Let us now denote by $F(\sigma_c(\hat{n})^*, \mu)$ the right hand side of equation (20). By the implicit function theorem,

$\frac{\partial \sigma_c(\hat{n})^*}{\partial \mu} = -\frac{\frac{\partial F(\sigma_c(\hat{n})^*,\mu)}{\partial \mu}}{\frac{\partial F(\sigma_c(\hat{n})^*,\mu)}{\partial \sigma_c(\hat{n})^*}},$

where,

$\frac{\partial F(\sigma_c(\hat{n})^*,\mu)}{\partial \mu} = \frac{\alpha(1-\gamma)}{(\alpha+(1-\alpha)(1-\sigma_c(\hat{n})^*))(\alpha+\gamma(1-\alpha)(1-\sigma_c(\hat{n})^*))} > 0,$

$\frac{\partial F(\sigma_c(\hat{n})^*,\mu)}{\partial \sigma_c(\hat{n})^*} < 0$ (shown in Corollary 1).

Consequently,

$\frac{\partial \sigma_c(\hat{n})^*}{\partial \mu} > 0.$ ■

<u>**Additional results**</u>

**Lemma 3.** *There exists $\bar{\alpha} \in (0,1)$ such that for all $\alpha > \bar{\alpha}$, $\sigma_c(\hat{n})^* = 1$.*

***Proof.***

First, Corollary 1 shows that $\sigma_c(\hat{n})^*$ is increasing in $\alpha$. Second, from the proof of Corollary 1, it follows that $\Delta_c\left[\sigma_n(\hat{n}) = 1\right] > 0 \iff F(\sigma_c(\hat{n}), \alpha) > 0$ .

Now, since

$F(\sigma_c(\hat{n}), \alpha = 0) = -\frac{2\sigma_c(\hat{n})}{(1-\sigma_c(\hat{n}))(1+\sigma_c(\hat{n}))} < 0$, and

$F(\sigma_c(\hat{n}), \alpha = 1) = 1 - ((1-\mu) + \mu\gamma) = \mu(1-\gamma) > 0,$

we have $\Delta_c\left[\sigma_n(\hat{n}) = 1, \alpha = 1\right] > 0$, and thus $\sigma_c(\hat{n})^* = 1$ for $\alpha = 1$. From here, the proof follows. ■

**Lemma 4.** $\frac{\partial \sigma_c(\hat{n})^*}{\partial \gamma} < 0$

***Proof.***

Let us now denote by $F(\sigma_c(\hat{n})^*, \gamma)$ the right hand side of equation (20). By the implicit function theorem,

$\frac{\partial \sigma_c(\hat{n})^*}{\partial \gamma} = -\frac{\frac{\partial F(\sigma_c(\hat{n})^*,\gamma)}{\partial \gamma}}{\frac{\partial F(\sigma_c(\hat{n})^*,\gamma)}{\partial \sigma_c(\hat{n})^*}},$

where,

$\frac{\partial F(\sigma_c(\hat{n})^*,\gamma)}{\partial \gamma} = -\frac{\mu\alpha}{(\alpha+\gamma(1-\alpha)(1-\sigma_c(\hat{n})^*))^2} < 0,$

$\frac{\partial F(\sigma_c(\hat{n})^*,\gamma)}{\partial \sigma_c(\hat{n})^*} < 0$ (shown in Corollary 1).

Consequently,

$\frac{\partial \sigma_c(\hat{n})^*}{\partial \gamma} < 0.$ ■

## A.2 Monopoly with unbalanced prior

In the main body of the paper, we differentiate two cases: $\theta < \frac{1}{2}$ and $\theta > \frac{1}{2}$. This was done for expositional purposes. However, here, there is no need for such a differentiation. Thus, next result (Proposition 2) considers the two cases together, and so holds for any $\theta \in (0,1)$. It then proves Theorem 2.

Before going into this proof, note that the only difference with respect to the monopoly scenario is that instead of considering beliefs (1)-(4), we now have to consider beliefs (7)-(10). As for the functions $\Delta_n$ and $\Delta_c$, they are those in (5) and (6), with $P(C \mid c) = \frac{\theta\gamma}{\theta\gamma+(1-\theta)(1-\gamma)}$ and $P(C \mid n) = \frac{\theta(1-\gamma)}{\theta(1-\gamma)+(1-\theta)\gamma}$.

**Proposition 2.** *Let $\theta \in (0,1)$. There exist $\bar{\theta}_1$, $\bar{\theta}_2$ and $\bar{\theta}_3$, with $0 < \bar{\theta}_1 < \bar{\theta}_2 < \bar{\theta}_3 < 1$. For each $\theta \in (1/2, 1)$ there is a unique equilibrium. In the equilibrium:*

1. *If $\theta \in (0, \bar{\theta}_1)$, $\sigma_n(\hat{n})^* = 1$ and $\sigma_c(\hat{n})^* = min\{1, x_1\} > 0$,*
2. *If $\theta \in (\bar{\theta}_1, \bar{\theta}_2)$, $\sigma_n(\hat{n})^* = 1$ and $\sigma_c(\hat{c})^* = 1$,*
3. *If $\theta \in (\bar{\theta}_2, \bar{\theta}_3)$, $\sigma_n(\hat{c})^* = x_2 > 0$ and $\sigma_c(\hat{c})^* = 1$,*
4. *If $\theta \in (\bar{\theta}_3, 1)$, $\sigma_n(\hat{c})^* = 1$ and $\sigma_c(\hat{c})^* = 1$*

*Where $x_1$ is such that $\Delta_c\left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 1 - x_1; \theta\right] = 0$, and $x_2$ is such that $\Delta_n\left[\sigma_n(\hat{n}) = 1 - x_2, \sigma_c(\hat{c}) = 1; \theta\right] = 0$*

***Proof***

The Proposition is proven through eight Lemmas.

**Lemma 5.** *The function $\Delta_n$ is strictly greater than $\Delta_c$.*

***Proof.***

$\Delta_n = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu P(C \mid n)\lambda(\hat{c}, C)) > \Delta_c = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu P(C \mid c)\lambda(\hat{c}, C))$
$\iff P(C \mid n) < P(C \mid c) \iff \frac{\theta(1-\gamma)}{\theta(1-\gamma)+(1-\theta)\gamma} < \frac{\theta\gamma}{\theta\gamma+(1-\theta)(1-\gamma)} \iff \gamma > \frac{1}{2}$. ♦

**Lemma 6.** *The functions $\Delta_n$ and $\Delta_c$ are decreasing in $\theta$.*

***Proof.*** From (5), (6) and (7)-(10), we obtain that, as $\frac{\partial\lambda(\hat{c},C)}{\partial\theta} = 0$, then $\frac{\partial\Delta_n}{\partial\theta} = \frac{\partial\lambda(\hat{n},0)}{\partial\theta} - \left((1-\mu)\frac{\partial\lambda(\hat{c},0)}{\partial\theta} + \mu\frac{\partial P(C|n)}{\partial\theta}\lambda(\hat{c}, C)\right)$ and $\frac{\partial\Delta_c}{\partial\theta} = \frac{\partial\lambda(\hat{n},0)}{\partial\theta} - \left((1-\mu)\frac{\partial\lambda(\hat{c},0)}{\partial\theta} + \mu\frac{\partial P(C|c)}{\partial\theta}\lambda(\hat{c}, C)\right)$,

with,

$\frac{\partial\lambda(\hat{n},0)}{\partial\theta} = \frac{-\alpha(1-\alpha)(\gamma\sigma_c(\hat{n})+(1-\gamma)\sigma_n(\hat{n}))}{(\alpha(1-\theta)+(1-\alpha)((1-\theta)(\gamma\sigma_n(\hat{n})+(1-\gamma)\sigma_c(\hat{n}))+\theta(\gamma\sigma_c(\hat{n})+(1-\gamma)\sigma_n(\hat{n}))))^2} < 0$,
$\frac{\partial\lambda(\hat{c},0)}{\partial\theta} = \frac{\alpha(1-\alpha)(\gamma\sigma_n(\hat{c})+(1-\gamma)\sigma_c(\hat{c}))}{(\alpha\theta+(1-\alpha)(\theta(\gamma\sigma_c(\hat{c})+(1-\gamma)\sigma_n(\hat{c}))+(1-\theta)(\gamma\sigma_n(\hat{c})+(1-\gamma)\sigma_c(\hat{c}))))^2} > 0$,
$\frac{\partial P(C|n)}{\partial\theta} = \frac{(1-\gamma)\gamma}{(\theta+\gamma-2\theta\gamma)^2} > 0$ and $\frac{\partial P(C|c)}{\partial\theta}\frac{(1-\gamma)\gamma}{(\theta+\gamma-2\theta\gamma-1)^2} > 0$.

Consequently,

$\frac{\partial\Delta_n}{\partial\theta} = \frac{\partial\Delta_c}{\partial\theta} < 0$. ♦

**Lemma 7.** *$\Delta_n [\theta = 1] < 0$ and $\Delta_c [\theta = 1] < 0$.*

***Proof.*** Note that $\Delta_n = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu P(C \mid n)\lambda(\hat{c}, C))$. Thus, $\Delta_n [\theta = 1] = 0 - ((1-\mu)\lambda(\hat{c}, 0) + \mu\lambda(\hat{c}, C)) < 0$, since $\lambda(\hat{c}, 0) > 0$, $\lambda(\hat{c}, C) > 0$ and $P(C \mid n) = 1$ for $\theta = 1$.

Analogously, we show $\Delta_c [\theta = 1] = -((1-\mu)\lambda(\hat{c}, 0) + \mu\lambda(\hat{c}, C)) < 0$. ♦

**Lemma 8.** *The function $\Delta_n$ is strictly decreasing in $\sigma_n(\hat{n})$.*

***Proof.*** Note that $\frac{\partial \Delta_n}{\partial \sigma_n(\hat{n})} = \frac{\partial \lambda(\hat{n},0)}{\partial \sigma_n(\hat{n})} - ((1-\mu)\frac{\partial \lambda(\hat{c},0)}{\partial \sigma_n(\hat{n})} + \mu P(C \mid n)\frac{\partial \lambda(\hat{c},C)}{\partial \sigma_n(\hat{n})})$, with

$\frac{\partial \lambda(\hat{n},0)}{\partial \sigma_n(\hat{n})} = \frac{-\alpha(1-\theta)(1-\alpha)(\gamma(1-\theta)+(1-\gamma)\theta)}{(\alpha(1-\theta)+(1-\alpha)((1-\theta)(\gamma\sigma_n(\hat{n})+(1-\gamma)\sigma_c(\hat{n}))+\theta(\gamma\sigma_c(\hat{n})+(1-\gamma)\sigma_n(\hat{n}))))^2} < 0,$

$\frac{\partial \lambda(\hat{c},0)}{\partial \sigma_n(\hat{n})} = \frac{\alpha\theta(1-\alpha)(\gamma(1-\theta)+(1-\gamma)\theta)}{(\alpha\theta+(1-\alpha)(\theta(\gamma\sigma_c(\hat{c})+(1-\gamma)\sigma_n(\hat{c}))+(1-\theta)(\gamma\sigma_n(\hat{c})+(1-\gamma)\sigma_c(\hat{c}))))^2} > 0,$

$\frac{\partial \lambda(\hat{c},C)}{\partial \sigma_n(\hat{n})} = \frac{\alpha(1-\alpha)(1-\gamma)}{(\alpha+(1-\alpha)(\gamma\sigma_c(\hat{c})+(1-\gamma)\sigma_n(\hat{c})))^2} > 0.$

Consequently, $\frac{\partial \Delta_n}{\partial \sigma_n(\hat{n})} < 0$. ♦

**Lemma 9.** *The function $\Delta_c$ is strictly increasing in $\sigma_c(\hat{c})$.*

***Proof.*** Note that $\frac{\partial \Delta_c}{\partial \sigma_c(\hat{c})} = \frac{\partial \lambda(\hat{n},0)}{\partial \sigma_c(\hat{c})} - ((1-\mu)\frac{\partial \lambda(\hat{c},0)}{\partial \sigma_c(\hat{c})} + \mu P(C \mid c)\frac{\partial \lambda(\hat{c},C)}{\partial \sigma_c(\hat{c})})$, with

$\frac{\partial \lambda(\hat{n},0)}{\partial \sigma_c(\hat{c})} = \frac{\alpha(1-\theta)(1-\alpha)((1-\gamma)(1-\theta)+\gamma\theta)}{(\alpha(1-\theta)+(1-\alpha)((1-\theta)(\gamma\sigma_n(\hat{n})+(1-\gamma)\sigma_c(\hat{n}))+\theta(\gamma\sigma_c(\hat{n})+(1-\gamma)\sigma_n(\hat{n}))))^2} > 0,$

$\frac{\partial \lambda(\hat{c},0)}{\partial \sigma_c(\hat{c})} = \frac{-\alpha\theta(1-\alpha)((1-\gamma)(1-\theta)+\gamma\theta)}{(\alpha\theta+(1-\alpha)(\theta(\gamma\sigma_c(\hat{c})+(1-\gamma)\sigma_n(\hat{c}))+(1-\theta)(\gamma\sigma_n(\hat{c})+(1-\gamma)\sigma_c(\hat{c}))))^2} < 0,$

$\frac{\partial \lambda(\hat{c},C)}{\partial \sigma_c(\hat{c})} = \frac{-\alpha(1-\alpha)\gamma}{(\alpha+(1-\alpha)(\gamma\sigma_c(\hat{c})+(1-\gamma)\sigma_n(\hat{c})))^2} < 0.$

Consequently, $\frac{\partial \Delta_c}{\partial \sigma_c(\hat{c})} > 0$. ♦

**Lemma 10.** *The equilibrium is unique.*

***Proof.*** This result is a consequence of Lemmas 8 and 9.

**Lemma 11.** *Let $\bar{\theta}_1$, $\bar{\theta}_2$, and $\bar{\theta}_3$ be thresholds such that*

$\Delta_c \left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 1; \theta = \bar{\theta}_1\right] = 0,$

$\Delta_n \left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 1; \theta = \bar{\theta}_2\right] = 0,$ *and*

$\Delta_n \left[\sigma_n(\hat{n}) = 0, \sigma_c(\hat{c}) = 1; \theta = \bar{\theta}_3\right] = 0.$

*Then,* $\frac{1}{2} < \bar{\theta}_1 < \bar{\theta}_2 < \bar{\theta}_3 < 1.$

***Proof.*** First, it is shown that $\bar{\theta}_1 > \frac{1}{2}$. If $\theta = \frac{1}{2}$, then $\Delta_c \left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 1; \theta = \frac{1}{2}\right] = \frac{\mu\alpha^2(1-\gamma)}{\alpha+\gamma(1-\alpha)} > 0$. Now, from Lemma 6, we know $\frac{\partial \Delta_c}{\partial \theta} < 0$. Then, $\bar{\theta}_1$ must be greater than $\frac{1}{2}$.

The inequality $\bar{\theta}_1 < \bar{\theta}_2$ follows, as $\Delta_n > \Delta_c$, $\frac{\partial \Delta_n}{\partial \theta} < 0$ and $\frac{\partial \Delta_c}{\partial \theta} < 0$ (by Lemma 5 and Lemma 6).

Now, from Lemmas 6 and 8, we have $\bar{\theta}_2 < \bar{\theta}_3$.

Last, since $\Delta_n [\theta = 1] < 0$ (by Lemma 7) and $\frac{\partial \Delta_n}{\partial \theta} < 0$ (by Lemma 6), threshold $\bar{\theta}_3$ must be strictly smaller than 1. ♦

**Lemma 12.** *Suppose $\sigma_c(\hat{c}) = 1$. Then:*

*1) If $\theta \in (0, \bar{\theta}_2)$, $\Delta_n > 0$.*

*2) If $\theta \in (\bar{\theta}_2, \bar{\theta}_3)$, $\Delta_n$ only has one inner root.*

*3) If $\theta \in (\bar{\theta}_3, 1)$, $\Delta_n < 0$.*

***Proof.*** Consider first $\theta \in (0, \bar{\theta}_2)$. As $\Delta_n \left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 1; \theta = \bar{\theta}_2\right] = 0$, $\frac{\partial \Delta_n}{\partial \theta} < 0$ and $\frac{\partial \Delta_n}{\partial \sigma_n(\hat{n})} < 0$ (see Lemmas 11, 6 and 8), we have $\Delta_n > 0$.

Consider now $\theta \in (\bar{\theta}_2, \bar{\theta}_3)$. As $\Delta_n \left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 1; \theta = \bar{\theta}_2\right] = 0$, $\Delta_n \left[\sigma_n(\hat{n}) = 0, \sigma_c(\hat{c}) = 1; \theta = \bar{\theta}_3\right] = 0$, $\frac{\partial \Delta_n}{\partial \theta} < 0$ and $\frac{\partial \Delta_n}{\partial \sigma_n(\hat{n})} < 0$ (see Lemmas 11, 6 and 8), we have that the function $\Delta_n$ has only one inner root (in $\sigma_n(\hat{n})$).

Last, consider $\theta \in (\bar{\theta}_3, 1)$. As $\Delta_n \left[\sigma_n(\hat{n}) = 0, \sigma_c(\hat{c}) = 1; \theta = \bar{\theta}_3\right] = 0$, $\frac{\partial \Delta_n}{\partial \theta} < 0$, $\frac{\partial \Delta_n}{\partial \sigma_n(\hat{n})} < 0$ and $\Delta_n [\theta = 1] < 0$ (see Lemmas 11, 6, 8 and 7), we have $\Delta_n < 0$. ♦

**Lemma 13.** *Suppose $\sigma_n(\hat{n}) = 1$. Then, if $\theta \in (\bar{\theta}_1, 1)$, $\Delta_c < 0$.*

***Proof.*** Consider $\theta \in (\bar{\theta}_1, 1)$. As $\Delta_c \left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 1; \theta = \bar{\theta}_1\right] = 0$, $\frac{\partial \Delta_c}{\partial \theta} < 0$, $\frac{\partial \Delta_c}{\partial \sigma_c(\hat{c})} > 0$ and $\Delta_c [\theta = 1] < 0$ (see Lemmas 11, 6, 9 and 7), we have $\Delta_c < 0$. ♦

Now, there are nine possible equilibrium configurations to analyze.

| | | | | | |
|---|---|---|---|---|---|
| 1. | $\sigma_c(\hat{c})^* = 1$ | $\sigma_n(\hat{n})^* = 1$ | $\iff$ | $\Delta_c \leq 0$ | $\Delta_n \geq 0$ |
| 2. | $0 < \sigma_c(\hat{c})^* < 1$ | $\sigma_n(\hat{n})^* = 1$ | $\iff$ | $\Delta_c = 0$ | $\Delta_n \geq 0$ |
| 3. | $\sigma_c(\hat{c})^* = 0$ | $\sigma_n(\hat{n})^* = 1$ | $\iff$ | $\Delta_c \geq 0$ | $\Delta_n \geq 0$ |
| 4. | $\sigma_c(\hat{c})^* = 1$ | $\sigma_n(\hat{n})^* = 0$ | $\iff$ | $\Delta_c \leq 0$ | $\Delta_n \leq 0$ |
| 5. | $0 < \sigma_c(\hat{c})^* < 1$ | $\sigma_n(\hat{n})^* = 0$ | $\iff$ | $\Delta_c = 0$ | $\Delta_n \leq 0$ |
| 6. | $\sigma_c(\hat{c})^* = 0$ | $\sigma_n(\hat{n})^* = 0$ | $\iff$ | $\Delta_c \geq 0$ | $\Delta_n \leq 0$ |
| 7. | $\sigma_c(\hat{c})^* = 1$ | $0 < \sigma_n(\hat{n})^* < 1$ | $\iff$ | $\Delta_c \leq 0$ | $\Delta_n = 0$ |
| 8. | $0 < \sigma_c(\hat{c})^* < 1$ | $0 < \sigma_n(\hat{n})^* < 1$ | $\iff$ | $\Delta_c = 0$ | $\Delta_n = 0$ |
| 9. | $\sigma_c(\hat{c})^* = 0$ | $0 < \sigma_n(\hat{n})^* < 1$ | $\iff$ | $\Delta_c \geq 0$ | $\Delta_n = 0$ |

Note that from Lemma 5, configurations 5, 6, 8, and 9 cannot be. Then, we next analyze the remaining equilibrium configurations (for each of the intervals of $\theta$ considered in Proposition 2). We do it taking into account the restriction $\Delta_n > \Delta_c$ imposed by Lemma 5.

a) Interval $\theta \in (0, \bar{\theta}_1)$. By Lemma 12, in this interval we have $\Delta_n[\sigma_c(\hat{c}) = 1] > 0$. Then, $\sigma_n(\hat{n})^* = 1$, and thus configurations 4 and 7 cannot be. Hence, only configurations 1, 2 and 3 are left. However, configuration 1 is neither possible. The reason is that if $\sigma_n(\hat{n})^* = 1$, then $\sigma_c(\hat{c})^* < 1$ (since $\Delta_c \left[\sigma_c(\hat{c}) = 1; \sigma_n(\hat{n}) = 1; \theta = \bar{\theta}_1\right] = 0$ and $\frac{\partial \Delta_c}{\partial \theta} < 0$, which implies $\Delta_c \left[\sigma_c(\hat{c}) = 1; \sigma_n(\hat{n}) = 1; \theta < \bar{\theta}_1\right] > 0$, and thus $\sigma_c(\hat{c})^* < 1$). Therefore, only configurations 2 and 3 are possible, and thus $\sigma_n(\hat{n})^* = 1$ and $0 \leq \sigma_c(\hat{c})^* < 1$. Additionally, as $\frac{\partial \Delta_c}{\partial \sigma_c(\hat{c})} > 0$ (see Lemma 9), there is only one equilibrium. Therefore, $\sigma_c(\hat{c})^*$ has to be either 0 or the root of equation $\Delta_c \left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}); \theta < \bar{\theta}_1\right] = 0$ in the interval (0,1). Let $\tilde{x}_1$ be that root. Then $\sigma_c(\hat{c})^* = \max\{0, \tilde{x}_1\}$ and consequently $\sigma_c(\hat{n})^* = \min\{1, x_1\}$, with $\tilde{x}_1 = 1 - x_1$.

b) Interval $\theta \in (\bar{\theta}_1, \bar{\theta}_2)$. The same argument above shows that configurations 4 and 7 can neither be here. Thus, in equilibrium, $\sigma_n(\hat{n})^* = 1$. In this case, if $\sigma_n(\hat{n})^* = 1$, then $\sigma_c(\hat{c})^* = 1$ (because by Lemma 13, if

$\sigma_n(\hat{n})^* = 1$, then $\Delta_c < 0$, and consequently $\sigma_c(\hat{c})^* = 1$).

c) Interval $\theta \in (\bar{\theta}_2, \bar{\theta}_3)$. Analogously to the previous point, by Lemma 13, if $\sigma_n(\hat{n})^* = 1$, then $\Delta_c < 0$, and consequently $\sigma_c(\hat{c})^* = 1$. Thus, configurations 2 and 3 cannot be. The only possible configurations that are left are 1, 4 and 7, which implies that in equilibrium $\sigma_c(\hat{c})^* = 1$. However, configurations 1 and 4 cannot be either. The reason is that by Lemma 12, in this interval, if $\sigma_c(\hat{c}) = 1$, then $\Delta_n$ has only one inner root. Let $\tilde{x}_2$ be that root. Thus, in equilibrium, $0 < \sigma_n(\hat{n})^* = \tilde{x}_2 < 1$ and consequently $0 < \sigma_n(\hat{c})^* = x_2 < 1$, with $x_2 = 1 - \tilde{x}_2$.

d) Interval $\theta \in (\bar{\theta}_3, 1)$. Again, from Lemma 13, if $\sigma_n(\hat{n})^* = 1$, then $\sigma_c(\hat{c})^* = 1$. Thus, only 1, 4 or 7 can be. However, from lemma 12, neither 1 nor 7 can hold. The reason is that in this interval, if $\sigma_c(\hat{c}) = 1$, then $\Delta_n < 0$, and thus $\sigma_n(\hat{n})^* = 0$. Consequently, in equilibrium, $\sigma_c(\hat{c})^* = 1$ and $\sigma_n(\hat{n})^* = 0$. ■

**Additional results**

**Lemma 14.** $\frac{\partial \sigma_c(\hat{n})^*}{\partial \theta} < 0$.

***Proof.***

Since $\frac{\partial \sigma_c(\hat{n})^*}{\partial \theta} = -\frac{\frac{\partial \Delta_c}{\partial \theta}}{\frac{\partial \Delta_c}{\partial \sigma_c(\hat{n})^*}}$, from Lemmas 6 and 9, the proof follows. ■

**Lemma 15.** *For any $\theta \in (0, \bar{\theta}_1)$, there exists $\bar{\alpha} \in (0,1)$ such that for all $\alpha > \bar{\alpha}$, $\sigma_c(\hat{n})^* = 1$*

***Proof.***

First note that from Proposition 2, if $\theta < \bar{\theta}_1$, then $\sigma_n(\hat{n})^* = 1$.

Now, we show that $\Delta_c\left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 0; \alpha\right]$ is increasing in $\alpha$. To this, note that $\Delta_c\left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 0\right] = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu + P(C \mid c)\lambda(\hat{c}, C)) = \frac{\alpha(1-\theta)}{1-\alpha\theta} + \mu + \frac{\gamma\theta\mu}{\gamma - 1 + \theta - 2\gamma\theta} - 1$, and

$\frac{\partial \Delta_c[\sigma_n(\hat{n})=1, \sigma_c(\hat{c})=0]}{\partial \alpha} = \frac{1-\theta}{(\theta\alpha - 1)^2} > 0$.

Finally, note that $\Delta_c\left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 0; \alpha = 0\right] = \mu + \frac{\gamma\theta\mu}{\gamma - 1 + \theta - 2\gamma\theta} - 1 < 0$, which implies that if $\alpha$ is small enough, then $\sigma_c(\hat{n})^* < 1$. Additionally, by Proposition 2, $\sigma_c(\hat{n})^* > 0$. Finally, $\Delta_c\left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 0; \alpha = 1\right] = \mu + \frac{\gamma\theta\mu}{\gamma - 1 + \theta - 2\gamma\theta} > 0$, which implies that if $\alpha$ is high enough, then $\sigma_c(\hat{n})^* = 1$. From here, the proof follows. ■

## A.3 High type plays strategically

In this section, we show that if the high type is strategic, then it is an equilibrium strategy for the high type to always report its signal.

We denote by $\sigma_{hs}(r) \in [0,1]$ the probability that, conditioned on its signal $s$, a high type newspaper takes action $r$. In addition, $\sigma_s(r)$ will continue denoting this probability for the normal type.

**Proposition 3.** *Let $\theta \in (0,1)$. There exist $\bar{\theta}_1$, $\bar{\theta}_2$ and $\bar{\theta}_3$, with $0 < \bar{\theta}_1 < \bar{\theta}_2 < \bar{\theta}_3 < 1$. For each $\theta \in (0,1)$ there is a unique equilibrium. In the equilibrium:*

1. *If* $\theta \in (0, \bar{\theta}_1)$, $\sigma_n(\hat{n})^* = 1$ *and* $\sigma_c(\hat{n})^* = min\{1, x_1\} > 0$,

2. *If* $\theta \in (\bar{\theta}_1, \bar{\theta}_2)$, $\sigma_n(\hat{n})^* = 1$ *and* $\sigma_c(\hat{c})^* = 1$,

3. *If* $\theta \in (\bar{\theta}_2, \bar{\theta}_3)$, $\sigma_n(\hat{c})^* = x_2 > 0$ *and* $\sigma_c(\hat{c})^* = 1$,

4. *If* $\theta \in (\bar{\theta}_3, 1)$, $\sigma_n(\hat{c})^* = 1$ *and* $\sigma_c(\hat{c})^* = 1$

*Where* $x_1$ *is such that* $\Delta_c\left[\sigma_n(\hat{n}) = 1, \sigma_c(\hat{c}) = 1 - x_1; \theta\right] = 0$, *and* $x_2$ *is such that* $\Delta_n\left[\sigma_n(\hat{n}) = 1 - x_2, \sigma_c(\hat{c}) = 1; \theta\right] = 0$

*In addition, if the high type plays strategically, the truthful strategy* ($\sigma_{hc}(\hat{c})^* = 1$ *and* $\sigma_{hn}(\hat{n})^* = 1$) *is an equilibrium strategy for the high type.*

***Proof.***

Proposition 2 shows that if the high type plays the truthful strategy ($\sigma_{hc}(\hat{c})^* = 1$ and $\sigma_{hn}(\hat{n})^* = 1$), the normal type's strategy described above is an equilibrium strategy. Therefore, we only have to show that if the normal type plays such a strategy, the truthful strategy is an equilibrium strategy for the high type. To this, we will assume that the high type plays the truthful strategy, ($\sigma_{hc}(\hat{c})^* = 1$ and $\sigma_{hn}(\hat{n})^* = 1$), and then show that this is indeed an equilibrium strategy.

First, we derive the payoff functions for the high type. As for the normal type, they are defined in equations (5) and (6).

Let $E_h\{\lambda(r, X) \mid s\}$ denote the expected payoff to the high type newspaper when it observes signal $s \in \{n, c\}$ and publishes $r \in \{\hat{n}, \hat{c}\}$.

$E_h\{\lambda(\hat{n}, X) \mid s\} = \lambda(\hat{n}, 0)$

$E_h\{\lambda(\hat{c}, X) \mid n\} = (1-\mu)\lambda(\hat{c}, 0) + \mu[\lambda(\hat{c}, N)] = (1-\mu)\lambda(\hat{c}, 0)$

$E_h\{\lambda(\hat{c}, X) \mid c\} = (1-\mu)\lambda(\hat{c}, 0) + \mu[\lambda(\hat{c}, C)]$

Now, we define the expected gain to reporting $\hat{n}$ rather than $\hat{c}$, after observing signal $s$, as $\Delta_s^h = E_h\{\lambda(\hat{n}, X) \mid s\} - E_h\{\lambda(\hat{c}, X) \mid s\}$.

Substituting, we obtain:

$$\Delta_n^h = \lambda(\hat{n}, 0) - (1-\mu)\lambda(\hat{c}, 0)$$
$$\Delta_c^h = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu\lambda(\hat{c}, C))$$

**Claim 1.** $\Delta_n^h > \Delta_n > \Delta_c > \Delta_c^h$.

***Proof.***

First, note that from Lemma 5, $\Delta_n > \Delta_c$.

Additionally,

$\Delta_n^h = \lambda(\hat{n}, 0) - (1-\mu)\lambda(\hat{c}, 0) > \Delta_n = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu P(C \mid n)\lambda(\hat{c}, C))$, and

$\Delta_c^h = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu\lambda(\hat{c}, C)) < \Delta_c = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu P(C \mid c)\lambda(\hat{c}, C))$.

Consequently, $\Delta_n^h > \Delta_n > \Delta_c > \Delta_c^h$. ♦

Next, we go into the analysis of the nine possible equilibrium configurations for the normal type, enumerated in the proof of Proposition 2. There, we showed that configurations 5, 6, 8 and 9 could not be in equilibrium (as $\Delta_n > \Delta_c$). This is also the case now. Then, we next analyze the equilibrium configurations that are left: 1, 2, 3, 4 and 7; and show that for none of them, the high type has an incentive to deviate from the truthful strategy.

Configuration 1: In this case, $\Delta_c \leq 0$. Then, from Claim 1, $\Delta_c^h < 0$, and thus $\sigma_{hc}(\hat{c})^* = 1$. In addition, $\Delta_n \geq 0$, consequently, $\Delta_n^h > 0$, and thus $\sigma_{hn}(\hat{n})^* = 1$.

Configuration 2: This case is analogous to the previous one.

Configuration 3: Since $\Delta_n \geq 0$, then $\Delta_n^h > 0$ and thus $\sigma_{hn}(\hat{n})^* = 1$. Because under this configuration, the normal type never sends $\hat{c}$, if $\hat{c}$ were to be reported, the newspaper would assigned a probability one of being the high type. Consequently, $\Delta_c^h = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu\lambda(\hat{c}, C)) = \lambda(\hat{n}, 0) - 1 < 0$, which implies $\sigma_{hc}(\hat{c})^* = 1$.

Configuration 4: Since $\Delta_c \leq 0$, then $\Delta_c^h < 0$, and thus $\sigma_{hc}(\hat{c})^* = 1$. Because under this configuration, the normal type never sends $\hat{n}$, if $\hat{n}$ were to be reported, the newspaper would assign a probability one of being the high type. Consequently, $\Delta_n^h = \lambda(\hat{n}, 0) - (1-\mu)\lambda(\hat{c}, 0) = 1 - (1-\mu)\lambda(\hat{c}, 0) > 0$, which implies $\sigma_{hn}(\hat{n})^* = 1$.

Configuration 7: This case is analogous to Configuration 1.

Then, the true strategy is an equilibrium strategy. ■

The next result shows that the equilibrium above is unique. To this, we make the following assumption: In equilibrium, the high type matches the state of the world more often than the normal type.[40] Formally, it implies $\frac{P(\hat{c}|\aleph,C)}{P(\hat{c}|H,C)} < \frac{P(\hat{c}|\aleph,N)}{P(\hat{c}|H,N)}$, where $P(\hat{c} \mid \aleph, C)$ is the probability that a normal type ($\aleph$) reports $\hat{c}$ when the state of the world is $C$. Analogously, $P(\hat{c} \mid H, C)$ is the probability that a high type ($H$) reports $\hat{c}$ when the state of the world is $C$ and so on, so forth. It is straightforward to prove that if $\frac{P(\hat{c}|\aleph,C)}{P(\hat{c}|H,C)} < \frac{P(\hat{c}|\aleph,N)}{P(\hat{c}|H,N)}$, then $\lambda(\hat{c}, C) > \lambda(\hat{c}, N)$.

**Corollary 6.** *If $\frac{P(\hat{c}|\aleph,C)}{P(\hat{c}|H,C)} < \frac{P(\hat{c}|\aleph,N)}{P(\hat{c}|H,N)}$, then the equilibrium described in Proposition 3 is unique.*

***Proof.***

First, note that from the proof of Proposition 2, we know that if the high type plays the true strategy, then the equilibrium strategy of the normal type is unique.

Then, we just have to show that the true strategy is the only equilibrium strategy for the high type. To this, we first rewrite the functions $\Delta_n$, $\Delta_c$, $\Delta_n^h$ and $\Delta_c^h$, to take into account the fact that the high type can now lie and report $\hat{c}$ when its signal indicates $n$ (in which case, the real state is $N$). They are:

[40] Note that this is a quite mild assumption. Nonetheless, if it were not the case, it would not make sense for a consumer to assign a reputational reward to a high type.

$\Delta_n = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu\,(P(C \mid n)\lambda(\hat{c}, C) + P(N \mid n)\lambda(\hat{c}, N)))$,
$\Delta_c = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu\,(P(C \mid c)\lambda(\hat{c}, C) + P(N \mid c)\lambda(\hat{c}, N)))$,
$\Delta_n^h = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu\lambda(\hat{c}, N))$, and
$\Delta_c^h = \lambda(\hat{n}, 0) - ((1-\mu)\lambda(\hat{c}, 0) + \mu\lambda(\hat{c}, C))$.

It is straightforward to show that $P(C \mid c) > P(C \mid n)$, with $P(N \mid c) = 1 - P(C \mid c)$ and $P(N \mid n) = 1 - P(C \mid n)$.

As $\frac{P(\hat{c}|\aleph,C)}{P(\hat{c}|H,C)} < \frac{P(\hat{c}|\aleph,N)}{P(\hat{c}|H,N)}$, then $\lambda(\hat{c}, C) > \lambda(\hat{c}, N)$, which implies:
$\lambda(\hat{c}, N) < P(C \mid n)\lambda(\hat{c}, C) + P(N \mid n)\lambda(\hat{c}, N) < P(C \mid c)\lambda(\hat{c}, C) + P(N \mid c)\lambda(\hat{c}, N) < \lambda(\hat{c}, C)$.

Consequently, $\Delta_n^h > \Delta_n > \Delta_c > \Delta_c^h$. The rest of the proof is analogous to the proof of Proposition 3. ■

## A.4 Competition and feedback power

In this section we consider the beliefs in (11)-(16) and the functions $\Delta_n$ and $\Delta_c$ defined in (17) and (18).

***Proof of Theorem 3***

The Theorem is proven through three Lemmas.

**Lemma 16.** *If $0 < \mu_J < \mu_{J+1} < 1$, then $\Delta_n > \Delta_c$.*

***Proof.*** $\Delta_n > \Delta_c$

$\iff$

$(1-\mu_J)\lambda(\hat{n}, 0) + \mu_J\gamma\lambda(\hat{n}, N) - ((1-\mu_{J+1})\lambda(\hat{c}, 0) + \mu_{J+1}(1-\gamma)\lambda(\hat{c}, C)) >$
$(1-\mu_J)\lambda(\hat{n}, 0) + \mu_J(1-\gamma)\lambda(\hat{n}, N) - ((1-\mu_{J+1})\lambda(\hat{c}, 0) + \mu_{J+1}\gamma\lambda(\hat{c}, C))$

$\iff$

$\mu_J\lambda(\hat{n}, N)(2\gamma - 1) > \mu_{J+1}\lambda(\hat{c}, C)(1 - 2\gamma)$.

Since $\gamma > 1/2$, the proof follows. ♦

**Lemma 17.** *If $0 < \mu_J < \mu_{J+1} < 1$ and $\sigma_c(\hat{c}) = 1$, then $\Delta_n > 0$.*

***Proof.***

$\Delta_n\,[\sigma_c(\hat{c}) = 1] =$
$= \frac{(1-\mu_J)\alpha}{\alpha+(1-\alpha)\sigma_n(\hat{n})} + \frac{\mu_J\gamma\alpha}{\alpha+(1-\alpha)\gamma\sigma_n(\hat{n})} - (\frac{(1-\mu_{J+1})\alpha}{\alpha+(1-\alpha)(1+\sigma_n(\hat{c}))} + \frac{\mu_{J+1}(1-\gamma)\alpha}{\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_n(\hat{c}))})$.

First, note that if $\mu_J = 0$, Lemma 2 implies $\Delta_n\,[\sigma_c(\hat{c}) = 1, \mu_J = 0] > 0$.

Now, we define $T = \frac{(1-\mu_J)\alpha}{\alpha+(1-\alpha)\sigma_n(\hat{n})} + \frac{\mu_J\gamma\alpha}{\alpha+(1-\alpha)\gamma\sigma_n(\hat{n})}$. Note that if $\frac{\partial T}{\partial \mu_J} < 0$, then $\frac{\partial \Delta_n[\sigma_c(\hat{c})=1]}{\partial \mu_J} < 0$. Consequently, as $\mu_J \in (0, \mu_{J+1})$, to show that $\Delta_n\,[\sigma_c(\hat{c}) = 1] > 0$, it is sufficient to prove that $\Delta_n\,[\sigma_c(\hat{c}) = 1; \mu_J = \mu_{J+1}] > 0$, where

$\Delta_n\left[\sigma_c(\hat{c})=1;\mu_J=\mu_{J+1}\right]=\frac{(1-\mu_{J+1})\alpha}{\alpha+(1-\alpha)\sigma_n(\hat{n})}+\frac{\mu_{J+1}\gamma\alpha}{\alpha+(1-\alpha)\gamma\sigma_n(\hat{n})}-(\frac{(1-\mu_{J+1})\alpha}{\alpha+(1-\alpha)(1+\sigma_n(\hat{c}))}+\frac{\mu_{J+1}(1-\gamma)\alpha}{\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_n(\hat{c}))})$.

Now, since $\gamma>\frac{1}{2}$ and $\sigma_n(\hat{n})\in[0,1]$, with $\sigma_n(\hat{c})=1-\sigma_n(\hat{n})$, we obtain $\frac{(1-\mu_{J+1})\alpha}{\alpha+(1-\alpha)\sigma_n(\hat{n})}>\frac{(1-\mu_{J+1})\alpha}{\alpha+(1-\alpha)(1+\sigma_n(\hat{c}))}$ and $\frac{\mu_{J+1}\gamma\alpha}{\alpha+(1-\alpha)\gamma\sigma_n(\hat{n})}>\frac{\mu_{J+1}(1-\gamma)\alpha}{\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_n(\hat{c}))}$. This completes the proof. ♦

Now, there are nine equilibrium configuration to analyze.

1. $\sigma_c(\hat{c})^*=1 \quad \sigma_n(\hat{n})^*=1 \iff \Delta_c\leq 0 \quad \Delta_n\geq 0.$
2. $\sigma_c(\hat{c})^*<1 \quad \sigma_n(\hat{n})^*=1 \iff \Delta_c=0 \quad \Delta_n\geq 0.$
3. $\sigma_c(\hat{c})^*=0 \quad \sigma_n(\hat{n})^*=1 \iff \Delta_c\geq 0 \quad \Delta_n\geq 0.$
4. $\sigma_c(\hat{c})^*=1 \quad \sigma_n(\hat{n})^*=0 \iff \Delta_c\leq 0 \quad \Delta_n\leq 0.$
5. $\sigma_c(\hat{c})^*<1 \quad \sigma_n(\hat{n})^*=0 \iff \Delta_c=0 \quad \Delta_n\leq 0.$
6. $\sigma_c(\hat{c})^*=0 \quad \sigma_n(\hat{n})^*=0 \iff \Delta_c\geq 0 \quad \Delta_n\leq 0.$
7. $\sigma_c(\hat{c})^*=1 \quad \sigma_n(\hat{n})^*<1 \iff \Delta_c\leq 0 \quad \Delta_n=0.$
8. $\sigma_c(\hat{c})^*<1 \quad \sigma_n(\hat{n})^*<1 \iff \Delta_c=0 \quad \Delta_n=0.$
9. $\sigma_c(\hat{c})^*=0 \quad \sigma_n(\hat{n})^*<1 \iff \Delta_c\geq 0 \quad \Delta_n=0.$

Note that from Lemma 16, configurations 5, 6, 8 and 9 cannot be. Similarly, from Lemma 17, configurations 4 and 7 can neither be. Consequently, $\sigma_n(\hat{n})^*=1$. Then, taking into account the restriction imposed by Lemma 16, the resulting possible configurations are:

1. $\sigma_c(\hat{c})^*=1 \quad \sigma_n(\hat{n})^*=1 \iff \Delta_c\leq 0 \quad \Delta_n\geq 0.$
2. $\sigma_c(\hat{c})^*<1 \quad \sigma_n(\hat{n})^*=1 \iff \Delta_c=0 \quad \Delta_n>0.$
3. $\sigma_c(\hat{c})^*=0 \quad \sigma_n(\hat{n})^*=1 \iff \Delta_c\geq 0 \quad \Delta_n>0.$

Let us now consider $\sigma_n(\hat{n})^*=1$ and analyze how the normal newspaper proceeds when it observes signal $c$.

$\Delta_c=(1-\mu_J)\lambda(\hat{n},0)+\mu_J(1-\gamma)\lambda(\hat{n},N)-((1-\mu_{J+1})\lambda(\hat{c},0)+\mu_{J+1}\gamma\lambda(\hat{c},C))$
$=\frac{(1-\mu_J)\alpha}{\alpha+(1-\alpha)(\sigma_c(\hat{n})+\sigma_n(\hat{n}))}+\frac{\mu_J(1-\gamma)\alpha}{\alpha+(1-\alpha)(\gamma\sigma_n(\hat{n})+(1-\gamma)\sigma_c(\hat{n}))}-(\frac{(1-\mu_{J+1})\alpha}{\alpha+(1-\alpha)(\sigma_c(\hat{c})+\sigma_n(\hat{c}))}+\frac{\mu_{J+1}\gamma\alpha}{\alpha+(1-\alpha)(\gamma\sigma_c(\hat{c})+(1-\gamma)\sigma_n(\hat{c}))})$.

$\Delta_c[\sigma_n(\hat{n})^*=1]=$
$=\frac{(1-\mu_J)\alpha}{\alpha+(1-\alpha)(\sigma_c(\hat{n})+1)}+\frac{\mu_J(1-\gamma)\alpha}{\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_c(\hat{n}))}-(\frac{(1-\mu_{J+1})\alpha}{\alpha+(1-\alpha)\sigma_c(\hat{c})}+\frac{\mu_{J+1}\gamma\alpha}{\alpha+(1-\alpha)\gamma\sigma_c(\hat{c})})$.

Now, let us suppose $\sigma_c(\hat{n})^*=0$. In this case,
$\Delta_c[\sigma_n(\hat{n})^*=1,\sigma_c(\hat{n})^*=0]=(1-\mu_J)\alpha+\frac{\mu_J(1-\gamma)\alpha}{\alpha+(1-\alpha)\gamma}-((1-\mu_{J+1})\alpha+\frac{\mu_{J+1}\gamma\alpha}{\alpha+(1-\alpha)\gamma})$
$=\frac{(1-\mu_J)\alpha(\alpha+(1-\alpha)\gamma)+\mu_J(1-\gamma)\alpha-(1-\mu_{J+1})\alpha(\alpha+(1-\alpha)\gamma)-\mu_{J+1}\gamma\alpha}{\alpha+(1-\alpha)\gamma}$

$= \frac{\alpha^2(\mu_{J+1}-\mu_J)+\alpha(1-\alpha)\gamma(\mu_{J+1}-\mu_J)+\alpha(\mu_J(1-\gamma)-\mu_{J+1}\gamma)}{\alpha+(1-\alpha)\gamma}$
$= \alpha(\mu_{J+1}-\mu_J) + \frac{\alpha(\mu_J(1-\gamma)-\mu_{J+1}\gamma)}{\alpha+(1-\alpha)\gamma} > 0 \Leftrightarrow \gamma < \frac{\mu_J+\alpha(\mu_{J+1}-\mu_J)}{2\mu_J+\alpha(\mu_{J+1}-\mu_J)}$.

Let $\hat{\gamma} = \frac{\mu_J+\alpha(\mu_{J+1}-\mu_J)}{2\mu_J+\alpha(\mu_{J+1}-\mu_J)}$, with $\hat{\gamma} \in (0,1)$. Hence, in equilibrium, $\sigma_c(\hat{n})^* > 0$ for $\gamma < \hat{\gamma} \in (0,1)$, and $\sigma_c(\hat{c})^* = 1$ for $\gamma > \hat{\gamma}$.

To conclude, note that function $\Delta_c[\sigma_c(\hat{c}), \sigma_n(\hat{n}) = 1]$ is strictly increasing in $\sigma_c(\hat{c})$,

$$\frac{\partial\Delta_c[\sigma_c(\hat{c}),\sigma_n(\hat{n})=1]}{\partial\sigma_c(\hat{c})} = \frac{\alpha(1-\alpha)(1-\mu_J)}{(\alpha+(1-\alpha)\sigma_c(\hat{c})-2)^2} + \frac{\alpha\mu_J(1-\alpha)(1-\gamma)^2}{((\alpha-1)\sigma_c+(1-\alpha)\gamma\sigma_c(\hat{c})+1)^2} + \frac{\alpha(1-\mu_{J+1})(1-\alpha)}{(\alpha+(1-\alpha)\sigma_c(\hat{c}))^2} + \frac{\alpha\gamma^2(1-\alpha)\mu_{J+1}}{(\alpha+(1-\alpha)\gamma\sigma_c(\hat{c}))^2} > 0. \quad (21)$$

Then, there is only one equilibrium. Now, if $\gamma < \hat{\gamma}$, $\sigma_c(\hat{c})^*$ has to be either 0 or the root of equation $\Delta_c[\sigma_c(\hat{c}), \sigma_n(\hat{n}) = 1] = 0$ in the interval (0,1). Let $\tilde{x}_3$ be that root. Then $\sigma_c(\hat{c})^* = \max\{0, \tilde{x}_3\}$ and consequently $\sigma_c(\hat{n})^* = \min\{1, x_3\}$, with $\tilde{x}_3 = 1 - x_3$. Additionally, if $\gamma \geq \hat{\gamma}$, there is only one equilibrium in which $\sigma_c(\hat{c})^* = 1$. ■

***Proof of Corollary 3***

$\frac{\partial\sigma_c(\hat{n})^*}{\partial\gamma} = -\frac{\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\gamma}}{\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\sigma_c(\hat{n})^*}}$, with

$\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\gamma} = -\frac{\mu_J\alpha}{(\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_c(\hat{n})))^2} - \frac{\mu_{J+1}\alpha^2}{(\alpha+(1-\alpha)\gamma\sigma_c(\hat{c}))^2} < 0$,

$\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\sigma_c(\hat{n})^*} < 0$, as $\frac{\partial\Delta_c[\sigma_n(\hat{n})=1]}{\partial\sigma_c(\hat{c})} > 0$ as shown in equation (21) and $\sigma_c(\hat{n})^* = 1 - \sigma_c(\hat{c})$.

Consequently, $\frac{\partial\sigma_c(\hat{n})^*}{\partial\gamma} < 0$. ■

***Proof of Corollary 4***

Note that $\frac{\partial\sigma_c(\hat{n})^*}{\partial\mu_J} = -\frac{\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\mu_J}}{\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\sigma_c(\hat{n})^*}}$ and $\frac{\partial\sigma_c(\hat{n})^*}{\partial\mu_{J+1}} = -\frac{\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\mu_{J+1}}}{\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\sigma_c(\hat{n})^*}}$, with

$\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\mu_J} = \frac{-\gamma\alpha^2+\alpha(1-\alpha)(1-2\gamma)}{(\alpha+(1-\alpha)(\sigma_c(\hat{n})+1))(\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_c(\hat{n})))} < 0$,

$\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\mu_{J+1}} = \frac{\alpha^2(1-\gamma)}{(\alpha+(1-\alpha)\sigma_c(\hat{c}))(\alpha+(1-\alpha)\gamma\sigma_c(\hat{c}))} > 0$,

$\frac{\partial\Delta_c[\sigma_n(\hat{n})^*=1]}{\partial\sigma_c(\hat{n})} = -\frac{\alpha(1-\alpha)(1-\mu_J)}{(\alpha+(1-\alpha)(\sigma_c(\hat{n})+1))^2} - \frac{\alpha\mu_J(1-\alpha)(1-\gamma)^2}{(\alpha+(1-\alpha)(\gamma+(1-\gamma)\sigma_c(\hat{n})))^2} - \frac{\alpha(1-\alpha)(1-\mu_J)}{(\alpha+(1-\alpha)(1-\sigma_c(\hat{n})))^2} - \frac{\alpha\gamma^2\mu_{J+1}(1-\alpha)}{(\alpha+(1-\alpha)\gamma(1-\sigma_c(\hat{n})))^2} <$ 0.

Consequently, $\frac{\partial\sigma_c(\hat{n})^*}{\partial\mu_J} < 0$ and $\frac{\partial\sigma_c(\hat{n})^*}{\partial\mu_{J+1}} > 0$. ■

# B Appendix: Analysis of a strategic duopoly

Appendix B describes and analyzes a game with exclusively two (strategic) scoop-firms, denoted by 1 and 2, that seek to maximize their reputation for high quality. Each newspaper $i \in \{1, 2\}$ is as the scoop-firm in Section 3. That is, each newspaper $i$ receives a signal $s_i \in \{n_i, c_i\}$ on the state of the world $w \in \{N, C\}$, and then chooses a report $r_i \in \{\hat{n}_i, \hat{c}_i\}$ to print. Each state occurs with equal probability. Newspapers can be either high type or normal type, with $\alpha$ being the probability that a newspaper is high ($H_i$). A high type

firm receives a perfectly informative signal and reports it honestly. A normal newspaper receives a signal of quality $\gamma > \frac{1}{2}$.

For any $i \in \{1, 2\}$, we denote by $\sigma^i_{s_i}(r_i)$ the probability that normal newspaper $i$ takes action $r_i \in \{\hat{n}, \hat{c}\}$, conditioned on having received signal $s_i \in \{n, c\}$. With respect to the probability of feedback, notice that with two newspapers we have three situations: (i) Both newspapers publish $\hat{n}$, in which case we stick to the formulation in Section 2 and assume there is no feedback, thus $X = 0$; (ii) One newspaper publishes $\hat{n}$ and the other $\hat{c}$, in which case we assume there is probability $\mu_1$ of feedback, then $X \in \{N, C, 0\}$; (iii) The two newspapers print $\hat{c}$, in which case the probability of feedback is $\mu_2$, with $\mu_2 > \mu_1 > 0$. Then $X \in \{N, C, 0\}$. We will further assume $\mu_2 = 2\mu_1$, with $\mu_1 \in (0, \frac{1}{2}]$, $\mu_2 \in (0, 1]$. Note that under this assumption, $\mu_2 - \mu_1 = \mu_1$. Thus, ceteris paribus firm $j$'s report, firm $i$'s feedback power is always $\mu_1$.

## B.1 Beliefs

We now turn to analyzing the beliefs. Let $\lambda_i(r_i, r_j, X)$, with $i, j \in \{1, 2\}$, $i \neq j$, be the posterior probability that the consumers place on newspaper $i$ being the high type. We differentiate two cases, with feedback and without feedback.

***With feedback:***

Let $r_i \in \{\hat{n}_i, \hat{c}_i\}$, with $i \in \{1, 2\}$, $i \neq j$, and $X \in \{N, C\}$. Then,

$$\begin{aligned}
\lambda_i(r_i, r_j, X) &= P(H_i \mid r_i, r_j, X) = \frac{P(H_i, r_i, r_j, X)}{P(r_i, r_j, X)} \\
&= \frac{P(r_j \mid r_i, X, H_i)P(r_i \mid X, H_i)P(X \mid H_i)P(H_i)}{P(r_j \mid r_i, X, H_i)P(r_i \mid X, H_i)P(X \mid H_i)P(H_i) + P(r_j \mid r_i, X, L_i)P(r_i \mid X, L_i)P(X \mid L_i)P(L_i)} \\
&= \frac{P(r_i \mid X, H_i)P(H_i)}{P(r_i \mid X, H_i)P(H_i) + P(r_i \mid X, L_i)P(L_i)}.
\end{aligned}$$

Applying this rule we obtain:

$$\lambda_i(\hat{n}_i, \hat{c}_j, N) = \frac{\alpha}{\alpha + (1-\alpha)(\gamma\sigma^i_n(\hat{n}) + (1-\gamma)\sigma^i_c(\hat{n}))} \tag{22}$$

$$\lambda_i(\hat{n}_i, \hat{c}_j, C) = 0 \tag{23}$$

$$\lambda_i(\hat{c}_i, r_j, N) = 0 \qquad \forall r_j \in \{\hat{n}_j, \hat{c}_j\} \tag{24}$$

$$\lambda_i(\hat{c}_i, r_j, C) = \frac{\alpha}{\alpha + (1-\alpha)(\gamma\sigma^i_c(\hat{c}) + (1-\gamma)\sigma^i_n(\hat{c}))} \qquad \forall r_j \in \{\hat{n}_j, \hat{c}_j\} \tag{25}$$

***Without feedback:***

The analysis is now a bit more complex, as in the absence of feedback, the reports sent by the two newspapers give information about the most likely state of the world. Let $r_i \in \{\hat{n}_i, \hat{c}_i\}$, with $i \in \{1, 2\}$, $i \neq j$, and $X = \{0\}$. Then,

$$\lambda_i(r_i, r_j, 0) = P(H_i \mid r_i, r_j) = \frac{P(r_j \mid r_i, H_i)P(r_i \mid H_i)P(H_i)}{P(r_j \mid r_i, H_i)P(r_i \mid H_i)P(H_i) + P(r_j \mid r_i, L_i)P(r_i \mid L_i)P(L_i)},$$

where,

$$\begin{aligned} P(r_j \mid r_i, L_i) &= P(r_j \mid r_i, L_i, C)P(C \mid r_i, L_i) + P(r_j \mid r_i, L_i, N)P(N \mid r_i, L_i) \\ &= P(r_j \mid C)P(C \mid r_i, L_i) + P(r_j \mid N)P(N \mid r_i, L_i), \end{aligned}$$

with,

$$\begin{aligned} P(C \mid r_i, L_i) &= \frac{P(r_i \mid L_i, C)P(L_i \mid C)P(C)}{P(r_i \mid L_i, C)P(L_i \mid C)P(C) + P(r_i \mid L_i, N)P(L_i \mid N)P(N)} \\ P(N \mid r_i, L_i) &= \frac{P(r_i \mid L_i, N)P(L_i \mid N)P(N)}{P(r_i \mid L_i, C)P(L_i \mid C)P(C) + P(r_i \mid L_i, N)P(L_i \mid N)P(N)}. \end{aligned}$$

Hence,

$$\begin{aligned} P(r_j \mid r_i, L_i) &= \frac{P(r_j \mid C)P(r_i \mid L_i, C)P(L_i \mid C)P(C) + P(r_j \mid N)P(r_i \mid L_i, N)P(L_i \mid N)P(N)}{P(r_i \mid L_i, C)P(L_i \mid C)P(C) + P(r_i \mid L_i, N)P(L_i \mid N)P(N)} \\ &= \frac{P(r_j \mid C)P(r_i \mid L_i, C) + P(r_j \mid N)P(r_i \mid L_i, N)}{P(r_i \mid L_i, C) + P(r_i \mid L_i, N)} \\ &= \frac{P(r_j \mid C)P(r_i \mid L_i, C) + P(r_j \mid N)P(r_i \mid L_i, N)}{2P(r_i \mid L_i)}. \end{aligned}$$

So,

$$\begin{aligned} \lambda_i(r_i, r_j, 0) =& \frac{P(r_j \mid r_i, H_i)\frac{1}{2}\alpha}{P(r_j \mid r_i, H_i)\frac{1}{2}\alpha + \frac{P(r_j|C)P(r_i|L_i,C)+P(r_j|N)P(r_i|L_i,N)}{2P(r_i|L_i)}P(r_i \mid L_i)(1-\alpha)} \\ =& \frac{P(r_j \mid r_i, H_i)\alpha}{P(r_j \mid r_i, H_i)\alpha + (P(r_j \mid C)P(r_i \mid L_i, C) + P(r_j \mid N)P(r_i \mid L_i, N))(1-\alpha)}. \end{aligned}$$

Now, let us consider $r_i = \hat{c_i}$ and $r_j = \hat{c_j}$. Since $P(\hat{c_j} \mid \hat{c_i}, H_i) = P(\hat{c_j} \mid C)$ in this case, we have:

$$\lambda_i(\hat{c_i}, \hat{c_j}, 0) = \frac{\alpha}{\alpha + (1-\alpha)(P(\hat{c_i} \mid L_i, C) + P(\hat{c_i} \mid L_i, N)\frac{P(\hat{c_j}|N)}{P(\hat{c_j}|C)})}.$$

Using a similarly argument we obtain

$$\begin{aligned} \lambda_i(\hat{c_i}, \hat{n_j}, 0) =& \frac{\alpha}{\alpha + (1-\alpha)(P(\hat{c_i} \mid L_i, C) + P(\hat{c_i} \mid L_i, N)\frac{P(\hat{n_j}|N)}{P(\hat{n_j}|C)})} \\ \lambda_i(\hat{n_i}, \hat{n_j}, 0) =& \frac{\alpha}{\alpha + (1-\alpha)(P(\hat{n_i} \mid L_i, C)\frac{P(\hat{n_j}|C)}{P(\hat{n_j}|N)} + P(\hat{n_i} \mid L_i, N))} \\ \lambda_i(\hat{n_i}, \hat{c_j}, 0) =& \frac{\alpha}{\alpha + (1-\alpha)(P(\hat{n_i} \mid L_i, C)\frac{P(\hat{c_j}|C)}{P(\hat{c_j}|N)} + P(\hat{n_i} \mid L_i, N)}, \end{aligned}$$

with:

$$
\begin{aligned}
P(\hat{c_i} \mid L_i, C) =& \gamma\sigma_c^i(\hat{c}) + (1-\gamma)\sigma_n^i(\hat{c}) \\
P(\hat{c_i} \mid L_i, N) =& \gamma\sigma_n^i(\hat{c}) + (1-\gamma)\sigma_c^i(\hat{c}) \\
P(\hat{n_i} \mid L_i, C) =& \gamma\sigma_c^i(\hat{n}) + (1-\gamma)\sigma_n^i(\hat{n}) \\
P(\hat{n_i} \mid L_i, N) =& \gamma\sigma_n^i(\hat{n}) + (1-\gamma)\sigma_c^i(\hat{n}) \\
P(\hat{c_j} \mid C) =& \alpha + (1-\alpha)(\gamma\sigma_c^j(\hat{c}) + (1-\gamma)\sigma_n^j(\hat{c})) \\
P(\hat{c_j} \mid N) =& (1-\alpha)(\gamma\sigma_n^j(\hat{c}) + (1-\gamma)\sigma_c^j(\hat{c})) \\
P(\hat{n_j} \mid N) =& \alpha + (1-\alpha)(\gamma\sigma_n^j(\hat{n}) + (1-\gamma)\sigma_c^j(\hat{n})) \\
P(\hat{n_j} \mid C) =& (1-\alpha)(\gamma\sigma_c^j(\hat{n}) + (1-\gamma)\sigma_n^j(\hat{n})).
\end{aligned}
$$

Putting all together, we obtain:

$$
\lambda_i(\hat{c_i}, \hat{c_j}, 0) = \frac{\alpha}{\alpha+(1-\alpha)(\gamma\sigma_c^i(\hat{c})+(1-\gamma)\sigma_n^i(\hat{c})+(\gamma\sigma_n^i(\hat{c})+(1-\gamma)\sigma_c^i(\hat{c}))\frac{(1-\alpha)(\gamma\sigma_n^j(\hat{c})+(1-\gamma)\sigma_c^j(\hat{c}))}{\alpha+(1-\alpha)(\gamma\sigma_c^j(\hat{c})+(1-\gamma)\sigma_n^j(\hat{c}))})} \tag{26}
$$

$$
\lambda_i(\hat{c_i}, \hat{n_j}, 0) = \frac{\alpha}{\alpha+(1-\alpha)(\gamma\sigma_c^i(\hat{c})+(1-\gamma)\sigma_n^i(\hat{c})+(\gamma\sigma_n^i(\hat{c})+(1-\gamma)\sigma_c^i(\hat{c}))\frac{\alpha+(1-\alpha)(\gamma\sigma_n^j(\hat{n})+(1-\gamma)\sigma_c^j(\hat{n}))}{(1-\alpha)(\gamma\sigma_c^j(\hat{n})+(1-\gamma)\sigma_n^j(\hat{n}))})} \tag{27}
$$

$$
\lambda_i(\hat{n_i}, \hat{n_j}, 0) = \frac{\alpha}{\alpha+(1-\alpha)(\gamma\sigma_c^i(\hat{n})+(1-\gamma)\sigma_n^i(\hat{n}))\frac{(1-\alpha)(\gamma\sigma_c^j(\hat{n})+(1-\gamma)\sigma_n^j(\hat{n}))}{\alpha+(1-\alpha)(\gamma\sigma_n^j(\hat{n})+(1-\gamma)\sigma_c^j(\hat{n}))}+\gamma\sigma_n^i(\hat{n})+(1-\gamma)\sigma_c^i(\hat{n}))} \tag{28}
$$

$$
\lambda_i(\hat{n_i}, \hat{c_j}, 0) = \frac{\alpha}{\alpha+(1-\alpha)((\gamma\sigma_c^i(\hat{n})+(1-\gamma)\sigma_n^i(\hat{n}))\frac{\alpha+(1-\alpha)(\gamma\sigma_c^j(\hat{c})+(1-\gamma)\sigma_n^j(\hat{c}))}{(1-\alpha)(\gamma\sigma_n^j(\hat{c})+(1-\gamma)\sigma_c^j(\hat{c}))}+\gamma\sigma_n^i(\hat{n})+(1-\gamma)\sigma_c^i(\hat{n})}. \tag{29}
$$

## B.2 Analysis

Let $E\{\lambda_i(r_i) \mid s_i\}$ be the expected payoff to newspaper $i \in \{1,2\}$ when it observes signal $s_i \in \{n_i, c_i\}$ and reports $r_i \in \{\hat{n_i}, \hat{c_i}\}$.

$$
\begin{aligned}
E\{\lambda_i(\hat{n_i}) \mid s_i\} &= P(\hat{n_j} \mid \hat{n_i}, s_i)E\{\lambda_i(\hat{n_i}, \hat{n_j}, X) \mid s_i\} + P(\hat{c_j} \mid \hat{n_i}, s_i)E\{\lambda_i(\hat{n_i}, \hat{c_j}, X) \mid s_i\} \\
E\{\lambda_i(\hat{c_i}) \mid s_i\} &= P(\hat{n_j} \mid \hat{c_i}, s_i)E\{\lambda_i(\hat{c_i}, \hat{n_j}, X) \mid s_i\} + P(\hat{c_j} \mid \hat{c_i}, s_i)E\{\lambda_i(\hat{c_i}, \hat{c_j}, X) \mid s_i\},
\end{aligned}
$$

where

$$
\begin{aligned}
P(\hat{n_j} \mid \hat{n_i}, n_i) &= P(\hat{n_j} \mid \hat{c_i}, n_i) = P(\hat{n_j} \mid n_i) \\
&= P(\hat{n_j} \mid n_i, C)P(C \mid n_i) + P(\hat{n_j} \mid n_i, N)P(N \mid n_i) \\
&= (1-\alpha)(\gamma\sigma_c^j(\hat{n}) + (1-\gamma)\sigma_n^j(\hat{n}))(1-\gamma) + (\alpha + (1-\alpha)(\gamma\sigma_n^j(\hat{n}) + (1-\gamma)\sigma_c^j(\hat{n})))\gamma, \\
P(\hat{n_j} \mid \hat{n_i}, c_i) &= P(\hat{n_j} \mid \hat{c_i}, c_i) = P(\hat{n_j} \mid c_i) \\
&= P(\hat{n_j} \mid c_i, C)P(C \mid c_i) + P(\hat{n_j} \mid c_i, N)P(N \mid c_i) \\
&= (1-\alpha)(\gamma\sigma_c^j(\hat{n}) + (1-\gamma)\sigma_n^j(\hat{n}))\gamma + (\alpha + (1-\alpha)(\gamma\sigma_n^j(\hat{n}) + (1-\gamma)\sigma_c^j(\hat{n})))(1-\gamma), \\
P(\hat{c_j} \mid \hat{n_i}, n_i) &= P(\hat{c_j} \mid \hat{c_i}, n_i) = P(\hat{c_j} \mid n_i) \\
&= P(\hat{c_j} \mid n_i, C)P(C \mid n_i) + P(\hat{c_j} \mid n_i, N)P(N \mid n_i) \\
&= (\alpha + (1-\alpha)(\gamma\sigma_c^j(\hat{c}) + (1-\gamma)\sigma_n^j(\hat{c})))(1-\gamma) + (1-\alpha)(\gamma\sigma_n^j(\hat{c}) + (1-\gamma)\sigma_c^j(\hat{c}))\gamma, \\
P(\hat{c_j} \mid \hat{n_i}, c_i) &= P(\hat{c_j} \mid \hat{c_i}, c_i) = P(\hat{c_j} \mid c_i) \\
&= P(\hat{c_j} \mid c_i, C)P(C \mid c_i) + P(\hat{c_j} \mid c_i, N)P(N \mid c_i) \\
&= (\alpha + (1-\alpha)(\gamma\sigma_c^j(\hat{c}) + (1-\gamma)\sigma_n^j(\hat{c})))\gamma + (1-\alpha)(\gamma\sigma_n^j(\hat{c}) + (1-\gamma)\sigma_c^j(\hat{c}))(1-\gamma),
\end{aligned}
$$

and

$$
\begin{aligned}
E\{\lambda_i(\hat{n_i}, \hat{n_j}, X) \mid s_i\} =& \lambda_i(\hat{n_i}, \hat{n_j}, 0), \\
E\{\lambda_i(\hat{n_i}, \hat{c_j}, X) \mid n_i\} =& (1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1(\gamma\lambda_i(\hat{n_i}, \hat{c_j}, N) + (1-\gamma)\lambda_i(\hat{n_i}, \hat{c_j}, C)) \\
=& (1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1\gamma\lambda_i(\hat{n_i}, \hat{c_j}, N), \\
E\{\lambda_i(\hat{n_i}, \hat{c_j}, X) \mid c_i\} =& (1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1(\gamma\lambda_i(\hat{n_i}, \hat{c_j}, C) + (1-\gamma)\lambda_i(\hat{n_i}, \hat{c_j}, N)) \\
=& (1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{n_i}, \hat{c_j}, N), \\
E\{\lambda_i(\hat{c_i}, \hat{n_j}, X) \mid n_i\} =& (1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1(\gamma\lambda_i(\hat{c_i}, \hat{n_j}, N) + (1-\gamma)\lambda_i(\hat{c_i}, \hat{n_j}, C)) \\
=& (1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{c_i}, \hat{n_j}, C), \\
E\{\lambda_i(\hat{c_i}, \hat{n_j}, X) \mid c_i\} =& (1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1(\gamma\lambda_i(\hat{c_i}, \hat{n_j}, C) + (1-\gamma)\lambda_i(\hat{c_i}, \hat{n_j}, N)) \\
=& (1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1\gamma\lambda_i(\hat{c_i}, \hat{n_j}, C), \\
E\{\lambda_i(\hat{c_i}, \hat{c_j}, X) \mid n_i\} =& (1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2(\gamma\lambda_i(\hat{c_i}, \hat{c_j}, N) + (1-\gamma)\lambda_i(\hat{c_i}, \hat{c_j}, C)) \\
=& (1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2(1-\gamma)\lambda_i(\hat{c_i}, \hat{c_j}, C), \\
E\{\lambda_i(\hat{c_i}, \hat{c_j}, X) \mid c_i\} =& (1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2(\gamma\lambda_i(\hat{c_i}, \hat{c_j}, C) + (1-\gamma)\lambda_i(\hat{c_i}, \hat{c_j}, N)) \\
=& (1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2\gamma\lambda_i(\hat{c_i}, \hat{c_j}, C).
\end{aligned}
$$

with $E\{\lambda_i(r_i, r_j, X) \mid s_i\}$ denoting the expected payoff to newspaper $i \in \{1, 2\}$ when it observes signal $s_i \in \{n_i, c_i\}$, publishes $r_i \in \{\hat{n_i}, \hat{c_i}\}$ and newspaper $j$ reports $r_j \in \{\hat{n_j}, \hat{c_j}\}$.

Now, let $\Delta_{s_i}[\sigma_n^1(\hat{n}), \sigma_c^1(\hat{c}), \sigma_n^2(\hat{n}), \sigma_c^2(\hat{c})]$ be the expected gain to newspaper $i$ from reporting $\hat{n_i}$ rather than $\hat{c_i}$, after observing signal $s_i \in \{n_i, c_i\}$:

$$\Delta_{n_i}[\sigma_n^1(\hat{n}), \sigma_c^1(\hat{c}), \sigma_n^2(\hat{n}), \sigma_c^2(\hat{c})] = E\{\lambda_i(\hat{n_i}) \mid n_i\} - E\{\lambda_i(\hat{c_i}) \mid n_i\} \tag{30}$$

$$\Delta_{c_i}[\sigma_n^1(\hat{n}), \sigma_c^1(\hat{c}), \sigma_n^2(\hat{n}), \sigma_c^2(\hat{c})] = E\{\lambda_i(\hat{n_i}) \mid c_i\} - E\{\lambda_i(\hat{c_i}) \mid c_i\}, \tag{31}$$

with

$$\begin{aligned}
\Delta_{n_i} =& E\{\lambda_i(\hat{n_i}) \mid n_i\} - E\{\lambda_i(\hat{c_i}) \mid n_i\} \\
=& (P(\hat{n_j} \mid n_i)E\{\lambda_i(\hat{n_i}, \hat{n_j}, X) \mid n_i\} + P(\hat{c_j} \mid n_i)E\{\lambda_i(\hat{n_i}, \hat{c_j}, X) \mid n_i\}) \\
-& (P(\hat{n_j} \mid n_i)E\{\lambda_i(\hat{c_i}, \hat{n_j}, X) \mid n_i\} + P(\hat{c_j} \mid n_i)E\{\lambda_i(\hat{c_i}, \hat{c_j}, X) \mid n_i\}) \\
=& P(\hat{n_j} \mid n_i)(E\{\lambda_i(\hat{n_i}, \hat{n_j}, X) \mid n_i\} - E\{\lambda_i(\hat{c_i}, \hat{n_j}, X) \mid n_i\}) \\
+& P(\hat{c_j} \mid n_i)(E\{\lambda_i(\hat{n_i}, \hat{c_j}, X) \mid n_i\} - E\{\lambda_i(\hat{c_i}, \hat{c_j}, X) \mid n_i\}) \\
=& P(\hat{n_j} \mid n_i)(\lambda_i(\hat{n_i}, \hat{n_j}, 0) - ((1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{c_i}, \hat{n_j}, C))) \\
+& P(\hat{c_j} \mid n_i)((1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1\gamma\lambda_i(\hat{n_i}, \hat{c_j}, N) - ((1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2(1-\gamma)\lambda_i(\hat{c_i}, \hat{c_j}, C)))
\end{aligned}$$

$$\begin{aligned}
\Delta_{c_i} =& E\{\lambda_i(\hat{n_i}) \mid c_i\} - E\{\lambda_i(\hat{c_i}) \mid c_i\} \\
=& (P(\hat{n_j} \mid c_i)E\{\lambda_i(\hat{n_i}, \hat{n_j}, X) \mid c_i\} + P(\hat{c_j} \mid c_i)E\{\lambda_i(\hat{n_i}, \hat{c_j}, X) \mid c_i\}) \\
-& (P(\hat{n_j} \mid c_i)E\{\lambda_i(\hat{c_i}, \hat{n_j}, X) \mid c_i\} + P(\hat{c_j} \mid n_i)E\{\lambda_i(\hat{c_i}, \hat{c_j}, X) \mid c_i\}) \\
=& P(\hat{n_j} \mid c_i)(E\{\lambda_i(\hat{n_i}, \hat{n_j}, X) \mid c_i\} - E\{\lambda_i(\hat{c_i}, \hat{n_j}, X) \mid c_i\}) \\
+& P(\hat{c_j} \mid c_i)(E\{\lambda_i(\hat{n_i}, \hat{c_j}, X) \mid c_i\} - E\{\lambda_i(\hat{c_i}, \hat{c_j}, X) \mid c_i\}) \\
=& P(\hat{n_j} \mid c_i)(\lambda_i(\hat{n_i}, \hat{n_j}, 0) - ((1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1\gamma\lambda_i(\hat{c_i}, \hat{n_j}, C))) \\
+& P(\hat{c_j} \mid c_i)((1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{n_i}, \hat{c_j}, N) - ((1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2\gamma\lambda_i(\hat{c_i}, \hat{c_j}, C))),
\end{aligned}$$

and beliefs given by equations (22)-(25) and (26)-(29).

## B.3 Results

We focus on symmetric equilibria, that is equilibria in which, conditioned on the same signal, the two newspapers 1 and 2 take the same action. This is $\sigma_n^1(\hat{n}) = \sigma_n^2(\hat{n}) = \sigma_n(\hat{n})$ and $\sigma_c^1(\hat{c}) = \sigma_c^2(\hat{c}) = \sigma_c(\hat{c})$, for $(\sigma_n(\hat{n}), \sigma_c(\hat{c})) \in [0,1]^2$. The definition of an equilibrium is as in Remark 1 in Section 2.

We next present the only analytical result we could derive for this case. It restricts attention to non-perverse equilibrium, satisfying $\sigma_c^i(\hat{c}) > \sigma_c^i(\hat{n})$ and $\sigma_n^i(\hat{n}) > \sigma_n^i(\hat{c})$, $\forall i \in \{1,2\}$. Note that this result extends Lemma 1 in Appendix A to the case of two strategic firms.

**Lemma 18.** *Consider a non-perverse equilibrium. If $\mu_2 > \mu_1 > 0$, then $\Delta_{n_i} > \Delta_{c_i}$ $\forall i \in \{1, 2\}$.*

***Proof.***

Note that $\Delta_n$ and $\Delta_c$ are given by expressions:

$$\begin{aligned}\Delta_{n_i} =& P(\hat{n_j} \mid n_i)(\lambda_i(\hat{n_i}, \hat{n_j}, 0) - ((1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{c_i}, \hat{n_j}, C))) \\ &+ P(\hat{c_j} \mid n_i)((1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1\gamma\lambda_i(\hat{n_i}, \hat{c_j}, N) - ((1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2(1-\gamma)\lambda_i(\hat{c_i}, \hat{c_j}, C)))\end{aligned}$$

$$\begin{aligned}\Delta_{c_i} =& P(\hat{n_j} \mid c_i)(\lambda_i(\hat{n_i}, \hat{n_j}, 0) - ((1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1\gamma\lambda_i(\hat{c_i}, \hat{n_j}, C))) \\ &+ P(\hat{c_j} \mid c_i)((1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{n_i}, \hat{c_j}, N) - ((1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2\gamma\lambda_i(\hat{c_i}, \hat{c_j}, C))).\end{aligned}$$

For convenience, we rewrite them as:

$$\begin{aligned}\Delta_{n_i} =& P(\hat{n_j} \mid n_i)\,(N1 - N2) + P(\hat{c_j} \mid n_i)\,(N3 - N4) \\ \Delta_{c_i} =& P(\hat{n_j} \mid c_i)\,(C1 - C2) + P(\hat{c_j} \mid c_i)\,(C3 - C4)\,.\end{aligned}$$

Clearly, $N1 - N2 > C1 - C2$ and $N3 - N4 > C3 - C4$.

We are now in position to show that if $\mu_2 > \mu_1 > 0$, $\Delta_{n_i} > \Delta_{c_i}$. We do it in several steeps.

1. First, we prove that $N1 - N2 > C3 - C4$. Note that this requires $\lambda_i(\hat{n_i}, \hat{n_j}, 0) - ((1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{c_i}, \hat{n_j}, C)) > (1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{n_i}, \hat{c_j}, N) - ((1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2\gamma\lambda_i(\hat{c_i}, \hat{c_j}, C))$. The following two conditions are sufficient for this inequality to hold:

   (a) $\lambda_i(\hat{n_i}, \hat{n_j}, 0) > (1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{n_i}, \hat{c_j}, N)$. Note that, in a non-perverse equilibrium, $\sigma_c^i(\hat{c}) > \sigma_c^i(\hat{n})$ and $\sigma_n^i(\hat{n}) > \sigma_n^i(\hat{c})$, which implies $\lambda_i(\hat{n_i}, \hat{n_j}, 0) > \lambda_i(\hat{n_i}, \hat{c_j}, 0)$ and $\lambda_i(\hat{n_i}, \hat{n_j}, 0) > \frac{1}{2}\lambda_i(\hat{n_i}, \hat{c_j}, N)$.

   (b) $(1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{c_i}, \hat{n_j}, C) < (1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2\gamma\lambda_i(\hat{c_i}, \hat{c_j}, C)$. Since $\lambda_i(\hat{c_i}, \hat{n_j}, 0) < \lambda_i(\hat{c_i}, \hat{c_j}, 0)$, $\lambda_i(\hat{c_i}, \hat{n_j}, C) = \lambda_i(\hat{c_i}, \hat{c_j}, C)$, $\mu_2 > \mu_1$ and $\gamma > \frac{1}{2}$, the condition holds.

2. Second, we prove that $C1 - C2 > C3 - C4$. Note that this requires $\lambda_i(\hat{n_i}, \hat{n_j}, 0) - ((1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1\gamma\lambda_i(\hat{c_i}, \hat{n_j}, C)) > (1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{n_i}, \hat{c_j}, N) - ((1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2\gamma\lambda_i(\hat{c_i}, \hat{c_j}, C))$. The following two conditions are sufficient for this inequality to hold:

   (a) $\lambda_i(\hat{n_i}, \hat{n_j}, 0) > (1-\mu_1)\lambda_i(\hat{n_i}, \hat{c_j}, 0) + \mu_1(1-\gamma)\lambda_i(\hat{n_i}, \hat{c_j}, N)$. This is condition 1.(a).

   (b) $(1-\mu_1)\lambda_i(\hat{c_i}, \hat{n_j}, 0) + \mu_1\gamma\lambda_i(\hat{c_i}, \hat{n_j}, C) < (1-\mu_2)\lambda_i(\hat{c_i}, \hat{c_j}, 0) + \mu_2\gamma\lambda_i(\hat{c_i}, \hat{c_j}, C)$. Since $\lambda_i(\hat{c_i}, \hat{n_j}, 0) < \lambda_i(\hat{c_i}, \hat{c_j}, 0)$, $\lambda_i(\hat{c_i}, \hat{n_j}, C) = \lambda_i(\hat{c_i}, \hat{c_j}, C)$ and $\mu_2 > \mu_1$, the condition holds.

Now, there are two possibilities: $N1 - N2 < N3 - N4$ and $N1 - N2 > N3 - N4$.

- If $N1 - N2 < N3 - N4$, it follows directly that $N3 - N4 > N1 - N2 > C1 - C2 > C3 - C4$, and so, that for any $P(r_j \mid s_i)$, with $r_j \in \{n_j\hat{,}c_j\}$ and $s_i \in \{n_i, c_i\}$, $\Delta_{n_i} > \Delta_{c_i}$.

- If $N1 - N2 > N3 - N4$, we also have $N1 - N2 > C1 - C2 > C3 - C4$. Note that in a non-perverse equilibrium, $P(\hat{n_j} \mid n_i) > P(\hat{n_j} \mid c_i) \Leftrightarrow (2\gamma - 1)(\alpha + (1-\alpha)(\sigma_n^j(\hat{n})(2\gamma - 1) + \sigma_c^j(\hat{n})(1 - 2\gamma))) > 0$ which, since $\gamma > \frac{1}{2}$, always holds. Similarly we obtain $P(\hat{c_j} \mid c_i) > P(\hat{c_j} \mid n_i)$. Taking this into account, we obtain $\Delta_{n_i} > \Delta_{c_i}$. ■

In the numerical calculus of the equilibria that we show next we consider $\mu_2 = 2\mu_1$, with $\mu_1 \in (0, \frac{1}{2}]$ and $\mu_2 \in (0, 1]$. This will allow us to analyze the effect of a variation in a firm's feedback power on its silence.

The analysis shows that, in equilibrium, normal newspapers always stick to a signal $n$. That is, $\sigma_n^1(\hat{n})^* = \sigma_n^2(\hat{n})^* = 1$. This is the same result than in the model considered in Section 3. Thus, the values below refer to the probability that in equilibrium a normal newspaper sticks to signal $c$, or in other words the probability that a firm prints a scoop, $\sigma_c(\hat{c})^*$. Notice that these values are unique; then, the equilibrium is unique.

[Table 1 should be placed about here]

[Table 2 should be placed about here]

There are two important conclusions we can derive from Tables 1 and 2.

First, with competition high quality newspapers stick to their signal, but not low quality ones (those with low values of $\gamma$). To see this, note that in Table 1 ($\gamma = 0.6$), there are parameters configuration for which $\sigma_c(\hat{c})^* < 1$; whereas in Table 2 ($\gamma = 0.8$) we always obtain $\sigma_c(\hat{c})^* = 1$ (same result if $\gamma$ is higher). That is, in the presence of a second strategic newspaper, only low quality firms may find it optimal to silence scandals, $\sigma_c(\hat{n})^* > 0$. Notice that this is the idea of the result in Theorem 3, Section 3.

Second, increasing the feedback power of a firm increases its silence. To see this, focus on Table 1 and note that an increase in $\mu_1$ unambiguously decreases $\sigma_c(\hat{c})^*$. Notice that this idea is also expressed in Corollary 4 in Section 3, that shows $\frac{\partial \sigma_c(\hat{n})^*}{\partial \mu_{J+1}} > 0$.

At this point, the reader may note that there is one result of Section 3 that Tables 1-2 alone cannot explain. Namely, that an increase in competition reduces media silence, i.e. $\frac{\partial \sigma_c(\hat{n})^*}{\partial \mu_J} < 0$. To this, we next present two tables with the numerical results for the monopoly scenario.

[Table 3 should be placed about here]

[Table 4 should be placed about here]

Comparing the numerical results in Tables 3-4 with those in Tables 1-2, we observe that competition disciplines all news organizations. To this, we compare the feedback of the industry in the monopoly and the duopoly, i.e., $\mu$ and $\mu_2$, respectively. Now, note that for any $\alpha$, $\gamma$ and $\mu_2 = \mu$, the value in the monopoly (Tables 3-4) is always smaller than the corresponding value in the duopoly (Tables 1-2). That is, the probability that

a firm lets a scandal go in the paper (conditioned on having received signal $c$) is higher in the duopoly than in the monopoly. Notice that this idea is also expressed in Corollary 4 in Section 3, that shows $\frac{\partial \sigma_c(\hat{n})^*}{\partial \mu_J} < 0$.

It is worth mentioning here that the numerical calculus done in Appendix B shows a very important disciplining effect of competition. It is our impression that this so important role of competition is rooted in the simplifying assumption made in Appendix B that newspapers are identical and equilibria are symmetric. Notice that this scenario reflects a situation in which the media industry is very competitive, as no firm dominates and the two newspapers have the same feedback power. If we were to consider firms with different feedback powers, we think the disciplining role of competition would be less extreme. We think this a really interesting exercise that however requires a less complex departing model from where to obtain clean analytical results.

# References


Anderson, William L (2004), 'Facts, fiction and the fourth state: The Washington Post and "Jimmy's World"', *American Journal of Economics and Sociology* **63**(5), 965–986.

Andina-Díaz, Ascensión (2009), 'Media competition and information disclosure', *Social Choice and Welfare* **33**(2), 261–280.

Andina-Díaz, Ascensión (2011), 'Mass media in economics: Origins and subsequent contributions', *Cuadernos de Ciencias Económicas y Empresariales* **61**(1), 89–101.

Baron, David P. (2006), 'Persistent media bias', *Journal of Public Economics* **90**(1-2), 1–36.

Besley, Timothy and Andrea Prat (2006), 'Handcuffs for the grabbing hand? Media capture and government accountability', *The American Economic Review* **96**(3), 720–736.

Bovitz, Gregory L., James N. Druckman and Arthur Lupia (2002), 'When can a news organization lead public opinion? Ideology versus market forces in decisions to make news', *Public Choice* **113**(1), 127–155.

Corneo, Giacomo (2006), 'Media capture in a democracy: The role of wealth concentration', *Journal of Public Economics* **90**(1-2), 37–58.

Cummins, Jason G. and Ingmar Nyman (2005), 'The dark side of competitive pressure', *RAND Journal of Economics* **36**(2), 361–377.

Djankov, Simeon, Caralee McLiesh, Tatiana Nenova and Andrei Shleifer (2003), 'Who owns the media?', *Journal of Law and Economics* **46**(2), 341–381.

Durante, Ruben and Brian Knight (2012), 'Partisan control, media bias, and viewer responses: Evidence from Berlusconi's Italy', *Journal of the European Economic Association* **10**(3), 451–481.

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

Ellman, Matthew and Fabrizio Germano (2009), 'What do the papers sell? A model of advertising and media bias', *Economic Journal* **119**(537), 680–704.

Entman, Robert M. (2012), *Scandal and silence: Media responses to presidential misconduct*, Malden, MA: Polity Press.

Fox, Justin and Richard Van Weelden (2012), 'Costly transparency', *Journal of Public Economics* **96**(1), 142–150.

Garoupa, Nuno (1999*a*), 'Dishonesty and libel law: The economics of the "chilling" effect', *Journal of Institutional and Theoretical Economics* **155**(2), 284–300.

Garoupa, Nuno (1999*b*), 'The economics of political dishonesty and defamation', *International Review of Law and Economics* **19**(2), 167–180.

Gentzkow, Matthew, Edward L. Glaeser and Claudia Goldin (2006), *In Corruption and Reform: Lessons from America's Economic History*, NBER, chapter The Rise of the Fourth Estate: How Newspapers Became Informative and Why it Mattered.

Gentzkow, Matthew and Jesse M. Shapiro (2006), 'Media bias and reputation', *Journal of Political Economy* **114**(2), 280–316.

Gentzkow, Matthew, Jesse M. Shapiro and Daniel F. Stone (2014), Media bias in the marketplace: Theory, Technical report, NBER Working Paper 19880.

Gratton, Gabriele (2015), 'The sound of silence: Political accountability and libel law', *European Journal of Political Economy* **37**(1), 266–279.

Heidhues, Paul and Johan Lagerlöf (2003), 'Hiding information in electoral competition', *Games and Economic Behavior* **42**(1), 48–74.

Hörner, Johannes (2002), 'Reputation and competition', *American Economic Review* **92**(3), 644–663.

Kovach, Bill and Tom Rosenstiel (2001), *The elements of journalism*, New Yok, NY: Crown Publishers.

Latham, Oliver (2015), 'Lame ducks and the media', *Economic Journal (forthcoming)* .

Levy, Gilat (1997), 'Decision making in committees: Transparency, reputation, and voting rules', *American Economic Review* **97**(1), 150–168.

Logan, Brian and Daniel Sutter (2004), 'Newspaper quality, Pulitzer Prizes, and newspaper circulation', *Atlantic Economic Journal* **32**(2), 100–112.

Milbourn, Todd T., Richard L. Shockley and Anjan V. Thakor (2001), 'Managerial career concerns and investments in information', *RAND Journal of Economics* **32**(2), 334–351.

Morris, Stephen (2001), 'Political correctness', *Journal of Political Economy* **109**(2), 231–265.

Mullainathan, Sendhil and Andrei Shleifer (2005), 'The market for news', *American Economic Review* **95**(4), 1031–1053.

Ottaviani, Marco and Peter Sorensen (2001), 'Information aggregation in debate: Who should speak first?', *Journal of Public Economics* **81**(3), 393–421.

Prat, Andrea (2005), 'The wrong kind of transparency', *American Economic Review* **95**(3), 862–877.

Prat, Andrea and David Stromberg (2005), Commercial television and voter information, Technical report, CEPR Discussion Paper 4989.

Prat, Andrea and David Strömberg (2011), The political economy of mass media, Technical report, CEPR Discussion Paper 8246.

Stanig, Piero (2014), 'Regulation of speech and media coverage of corruption: An empirical analysis of the mexican press', *American Journal of Political Science (forthcoming)* .

Stromberg, David (2004), 'Mass media competition, political competition, and public policy', *Review of Economic Studies* **71**(1), 265–284.

Suurmond, Guido, Otto H. Swank and Bauke Visser (2004), 'On the bad reputation of reputational concerns', *Journal of Public Economics* **88**(12), 2817–2838.

Tella, Rafael Di and Ignacio Franceschelli (2011), 'Government advertising and media coverage of corruption scandals', *American Economic Journal: Applied Economics* **3**(4), 119–151.

Vaidya, Samarth (2005), 'Corruption in the media's gaze', *European Journal of Political Economy* **21**(3), 667–687.

| | $\alpha = 0.1$ | $\alpha = 0.3$ | $\alpha = 0.5$ | $\alpha = 0.7$ | $\alpha = 0.9$ |
|---|---|---|---|---|---|
| $\mu_1 = 0.1; \mu_2 = 0.2$ | 1 | 1 | 1 | 1 | 1 |
| $\mu_1 = 0.2; \mu_2 = 0.4$ | 1 | 0.992996 | 0.994957 | 1 | 1 |
| $\mu_1 = 0.3; \mu_2 = 0.6$ | 0.994058 | 0.969039 | 0.950454 | 0.944267 | 0.955871 |
| $\mu_1 = 0.4; \mu_2 = 0.8$ | 0.986915 | 0.942498 | 0.898715 | 0.86336 | 0.850705 |
| $\mu_1 = 0.5; \mu_2 = 1$ | 0.97933 | 0.913114 | 0.838691 | 0.761851 | 0.702157 |

Table 1: Duopoly with $\gamma = 0.6$

| | $\alpha = 0.1$ | $\alpha = 0.3$ | $\alpha = 0.5$ | $\alpha = 0.7$ | $\alpha = 0.9$ |
|---|---|---|---|---|---|
| $\mu_1 = 0.1; \mu_2 = 0.2$ | 1 | 1 | 1 | 1 | 1 |
| $\mu_1 = 0.2; \mu_2 = 0.4$ | 1 | 1 | 1 | 1 | 1 |
| $\mu_1 = 0.3; \mu_2 = 0.6$ | 1 | 1 | 1 | 1 | 1 |
| $\mu_1 = 0.4; \mu_2 = 0.8$ | 1 | 1 | 1 | 1 | 1 |
| $\mu_1 = 0.5; \mu_2 = 1$ | 1 | 1 | 1 | 1 | 1 |

Table 2: Duopoly with $\gamma = 0.8$

| | $\alpha = 0.1$ | $\alpha = 0.3$ | $\alpha = 0.5$ | $\alpha = 0.7$ | $\alpha = 0.9$ |
|---|---|---|---|---|---|
| $\mu = 0.2$ | 0.99297 | 0.975429 | 0.947664 | 0.887794 | 0.6 |
| $\mu = 0.4$ | 0.985762 | 0.94918 | 0.889957 | 0.760983 | 0.139113 |
| $\mu = 0.6$ | 0.978366 | 0.920978 | 0.825492 | 0.614665 | 0 |
| $\mu = 0.8$ | 0.970771 | 0.89046 | 0.75219 | 0.440527 | 0 |
| $\mu = 1$ | 0.962963 | 0.857143 | 0.666667 | 0.222222 | 0 |

Table 3: Monopoly with $\gamma = 0.6$

| | $\alpha = 0.1$ | $\alpha = 0.3$ | $\alpha = 0.5$ | $\alpha = 0.7$ | $\alpha = 0.9$ |
|---|---|---|---|---|---|
| $\mu = 0.2$ | 0.997277 | 0.989896 | 0.977296 | 0.948928 | 0.809884 |
| $\mu = 0.4$ | 0.994527 | 0.979507 | 0.953565 | 0.894745 | 0.605452 |
| $\mu = 0.6$ | 0.99175 | 0.968814 | 0.928697 | 0.836997 | 0.384072 |
| $\mu = 0.8$ | 0.988945 | 0.957795 | 0.902561 | 0.775107 | 0.142222 |
| $\mu = 1$ | 0.986111 | 0.946429 | 0.875 | 0.708333 | 0 |

Table 4: Monopoly with $\gamma = 0.8$