The London School of Economics and Political Science

Three Essays on Media Politics in Democracies and Autocracies

Greg Chih-Hsin Sheen

A thesis submitted to the Department of Government of the London School of Economics for the degree of Doctor of Political Science, London, May 2019

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I declare that my thesis consists of 32,514 words.

Statement of conjoint work

I confirm that Chapter 3 was jointly co-authored with Hans H. Tung and Wen-Chin Wu and I contributed 60% of this work.

Acknowledgements

I am deeply grateful to my supervisors Torun Dewan, Stephane Wolton and Rafael Hortala-Vallve for their invaluable guidance and constant support. Without their comments on my works and advice on professionalization, my PhD is not possible.

I thank Stephane in particular for patiently enduring all my very rough ideas and drafts that I now find very embarrassing, and for constantly providing me important comments.

I would like to thank all the Political Science and Political Economy group members for accompanying me during this journey. In particular, I would like to thank Weihan Ding, Federica Izzo, Clement Minaudier, Nelson Ruiz, Kiwi Ting, Arduino Tomasi, and Tinghua Yu.

I also thank my advisors and friends hosting me when I was visiting the Washington University in St. Louis, in particular, Justin Fox, Keith Schnakenberg, and Myunghoon Kang.

Finally, I would like to thank my coauthors Fang-Yu Chen, Selina Hofstetter, Weihao Huang, Hans H. Tung, and Wen-Chin Wu for their engaging efforts in our works.

My family and friends have been a source of constant support, and I am grateful to my parents, my brother, and my wife Sharon.

Abstract

My dissertation discusses information problems encountered by dictators and voters. In the first essay titled *Media with Reputational Concerns: Yes Men or Watchdogs?* I consider why the media tend to endorse the government instead of reporting critically. Traditionally this is explained by media capture or the policy bias of the media. Analyzing a cheap-talk model, I suggest that the media outlets reputational concerns can on its own cause such media behaviour. This is true even when the media outlets are on average believed to be more competent than the government, and when the media market is perfectly competitive.

In the second essay titled *Tell Me the Truth? Dictatorship and the Commitment to Media Freedom* (jointly with Hans H. Tung and Wen-Chin Wu), we formally show that the dictator faces a commitment problem to uphold a promised level of media freedom. Anticipating the threat after truth-reporting, the media might self-censor their reports in advance. The dictator thus suffers from information insufficiency. This paper further characterizes the situations when the commitment problem is more severe and provides empirical implications that can help reinterpret the recent conclusion on the censorship strategy of the Chinese government, in particular, King, Roberts and Pan (2013).

In the third essay titled *Reputation and Media Selection*, by applying the analytic framework proposed by Tadelis (1999, 2002) that reputation is a tradable asset, I argue that providing reputational information about media outlets to the public, a commonly recommended remedy for addressing low-quality reporting, might not be effective because of the possibility of media ownership transactions. Low-quality editorial teams can survive by acquiring names from reputable newspapers through acquisitions or mergers. I discuss potential institutional and governmental interventions that could improve social welfare once introduced along with provision of media reputational information.

Contents

1 Introduction **9**

2 Media with Reputational Concerns: Yes Men or Watchdogs? **12**

2.1 Introduction 12

2.2 Related Literature 16

2.3 The Model 19

2.4 Analysis 24

2.5 Multiple Media Outlets 37

2.6 Concluding Remarks 42

3 Tell Me the Truth? Dictatorship and the Commitment to Media Freedom **49**

3.1 Introduction 49

3.2 Literature Review 54

3.3 Model 57

3.4 Extensions 71

3.5 Empirical Implications 80

3.6 Conclusion 83

4 Reputation and Media Selection **92**

4.1 Introduction 92

4.2 Related Literature . . . 95
4.3 Model Setup . . . 98
4.4 Analysis . . . 102
4.5 Market Restrictions . . . 106
4.6 Scrutiny of Ownership Transactions . . . 110
4.7 Discussion on the Media's Payoff Structure . . . 111
4.8 Considering Moral Hazard . . . 112
4.9 Concluding Remarks . . . 113

A Appendix for Chapter 2 **120**

A.1 Proofs . . . 120

B Appendix for Chapter 3 **126**

B.1 Proofs . . . 126
B.2 Private Signal . . . 127
B.3 Selection Bias . . . 131
B.4 EM Algorithm . . . 134

List of Figures

2.1 Timeline 23

2.2 Conditions for Truthful Reporting 30

2.3 Comparative Statics 1 31

2.4 Comparative Statics 2 33

2.5 Comparative Statics 3 34

3.1 Process of censorship 53

3.2 Timeline 61

3.3 Commitment Problem 65

3.4 The location of π on μ^P 70

3.5 The incentive for censorship when a society is stable 71

3.6 The incentive for censorship when a society is unstable 71

3.7 The incentive for censorship when a society is stable 72

3.8 The incentive for censorship when a society is unstable 72

3.9 Timeline 78

4.1 Generations of the Editorial Teams 98

Chapter 1

Introduction

This dissertation consists of three essays discussing some information problems encountered by dictators and voters, i.e., whether they can obtain valuable information from the media. In the first essay titled *Media with Reputational Concerns: Yes Men or Watchdogs?*, I consider why the media tend to endorse the government instead of reporting critically. Traditionally this is explained by media capture or the policy bias of the media. Analyzing a cheap-talk model, I suggest that the media outlets' reputational concerns can be an important factor to explain such media behavior. Because a news report does not only indicate the quality of the government but also reflect that of the media outlet, endorsing government decisions usually improves the reputation of the media outlet. This is true even when the media outlets are on average believed to be more competent than the government, and when the media market is perfectly competitive. As a result, when the government is considered to be reasonably competent, the public is unlikely to obtain useful information from the media. Moreover, the model shows that when the quality gap between the high- and low-quality media outlets is widened, the probability of having truthful reporting decreases.

The second essay titled *Tell Me the Truth? Dictatorship and the Commitment to Media Freedom* (jointly with Hans H. Tung and Wen-Chin Wu) formalizes a more detailed

censorship procedure than those analyzed by the existing literature. In the literature, researchers assume that the media will truthfully report information whenever a dictator allows media freedom. However, this assumption is rather unrealistic considering that the constitutions of many authoritarian regimes guarantee freedom of speech, but censorship prevails. In our framework, a dictator can always censor and punish media outlets that report inconvenient truths after news is uncovered but before it is publicized. Studying this important phase in censorship, we show that the dictator faces a commitment problem to uphold a promised level of media freedom. Anticipating the threat after truth-reporting, the media might self-censor their reports in advance. The dictator thus suffers from information insufficiency. Importantly, censorship and self-censorship yield very different welfare outcomes to the dictator. The dictator learns the information before censorship, but does not learn that if it is self-censored by the media outlet. This paper characterizes the situations when the commitment problem is more severe. For example, when the society is rather stable, the dictator does not need to rely on positive news from credible media to diffuse public dissents, and hence has less incentive to commit to media freedom. It turns out that the dictator's information problem is more severe when in a stable society than when in a rather volatile one. Finally, this analysis provides empirical implications that can help reinterpret the recent conclusion on the censorship strategy of the Chinese government, in particular, King, Roberts and Pan (2013). We argue that attempting to infer the censorship strategy of the Chinese government by looking at the observed data could be problematic due to the missing data problem resulting from self-censorship.

In the third essay titled *Reputation and Media Selection*, I evaluate whether provision

of reputational information about media outlets to the public could be a promising remedy for addressing low-quality reporting. With respect to the prevalence of low-quality news reporting, many suggest that if news consumers are provided with information about the quality of media outlets, they can avoid fabricated information by choosing media outlets wisely. It is also believed that reputational information could encourage quality reporting. The underlying logic is that if the public can monitor the performance of media outlets, then the media outlets with reputational concerns can be incentivized to produce high-quality reports. I analyze a media market and assess whether knowing the past performance of newspapers can help the readers to avoid low-quality journalism by applying the framework proposed by Tadelis (1999, 2002) that reputation is a tradable asset. I show that provision of reputational information *per se* could not eliminate low-quality journalism, and could instead create a subtle opportunity for information manipulation through media ownership. The analysis provides a new theoretical perspective arguing for regulations on media market aiming at preserving pluralism.

Chapter 2

Media with Reputational Concerns: Yes Men or Watchdogs?

2.1 Introduction

Information is essential for the electorate to assess the quality of their government. Among all information sources, the media is primary. At the same time, media outlets also care about their reputations. Indeed, almost all media outlets claim that they are able to gather more—and more accurate—information than other media outlets. However, when popular opinion of a government is not too low, media outlets endorse government decisions instead of reporting critical opinions. Despite the principles of watchdog journalism being widely accepted in democracies, journalists do not always follow them (Boydstun, 2013; Norris, 2014).

This lack of independent and critical reporting could be an especially significant issue in areas in which the government is believed to have no strong partisan bias but have access to secret information. A classic example is the media coverage of the Gulf of Tonkin Incident in 1964, which was the basis for President Johnson ordering a retaliatory action against North Vietnam. In 1964, the U.S. media did not challenge the government but reported the incident as fact. Yet, in 2007, when classified information on the incident was made public following a Freedom of Information Act

request, the public finally learned that this attack was imaginary. This cannot be explained by lack of accessible information alone. In a seminal book about this event, Hallin (1986) asserts that "[government] control of information by itself, however, by no means explains the effectiveness of their efforts. There was, in fact, a great deal of information available which contradicted the official account; it simply wasn't used" (p. 31). He considers the fact that the French newspaper *Le Monde* reported a story on the incident very different from those of the U.S. news outlets as evidence of the media's access to (non-reported) information.[1] Indeed, there are also many similar discussions surrounding, for example, the contemporary Iraq War.

In the political economy literature, explanations about the conformist reporting include "grabbing hand" (for example, Besley and Prat, 2006)—suggesting that the government may use its power to intervene in media reporting—and also include the political motives of media owners (for example, Larcinese et al., 2011; Anderson and McLaren, 2012). This paper argues that neither of these conditions is necessary. Instead, a media outlet may report untruthfully solely because of its reputational concerns. Using a game theoretical model, I show that a reputation-concerned media outlet usually refrains from truthfully reporting information. Instead of carrying out its role as a watchdog, a media outlet tends to act as a yes man,[2] which means it endorses the government's decisions despite possessing contradictory information. With access to only yes-man reports, the public cannot acquire relevant information about the government's competence.

[1]This is not saying that the U.S. public should have believed in a French newspaper or other foreign media more than the American ones, but because the issue was less salient in France than in the United States, the French media outlet was subject to less reputational concerns and was able to transmit its information without the same level of strategic consideration as its U.S. counterparts.

[2]Prendergast (1993) is probably the first to study the "yes man" behavior. Prendergast (1993) shows that when firms use subjective performance evaluation, workers have an incentive to cater to the opinion of their managers.

This paper analyzes a game with three actors: the government, the media outlet, and the representative voter. The government makes a decision to implement a policy. Following the policy decision, the media outlet reports the information about it to the voter. In order to decide whether to retain the incumbent, the voter wants to determine the competence of the government from both the observed policy decision and the media report. However, the news report also reveals information about the quality of the media outlet, so the voter forms opinions not only about the government but also about the media outlet by Bayesian updating. The media outlet that is concerned about its own reputation thus has to make its reporting decision strategically.

With this model, I first consider the case when a single media outlet is the only source of information for the voter. The model shows that the media outlet never truthfully reports information because of its reputational concerns. The intuition is that when there is no other source of information that can reveal yes-man behavior, it is always beneficial for the media outlet to endorse the government's decision because doing so never decreases the media outlet's reputation: When the reader believes that the media outlet is reporting truthfully, a report matching the government policy improves its reputation. On the other hand, a report conflicting the government policy does not improve the media outlet's reputation since either the policy decision or the report can be wrong.

I next study the general case, in which the voter can also possibly receive a perfect signal from other sources of information. For example, the voter can receive information from foreign media or a natural disaster may take place that reveals the truth of low-quality governance. The analysis shows that, with the imperfect feedback, the

media outlet becomes more likely to report truthfully. However, the conditions for a truthful reporting equilibrium to occur are still restrictive: The government has to be considered incompetent and the probability of the exogenous revelation has to be high. Otherwise, endorsing the government's decision usually yields higher reputation for the media outlet. This means, in general, that the voter is unlikely to acquire truthful information about the quality of the government from media reports. It is particularly true when the government is expected to be reasonably competent because in this case, an incompetent media outlet has a higher incentive to pander to the government policy in its reporting in order to maximize the probability of *being correct*. To have a truthful reporting equilibrium, two conditions must hold simultaneously: The expected competence of the government is sufficiently low, and the probability for the voter to learn from other information sources is sufficiently high.

Because media outlets in this model care only about reputation, do not have political agendas, and are not under the threat of media capture, the setup can be considered the most likely case for independent reporting. Importantly, the yes-man problem persists even when the electorate will know with certainty the true state *ex post*, as well as when there exist multiple media outlets. The results also hold when the media outlet cares not only about reputation but ideology.

The paper proceeds as follows. I first relate this topic to the existing literature. Then, I introduce the model and analyze the conditions for the existence of a truthful reporting equilibrium and present comparative statics in the case of one media outlet. I further analyze the cases when the media outlet is also concerned about political ideology and when there exist multiple media outlets. I then conclude the paper. All proofs are presented in the appendix.

2.2 Related Literature

This paper relates to the literature of strategic media reporting, media reputational concerns, and media's incumbency bias, particularly Gentzkow and Shapiro (2006) and Ashworth and Shotts (2010). This paper is connected more generally with the literature of strategic information transmission.

Because the electorate relies on media reports to learn information for political selection and to discipline politicians, there is a natural concern about the quality of the reports. Besley (2005) remarks that "the media can assist voters in identifying the quality of candidates and conversely, political selection will often work poorly in countries where the media is repressed" (p. 56). Chiang and Knight (2011) show empirically that endorsements from newspapers are influential in voters' decisions, but the level of influence is conditioned on the perceived credibility of the media. Bruns and Himmler (2016) consider a model in which the media outlet is a profit maximizer, showing that voters are willing to incentivize the media to report valuable information in order to hold politicians accountable. Politicians also respond to the media environment of a society (Wolton, 2019).

However, this paper is one of several that argues that the media may fail to deliver on its promise of competency and truthfulness. In particular, a media outlet that wants to be regarded as high-quality may conceal information that endangers its reputation.

Gentzkow and Shapiro (2006) show that media reports tend to pander toward voters' prior beliefs to build their reputations. In this paper, I show that, to bolster its reputation, a media outlet may become a yes man to the government. The two papers are similar in three ways: (1) the interim belief for the voter of what is the true

policy (depending on the believed competence of the government) is similar to the prior in the model; (2) both papers characterize pandering reporting equilibrium of media outlets; (3) both papers show that availability of feedback and competition between media market can induce truthful reporting. However, the pandering behavior characterized in the present paper is different from that in Gentzkow and Sharpiro (2006). With a richer policy space adopted from Caillaud and Tirole (2002), instead of a binary policy choice, a competent media outlet does not always have an incentive to report truthfully (which is assumed in Gentzkow and Sharpiro, 2006), and the media outlet does not always report truthfully even when the voter receives feedback with certainty. My analysis further shows the upper limit of the competition effect: Even when there are an infinite number of media outlets, the existence of a truthful reporting equilibrium is not guaranteed. Hence the present paper shows the existence of a more substantial yes-man problem of the media when the policy option is not binary.

In terms of the media's reputational incentive, this paper focuses on media's reputation for competence; previous research has considered reputation for neutrality. Shapiro (2016) argues that a media outlet that wants to be viewed as unbiased may not report the full set of facts to the voter, and, thus, the voter can only form an ambiguous impression about the debate even when there is a clear consensus among experts. Morris (2001) suggests that an information provider may conceal information in order to be considered as unbiased.

Ashworth and Shotts (2010) discuss whether informative media commentary can help solve the problem of a government choosing to implement a popular, but incorrect, policy even when the policy-maker does know that this is not the correct decision. Their model shows that sometimes, a media report makes the problem of

pandering worse, especially when the media outlet acts as a yes man. This paper is closely related to Ashworth and Shotts (2010), but there are at least two significant differences: (1) This paper shows that the media outlet may act as a yes man even when it is more competent than the government, and (2) this model allows analysis of media competition, as well as of cases with various other information sources.

The media's reputational concerns thus lead to yes-man behavior, and media reports tend to be biased in favor of the incumbent government. In the political economy literature, there are other explanations of conformist reporting. A prominent explanation, "grabbing hand," suggests that the government may use power to intervene into media reporting. Besley and Prat (2006) and Gehlbach and Sonin (2014) belong to this strand. By contrast, Baron (2006), Larcinese et al. (2011), and Anderson and McLaren (2012) explain the existence of incumbency bias by political motives of media owners.

In the political communication literature, many studies aim to explain the lack of independent reports by "indexing theory" (Bennett, 1990); that is, the news coverage reflects the range of views that exists within the government. If there is a consensus within the government, there will be few critical reports, and non-official views are not covered (Zaller and Chiu, 1996). Hallin (1986) suggests that patriotism or principles of professional journalism led news outlets to conceal some information during wartime. Althaus et al. (1996) suggest that the norm of objectivity that encourages journalists to present conflict and balance in debate may contribute to the overemphasis on a one-sided elite voice. Boydstun (2013) explains why media cover some issues in detail but not others by the competition for attention across issues depending on the nature of the stories, along with a range of institutional incentives, such

as policy-maker attention and public concern. Selective reporting across issues is an important research topic but beyond the scope of this paper.

In this paper, media reporting is cheap talk, and, thus, this paper is connected with the literature of strategic information transmission that begins with Crawford and Sobel (1982). It is especially close to Ottaviani and Sørensen (2006), who discuss a reputation-concerned expert using a cheap-talk message to signal its quality to the evaluator. Unlike most cheap talk models, the media outlet in this paper is policy-wise unbiased. Nevertheless, due to the media outlet's reputational concerns, its report might still not reveal the truth to the voter.

2.3 The Model

Players There are three players in this model: a government G, a media outlet M and a voter V. The government G and the media outlet M are either competent h or incompetent l. The type of the government ($\tau^G = J$) and the type of the media outlet ($\tau^M = J$), where $J \in \{h, l\}$, are each organizations private information. It is common knowledge, however, that $Pr(\tau^G = h) = p$ and $Pr(\tau^M = h) = q$. The competent actors are more likely to learn the true state of the world than are the incompetent actors.

Policy Choice and Media Report Following the setup of Caillaud and Tirole (2002), the government chooses a policy x from an infinite number of policy choices. Among them, only one unique policy is correct, meaning that only one policy matches the true state of the world ω, which is determined by Nature. All other options are incorrect. I apply this policy structure to clearly capture the event that the second mover herds its action to that of the first mover. It is noteworthy that, as in Caillaud and

Tirole (2002), the policy options are not assumed to be on any particular distribution. Instead, the existence of an indefinite number of wrong choices makes the situation in which multiple actors receive private signals suggesting a *same* wrong choice to be a probability 0 event.

The government observes a signal s_J^G indicating the correct policy, where $J \in \{h, l\}$. The signal may be noisy. When the government is competent, the probability of the signal matching the true state $Pr(s_h^G = \omega)$ is $\overline{\gamma}$. With probability $1 - \overline{\gamma}$, s_h^G is randomly drawn from the infinitely many wrong options, and hence $Pr(s_h^G = \omega) = 0$. For an incompetent government, the probability $Pr(s_l^G = \omega)$ is $\underline{\gamma}$, and $0 \leq \underline{\gamma} \leq \overline{\gamma} \leq 1$.

Similar to the government's decision, the media outlet chooses its report r from a set of infinite options. The media outlet observes x and also a private signal s_J^M, where $J \in \{h, l\}$ before reporting. The probability that the competent media outlet learns the correct policy is $Pr(s_h^M = \omega) = \overline{\mu}$. For an incompetent media outlet, the probability $Pr(s_l^M = \omega)$ is $\underline{\mu}$, and $0 \leq \underline{\mu} \leq \overline{\mu} \leq 1$. The media outlet then decides what information r to report to the voter. When $x \neq s_J^M$, I call $r = s_J^M$ truthful reporting, and $r = x$ yes man reporting.

To simplify exposition, I denote the *ex-ante* expected quality of the government as $E(\gamma) = p\overline{\gamma} + (1 - p)\underline{\gamma}$, and the expected quality of the media outlet as $E(\mu) = q\overline{\mu} + (1 - q)\underline{\mu}$. To focus on the interesting cases, I consider that the qualities of the competent media outlet, the government, and the incompetent media outlet are ordered as $\overline{\mu} \geq E(\gamma) \geq \underline{\mu}$. For numerical analyses in the comparative statics section, we only need $\overline{\mu} \geq \underline{\mu}$.

An important attribute of this model is that, policy-wise, the government and the media are unbiased. This is because I aim to show that conformist reporting

is possible even when no political bias is involved. This can be understood as all members of the society share a common interest in issues such as national security or economic stability.

As yes-man behavior is most likely to be detected in this situation than in other environments such as a binary choice setting, the Caillaud and Tirole setup, combined with *the most severe punishment* imposed by the voter (discussed below), means that the model is analyzing the most unlikely case that the media outlet will deviate from truthful reporting. After the analysis of the general model, I discuss how the equilibria change only slightly when there is also uncertainty about the ideological position of the media outlet.

Learning, Reputation, and Reelection The voter decides whether to retain the incumbent by observing the policy implemented x and learning information from the media. The voter does not observe the consequences of the policy directly. The justification is that the consequences of a policy are sometimes witnessed long after it is implemented (Besley 2007, p. 80). Instead, the voter reads the related report from the media to determine the quality of the government and decides whether to reelect the incumbent. At the same time, the media outlet builds (or ruins) its reputation via its report.

This paper also considers the case that, with probability ρ, information from other exogenous sources, such as foreign media reports or natural disasters, reveal a perfect signal of the true state $s_V = \omega$ to the voter. In other words, the voter could potentially, with probability ρ, receive perfect *ex-post* feedback about the state of the world. Formally, $Pr(s_V = \omega) = \rho$ and $Pr(s_V = \emptyset) = 1 - \rho$. With the policy decision x, the media report r, and the exogenous information s_V, the voter updates her beliefs about

the type of the government and the media outlet, and makes her reelection decision.[3] The voter chooses the new government G' either by retaining the incumbent or by a random draw from the candidate pool. In the latter case, I suppose $Pr(\tau^{G'} = h) = p$. Also, I assume that if indifferent, the voter retains the incumbent.

It is assumed that the voter applies *the most severe punishment* to the media outlet once she finds it does not report truthfully. That is, once the media outlet is caught cheating, the voter believes it is incompetent, or formally, $Pr(\tau^M = h | s_V = \omega \neq (r = x)) = 0$. Relaxing this assumption only disincentivizes the media outlet to report truthfully.

The Payoffs The model incorporates the idea of reputational payoff used in Kartik and Van Weelden's (2019) analysis of cheap talk in elections. The payoff for the media outlet U_M is a strictly increasing mapping of its reputation for competence. Formally, I denote $U_M = \Pi(\hat{q})$, where $Pr(\tau^M = h | x, r, s_V) = \hat{q}$, and $\Pi : \hat{q} \to \mathbb{R}_+$ can be any strictly increasing function. The media outlet enjoys utility from an increase in its reputation. For example, higher reputation could lead to an expansion of readership or an increase in advertising revenue.

Similarly, I define the payoff of the government as $U_G = \Delta(\hat{p})$, where $Pr(\tau^G = h | x, r, s_V) = \hat{p}$. $\Delta : \hat{p} \to \mathbb{R}_+$ can be any strictly increasing function. The government's increasing utility from reputation comes from the rising probability of winning reelection, increasing political donations due to popularity, etc.

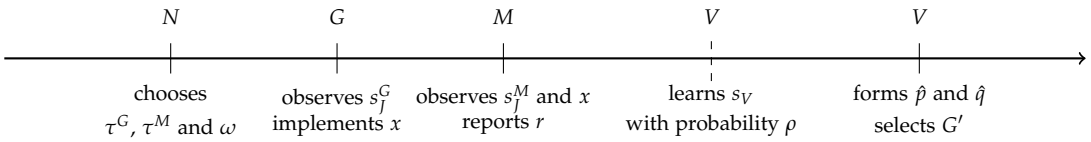
[3] Letting s_V come after the reelection does not affect the main results of this paper. This change affects the incumbent's reelection probability because the voter can no longer update the incumbent's posterior with s_V. If the media outlet is only concerned about its own reputation, its reporting strategy, which this paper focuses on, does not change. If the media outlet has its own political bias and cares more about the incumbent's reelection than its own reputation, this change incentivizes the outlet to untruthfully endorse or criticize the government. Overall, the present setup does not overestimate the problem of not reporting truthfully.

The voter wants to know the quality of the government to facilitate her political selection. The payoff for the voter is $U_V = \mathbb{1}_{\{\tau^{G'}=h\}}$, which means the voter enjoys a utility normalized to one when she selects a competent government G'.

Timing The timing of this game is as follows and can be summarized as in Figure 2.1:

1. Nature chooses the types of actors $\tau^G = J$ and $\tau^M = J, J \in \{h, l\}$ and the state of the world ω
2. G observes s_J^G and implements a policy x
3. M observes s_J^M and x, and reports r
4. With probability ρ, V learns the true state via other information sources
5. V forms the posterior belief about the types of G ($\hat{p}$) and M ($\hat{q}$) and selects G'
6. The payoffs are realized, and the game ends.

Figure 2.1: Timeline



Equilibrium Concept The equilibrium concept applied is pure-strategy Perfect Bayesian Equilibrium, and weakly dominated strategies are excluded. Similar to Ottaviani and Sørensen (2006), the main goal of the analysis is to check whether there could exist truthful equilibrium in which both types of media outlet report truthfully. Formally, *truthful reporting equilibrium* means the media outlet reports the information it learnt truthfully to the voter, i.e., $r = s_J^M$. When a truthful reporting equilibrium does not exist, I characterize a *yes-man equilibrium*, which is a pooling equilibrium in which

both types of the media outlet report an endorsement of the government's decision independent of the private information, formally $r = x$. Discussing the yes-man equilibrium is reasonable because media outlets are not likely to challenge the government without very good reason. A good relationship with officials is usually critical for future news reporting (Eraslan and Özertürk, 2017). Hence, intuitively, we can think that if a truthful reporting equilibrium fails to exist, the media outlet will automatically switch to the yes-man reporting strategy.

With a truthful reporting equilibrium, the voter updates her belief about the quality of the government based on the media's report. With a yes-man equilibrium, the voter does not update her belief. I have also checked for the existence of separating equilibrium or semi-separating equilibrium, but neither exists.

2.4 Analysis

Ideal Case: When Truthful Reporting Exists

The voter wants the media outlet to be an effective watchdog. The ideal world for the voter is one in which only a competent media outlet exists and it always truthfully reports its information. Thus, when the media report endorses the government's policy, the voter can infer that the government is more likely to be of high competence, and when the media report conflicts with the government's choice, then the voter knows the government is more likely to be of low quality. The voter can therefore decide whether to reelect the incumbent following the media report.

Formally, this first-best situation means

$$Pr(\tau^G = h | r = x) \geq p \geq Pr(\tau^G = h | r \neq x).$$

Suppose $Pr(\tau^M = h) = 1$ and the media outlet reports information truthfully. By

Bayes' Rule,

$$Pr(\tau^G = h|r = x) = \frac{p\overline{\gamma}\overline{\mu}}{p\overline{\gamma}\overline{\mu} + (1-p)\underline{\gamma}\overline{\mu}} = \frac{p\overline{\gamma}}{p\overline{\gamma} + (1-p)\underline{\gamma}}, \tag{2.1}$$

and $Pr(\tau^G = h|r = x)$ is always greater than p when $\overline{\gamma} > \underline{\gamma}$.

Similarly, we have

$$Pr(\tau^G = h|r \neq x) = \frac{p(1-\overline{\gamma}\overline{\mu})}{p(1-\overline{\gamma}\overline{\mu}) + (1-p)(1-\underline{\gamma}\overline{\mu})}, \tag{2.2}$$

and p is always greater than $Pr(\tau^G = h|r \neq x)$ when $\overline{\gamma} > \underline{\gamma}$.

Lemma 2.4.1. *A media report is a meaningful signal of the government's competence if and only if the media outlet reports truthfully, independent of the quality of the media outlet.*

As such, we know that a media report is a meaningful signal of the government's competence if the media outlet[4] reports information truthfully (instead of reporting as a yes man), but in the following sections, I show that because of the media outlet's reputational concerns, the voter is not likely to learn the type of the government from the media report.

Case 1: No Other Source of Information

Following the practice of Gentzkow and Shapiro (2006), the analysis begins with a case with a single media outlet. In this case, there are no other sources of information available for the voter, so she relies solely on one media outlet's report to learn the quality of the government. In a democratic context, this happens in a country with a media monopoly, such as Brazil (see Amaral and Guimarães, 1994), or with a

[4]Whether the media outlet is competent or not is irrelevant because $\overline{\mu}$ will be cancelled out during the calculation of (2.1) $\geq p$ and $p \geq$ (2.2). The intuition is that as long as the reporting strategy is consistently truthful, regardless of the competence of the media outlet, a competent government is more likely to be endorsed. Thus, the media report is meaningful.

very homogeneous media market evolved after a series of mergers (as discussed in Anderson and McLaren, 2012).

I solve this game using backward induction and checking if there is a profitable deviation for the media outlet from the truthful reporting equilibrium. The voter updates her beliefs and makes her electoral decision based on the media report when the media outlet reports truthfully, instead of sending a babbling signal. It is obvious that once the media outlet receives a signal $s_J^M = x$, it will transmit the signal truthfully. Whereas, if the media outlet receives a signal $s_J^M \neq x$, the media outlet needs to decide to report either $r = x$ or $r \neq x$. If the voter believes the media outlet reports truthfully, by Bayes' rule, the voter's posterior belief about the media outlet's competence when $r \neq x$ is:

$$Pr(\tau^M = h | r \neq x) = \frac{q(1 - \overline{\mu}E(\gamma))}{q(1 - \overline{\mu}E(\gamma)) + (1-q)(1 - \underline{\mu}E(\gamma))}. \tag{2.3}$$

If $r = x$, again supposing that the voter believes the media outlet reports truthfully, the posterior is:

$$Pr(\tau^M = h | r = x) = \frac{qE(\gamma)\overline{\mu}}{qE(\gamma)\overline{\mu} + (1-q)E(\gamma)\underline{\mu}} = \frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}}. \tag{2.4}$$

Comparing the two posteriors, we can see that, except when the media outlet does not have to signal its type by its report,[5] endorsing the government ($r = x$) leads to strictly higher reputation for the media outlet. Thus, when receiving a signal that is different from the government's decision, the media outlet will deviate from truthful reporting. The mechanism driving this result is simple: Being a yes man yields higher posterior reputation for a media outlet if the voter believes the report is truthful. Thus, the truthful reporting equilibrium does not exist.

[5] $Pr(\tau^M = h | r = x) = Pr(\tau^M = h | r \neq x)$ when $q = 1$, $q = 0$, or $\overline{\mu} = \underline{\mu}$. In these three cases, there is essentially only one type of media outlet, and thus the media outlet does not need to signal its competence.

When it is in yes-man equilibrium, because the report is a babbling signal and the voter does not update beliefs with it, the posterior is $Pr(\tau^M = h|r = x) = Pr(\tau^M = h|r \neq x) = q$. Thus, there is no profitable deviation for the media to report $r \neq x$ instead of $r = x$. As noted earlier, it is intuitive to understand that the media outlet will not issue a report challenging the governments decision for no obvious reason.

The government's action is very straightforward in this model. When the government is making the policy decision, the only information on hand is the signal s_J^G. Thus, irrespective of the media outlet's reporting strategy, the probability that the decision and the report match ($r = s_J^G$) is always higher than the probability that they don't match ($r \neq s_J^G$).[6] Assuming a truthful reporting equilibrium, when observing $r = s_J^G$, the voter believes the government is more likely to be competent. Therefore, it is a dominant strategy for the government to set the policy $x = s_J^G$, and this is true in all cases discussed in this paper.

The yes-man equilibrium is the unique equilibrium in this case. It is robust to Cho and Kreps's Intuitive Criterion because the competent media outlet cannot do better by reporting $r \neq x$. On the other hand, any separating equilibrium—for example *only the competent media outlet reports truthfully*—cannot survive the Intuitive Criterion. In this situation, the incompetent media outlet has an incentive to always report $r \neq x$ in order to be viewed as competent.

The following example can demonstrate the mechanism clearly: Suppose the government and the media outlet received different signals, and $\{\overline{\mu} = 0.6, \underline{\mu} = 0.4, E(\gamma) = 0.6, q = 0.6\}$. Then, if the voter believes the news report is truthful, the posterior about the media outlet's competence is 0.558. If, instead of reporting

[6]When the media outlet reports truthfully, it is more likely to have $r = x$ by choosing $x = s_J^G$ than any $x \neq s_J^G$. When the media outlet reports untruthfully, it is always $r = x$. Hence, choosing $x = s_J^G$ is always optimal for the government.

$r \neq x$, the media outlet reports $r = x$, the posterior increases to 0.692. Thus, it is a profitable deviation for the media outlet. In this case, the media outlet never has an incentive to report truthfully. The media outlet is a yes man instead of a watchdog, and the voter cannot infer the quality of the government from the media report. Following the assumed tie-break rule, the voter retains the incumbent because she does not find the incumbent and challenger discernibly different from one another ($\hat{p} = p$).

Proposition 2.4.1. *When the media outlet is the only source of information for the voter, there is no truthful reporting equilibrium. Even when the media outlet is more competent than the government, it does not report truthfully.*

Case 2: When the Voter May Learn from Other Sources

Now I analyze the general case of the model where, with some probability $\rho \in (0, 1)$, the voter learns the true state from other information sources.

When $\rho = 1$ To find the conditions for the media outlet to report truthfully when $\rho \in (0, 1)$, we need to find the posteriors attainable with different actions when $\rho = 0$ and $\rho = 1$. We already have the posteriors for $\rho = 0$, so now we need to find the posteriors for $\rho = 1$. The detailed calculations are in the appendix. The analysis shows that when the voter can perfectly learn the true state *ex post*, the competent media outlet will always report truthfully because the probability for the competent media outlet to be correct is higher than that for the government ($\overline{\mu} > E(\gamma)$). However, the incompetent media outlet could still have an incentive to report untruthfully even when yes-man behavior is very likely to be detected.[7] To have a truthful reporting

[7]Whenever the voter observes $r = x \neq \omega$, she knows the media outlet is a yes man, because the probability for the government and media outlet to receive the same wrong signal is 0 (Caillaud and Tirole, 2002).

equilibrium, we need both types of media outlets to report truthfully. Thus, a truthful reporting equilibrium is not guaranteed despite the fact that the voter can perfectly learn the true state *ex post*. If the truthful reporting equilibrium is not guaranteed even in this extreme case,[8] we can expect that the truthful reporting equilibrium is less likely to exist in the general case.

When $\rho \in (0,1)$ With the posteriors in different situations when $\rho = 0$ and $\rho = 1$, we can analyze the general case and find the conditions for the existence of the truthful reporting equilibrium. The detailed calculations are in the appendix.

The condition for the competent media to report truthfully is:

$$\rho \geq \frac{\frac{\overline{\mu}}{E(\mu)} - \frac{1-\overline{\mu}E(\gamma)}{T}}{(1+\overline{\mu}-E(\gamma))\frac{\overline{\mu}}{E(\mu)} + (1-\overline{\mu})\frac{1-\overline{\mu}}{1-E(\mu)} - \frac{1-\overline{\mu}E(\gamma)}{T}} \equiv \rho_h, \tag{2.5}$$

where $T = q(1-\overline{\mu}E(\gamma)) + (1-q)(1-\underline{\mu}E(\gamma))$.[9]

The condition parallel to (2.5) for the incompetent media is:

$$\rho \geq \frac{\frac{\overline{\mu}}{E(\mu)} - \frac{1-\overline{\mu}E(\gamma)}{T}}{(1+\underline{\mu}-E(\gamma))\frac{\overline{\mu}}{E(\mu)} + (1-\underline{\mu})\frac{1-\overline{\mu}}{1-E(\mu)} - \frac{1-\overline{\mu}E(\gamma)}{T}} \equiv \rho_l. \tag{2.6}$$

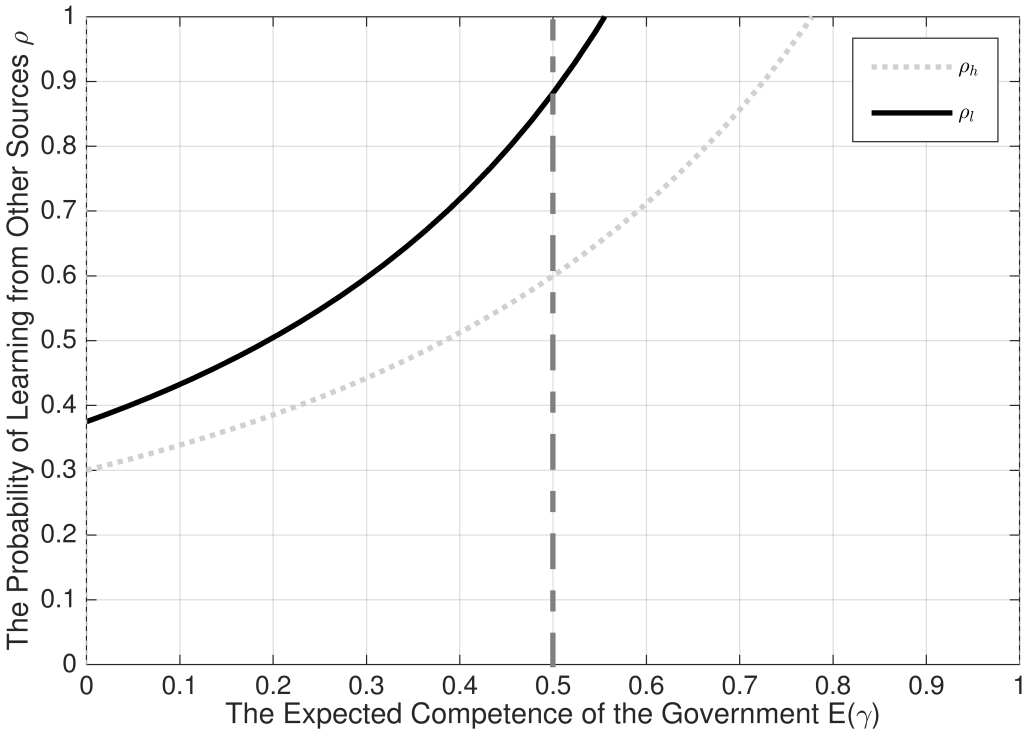
Given $\overline{\mu} \geq \underline{\mu}$, we have $\rho_l \geq \rho_h$. As expected, an incompetent media outlet is less likely to report truthfully than a competent one. Hence, ρ_l is the binding condition for the truthful reporting equilibrium. A numerical example is when $\{E(\gamma) = 0.5, \overline{\mu} = 0.6, \underline{\mu} = 0.2, q = 0.5\}$, we have $\rho_h = 0.6$ and $\rho_l \approx 0.882$. Hence, the truthful reporting equilibrium exists if and only if $\rho \geq \rho_l \approx 0.882$. This means that unless the probability of receiving information from other sources is very high, a truthful reporting equilibrium does not exist.

[8] In this case, the media report is actually made redundant by the availability of exogenous information.

[9] T is the *ex-ante* probability that the media outlet and the government receive different signals.

When fixing $\{\overline{\mu} = 0.6, \underline{\mu} = 0.2, q = 0.5\}$, the conditions for the two types of media outlet to report truthfully ρ_h and ρ_l can be graphically shown as in Figure 2.2.[10] The area between the horizontal line $\rho = 1$ and ρ_l is where the truthful reporting equilibrium exists. If the government is expected to be reasonably competent, that is $E(\gamma) \geq 0.5$, the truthful reporting equilibrium only exists in a very restricted area, and it requires a very high probability for the voter to learn from other sources to sustain this equilibrium. Figure 2.2 also demonstrates that even when the quality of the government is expected to be very low, the truthful reporting equilibrium can exist only with a sufficient probability for the voter to learn from other information sources.

Figure 2.2: Conditions for Truthful Reporting



Proposition 2.4.2. *Given the expected competence of the government ($E(\gamma)$), when the voter*

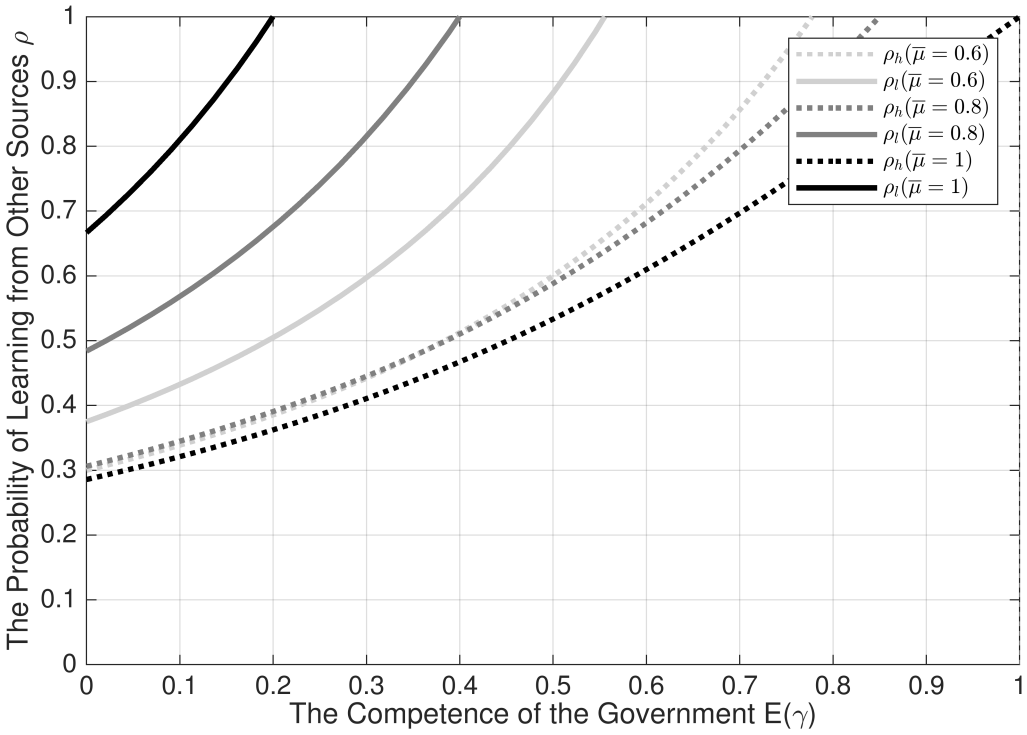
[10] In Figure 2.2, the solid line ρ_l is the binding condition, and the dotted line ρ_h is satisfied automatically once ρ_l is satisfied, but to illustrate the dynamic clearly, ρ_h is still shown in all of the graphs.

observes x, receives r from the media outlet and, with probability ρ, learns s_V from other sources, the existence of the truthful reporting equilibrium requires the probability for the voter to learns from other sources to be $\rho \geq \rho_l$.

Comparative Statics

Based on Proposition 2.4.2, I perform three numerical comparative statics exercises.[11] First of all, I set the quality of the competent media outlet $\overline{\mu}$ as a variable and fix the quality of the incompetent outlet $\underline{\mu}$ and the probability to have the competent outlet q. Figure 2.3 shows that, counterintuitively, when the quality of the competent media outlet increases, the truthful reporting equilibrium becomes less likely to occur.

Figure 2.3: Comparative Statics 1



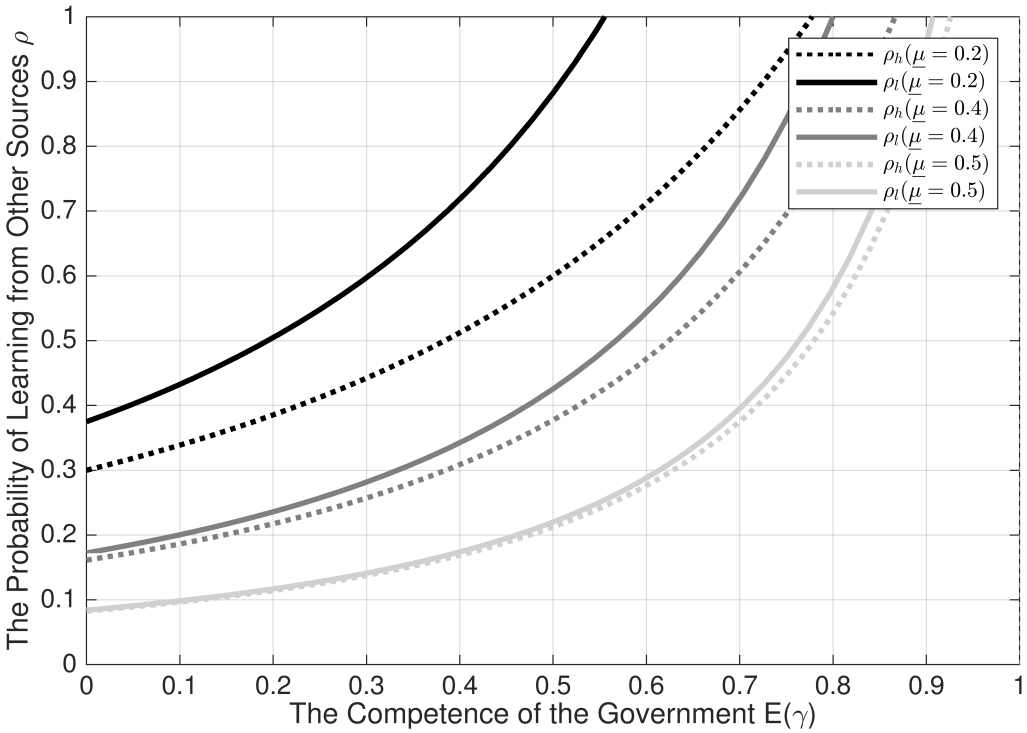
Two forces drive this surprising result. First, as the quality gap between the two types of media outlets widens, if the true state is not revealed *ex post* but the voter

[11] Solving for closed-form solutions is impossible as the posterior functions are a very complicated composition of functions.

believes the media report is truthful, as before, the media outlet has a higher posterior when its report matches the policy decision and a lower posterior when its report does not match the policy decision. However, now the difference between these two posteriors becomes larger. Thus, matching the report to the policy leads to higher return from reputation. Second, if *ex post* the report is verified to be correct, the media outlet is more likely to be competent, and because the difference in posteriors of being correct and being incorrect increases when the quality gap widens, *being correct* becomes a strong signal of competence and leads to higher reputational returns. The incompetent outlet becomes more willing to take the risk of being caught cheating by reporting as a yes man.

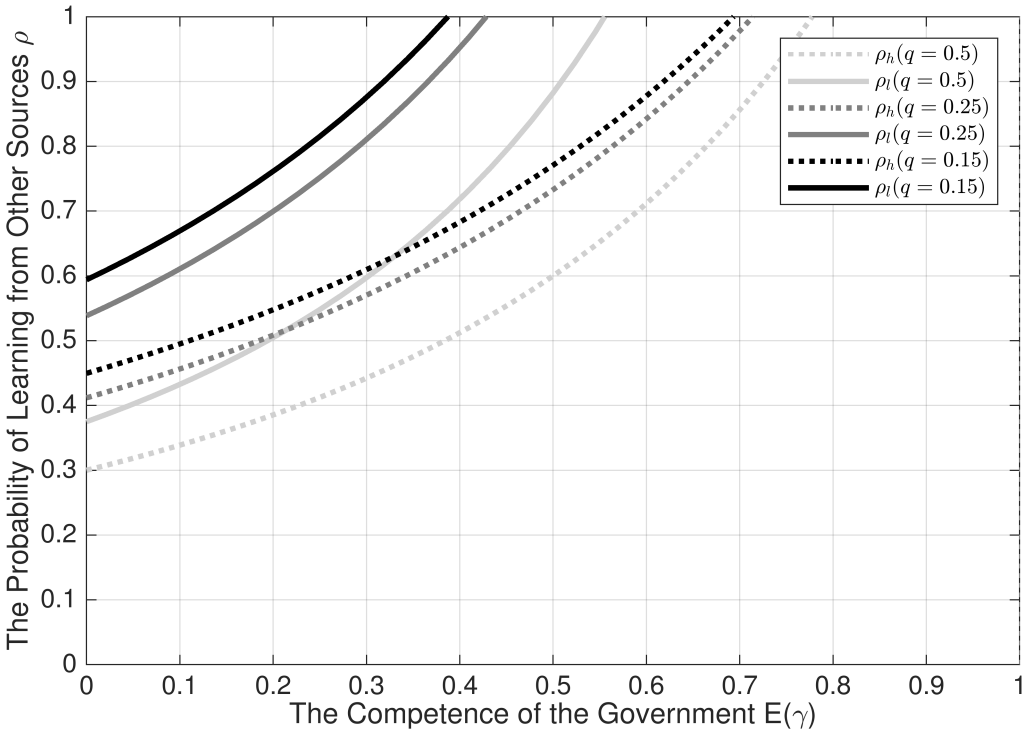
For the second set of comparative statics, I fix the quality of the competent media outlet $\overline{\mu}$ and the probability to have the competent media outlet q, and vary the quality of the incompetent outlet $\underline{\mu}$. From Figure 2.4, we know that when $\underline{\mu}$ increases, the required ρ_l to support the truthful reporting equilibrium decreases. The result again indicates that the likelihood of the truthful reporting equilibrium decreases as the difference between the types of media outlets increases. When the two types are of similar quality, such as the case where $\overline{\mu} = 0.6$ and $\underline{\mu} = 0.5$, a truthful reporting equilibrium is more likely to exist. Even in this situation, though, the quality of the government cannot be too high, and the probability of learning from other sources of information cannot be too low. The driving force for this result is that deviating from truthful reporting is more rewarding for a media outlet with a lower competence $\underline{\mu}$. This is because when the incompetent media outlet is less likely to receive a correct signal, it is more willing to pander to the government policy in order to increase the probability the media report is correct.

Figure 2.4: Comparative Statics 2



For the final set of comparative statics, I fix the qualities of the two types of media outlets $\overline{\mu}$ and $\underline{\mu}$, and allow the probability of having the competent media outlet q to vary. Figure 2.5 shows that when the media outlet is more likely to be competent, the likelihood of truthful reporting increases. So, although the previous results show that homogeneity of the types of media outlet leads to truthful reporting, *ceteris paribus*, it is more likely for the voter to receive a truthful report when the probability of having a competent media outlet is higher.

Figure 2.5: Comparative Statics 3



It is quite intuitive to think that critical reporting is less likely to occur when the quality of the incompetent outlet declines, but it is surprising that when the quality of the competent outlet is further improved, the probability of critical reporting also decreases. As government subsidy is usually considered as a promising means of improving the quality of news reporting (see Sunstein, 2017), the first set of comparative statics serves as a warning. Those rewards and supports for competent news outlets, aiming at further improving their quality, may turn out to make critical reporting even more unlikely. Improving the quality of incompetent news outlets, as the second set of comparative statics shows, could be a more effective way to encourage critical reporting.

Political Ideology of the Media Outlet

The baseline analysis intentionally avoids discussion about media ideological concerns in order to show that even without political agenda and media capture, media outlets may not report truthfully. However, it is interesting to see how the equilibrium reporting behavior changes when the media outlet is concerned not only with its own reputation but also with ideology. Specifically, I look at how the likelihood to reach the truthful reporting equilibrium changes once the media outlet's utility depends not only on its own posterior but also the incumbent's, because the outlet cares also about the popularity and reelection of the incumbent. Formally, the payoff functions are $U_M^f = \Pi(\hat{q}) + I(\hat{p})$ for a pro-incumbent outlet, and $U_M^e = \Pi(\hat{q}) - I(\hat{p})$ for an anti-incumbent outlet. As in the baseline model, returns from the outlet's reputation $\Pi(\hat{q})$ is a strictly increasing function of $\hat{q}$, and returns from the incumbent's reputation $I(\hat{p})$ is a strictly increasing function of $\hat{p}$. The superscript e stands for enemy (of the incumbent) and f stands for for friend (of the incumbent).

To simplify exposition, I denote the following payoffs: $\Pi(Pr(\tau^M = h|r = x)) \equiv \overline{\Pi}$; $\Pi(Pr(\tau^M = h|r \neq x)) \equiv \underline{\Pi}$; $I(Pr(\tau^G = h|r = x)) \equiv \overline{I}$; $I(Pr(\tau^G = h|r \neq x)) \equiv \underline{I}$. The first two terms are the returns for the media outlet from its reputation, and the last two terms are the returns for the media outlet from the government's reputation. Both are conditional on whether the report matches with the policy decision. From the baseline analysis, we know that when the voter believes the media outlet is reporting truthfully, matching the report to the government policy yields higher reputation and, thus, higher returns than when the report does not agree with the policy. Hence, $\overline{\Pi} > \underline{\Pi}$. On the other hand, the voter's posterior belief about the government's competence is higher when the report matches with the policy than when they conflict. Hence,

for a pro-incumbent media outlet, the return $\bar{I}$ is greater than $\underline{I}$. An anti-incumbent outlet, on the other hand, prefers $-\underline{I}$ to $-\bar{I}$.

When the media outlet is the only source of information for the voter ($\rho = 0$), and supposing that the voter believes the media outlet is reporting truthfully, a pro-incumbent outlet will always be a yes man as in Case 1, because $(\overline{\Pi} - \underline{\Pi}) + (\bar{I} - \underline{I}) > 0$. For an anti-incumbent outlet, it will report as a yes man if $(\overline{\Pi} - \underline{\Pi}) - (\bar{I} - \underline{I}) > 0$. That means, if the outlet cares about its reputation over its ideology, it will still report as a yes man. If the media outlet values its ideology over its reputation ($(\overline{\Pi} - \underline{\Pi}) - (\bar{I} - \underline{I}) < 0$), it will always report a government criticism. Even when the media outlet receives a signal $s_J^M = x$, it will deviate to report $s_J^M \neq x$.

When the voter is able to learn the true state from other sources of information (s_V) with certainty ($\rho = 1$), the voter can update the posterior of the government using s_V. As a result, a fabricated critique or endorsement neither harms nor helps the government. Therefore, the media outlet's decision depends only on whether matching its report with the government policy improves its own reputation. The reasoning is entirely the same as in the baseline analysis. A competent media outlet will always transmit its signal truthfully, and the incompetent media outlet may pander its reporting to the government in order to maximize the media outlet's own reputation.

The general case is again in between of the two extremes, $\rho = 0$ and $\rho = 1$. Because the results when $\rho = 0$ or $\rho = 1$ only differ from the baseline results when the media outlet is anti-incumbent and cares more about ideology, a pro-incumbent media outlet and an anti-incumbent outlet that cares more about its own reputation behave identically to the baseline results. When the media outlet is anti-incumbent

and cares more about ideology, the outlet is more likely to report government criticism. However, unfounded critical reporting is not equivalent to truthful reporting. The information problem for the voter still prevails.

It is worth noting that whether the voter knows if the media outlet is supporting the incumbent or not does not affect the media reporting strategy. When the voter is uncertain about the media outlet's ideology, she knows that a government criticism is more likely to be from an anti-incumbent outlet, and an endorsement is more likely to be from a pro-incumbent one. However, an anti-incumbent media outlet has no incentive to pretend to be a pro-incumbent one by endorsing the government (otherwise it suffers $-\bar{I}$), and vice versa. Hence, reputational concerns dominate a given media outlet's decision-making.

2.5 Multiple Media Outlets

I now extend the analysis from the baseline case of a unique outlet to cases with two or more media outlets. The media outlets are indexed as: M_1, M_2, ..., M_n. The type of each media outlet is private information, and the outlets report news r_1, r_2, ..., r_n simultaneously and independently, conditional on the true state of the world.[12] Without loss of generality, I analyze the game from the perspective of M_1.

Two Media Outlets First, consider the case where the voter can learn the information from two media outlets M_1 and M_2. Suppose that the voter believes the two media outlets play the truthful reporting equilibrium, and the true state is not revealed

[12]If the media outlets report sequentially, the government's decision and all earlier news reports create multiple focal points for the upcoming media outlets to pander to. Thus, a true state will not be revealed by the fact that at least two media outlets report the same story. Hence, yes-man behavior is harder to detect. In other words, the simultaneous case constitutes the most likely case for truthful reporting.

by other sources ($\rho = 0$). The posteriors given different actions and the probabilities for these different situations to occur are summarized in Table 2.1. Throughout this section, to simplify the notation regarding the probability of occurrence of different situations, I use μ (without overline or underline) to denote that it is either of two cases: $\underline{\mu}$ when the media outlet M_1 is incompetent and $\overline{\mu}$ when M_1 is competent.

Table 2.1: Posteriors Given Different Actions: Two Media Outlets ($\rho = 0$)

	Probability of Occurrence	$\hat{q}$ When $\rho = 0$
$x = r_1 = r_2$	$E(\gamma)E(\mu)\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$x \neq (r_1 = r_2)$	$[1 - E(\gamma)]E(\mu)\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$(x = r_1) \neq r_2$	$E(\gamma)[1 - E(\mu)]\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$(x = r_2) \neq r_1$	$E(\gamma)E(\mu)(1 - \mu)$	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
$x \neq r_1 \neq r_2$	$1 - E(\gamma)E(\mu) - E(\gamma)\mu - E(\mu)\mu + 2E(\gamma)E(\mu)\mu$	Ψ_2

Note: Denote $\frac{q(1-E(\gamma)E(\mu)-E(\gamma)\overline{\mu}-E(\mu)\overline{\mu}+2E(\gamma)E(\mu)\overline{\mu})}{q(1-E(\gamma)E(\mu)-E(\gamma)\overline{\mu}-E(\mu)\overline{\mu}+2E(\gamma)E(\mu)\overline{\mu})+(1-q)(1-E(\gamma)E(\mu)-E(\gamma)\underline{\mu}-E(\mu)\underline{\mu}+2E(\gamma)E(\mu)\underline{\mu})}$ as Ψ_2.

When deviating from truthful reporting, the posterior will be $\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$, which is greater or equal to both Ψ_2 and $\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$ (unless $q = 0$, $q = 1$ or $\overline{\mu} = \underline{\mu}$). Thus, the media outlet has an incentive to deviate from the truthful reporting equilibrium, and there is no truthful reporting equilibrium.

Now, consider the situation where the voter can learn the true state of the world from other sources of information with certainty ($\rho = 1$). Table 2.2 summarizes the posteriors given different actions and the probabilities for these different situations to occur when the voter believes the media outlets play the truthful reporting equilibrium.

Table 2.2: Posteriors Given Different Actions: Two Media Outlets ($\rho = 1$)

	Probability of Occurrence	$\hat{q}$ when $\rho = 1$
$x = r_1 = r_2$	$E(\gamma)E(\mu)\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$x \neq (r_1 = r_2)$	$[1-E(\gamma)]E(\mu)\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$(x = r_1) \neq r_2$	$E(\gamma)[1-E(\mu)]\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$(x = r_2) \neq r_1$	$E(\gamma)E(\mu)(1-\mu)$	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
$x \neq r_1 \neq r_2 \neq \omega$	$[1-E(\gamma)][1-E(\mu)](1-\mu)$	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
$r_1 \neq r_2 \neq x = \omega$	$E(\gamma)[1-E(\mu)](1-\mu)$	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
$x \neq r_2 \neq r_1 = \omega$	$[1-E(\gamma)][1-E(\mu)]\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$x \neq r_1 \neq r_2 = \omega$	$[1-E(\gamma)]E(\mu)(1-\mu)$	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$

If a media outlet deviates from reporting truthfully when the voter will observe the true state with certainty, its posterior reputation is

$$E(\gamma)(1-\mu)\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}.$$

As in the baseline case, the competent media outlet has no incentive to deviate from truthful reporting, but the incompetent type will only report truthfully if the following condition is satisfied:

$$(1-\underline{\mu})\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})} \geq [E(\gamma)[1-\underline{\mu}+E(\mu)\underline{\mu}]-\underline{\mu}]\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}.$$

This condition does not always hold. For example, when $\{q = 0.2, \overline{\mu} = 0.8, \underline{\mu} = 0.2\}$, the condition is only satisfied when $E(\gamma) \leq 0.34$. Hence we know that $\rho = 1$, again, does not guarantee the existence of a truthful reporting equilibrium.

The general case we are interested in is a convex combination of the situations in terms of ρ, summarized in the two tables. That is, the voter may sometimes learn the true state from the perfect feedback, but not always. The analysis shows that, as in the baseline case, an incompetent media outlet does not always follow the truthful

reporting strategy. Essentially, when there are two media outlets, they play the *same* signaling game simultaneously, so that one report cannot discredit the other since the voter cannot learn the state of the world from other sources. The decision made by the government then serves as a focal point for the media outlets to tend toward. The truthful reporting equilibrium exists only when the voter can learn the true state of the world from other sources with a sufficiently high probability.[13]

Three or More Media Outlets To analyze cases with three or more media outlets, I use a general approach and categorize n media outlets into three groups: M_1, M_i, and M_j. M_i and M_j partition the $n-1$ media outlets other than M_1 into two groups, each of which consists of at least one media outlet. Each media outlet within a group does not necessarily receive the same signal.

Again, the analysis should focus on the situation where M_1 receives a signal different from the government's decision x and compare posteriors from $r \neq x$ and $r = x$ in different scenarios, presuming every other actor follows the truthful reporting strategy. The different scenarios, their probabilities of occurrence, and the corresponding posterior beliefs, are summarized in the two tables below. Table 2.3 is for the situation $\rho = 0$, and Table 2.4 is for $\rho = 1$.

The difference between this case and those previously analyzed is that if some M_i and M_j receive the correct signal and report truthfully, yes-man behavior will be detected. Hence, supposing all other media outlets follow the truthful reporting strategy, even when $\rho = 0$, yes-man reporting is harmful (to the media outlet) if: (1) M_1 actually received a correct signal but deviates from truthful reporting (because it

[13]Because in this model, the media outlets work on one dimension and are not biased, either mechanism suggested in Krishna and Morgan (2001) or Battaglini (2002), in which information could be fully revealed because the information senders have biases with different directions or on different policy dimensions, does not apply.

is possible to match reports with one or more other news outlets on a non-x option, and such a report is perfect evidence of being correct) or (2) there are at least two media outlets that receive the correct signal (because the true state is revealed, and the yes-man behavior is detected).

When the number of media outlets becomes large, it is certain that at least some media outlets will receive the correct signal, but truthful reporting equilibrium is *still* not guaranteed. The analysis shows that when the media market is in perfect competition ($n \to \infty$), the role of exogenous information sources (ρ) can be replaced by the existence of a large number of media outlets because it is certain that at least two media outlets will receive the correct signal. Thus, supposing that the media outlets report truthfully, the voter is going to learn the true state for sure *ex post*. Indeed, when $n \to \infty$, the case with $\rho = 0$ is equivalent to the case with $\rho = 1$. Even so, an incompetent media outlet still has an incentive to be a yes man whenever $[E(\gamma) - \underline{\mu}]\overline{\mu} - (1 - \overline{\mu})E(\mu) \geq 0$ (for example, when $\{E(\gamma) = 0.5, \overline{\mu} = 0.8, \underline{\mu} = 0.2, q = 0.5\}$). Thus, a truthful reporting equilibrium is still not guaranteed and the baseline result is robust in this extreme case.

Table 2.3: Posteriors Given Different Actions: More than Two Media Outlets ($\rho = 0$)

	Probability of Occurrence	$\hat{q}$ when $\rho = 0$
$x = r_1$	$E(\gamma)\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$x \neq (r_1 = r_i) \neq r_j$	$\Sigma_1 + \Sigma_2$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$x = r_i \neq r_1 \neq r_j$	$E(\gamma)[1-(1-E(\mu))^{n-1}](1-\mu)$	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
$x \neq (r_i = r_j) \neq r_1$	Σ_3	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
$x \neq r_1 \neq r_i \neq r_j$	Φ_n	Ψ_n

Note: Denote $\Sigma_1 \equiv [1-E(\gamma)][(n-1)E(\mu)(1-E(\mu))^{n-2}]\mu$;
$\Sigma_2 \equiv [1-E(\gamma)][1-(1-E(\mu))^{n-1}-(n-1)E(\mu)(1-E(\mu))^{n-2}]\mu$;
$\Sigma_3 \equiv [1-E(\gamma)][1-(1-E(\mu))^{n-1}-(n-1)E(\mu)(1-E(\mu))^{n-2}](1-\mu)$;
$\Phi_n \equiv [1-E(\gamma)][1-E(\mu)]^{n-1}(1-\mu)+E(\gamma)[1-E(\mu)]^{n-1}(1-\mu)+(n-1)E(\mu)[1-E(\gamma)](1-E(\mu))^{n-2}(1-\mu)+[1-E(\gamma)][1-E(\mu)]^{n-1}\mu$;
$\Psi_n \equiv \frac{q\Phi_n^h}{q\Phi_n^h+(1-q)\Phi_n^l}$.

Table 2.4: Posteriors Given Different Actions: More than Two Media Outlets ($\rho = 1$)

	Probability of Occurrence	$\hat{q}$ when $\rho = 1$
$x = r_1$	$E(\gamma)\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$x \neq (r_1 = r_i) \neq r_j$	$[1-E(\gamma)][1-E(\mu)]^{n-1}\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$
$x = r_i \neq r_1 \neq r_j$	$E(\gamma)[1-E(\mu)]^{n-1}(1-\mu)$	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
$x \neq (r_i = r_j) \neq r_1$	Σ_3	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
Every actor wrong	$[1-E(\gamma)][(1-E(\mu))^{n-1}](1-\mu)$	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
G correct	$E(\gamma)[1-E(\mu)](1-\mu)$	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
At least a M_i correct	$E(\gamma)[1-E(\mu)]^{n-1}(1-\mu)$	$\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$
M_1 correct	$[1-E(\gamma)][1-E(\mu)]^{n-1}\mu$	$\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$

Proposition 2.5.1. *The existence of multiple media outlets does not guarantee truthful reporting. Even when there are an infinite number of media outlets, the truthful reporting equilibrium exists if and only if* $[E(\gamma)-\underline{\mu}]\overline{\mu} - (1-\overline{\mu})E(\mu) < 0$.

2.6 Concluding Remarks

In democracies, voters need to know whether the incumbent government is competent in order to make election decisions. Therefore, apart from observing the government's policy decisions, voters also acquire related information from various information sources, of which media is the primary one. However, all too often, media reports are neither critical nor independent. In this paper, I study a model to see why media outlets cannot dutifully serve as watchdogs.

I argue that reputational concerns can sufficiently lead to a lack of independent reporting; any coercive power or policy bias is not necessary. The mechanism is that the voter wants to learn the competence of the government, but news reports also reflect the quality of the media outlets, not just the quality of the government. Reputational concerns thus make the media outlets unable to commit to reporting

information truthfully. Instead, the media outlets tend to endorse the government policy as yes men, because matching (or pandering) reports to the government's decisions weakly increases media reputation in most situations. Therefore, the voter is not able to learn valuable information from media reports even though media outlets may not be partisan or subject to coercive power. This result is robust under the different scenarios discussed in the paper: when the electorate can perfectly learn the true state *ex post*, when the media outlet cares also about political ideology, and when there exist multiple—even an infinite number of—media outlets.

The main implication of this paper is that when the government is believed to be at least reasonably competent, truthful reporting is very unlikely to exist. This is a serious problem for democratic accountability because we want to accurately assess the performance of the government, not be misled by its relative popularity. However, we will not have sufficient information at hand when the government seems to be competent. Essentially, to have a truthful reporting equilibrium, the expected competence of the government must be low enough, while the probability for the voter to learn from other information sources must be high enough.

Comparative statics of this model indicate that widening the quality gap between types of media is detrimental. It is intuitive to think that increasing the quality of competent media outlets could increase the probability of having truthful reporting. Consequently, there are many schemes and rewards aiming at supporting high-quality media outlets to further improve their quality. However, the model shows that, other things being equal, it worsens the yes-man problem as the quality gap between media types is widened. Instead, it would be helpful to encourage the low-quality media outlets to improve, thus narrowing the quality gap.

Future research can empirically test the argument by estimating the perceived competence of the government (i.e., the popularity of the government) and the media in different issue areas and applying a keyword based method adopted by Larcinese et al. (2011), or performing quantitative textual analysis (Laver et al., 2003; Gentzkow and Shapiro, 2010), to identify the reporting stances of the media outlets on relevant issues. The relationship between the perceived competence of the actors and the criticalness of reports in different issue areas could then be evaluated.

Bibliography

[1] Althaus, Scott L., Jill A. Edy, Robert M. Entman and Patricia Phalen. 1996."Revising the Indexing Hypothesis: Officials, Media, and the Libya Crisis." *Political Communication* 13 (4): 407-421.

[2] Amaral, Roberto, and Cesar Guimarães. 1994. "Media Monopoly in Brazil." *Journal of Communication* 44 (4): 26-38.

[3] Anderson, Simon P., and John McLaren. 2012. "Media Mergers and Media Bias with Rational Consumers." *Journal of the European Economic Association* 10 (4): 831-859.

[4] Ashworth, Scott, and Kenneth W. Shotts. 2010. "Does Informative Media Commentary Reduce Politicians' Incentives to Pander?." *Journal of Public Economics* 94 (11-12): 838-847.

[5] Baron, David. 2006. "Persistent Media Bias." *Journal of Public Economics* 90: 1-36.

[6] Battaglini, Marco. 2002. "Multiple Referrals and Multidimensional Cheap Talk." *Econometrica* 70 (4): 1379-1401.

[7] Bennett, W. Lance. 1990. "Toward a Theory of Press-State Relations in the United States." *Journal of Communication* 40 (2):103-127.

[8] Besley, Timothy. 2005. "Political Selection." *Journal of Economic Perspectives* 19 (3): 43-60.

[9] Besley, Timothy. 2007. *Principled Agents?: The Political Economy of Good Government.* Oxford: Oxford University Press.

[10] Besley, Timothy, and Andrea Prat. 2006. "Handcuffs for the Grabbing Hand?: Media Capture and Government Accountability." *American Economic Review* 96 (3): 720-736.

[11] Boydstun, Amber E.. 2013. *Making the News: Politics, the Media, and Agenda Setting.* Chicago: The University of Chicago Press.

[12] Bruns, Christian and Oliver Himmler. 2016. "Mass Media, Instrumental Information, and Electoral Accountability." *Journal of Public Economics* 134: 75-84.

[13] Caillaud, Bernard and Jean Tirole. 2002. "Parties as Political Intermediaries." *The Quarterly Journal of Economics* 117(4):1453-1489.

[14] Chiang, Chun-Fang and Brian Knight. 2011. "Media Bias and Influence: Evidence from Newspaper Endorsements." *The Review of Economic Studies* 78 (3): 795-820.

[15] Crawford, Vincent P. and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6): 1431-1451.

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

[17] Gentzkow, Matthew, and Jesse M. Shapiro. 2006. "Media Bias and Reputation." *Journal of Political Economy* 114 (2): 280-316.

[18] Gentzkow, Matthew, and Jesse M. Shapiro. 2010. "What Drives Media Slant? Evidence from U.S. Daily Newspapers." *Econometrica* 78 (1): 35-71.

[19] Hallin, Daniel C. 1986. *The "Uncensored War": The Media and Vietnam.* Oxford: Oxford University Press.

[20] Hamilton, John Maxwell, Regina G. Lawrence and Raluca Cozma. 2010. "The Paradox of Respectability: The Limits of Indexing and Harrison Salisbury's Coverage of the Vietnam War." *International Journal of Press/ Politics* 15 (1): 77-103.

[21] Hamilton, John Maxwell, Regina G. Lawrence and Raluca Cozma. 2010. "The Paradox of Respectability: The Limits of Indexing and Harrison Salisbury's Coverage of the Vietnam War." *International Journal of Press/ Politics* 15 (1): 77-103.

[22] Eraslan, Hülya, and Saltuk Özertürk. 2017. "Information Gatekeeping and Media Bias." *Rice University Working Paper* available at `http://he6.web.rice.edu/Papers/media2017-12-05.pdf`

[23] Kartik, Navin, and Richard Van Weelden. 2019. "Informative Cheap Talk in Elections." *The Review of Economic Studies* 86 (2): 755-784.

[24] Krishna, Vijay and John Morgan. 2001. "A Model on Expertise." *The Quarterly Journal of Economics* 116 (2): 747-775.

[25] Larcinese, Valentino, Riccardo Puglisi, and James M. Snyder Jr. 2011. "Partisan Bias in Economic News: Evidence on the Agenda-setting Behavior of U.S. Newspapers." *Journal of Public Economics* 95 (9-10): 1178-1189.

[26] Laver, Michael, Kenneth Benoit, and John Garry. 2003. "Extracting Policy Positions from Political Texts Using Words as Data." *American Political Science Review* 97 (2): 311-331.

[27] Morris, Stephen. 2001. "Political Correctness." *Journal of Political Economy* 109 (2): 231-265.

[28] Norris, Pippa. 2014. "Watchdog Journalism." In Mark Bovens, Robert E. Goodin, and Thomas Schillemans (eds.) *The Oxford Handbook of Public Accountability*. Oxford: Oxford University Press.

[29] Ottaviani, Marco, and Peter Norman Sørensen. 2006. "Reputational Cheap Talk." *RAND Journal of Economics* 37 (1): 155-175.

[30] Prendergast, Canice. 1993. " A Theory of "Yes Men"." *American Economic Review* 83 (4): 757-770.

[31] Shapiro, Jesse M. 2016. "Special Interests and the Media: Theory and An Application to Climate Change." *Journal of Public Economics* 144 (December): 91-108.

[32] Sunstein, Cass R.. 2017. *#Republic: Divided Democracy in the Age of Social Media.* Princeton: Princeton University Press.

[33] Wolton, Stephane. 2019. "Are Biased Media Bad for Democracy?." *American Journal of Political Science* forthcoming.

[34] Zaller, John, and Dennis Chiu. 1996. " Government's Little Helper: U.S. Press Coverage of Foreign Policy Crises, 1945-1991." *Political Communication* 13 (4): 385-405.

Chapter 3

Tell Me the Truth? Dictatorship and the Commitment to Media Freedom

"It was terribly dangerous to let your thoughts wander when you were in any public place or within range of a telescreen. The smallest thing could give you away .... In any case, to wear an improper expression on your face (to look incredulous when a victory was announced, for example) was itself a punishable offense. There was even a word for it in Newspeak: facecrime, it was called. (65)"

— George Orwell, *1984*

3.1 Introduction

As in George Orwell's *1984*, dictatorships have long been viewed as an antithesis of freedom of expression. The burgeoning literature on comparative authoritarianism, however, finds a wide variation in media freedom among nondemocratic regimes and some dictators allow a higher degree of it in hopes of collecting local information (Egorov et al. 2009, Gehlbach and Sonin, 2014; Lorentzen, 2014, Qin et al., 2017). As the "authoritarian resilience" thesis has it, the understanding of social problems is the first-order task for an authoritarian government to correctly respond to public demands (Nathan, 2003: 14). Henceforth, when Mikhail Gorbachev tried to bring a

new lease of life to the regime in late 1980s, one of the major reforms he introduced was to lift the media censorship. According to his former chief spokesman, Gennadi Gerasimov,

> "*Hoping to use the media to help identify his nation's problems in order to solve them, Gorbachev gradually lifted Communist Party control of the mass media starting in 1985. In a matter of months, he introduced a degree of freedom unheard of before in the Russian press, or, to use his term, glasnost. He viewed this opening not only as a window on what was happening in the country, but also as a chance to ensure feedback as he tackled economic and political problems facing the nation.*" (Gerasimov, 1998: 2-3)

In other words, while media freedom can sometimes be threatening to dictators' survival for exposing politically sensitive information, it can also be instrumental for them to acquiring the information elusive to their private sources (e.g., secret police). As a result, when the information asymmetry between the ruler and the ruled is severe enough, dictators will have incentive to relax their control over the media "to learn from bottom-up information and to address social problems before they become threatening" (Qin et al., 2017: 137).

The discussions based on this instrumental view of authoritarian media freedom, however, treat media outlets under a dictatorship as a non-strategic actor and entirely neglect the case of *interim censorship* that dictators can censor a piece of news after it is published but before it reaches the general public. As such, media outlets are shadowed by the possibility of being punished for their truth-telling, especially when the truth could destabilize the regime. Even if an authoritarian government explicitly grants the freedom of expression to people *ex ante*, the fear of being punished *ex post*

can still make them remain silent or self-censor the information unfavorable to the dictator (see, for example, Stern and Hassid, 2012). However, as it will be discussed in the literature review, most existing works hold the view that *once the government sets a certain level of media freedom, information will be automatically revealed accordingly.*

As a matter of fact, students of dictatorships have long understood very well that dictators face a serious informational problem for being unable to know "whether the population genuinely worships them or worships them because they command such worship" (Wintrobe, 1998: 20). While, on the outside, dictators' commands are met with obedience, the truth on the ground—e.g., the true level of people's support for the regime or real problems of their governance—might still elude them if self-censorship prevails. In other words, unless dictators are able to make a *credible commitment* to not censoring any news reports, the non-sustainable media freedom policy would not be able to quench their thirst for information.[1]

In this paper, we make media a strategic actor and introduce the dictator's commitment problem to the authoritarian politics of media freedom. Figure 3.1 illustrates the *complete* process of censorship for highlighting our contributions vis-á-vis the existing literature. Most existing papers only consider the media freedom policy proposed by dictators, but neglect 1) the existence of *interim censorship* which enables the dictators to deviate from the announced media freedom policy and 2) the me-

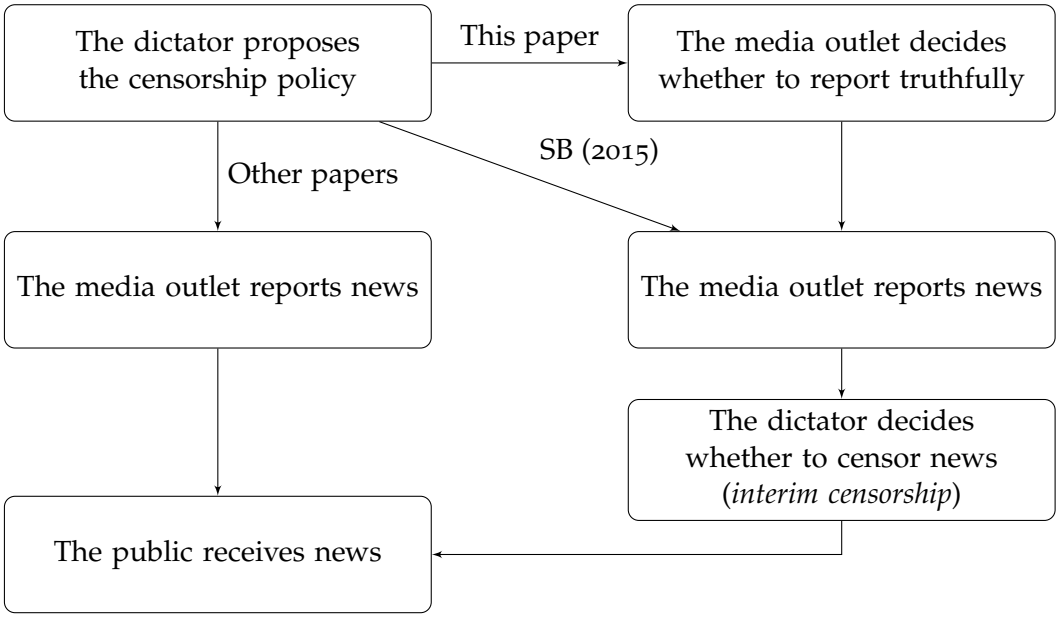
[1]This is not merely a theoretical speculation. The credible commitment from the government indeed determines whether journalists are willing to report truthfully. For example, Shawn W. Crispin, a journalist based in Southeast Asia and a member of the Committee to Protect Journalists, made the following comment on media freedom in Myanmar:

> *"What the [Myanmar] government doesn't understand is how professional reporting happens. If they want to get the truth out, they need to open everything up, let reporters in, and let them publish what they find ... They're not doing that, and everyone knows they won't have a shred of credibility until they do."*

quoted from "Myanmar Government under Suu Kyi Cracks Down on Journalists," Associated Press, February 15, 2018. Retrieved from `https://www.reuters.com/article/us-myanmar-press/private-dailies-re-emerge-in-myanmar-face-difficulties-idUSBRE93007320130401`.

dia's strategic reporting behavior. We consider interim censorship an important stage of the complete process of censorship. For instance, in contemporary China, censorship guidelines are circulated weekly from the Communist Party's propaganda department and the government's Bureau of Internet Affairs to prominent editors and media providers. The editors then self-regulate the reports. For the Internet platforms, the government screens the posts by the *Great Firewall* before publication. However, the rather opaque guidelines allow authorities, such as the provincial or municipal information offices, "to crack down on news stories by claiming that they expose state secrets and endanger the country" after their publication (Xu and Albert, 2017). As a result, we observe episodes, for instance, the former Chinese Premier Wen Jiabao actively urged the media to "fully play their oversight role" and help combat corruption, but those news outlets that actually uncovered scandals were usually punished by the government afterwards (Stern and Hassid, 2012: 1239). In the USSR, the Glavlit executed censorship following the censorship instructions and regulations proposed by the central government, but they also made decisions based on the ideological line of the Communist Party. Published works could be retracted under the Glavlit's discretion post-publishing and before distribution (Lauk, 1999). Despite the importance of the interim censorship, Shadmehr and Bernhardt (2015) (called as SB (2015) in Figure 3.1) is the only exception in the political economy literature that analyzes the interim censorship and its subsequent commitment problem. We model censorship with the process close to theirs, but explicitly consider the media outlet as a strategic actor.

Figure 3.1: Process of censorship



As such, we are able to formalize the concept of self-censorship by the media. Our baseline model establishes that self-censorship by the media is not equivalent to censorship operated by the dictators. Hence, media freedom under dictatorships can only improve their governance and stability by collecting more information when dictators can credibly commit to no censorship and avoid self-censorship. The extensions of the basic model contextualize the dictator's commitment problem in various different scenarios. First, when regime survival is less vulnerable to media information or when dictators can even manipulate it to their favor, counterintuitively, our model finds that their regimes might not be made stabler by doing so since they could even be more blinded by media outlets' self-censorship. Second, from the media's perspective, understanding what constitutes threatening news to the dictator is critical for them to know how they should muzzle themselves. This knowledge, however, is not always unambiguous. Therefore, another extension in our paper relaxes the assumption that the distinction between good and bad news is common knowledge among media outlets. Finally, in the Appendix we show that our results

remain unchanged when dictators also enjoy access to a separate and private channel of information.

The rest of the paper is structured as follows. Section 3.2 reviews the relevant literature and identifies what distinguishes our paper from other studies. Section 3.3 develops a Bayesian game with three actors, the dictator, the media, and the public. Section 3.4 then presents several extensions based on the baseline model. Section 3.5 discusses how this theoretical analysis helps reinterpret the recent empirical findings on the censorship policy of the Chinese government. Section 3.6 concludes with a summary of results and possible avenues for future research.

3.2 Literature Review

To our knowledge, this paper is the first formal analysis on how the dictator's commitment problem affects the authoritarian politics of media freedom. We develop a model of truth-reporting under a dictatorship with testable hypotheses. Our paper makes contributions to various strands of literature on comparative authoritarianism. In addition to the information paucity owing to the lack of mechanisms such as electoral competition in democracies, the dictatorship literature also finds that dictators tend to suppress negative information about the regime so their challengers have nothing to count on for coordinating with potential rebels (Hollyer et al. 2015; Lohnmann,1994). On the empirical front, for instance, King, et al. (2013, 2014, 2017) investigate how the Chinese government implemented censorship and its strategy on information manipulation by applying big data techniques. In the formal theory literature, information manipulation by censorship and propaganda is usually considered as a policy tool to maintain regime stability and survival (e.g., Guriev and

Treisman, 2015).

As the conventional wisdom suggests, censorship can be applied to hinder public coordination (as in the model of Egorov et al., 2009). In a more sophisticated sense, dictators can strategically choose an optimal level of censorship and propaganda to mobilize citizens to take an action (e.g., Edmond, 2013; Gehlbach and Sonin, 2014). We complement this literature by showing that the information paucity under dictatorships is not merely caused by the government censorship, but it also arises when people are too afraid to speak up and censor themselves from revealing the truth. In this paper, we make it clear when dictators will impose censorship, and the conditions under which the media, anticipating the potential censorship, will "tell the truth." As self-censorship is an important but neglected issue in the literature,[2] this paper makes a contribution in providing a theoretical prediction about the characteristics of censored and self-censored information, and discusses the fundamental problem of authoritarian commitment: *How can dictators credibly commit to policies that may endanger themselves?*

Apart from the need to censor unfavorable information, dictators also need the information about the true level of their regime support in order to assess the regime strength and fend off revolution threats from below. In the literature, Egorov et al. (2009), Lorentzen (2014), Shadmehr and Bernhardt (2015), and Huang et al. (2016), have all shown that dictators can sometimes benefit from free media. Empirically, we do observe that there are different levels of media freedom across autocracies. Moreover, Lorentzen (2014), Huang et al. (2016), and Chen and Xu (2017a, 2017b) develop formal models to discuss censorship in the Chinese politics. Most of the works cited

[2]Germanoa and Meier's (2013) study on self-censorship of commercial media in democracies is among the few exceptions.

above, however, do not consider *interim censorship*. Hence, they assume that once the government sets a certain level of media freedom, information will be automatically revealed. They also do not consider how media outlets might strategically respond to dictators' policy given the lack of any institutional checks on their power. We argue that when considering the *complete* process of censorship (as illustrated in Figure 3.1), dictators might face a commitment problem that the pre-announced level of censorship may not be honored. Failing to consider dictators' commitment problem could be an especially critical issue for the papers assuming that dictators manipulate information by Bayesian persuasion (e.g., Gehlbach and Sonin, 2014; Chen and Xu, 2017a), as such technology requires dictators to commit to a particular media freedom policy.

In the literature, Myerson (2008) and Svolik (2012) scrutinize the commitment problem on power building and power sharing in autocratic regime. Our analysis is the first attempt focusing on the commitment problem on media freedom. With regard to this discussion, Shadmehr and Bernhardt (2015) is the most closely related paper to our project as we both recognize the existence of interim censorship and discuss the dictator's commitment problem to optimal amount of censorship. However, just as the bulk of literature in this field, their model focuses exclusively on dictators' decision over censorship and the media are entirely nonstrategic. Moreover, the dictator in their model does not face the informational problem that we focus on.

In addition to investigating the dictator's commitment problem, we further engage our formal analyses with other studies on media freedom. Specifically, we analyze whether the dictator's commitment problem is mitigated or deteriorates when other determinants of media freedom are also incorporated into our baseline model. First, previous studies argue that dictators allow freedom of speech to enhance regime

stability (Chen and Xu, 2017a). Yet, we show that it is more difficult for dictators in stable regimes to overcome the commitment problem than their counterparts in unstable regimes. In addition, we consider the role of propaganda under dictatorships. When a dictator has the capacity of manipulating information via propaganda (Huang 2015; Qin et al. 2017), the media's expectation of being censored would induce them to self-censor in the first place. Furthermore, we consider the case in which the dictator has heterogeneous preferences toward different issues reported by the media. For instance, the dictator may allow the media to report misbehavior of local officials (e.g., Egorov et al., 2009; Huang et al., 2016), but they may not want the media to report negative news against the central government (Qin et al. 2017).

3.3 Model

We consider a one-period model with three actors: a dictator A, a media outlet M, and the public P. There are two states of the world: good or bad ($\omega \in \{B, G\}$). The prior probability for the state being bad is $Pr(\omega = B) = p$.

The media outlet has a private signal $s \in \{b, g\}$ about the state of the world. The precision of the media outlet's information is q, meaning that $Pr(s = g|\omega = G) = Pr(s = b|\omega = B) = q$. We set $q \in (\frac{1}{2}, 1]$ such that the media signal is informative in a straightforward way. The media outlet reports a piece of news $m : \{b, g\} \to \{b, g\}$. Without loss of generality, we restrict the media outlet's options to reporting either a positive or negative piece of news.

The dictator decides whether to censor the report or not ($c \in \{0, 1\}$). Formally, the dictator's strategy is a mapping from m to actions: $c : \{b, g\} \to \{0, 1\}$. Censorship costs k to the dictator.

The media outlet wants to report truthfully and not to be censored. Thus the media outlet's utility is R if it reports truthfully ($m(s) = s$) and is 0 if it reports untruthfully ($m(s) \neq s$), where $R > 0$.[3] If the dictator censors the report, the media outlet incurs a cost (punishment) K. To exclude trivial results, we set $R - K < 0$.

Hence the utility of the media outlet U_M is:

$$U_M = R\mathbb{1}_{\{m(s)=s\}} - Kc.$$

If there is no censorship, the public learn the news report ($\hat{m} = m$) with probability β. With probability $1 - \beta$, the public fail to learn the news report ($\hat{m} = \emptyset$). β can depend on the penetration of the media outlet, the quality of news broadcasting system of a country, and can also be understood as the public's intention toward acquiring political information.[4]

For the baseline model, we analyze a censorship technology such that once implemented, the public will not learn the report.[5] We will discuss a powerful information manipulation technology that can move posteriors further as an extension. The state of the world is realized at the end, so are the payoffs of the dictator and the media outlet.

The dictator may suffer a damage in two ways, from the public's reaction against a negative report or from the realization of the bad state of the world. An example for the former can be the damage from public protests against local corruption, and for the latter can be the loss from the misconducts of corrupt bureaucrats.

[3] We only need to require R to be slightly greater than 0, so that, *ceteris paribus*, the media outlet prefers telling the truth to lying.

[4] β must be less than 1, otherwise the public can infer that no news is bad news, because no news can only happen when bad news is reported but censored. If so, censorship is not useful anymore. In reality, the public are not going to read all the news reported.

[5] We assume that censorship is always successful and the state of the world is always realized. Relaxing these assumptions does not change any results qualitatively.

To highlight the strategic interaction between the dictator and the media outlet, we do not explicitly model the collective action problem of the public. Since it has been well-established in the literature that a publicly broadcasted signal, here the bad news, can facilitate collective actions, among them protests (see, for example, Dewan and Myatt, 2007, 2008, and Casper and Tyson, 2014), we build upon this literature and treat ordinary citizens, i.e., news receivers, as a unitary representative actor.

By protesting against the government, the public can enjoy a utility Y once the action is successful, e.g., forcing the dictator to make some concessions on an issue. The utility from a non-successful action is normalized to 0. The cost for the action is E. Suppose the protest is going to succeed when the state of the world is bad. Formally, the public decides whether to protest $f : \{b, g\} \to \{0, 1\}$. The public's utility function U_P is thus:

$$U_P = \begin{cases} Pr(\omega = B)Y - E & \text{if } f = 1, \\ 0 & \text{if } f = 0. \end{cases}$$

Whenever there is a public protest, the dictator suffers a damage. If there is no public protest, the state of the world is realized at the end of the game. The dictator suffers a damage once the realized state is bad. The utility function of the dictator U_A is therefore:

$$U_A = -D(m, f = 1, \omega)f - D(m, f = 0, \omega = B)(1 - f) - kc.$$

We impose the following assumptions:

Assumption 1:

$$-D(m = b, f = 1, \omega) = -D(m = b, f = 0, \omega = B) \equiv -D_b$$

and

$$-D(m = g, f = 1, \omega) = -D(m = g, f = 0, \omega = B) \equiv -D_g.$$

This means we do not differentiate the damages resulted from a protest or a realization of the state of the world. The damage severity only depends on the information received by the dictator, and the two possible values are denoted as $-D_b$ and $-D_g$. This assumption helps simplify the exposition and analysis significantly, but does not lead to any loss of interesting result.[6]

Assumption 2:

$$D_g > D_b.$$

A piece of bad news from the media outlet serves as a warning message that enables the autocrat to respond to a challenge better. For example, a news report could give the dictator a heads-up for where an upcoming insurrection might take place. Hence, the damage severity is smaller when receiving the negative report from the media outlet.[7]

Assumption 3:

$$D(m, f = 0, \omega = G) = 0.$$

This means if the good state is realized and $f = 0$, there is no damage for the dictator.

Table 3.1 summarizes the damages the dictator suffers in different situations.

[6]Differentiating the values of damage from the public protest and the realization of the state of the world does not provide qualitatively interesting result. Similarly, we do not differentiate the $-D_g$ resulting from a truthful but wrong report and that from a babbling report. Treating them as the same simplifies the analysis significantly, while interesting results remain. Particularly, if we think that the damage arising from a protest $-D(m = b, f = 1)$ is greater than that from, for example, corruption $-D(m, f = 0, \omega = B)$, the commitment problem could only be more severe than under our assumptions.

[7]Because the media report is informative regarding the state of the world, even when the public protest is not going to take place, the dictator can still benefit from having the information in order to better prepare for a possible realization of the bad state of the world.

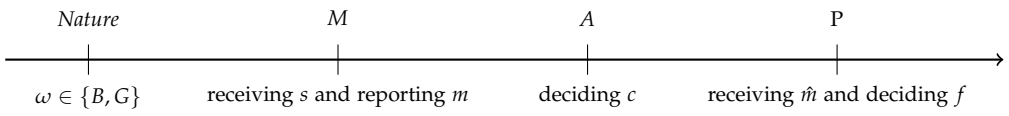
Table 3.1: Damages to the dictator depending on the report he receives $\{m = b, m = g\}$

	$f = 0$	$f = 1$
$\omega = B$	$\{-D_b, -D_g\}$	$\{-D_b, -D_g\}$
$\omega = G$	$\{0, 0\}$	$\{-D_b, -D_g\}$

In summary, the game proceeds as follows (Figure 3.2 is the timeline of the game):

1. N decides the state of the world ($\omega \in \{B, G\}$) and reveals a signal $s \in \{b, g\}$ to M
2. M sends a report m
3. A decides whether to censor the news report
4. If not censored, the public receive the report with probability β
5. The public decide whether to protest
6. If there is a public protest, the game ends and payoffs are realized
7. If there is no public protest, the state of the world and payoffs are realized.

Figure 3.2: Timeline



The equilibrium concept is Perfect Bayesian Equilibrium (PBE) in pure strategies. We characterize two types of equilibria: truth-reporting equilibrium and self-censoring equilibrium. A truth-reporting equilibrium means that the media outlet sends truthful signal $m(s) = s$, and the other actors update their beliefs about the state of the world with the report; a self-censoring equilibrium means that the media outlet always sends a positive signal $m(b) = m(g) = g$, and the other actors do not update their beliefs with the report. To avoid discussion of uninteresting equilibria, weakly dominated strategies are excluded.

Analysis

We begin the analysis by defining the dictator's and the public's posterior beliefs, supposing the media outlet follows the truth-reporting strategy and the dictator does not censor the report. The dictator's posterior belief is denoted as $Pr(\omega = B|m) = \mu_m^A$. The public's posterior belief is denoted as $Pr(\omega = B|\hat{m}) = \mu_{\hat{m}}^P$. There are three possible posteriors for the public:

$$\begin{cases} Pr(\omega = B|\hat{m} = g) = \frac{p(1-q)}{p(1-q)+(1-p)q} & \equiv \mu_g^P, \\ Pr(\omega = B|\hat{m} = \emptyset) = p & \equiv \mu_n^P, \\ Pr(\omega = B|\hat{m} = b) = \frac{pq}{pq+(1-p)(1-q)} & \equiv \mu_b^P. \end{cases}$$

μ_g^P is the posterior if receiving a truthful positive media report, and μ_b^P is the posterior if receiving a truthful negative media report, and μ_n^P is when the public do not receive the information. Given $q > \frac{1}{2}$, we have $\mu_b^P > \mu_n^P > \mu_g^P$. Essentially, μ_g^P and μ_b^P are the lower and upper bound of the posterior possible to form with the media report. Similarly, we define the posteriors for the dictator:

$$\begin{cases} Pr(\omega = B|m = g) = \frac{p(1-q)}{p(1-q)+(1-p)q} & \equiv \mu_g^A, \\ Pr(\omega = B|m = b) = \frac{pq}{pq+(1-p)(1-q)} & \equiv \mu_b^A. \end{cases}$$

We solve this game by supposing truth-reporting equilibrium, and checking if every player has no incentive to deviate from it. Following backward induction, we consider the public's action first. As a protest is only successful when the state of the world is bad, the public believe that the probability for their protest to succeed is their posterior μ^P. Formally, the public decides whether to protest by comparing the expected payoffs with and without protest:

$$EU_P = \begin{cases} \mu_{\hat{m}}^P Y - E & \text{if } f = 1, \\ 0 & \text{if } f = 0. \end{cases}$$

As such we can further define a social stability threshold $\pi = \frac{E}{Y}$. Once $\mu_{\hat{m}}^P > \pi$, the public decide to protest against the regime.

In this section, we assume that $\mu_b^P \geq \pi > \mu_n^P$ to introduce our main results, and we discuss the other cases in Section 3.3. When $\mu_b^P \geq \pi > \mu_n^P$, the public decide to protest only after reading a bad news report. Recall that the disutility for the dictator could come from either the public's response against a negative report or the realization of the bad state of the world. The negative report is useful for the dictator as it can help him prepare for challenges and hence decreases subsequent damage. Censorship, on the other hand, can help him avoid damage brought by the negative report. As such, suppose truthful equilibrium, we can calculate the utility of the dictator U_A in different situations as below:

$$U_A = \begin{cases} 0 & \text{if } m = g \text{ and } \omega = G, \\ -D_g & \text{if } m = g \text{ and } \omega = B, \\ -k & \text{if } m = b \text{ and } c = 1 \text{ and } \omega = G, \\ -D_b - k & \text{if } m = b \text{ and } c = 1 \text{ and } \omega = B, \\ -\beta D_b - (1-\beta)0 = -\beta D_b & \text{if } m = b \text{ and } c = 0 \text{ and } \omega = G, \\ -\beta D_b - (1-\beta)D_b = -D_b & \text{if } m = b \text{ and } c = 0 \text{ and } \omega = B. \end{cases}$$

Note that when $m = b$ and $c = 0$ and $\omega = G$, the expected damage is from the protest; when $m = b$ and $c = 0$ and $\omega = B$, the expected damage is from both the protest and the realization of the bad state of the world.

Ex ante, the dictator may want the media outlet to report truthfully, as a bad report that serves as a warning message can decrease the damage from $-D_g$ to $-D_b$ once a protest take place or a bad state is realized. Given the prior for the state to be bad is p, the information value is the difference between the expected payoffs with and without truthful media reports (suppose censorship is not an option for the dictator). The former is

$$Pr(m = b|m(s) = s)(-\mu_b^A D_b - (1-\mu_b^A)\beta D_b) + Pr(m = g|m(s) = s)\mu_g^A(-D_g) =$$
$$-(1-p-q+2pq)(\mu_b^A D_b + (1-\mu_b^A)\beta D_b) - (p+q-2pq)\mu_g^A D_g,$$

and the latter is

$$Pr(\omega = B)(-D_g) = -pD_g.$$

Hence the information value is

$$\begin{aligned} V &= pD_g - (1-p-q+2pq)(\mu_b^A D_b + (1-\mu_b^A)\beta D_b) - (p+q-2pq)\mu_g^A D_g \\ &= pq(D_g - D_b) - (1-p-q+pq)\beta D_b. \end{aligned} \tag{3.1}$$

When it is greater than 0, the dictator will have incentive to promise media freedom.

However, once the media report a piece of bad news truthfully,[8] the dictator must decide whether to censor the news. The expected payoff for the dictator if not censoring the report is

$$\underbrace{-\beta D_b}_{\hat{m}=b} \overbrace{-\mu_b^A(1-\beta)D_b}^{\hat{m}=\emptyset,\omega=B}.$$[9]

If he censors the report, the expected payoff is

$$\underbrace{-k}_{c=1} \overbrace{-\mu_b^A D_b}^{\hat{m}=\emptyset,\omega=B}.$$

Overall, the dictator chooses between *no censorship and probably suffering an immediate damage from public protest* $(-\beta D_b)$ and *committing the costly censorship* $(-k)$ *and waiting for the realization of the state of the world*. Thus, once $(1-\mu_b^A)\beta D_b \geq k$, the dictator will censor a bad report. In other words, the commitment constraint for the dictator to not censor bad news is

$$(1-\mu_b^A)\beta D_b < k. \tag{3.2}$$

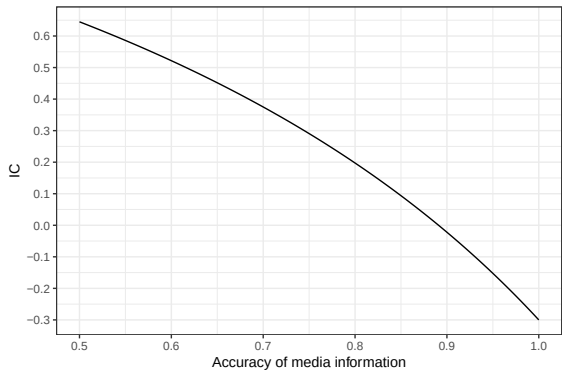
When the incentive for censorship is positive to the dictator ($IC > 0$ as shown in Figure 3.3), he faces the commitment problem.

[8] We know that whenever the media outlet receives a good signal $s = g$, it will report this signal truthfully, so we only need to consider its action when $s = b$.

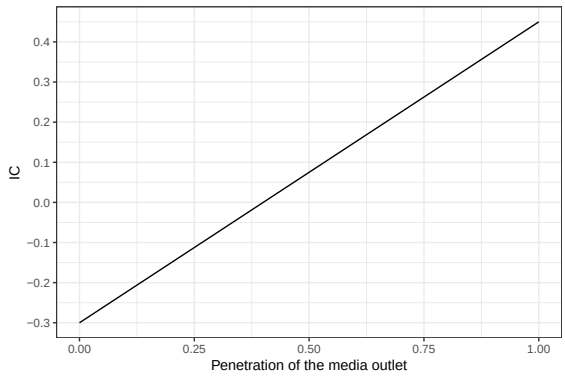
[9] The expected payoff is derived from $-\beta D_b - \mu_b^A(1-\beta)D_b$. The first term is the immediate damage arising from the public's reaction against the negative report, and the second term is the damage when the public do not receive the report but the bad state of the world is eventually realized.

Figure 3.3: Commitment Problem

(a) Fixed $\beta = 0.9$



(b) Fixed $q = 0.7$



If this constraint is not satisfied, the dictator will censor bad news, and expecting to be censored, the media outlet will choose self-censorship in advance because $R - K < 0$, and the equilibrium is self-censoring. If the commitment constraint is satisfied, the equilibrium will be truth-reporting instead.[10]

Proposition 3.3.1. *When the information value of truthful reporting V is greater than* 0 *for the dictator, he benefits from truthful reporting. Nevertheless, he can only commit to not censoring bad news if* $(1 - \mu_b^A)\beta D_b < k$ (truth-reporting equilibrium: {c=0, m(s)=s}). *When he cannot commit to no censorship, the media outlet does not report truthfully* (self-

[10]One may wonder whether the results are robust once the game is repeated. Suppose an infinitely repeated game and the dictator can rationally expect the future. At the period one, when he makes his censorship decisions, reputational concerns have to be taken into account as a component of the cost for censorship k. Therefore, everything remains almost the same as in the one-shot game, except that the patience of the dictator also partially determines the cost for censorship.

censoring equilibrium: {c=1, m=g}).

We therefore know that the severity of the commitment problem is:

1. increasing in β,

2. decreasing in μ_b^A *(and hence decreasing in* p *and* q*),*

3. increasing in D_b*, and*

4. decreasing in k.

The significance of Proposition 3.3.1 is twofold. First of all, the bulk of autocratic media freedom models in the literature (Egorov et al. 2009; Lorentzen, 2014; Qin et al., 2017) treat media as a non-strategic actor, and therefore assume away the possibility of self-censorship. The dictator, however, cannot easily obtain the true information when he cannot commit himself to not censoring a bad report *ex post* even if truthful reporting is valuable to him. In this self-censoring equilibrium, the dictator's expected payoff is $-pD_g$. The implication from this result is that, once the commitment constraint is not satisfied, the media outlet self-censors its reports even if the dictator openly vows to promote greater media freedom.

Censorship and self-censorship may seem to be very similar in the stylized models, but the consequences are substantially different. When only considering censorship, as in the existing literature, the theoretical framework assumes that the dictator can decide his optimal information manipulation strategy, and thus the dictator can learn the truthful information before deciding whether to censor the report. However, as our analysis suggests, when self-censorship is involved, the dictator does not learn the information that is blocked by the media outlet, and, as a result, he cannot enjoy the information value V.

Second, the comparative statics derived in Proposition 3.3.1 also allow us to make

various meaningful predictions for future empirical implementations. Analyzing the constraint, we know that the commitment problem is more severe when: 1) β is higher, 2) μ_b^A is smaller, 3) D_b is larger, and 4) k is smaller. Intuitively, higher β means that once the media outlet reports bad news, the likelihood for the public to receive the news is also higher and people know the dictator would have stronger incentive to impose censorship. It therefore predicts that the dictator has harder time making a commitment to no censorship when the information transmission is easier from media to ordinary citizens. With the recent advance in information and communication technology, the issue of self-censorship in a dictatorship could therefore be even more salient than ever before.[11]

Moreover, a smaller μ_b^A gives the autocrat stronger incentive to gamble on the realization of the state of the world, hoping that the media outlet's report is a false alarm. In other words, when the dictator is more uncertain about the true state of the world *ex post*, his commitment ability is also undermined. In addition, when the damage (D_b) caused by bad news is more severe, the dictator will certainly have stronger incentive to censor it. Naturally, it will also make it more difficult for him to commit to no censorship to media. Finally, since k is the upper bound for Inequality (3.2), as k gets larger, the range of values for the other parameters can take for the inequality to be satisfied is also larger, thus strengthening dictator's commitment ability.

Corollary 3.3.1. *When the information value of truthful reporting for the dictator is positive and he suffers from the commitment problem, the cost of censorship k can serve as a commitment device. If the transaction cost for adjusting k to a new k* (e.g., the cost for introducing*

[11]The prevalence of social media may or may not make censorship *per se* less effective. On the one hand, the news contents online could be duplicated and spread rapidly, but on the other hand, as Crook (2017) argues, they can also be removed more quickly and quietly.

foreign media) is smaller than the information value of truth-reporting

$$V = pq(D_g - D_b) - (1 - p - q + pq)\beta D_b,$$

the dictator may want to choose the k^* high enough to induce truth-reporting.

Myerson (2008) suggests that when a dictator faces commitment problem, he has incentive to tie his hands in order to make his promise credible. Here we discuss how a dictator could address this commitment problem by making himself less able to implement censorship. Suppose before the media outlet reports any news, the dictator can choose the cost of censorship k^* such that he cannot adjust it back and forth easily, as long as the cost of setting the parameter is less than the information value of truthful reporting, it pays for the dictator to use k^* as a commitment device for more easily inducing the media to tell the truth. Note that when making this decision, the dictator and the public have same information (the common prior belief and other parameters), so changing k does not serve as a signal about the state of the world to the public.

For example, when Myanmar's authoritarian government decided to relax its media control in 2012, not only did it eliminate in August the pre-publication censorship that had been imposed on Burmese print media for almost 50 years, but it also, in the following year, formally dissolved the censorship authority, the Press Scrutiny and Registration Division, and granted publishing licenses to private media—including "D-Wave," the official paper of the then opposition leader Aung San Suu Kyi's National League for Democracy (NLD).[12] More critically, the government also allowed

[12]"Chief Censor in Myanmar Caps His Red Pen," New York Times, September 21, 2013. Retrieved from http://www.nytimes.com/2012/09/22/world/asia/myanmars-chief-censor-is-closing-his-office.html; "Private Dailies Re-emerge in Myanmar, Face Difficulties," Reuters, April 1, 2013. Retrieved from https://www.reuters.com/article/us-myanmar-press/private-dailies-re-emerge-in-myanmar-face-difficulties-idUSBRE93007320130401.

two foreign media, Associated Press and Japan Broadcasting Corporation (NHK) to establish local news bureaus in Yangon.[13] These new policies brought significant changes to Myanmar's media environment. As argued in Corollary 3.3.1, the introduction of the two foreign news agencies could have served as a commitment device for the dictator, considering the higher cost (k) of censoring them vis-à-vis domestic ones. First of all, according to the annual reports published by the Freedom House before and after the changes, Myanmar's Freedom of the Press score rose substantially from an embarrassingly poor score of 94 (with 100 being the worst) in 2011 to 70 in 2014.[14] Second, while a score of 70 still placed Myanmar's press freedom in the category of "Not Free" in 2014, in the subsequent years after the 2012 reform, local private print media and professional associations for journalists mushroomed and more "truthful" reports criticizing the government for violating the freedom of press also became prevalent (Kean, 2017: 150-151). The inclusion of foreign media as a commitment device in the reform package seems to have made Myanmar's local media react more proactively to the government press liberalization and forthcoming about what they saw on the ground.

Social Stability

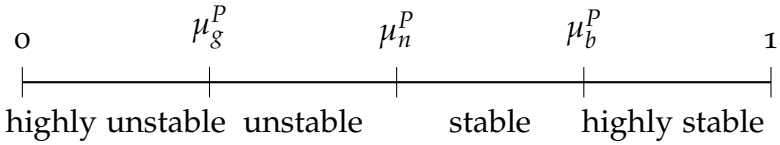
Now we generalize the analysis by considering different levels of social stability, i.e., different values of π. Recall that we have three possible posteriors for the public when on the truth-reporting equilibrium: μ_g^P, μ_n^P, and μ_b^P, and therefore we can divide a society into four levels of stability by these three posteriors as shown in Figure 3.4.

[13] "6 Multiformat Journalists to Staff New AP Bureau in Myanmar," Associated Press, March 30, 2013. Retrieved from `https://www.ap.org/press-releases/2013/6-multiformat-journalists-to-staff-new-ap-bureau-in-myanmar`.

[14] What has to be noted here is that the state of the media freedom in Myanmar has not made any progress since then. From the most recent Freedom House report (Freedom House, 2017), the score even went down slightly to 73.

We call a society highly stable if $\pi > \mu_b^P$, which means the dictator will not be hurt by bad news even when the public receive a truthful negative report. If π lies in the interval of μ_n^P and μ_b^P ($\mu_b^P \geq \pi > \mu_n^P$), which means the dictator will be hurt if the public receive a truthful negative report, we call the society stable. This is essentially the baseline case. Likewise, we call a society unstable if the autocrat is subject to a damage unless the public receive a truthful positive report ($\mu_n^P \geq \pi > \mu_g^P$). A society is highly unstable if the autocrat is always damaged no matter what information the public receive ($\mu_g^P \geq \pi$).[15]

Figure 3.4: The location of π on μ^P



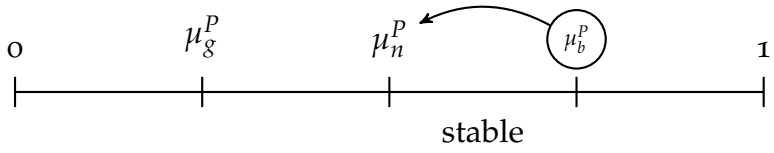
When a society is stable ($\mu_b^P \geq \pi > \mu_n^P$), the autocrat can avoid a loss from bad news by censoring news report once the media outlet produces a negative report as the censorship could decrease the posterior from μ_b^P to μ_n^P. The dictator's censorship decision is based on Inequality (3.2).

Figures 3.5-3.8 illustrate the impact of dictator's action on the public's posterior belief in different scenarios. The public's posterior beliefs with circle indicate the tipping points for the public's protest decisions. The solid arrows indicate the effects of the dictator's information manipulation technology and the public's decisions on protest will be changed subsequently; the dashed arrows, on the other hands, are the cases that, although the dictator can manipulate the public's posterior beliefs, but

[15]One may ask why, when the society is highly unstable, the dictator still needs information to prepare for an upcoming challenge when he already anticipates it is going to happen for sure. A rationale could be, when there are many localities where the crisis could take place and the dictator does not know exactly which is the hotspot, the media report provides this information to him.

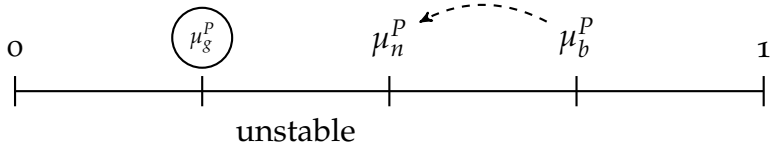
cannot change the decisions on protest.

Figure 3.5: The incentive for censorship when a society is stable



On the other hand, when a society is unstable ($\mu_n^P \geq \pi > \mu_g^P$), changing μ_b^P to μ_n^P is not useful, as the dictator will be hurt when either bad news or no news is received by the public.

Figure 3.6: The incentive for censorship when a society is unstable



Thus we know that when the dictator can only attempt to decrease the posterior from μ_b^P to μ_n^P, the commitment problem only arises when the society is stable. When the society is unstable, highly stable or highly unstable, information manipulation is not helpful, and thus the dictator can commit not to censoring any news report.

Proposition 3.3.2. *If the dictator's information manipulation technology can only decrease the posterior from μ_b^P to μ_n^P, the commitment problem arises when the society is stable ($\mu_b^P \geq \pi > \mu_n^P$).*

3.4 Extensions

Censorship with Propaganda

Now suppose the dictator can engage in a *strong* censorship technology which can turn a negative report to a positive one (changing $m = b$ to $m = g$[16]), i.e., by pro-

[16]The public will not receive the bad news for sure, but the probability for them to read the good news is again β.

paganda. For example, Stockmann and Gallagher (2011) suggest that the Chinese government could turn a disaster to a positive message to stabilize the regime by propaganda. If this information manipulation technology is successfully implemented, the public will not learn the negative report and will receive a positive report instead. That is, now the dictator has the ability to turn the posterior from μ_b^P to μ_g^P.

Figure 3.7: The incentive for censorship when a society is stable

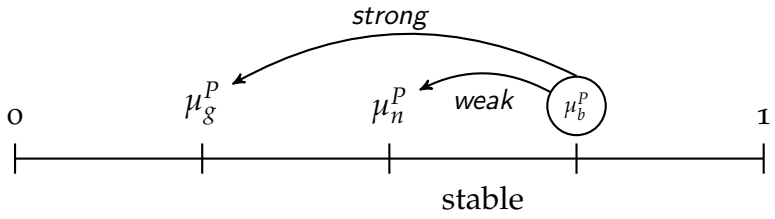
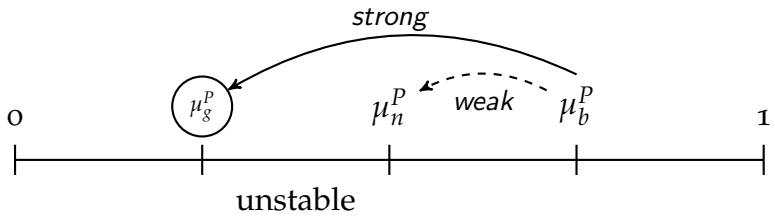


Figure 3.8: The incentive for censorship when a society is unstable



From Figures 3.7 and 3.8 above, it is clear that now the dictator faces the commitment problem not only when the society is stable, but also when it is unstable. That is, with a stronger capacity in manipulating information, the commitment problem is now more severe.

Without truthful reporting, the public always protest, and the payoff for the dictator is $-D_g$, because he is not aware where the problem is. Suppose truthful equilibrium, the utility of the dictator U_A in different situations as below:

$$U_A = \begin{cases} -(1-\beta)D_g & \text{if } m = g \text{ and } \omega = G, \\ -D_g & \text{if } m = g \text{ and } \omega = B, \\ -(1-\beta)D_b - k & \text{if } m = b \text{ and } c = 1 \text{ and } \omega = G, \\ -D_b - k & \text{if } m = b \text{ and } c = 1 \text{ and } \omega = B, \\ -D_b & \text{if } m = b \text{ and } c = 0 \text{ and } \omega = G, \\ -D_b & \text{if } m = b \text{ and } c = 0 \text{ and } \omega = B. \end{cases}$$

Without censorship, the *ex ante* value of truthful reporting is

$$V_u = D_g - (1-p-q+2pq)D_b - (p+q-2pq)(\beta\mu_g^A + (1-\beta))D_g.^{17} \tag{3.3}$$

Recall that in an unstable society, the dictator suffers the damage unless the public receive a credible good report. When the media outlet reports a negative piece of news, if it is not censored, the payoff is $-D_b$. If the dictator censors the news and reverts it to be a positive report, the payoff is

$$\underbrace{\mu_b^A}_{\omega=B}\overbrace{(-D_b)}^{\hat{m}\in\{\emptyset,g\}} + \underbrace{(1-\mu_b^A)}_{\omega=G}\overbrace{(-(1-\beta)D_b)}^{\hat{m}=\emptyset}\underbrace{-k}_{c=1}.$$

The commitment constraint for no censorship is thus

$$(1-\mu_b^A)\beta D_b < k,$$

which is the same as when the society is stable (Inequality 3.2).[18]

We conclude the above discussion in Proposition 3.4.1.

Proposition 3.4.1. *If the dictator's information manipulation technology is moving μ_b^P to μ_g^P, the commitment problem arises not only when the society is stable ($\mu_b^P \geq \pi > \mu_n^P$) but also when it is unstable ($\mu_n^P \geq \pi > \mu_g^P$).*

By comparing the values of truthful reporting, V and V_u, we can find that a dictator with a strong capability in information manipulation is more willing to commit

[17] Because when the society is unstable, the dictator suffers the damage whenever the public does not receive a credible positive news report. See the proof of Corollary 2 for details.

[18] The dictator compares the same expected payoffs from censorship and no censorship in both cases.

to no censorship when the society is unstable than when the society is stable. This is because, in a unstable society, the dictator is not only more in need of a warning message, a credible positive report can also "correctly guide public opinion" (Dimitrov, 2017). In other words, committing to no censorship is particularly beneficial for the dictator when the society is unstable, as an informative negative signal helps the dictator to prepare for the upcoming challenge and a promising positive signal is needed to diffuse public grievance.

Corollary 3.4.1. *Given $V_u > V$ for all possible parameter values, a dictator with a strong capability in information manipulation is more willing to commit to no censorship when the society is unstable than when the society is stable.*

Proof. See the appendix. □

We focus our discussion on the cases in which the society is either stable or unstable. It is because when the society is highly stable, the dictator would not be hurt by negative reports, and accordingly, the dictator has no incentive to censor them. Similarly, the dictator will not censor any news in a highly unstable society and should commit to not censoring any reports to ensure truthful reporting and minimize any potential damages. The key difference between a stable society from an unstable one is that the dictator in the former does not need any positive news to sustain his legitimacy.

Recognizing this difference clarifies the relationships between the severity of the commitment problem and 1) the dictator's *capability* in information manipulation and 2) his *willingness* to commit to no censorship. The dictator in an unstable society suffers the commitment problem only when he is more able to manipulate information. However, as he needs positive reports to boost his legitimacy, in an *unstable* society,

he is more willing to commit to no censorship to convince media outlets not to self-censor than when in a stable society. Otherwise self-censorship leads to the babbling equilibrium in which any positive reports are useless in convincing the public. It turns out that, due to the dictator's lower willingness to commit to no censorship, the regime in a *stable* society could suffer more severe damage once the true state of the world is bad.[19]

This finding suggests that *social instability* is an important determinant of the dictator's willingness to commit to media freedom. The case of Taiwan provides an illustrating example to this point. After the KMT retreated to Taiwan in 1949, some liberal intellectuals affiliated with it published a fortnightly periodical, *Free China Review*. With the goal of anti-communism, *Free China Review* was employed by the government to propagandize its support for liberal democratic values in order to earn political and economic sponsorship from the U.S. At the same time, *Free China Review* commented on domestic political and economic issues to challenge the KMT's authoritarian rule (Rigger, 1999: 104). The outbreak of the Korean War in the early 1950s consolidated the relationship between the KMT and the U.S. As Taiwans society was more stabilized, the KMT was less willing to accept criticisms raised by liberal intellectuals. Eventually, in 1960, the government shut down *Free China Review* and arrested Lei Zhen, a prominent member of its editorial board, who was jailed for ten years. The arrest of Lei Zhen also "put to rest any lingering doubts about the regime's willingness to suppress its critics" (Rigger, 1999: 106).

[19]This point could be related to Kuran (1991) about why the collapse of "stable" East European communist regimes, where "people routinely applauded speakers whose message they disliked, joined organizations whose mission they opposed" (26), was so surprising. The realization of the true state of the world that initiated the whole event was, however, rather hard to be identified, as Kuran contends: "What specific events set the revolutionary bandwagon in motion? One must recognize that attempting to answer this question is akin to trying to identify the spark that ignited a forest fire or the cough responsible for a flu epidemic" (37).

It was not until the late 1980s that the government committed to media freedom to rebuild its legitimacy. As the opposition force challenged the authoritarian rules via social movements during the late 1980s, the KMT also faced internal power struggle. In addition, many core party members were involved in political and economic scandals that hurt the KMT's legitimacy, such as the government-ordered assassination of Henry Liu in the U.S. and the bank run of the Tenth Credit Cooperation of Taipei (Hsiung 1986; Nathan and Ho 1993). As a result, in a meeting with the chairperson of the *Washington Post* in 1986, President Chiang announced that the KMT government would repeal the martial law and lift the ban on parties and media. Subsequently, the number of registered newspapers increased from 31 to 122 in 1988, and the government no longer censored news reports. In short, the social instability induced KMT to commit to media freedom that it hoped could help regain its reputation.

Selective Authoritarian Responsiveness

In this section, we return to the baseline model but allow the dictator to choose between actively addressing the problem raised by the media, censoring the negative report, or not taking any action at all. This extension goes beyond the recent idea of "responsive authoritarianism", of which China is often regarded as an example (Huang et al., 2016; Chen and Xu, 2017b), in making the dictator's responsiveness to (social) problems a choice variable. Among all the issues the dictator could encounter, some of them are *softer* than the others, and the determinants could be, for example, the relationship between the dictator and the officials concerned with regard to a particular problem, or the nature of the problem. It is well-known that the Chinese government is particularly responsive to environmental issues but still some of them are more politically sensitive and are stalled in the Chinese government's

policy agenda.[20]

There are therefore two twists to the baseline model. First, the dictator now must simultaneously decide 1) if the report should be censored, and 2) if the issue should be addressed. We denote the latter decision as $r \in \{0,1\}$. Second, the dictator has preferences over addressing some issues more than the others. Formally, denote the type of issues $i \in \{L,H\}$. The dictator will get a disutility $-D_r$ if addressing such issues ($i = H$) undermines his power base. On the other hand, dealing with the issues that push the balance to the dictator's favor ($i = L$) will give him a payoff $-D_r + e$. For example, the dictator might use media reports as a justification to purge a political enemy from the government. Overall, for the dictator, addressing type-H issues is more costly than addressing type-L issues. When $e > D_r$, the reform can *even* yield positive payoff for the dictator.

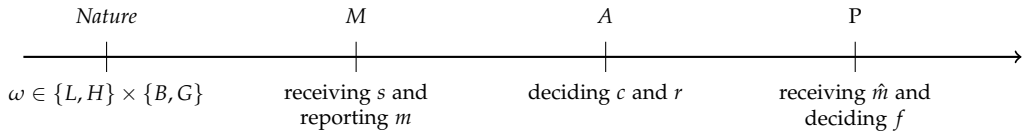
The dictator sets the priority for his (policy/reform) agenda, while the media outlet does not know it precisely. The media outlet is aware of the probability for a specific issue belonging to type H is $Pr(i = H) = \alpha$. Formally the dictator's strategy is a mapping from i and m to actions: $c : \{L,H\} \times \{b,g\} \to \{0,1\}$ and $r : \{L,H\} \times \{b,g\} \to \{0,1\}$. Overall, the tuple of strategies for the dictator is $X_i : \{b,g\} \to \{0,1\}^2$, and we collect X_i as $X = \{X_L, X_H\}$. For the media outlet, it makes its reporting decision after observing α, conditional on its anticipation of the dictator's action. Other things remain the same as in the baseline model. Figure 3.9 is the

[20]For instance, in a 2006 report the Eurasia Group prepared for the US-China Economic and Security Review Commission, we see how some of the issues (reforms) were removed from the Chinese government's policy agenda owing to certain political complications.

> "*While recent reforms to China's energy regulation structure will undoubtedly change the dynamics of bureaucratic decision-making, ... top-level attention to energy issues has not helped the government to overcome turf wars between different ministries, and China's domestic energy reform, such as liberalizing pricing, remains bogged down by political sensitivities.*" (6-7)

timeline of the game.

Figure 3.9: Timeline



The payoffs for the dictator given different actions when $m = b$ is summarized as Table 3.2:

Table 3.2: U_A when $m = b$

$X_i = \{c_i, r_i\}$	$i = L$	$i = H$
$X_i = \{0,0\}$	$-\mu_b^A D_b - (1-\mu_b^A)\beta D_b$	$-\mu_b^A D_b - (1-\mu_b^A)\beta D_b$
$X_i = \{1,0\}$	$-\mu_b^A D_b - k$	$-\mu_b^A D_b - k$
$X_i = \{0,1\}$	$-D_r + e$	$-D_r$
$X_i = \{1,1\}$	$-D_r + e - k$	$-D_r - k$

It is clear that $\{1,1\}$ is dominated by $\{0,1\}$. By comparing the payoffs of the dictator between whether to address an issue and censor a bad report depending on issue types, we characterize the dictator's best responses in Table 3.3.

Table 3.3: The dictator's best responses

$X = \{c_L, r_L, c_H, r_H\}$	Conditions
$X^A = \{1,0,1,0\}$	(1), (3)
$X^B = \{0,1,1,0\}$	(1), (4), (5)
$X^C = \{0,1,0,1\}$	(6), (7)
$X^D = \{0,1,0,0\}$	(2), (8), (9)
$X^E = \{0,0,0,0\}$	(2), (10)

Note: Conditions: (1) $(1-\mu_b^A)\beta D_b > k$; (2) $(1-\mu_b^A)\beta D_b < k$; (3) $k + e < D_r - \mu_b^A D_b$; (4) $k + e > D_r - \mu_b^A D_b$; (5) $k < D_r - \mu_b^A D_b$; (6) $k > D_r - \mu_b^A D_b$; (7) $\mu_b^A D_b + (1-\mu_b^A)\beta D_b > D_r$; (8) $\mu_b^A D_b + (1-\mu_b^A)\beta D_b < D_r$; (9) $\mu_b^A D_b + (1-\mu_b^A)\beta D_b + e > D_r$; (10) $\mu_b^A D_b + (1-\mu_b^A)\beta D_b + e < D_r$.

When expecting the dictator to choose X^A, the media outlet will self-censor its report because the expected payoff for the news outlet is $R - K < 0$. That is, when reporting a bad news, none of the issues will be addressed but a report will be

censored for sure. If X^C, X^D, or X^E is expected, the media outlet will truthfully report its information because the dictator can commit himself not to censoring news reports in this case.

When X^B is expected, the result is more interesting. In this case, the media outlet will report truthfully when $\alpha < \frac{R}{K}$. As the media outlet cannot perfectly recognize whether an issue is among those the dictator is willing to address, the media outlet might report something unwanted by the dictator and conceals something wanted. It is worth noting that this is not a media outlet's mixed strategy, as the origin of this strategy is not the dictator's indifference between censoring and not censoring a report, but the media outlet's uncertainty about the dictator's preference on the issues.

By combining the conditions for the dictator's best responses above, we can summarize the equilibria as Table 3.4 and Proposition 3.4.2.

Table 3.4: Equlibria

X	Equlibrium
X^A	Self-censoring
X^B	Truth-reporting when $\alpha < \frac{R}{K}$
X^C, X^D, X^E	Truth-reporting

Proposition 3.4.2. *When $k < min\{D_r - \mu_b^A D_b - e, (1 - \mu_b^A)\beta D_b\}$, the media outlet always self-censors its report* (self-censoring).

When $D_r - \mu_b^A D_b - e < k < min\{D_r - \mu_b^A D_b, (1 - \mu_b^A)\beta D_b\}$, the media outlet reports truthfully when $\alpha < \frac{R}{K}$; when $\alpha \geq \frac{R}{K}$, the media outlet self-censors its report (truth-reporting with some censorship).

For other possible parameter values, the media outlet reports truthfully (truth-reporting with no censorship).

3.5 Empirical Implications

In addition to deriving useful comparative statics for understanding the relationship between dictator's commitment constraint and media outlets' strategic behavior from the propositions above, our formal analysis also helps advance the recent empirical research on the censorship under dictatorships. For example, in one of their seminal studies on the Chinese internet censorship,[21] King et al. (2013) use a data-driven method to find that "posts are censored if they are in a topic area with collective action potential and not otherwise" (13). However, they also note that

> *...our methodology reveals a great deal about the goals of the Chinese leadership, but it misses self-censorship and censorship that may occur before we are able to obtain the post in the first place.* (5)

In our theoretical framework, the observed information in their study belongs to either "not self-censored but censored" (from the *truth-reporting with some censorship* equilibrium) or "not censored" (from the *truth-reporting with no censorship* equilibrium) categories. If the data truncation was random, their conclusion reached by counting "not self-censored but censored" posts on the social media would not be biased. However, as our model shows, the truncation caused by self-censorship was very likely to be a result of the media's (or netizens') strategic behavior, and therefore was anything but random. As the following analysis is going to show, our empirical contributions can be easily explicated through a simple model of sample selectivity.[22]

[21] What is worth noting here is that our arguments can also travel to the behavior on social media. While the information disclosed by social media users can be valuable to the dictator, it can also easily fall prey to his censorship given its potential for facilitating contentious collective actions even among non-users (Little, 2016). Naturally, just as traditional media, social media users will also self-censor themselves in the presence of the dictator's commitment problem.

[22] Given the limit on paper length, please see Appendix for a more formal discussion of how our model can help address the sample selectivity in King et al. (2013) and an extension using EM algorithm to deal with the issue of missing information on self-censored messages.

Specifically, the comparative statics derived from our theoretical model can make it relatively easy to specify a function for the determinants of one's self-censorship behavior since the commitment constraint shown in our model can all be viewed as issue and media outlet specific. From a Maximum Likelihood Estimation perspective, this simple statistical model also allows us to derive a full log-likelihood function. More importantly, the parameters in the following discussion are all observables.

Consider that the equilibrium conditions characterized above are specific to each issue and each media outlet.[23] From Proposition 3.4.2, we know that when $k < min\{D_r - \mu_b^A D_b - e, (1-\mu_b^A)\beta D_b\}$, the media outlet self-censors the information and this kind of information is thus not observed in King et al. (2013).[24] When $D_r - \mu_b^A D_b - e < k < min\{D_r - \mu_b^A D_b, (1-\mu_b^A)\beta D_b\}$, the media outlet reports truthfully when $\alpha < \frac{R}{C}$. Some of the reports however will be censored. They are the censored information observed by King et al. (2013). Those reported on the truth-reporting equilibrium path should also have been observed in their data. Analyzing the conditions further, we can categorize the following situations:

Table 3.5: Empirical implication

Situations	Censorship observable?
A. $(a) < (b)$ and $(b) < (c)$	Some censorship observed when $(a) < k < (b)$
B. $(a) < (b)$ and $(c) < (b)$	Some censorship observed when $(a) < k < (c)$
C. $(b) < (a)$	No censorship observed (information either self-censored or not censored)

Note: Denote $(a) = D_r - \mu_b^A D_b - e$, $(b) = (1-\mu_b^A)\beta D_b$, $(c) = D_r - \mu_b^A D_b$

[23] Extending our model to explain the behavior of multiple media outlets is straightforward and even their interactions do not really concern us. This is because the media outlet only cares about its own behavior and the dictator's reaction to its reports. Strategic interactions between the news outlets could be interesting for future research.

[24] In the model, we assume that the media outlet must report a piece of news, and therefore, it self-censors its information by sending out a positive report. In reality, the self-censoring media outlet can choose to remain silent. In either case, the true information obtained by the media outlet is concealed and hence not observable by researchers.

The reports that are not self-censored but censored are either from Situation A or B. In Situation A, the media outlets anticipate that some information might be censored once reported but they have a sufficiently strong belief that the dictator would address the related problems, so they do not self-censor their reports. In Situation B, the commitment constraint (3.2) does not hold for all issues, but again the media outlets believe the dictator might welcome their truthful information in order to address social problems, and hence may still publish the news. In Situation C, the dictator is not going to carry out any reforms. Hence, the media outlets make reporting decisions based on the commitment constraint (3.2). No reports that were censored and observed by King et al. (2013) belong to this category, as the information is either self-censored or not censored at all.

It turns out that the reports which are not self-censored but censored (from Situation A or B in Table 3.5) might well be those the dictator potentially has incentive to address. Then it is not surprising that they are highly likely to have "collective action potential," since the collective actions that can be provoked *explicitly* online tend to aim at asking for policy reforms. While some mild government criticisms could satisfy the commitment constraint (belonging to Situation C in Table 3.5), serious ones related to long-held political taboos such as 1989 Tiananmen Square protests would have been self-censored in advance if the media do not expect the government will address them. In other words, instead of allowing all kinds of government criticisms, the existence of censorship stifles most serious types even before they can be seen. This then gives rise to a potential issue of sample selection where observable data only contain the *politically admissible* criticisms.

The analysis above provides a potentially fruitful way to move forward the re-

search agenda set by King et al.'s (2013) paper. Based on the variations in the measurable characteristics of the media outlets, a fuller picture of missing information because of self-censorship could be reconstructed. With a better understanding of the information environment, researchers then can more correctly infer the dictator's censorship strategy.

3.6 Conclusion

The recent rise of scholarly interests in comparative authoritarianism finds that there is a wide range of variation in media freedom in dictatorships. Yet, this line of research assumes that the media will report truthfully when the dictator sets a policy to allow for the media freedom. Most papers, apart from Shadmehr and Bernhardt (2015), do not consider the fact that the dictator can hide a piece of news before it reaches the public by *interim censorship*. Moreover, existing studies treat the media as a non-strategic actor, thus assuming away the possibility of their strategic self-censorship. This paper takes issue with this assumption and develops a game theoretic model to investigate the strategic interaction among the dictator, the media, and the public.

We study determinants for the dictator to successfully commit media freedom without inducing self-censorship of the media. As the dictator can always punish the media outlet for truthful reports that may humiliate or even destabilize the regime, his proposal to media freedom may not be credible. Accordingly, we demonstrate that the commitment problem could prevent the media from "telling the truth" even when the dictator has a clear demand for truthful information. The dictator fails to gain the potential information value due to media self-censorship. This is a key

difference between censorship and self-censorship, because when deciding whether to censor a report, the dictator would still learn the information; when the media outlet self-censors its report, the dictator is not able to learn the information.

The commitment problem becomes more severe under when: 1) there is higher quality of information transmission or the public have a higher intention to acquire political information, 2) if the dictator has a stronger ability of manipulating information to his favor (e.g., propaganda). Due to the self-censorship induced by the commitment problem, it is very likely that total amount of information censored should be much more than the observed amount as the recent studies claim.

Our analysis further argues that, when in a stable society, a dictator faces a more serious information insufficiency due to the commitment problem. Nevertheless, it also demonstrates that the dictator may have incentive to commit to no censorship, especially when the society is unstable. Making censorship more costly can serve as a commitment device, which can be achieved by, for example, introducing foreign media into the country. In addition, uncertainty about whether the dictator welcomes a truthful report leads to situations that unwanted reports are reported and desirable information is self-censored.

Our results provide both theoretical and empirical implications. Theoretically, our results point out that the dictator's commitment problem is much more substantially important than it was assumed in previous studies of authoritarian media politics. In particular, the dictator's information insufficiency should be much more severe than previously expected when considering self-censorship. Future studies should consider this commitment problem when investigating the emergence and determinants of media freedom in dictatorships. Empirically, our findings suggest

that further investigation on the linkage between censorship and self-censorship can advance the understanding of authoritarian media politics. Current studies delineate the preferences of authoritarian governments to policies or governance by collecting the censored information (before it is censored). Yet, the self-censored information may be more critical to the dictator, because the contents are probably more sensitive and (or) the information outlets are with higher quality or penetration. Future empirical studies may endeavor to identify the media's self-censored information in order to infer the dictator's censorship strategy correctly with a more holistic picture of missing information in autocracies.

Bibliography

[1] Casper, Brett Allen, and Scott A. Tyson. 2014. "Popular Protest and Elite Coordination in a Coup d'état." *The Journal of Politics* 76 (2): 548-564.

[2] Chen, Jidong, and Yiqing Xu. 2017a. "Information Manipulation and Reform in Authoritarian Regimes." *Political Science Research and Methods* 5 (1):163-178.

[3] Chen, Jidong, and Yiqing Xu. 2017b. "Why Do Authoritarian Regimes Allow Citizens to Voice Opinions Publicly?." *The Journal of Politics* 79 (3): 792-803.

[4] Crook, Tim. 2017. "If Print Dies So Does Freedom." In John Mair, Tor Clark, Neil Fowler, Raymond Snoddy, and Richard Tait (Eds.). *Last Words? How Can Journalism Survive the Decline of Print?*. Suffolk, UK: Abramis Academic.

[5] Davidson, Russell and James G. MacKinnon. 2004. *Econometric Theory and Methods*. New York : Oxford University Press.

[6] Dewan, Torun, and David P. Myatt. 2007. "Leading the Party: Coordination, Direction and Communication." *American Political Science Review* 101 (4):825-843.

[7] Dewan, Torun, and David P. Myatt. 2008. "The Qualities of Leadership: Direction, Communication, and Obfuscation." *American Political Science Review* 102 (3):351-368.

[8] Dimitrov, Martin K. 2017. "The Political Logic of Media Control in China." *Problems of Post-Communism* 64 (3-4): 121-127.

[9] Edmond, Chris. 2013. "Information Manipulation, Coordination, and Regime Change." *The Review of Economic Studies* 80 (4): 1422-1458.

[10] Egorov, Georgy, Sergei Guriev, and Konstantin Sonin. 2009. "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *American Political Science Review* 103 (4): 645-668.

[11] Egorov, Georgy, and Konstantin Sonin. 2011. "Dictators and Their Viziers: Endogenizing the Loyalty-Competence Trade-off." *Journal of the European Economic Association* 9 (5): 903-930.

[12] Eurasia Group. 2006. *China's Overseas Investments in Oil and Gas Production*. Report Prepared for the US-China Economic and Security Review Commission.

[13] Freedom House. 2011. *Freedom of the Press 2011*. Available at: `https://freedomhouse.org/report/freedom-press/freedom-press-2011`.

[14] Freedom House. 2014. *Freedom of the Press 2014*. Available at: `https://freedomhouse.org/report/freedom-press/freedom-press-2014`.

[15] Freedom House. 2017. *Freedom of the Press 2017*. Available at: `https://freedomhouse.org/report/freedom-press/freedom-press-2017`.

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

[17] Gerasimov, Gennadi. 1998. "Russia's Media Revolution: From Party Control to Money Control." Asia Pacific Issues No. 37.

[18] Germano, Fabrizio, and Martin Meier. 2013. "Concentration and Self-Censorship in Commercial Media." *Journal of Public Economics* 97: 117-130.

[19] Guriev, Sergei, and Daniel Treisman. 2015. "How Modern Dictators Survive: An Informational Theory of the New Authoritarianism." *NBER Working Paper* No. 21136.

[20] Hollyer, James R., B. Peter Rosendorff, and James Raymond Vreeland. 2015. "Transparency, Protest, and Autocratic Instability." *American Political Science Review* 109 (4): 764-784.

[21] Hsiung, James C. 1986. "Taiwan in 1985: Scandals and Setbacks." *Asian Survey* 26 (1): 93-101.

[22] Huang, Haifeng. 2015. "Propaganda as Signaling." *Comparative Politics* 47 (4): 419-437.

[23] Huang, Haifeng, Serra Boranbay-Akan, and Ling Huang. 2016. "Media, Protest Diffusion, and Authoritarian Resilience." *Political Science Research and Methods* 1-20.

[24] Kean, Thomas. 2017. "Public Discourse." In Adam Simpson, Nicholas Farrelly and, Ian Holliday (Eds.). *Routledge Handbook of Contemporary Myanmar*. 1st Edition. Abingdon, Oxon ; New York, NY: Routledge.

[25] King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107 (2): 326-343.

[26] King, Gary, Jennifer Pan, and Margaret E. Roberts. 2014. "Reverse-engineering Censorship in China: Randomized Experimentation and Participant Observation." *Science* 345 (6199): 1-10.

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

[28] Kuran, Timur. 1991. "Now Out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics* 44 (1): 7-48.

[29] Lauk, Epp. 1999. "Practice of Soviet Censorship in the Press: The Case of Estonia." *NORDICOM Review* 20 (2): 27-39.

[30] Lee, Francis L. F., and Joseph Chan. 2009. "Organizational Production of Self-Censorship in the Hong Kong Media." *International Journal of Press/Politics* 14(1): 112-133.

[31] Little, Andrew T. 2016. "Communication Technology and Protest." *Journal of Politics* 78(1): 152-166.

[32] Little, Roderick J. A. and Donald B. Rubin. 2002. *Statistical Analysis with Missing Data.* 2nd Edition. New York: John Wiley & Sons.

[33] Lohmann, Susanne. 1994. "Information Aggregation Through Costly Political Action." *The American Economic Review* 84 (3): 518-530.

[34] Lorentzen, Peter. 2014. "China's Strategic Censorship." *American Journal of Political Science* 58 (2): 402-414.

[35] Myerson, Roger B.. 2008. "The Autocrat's Credibility Problem and Foundations of the Constitutional State." *American Political Science Review* 102 (1): 125-139.

[36] Nathan, Andrew J., and Helena Ho. 1993. "Chiang Ching-kuo's Decision for Political Reform." In Shao-Chuan Leng, ed., *Chiang Ching-kuo's Leadership in the Development of the Republic of China on Taiwan*. Lanham, M.D.: University Press of American.

[37] Nathan, Andrew J.. 2003. "Authoritarian Resilience." *Journal of Democracy* 14(1): 6-17.

[38] Qin, Bei, David Strömberg, and Yanhui Wu. 2017. "Why Does China Allow Freer Social Media? Protests versus Surveillance and Propaganda." *Journal of Economic Perspectives* 31(1): 117-140.

[39] Rigger, Shelley. 1999. *Politics in Taiwan: Voting for Democracy*. New York, NY: Routledge.

[40] Shadmehr, Mehdi, and Dan Bernhardt. 2015. "State Censorship." *American Economic Journal: Microeconomics* 7 (2): 280-307.

[41] Shiraev, Eric. 2013. *Russian Government and Politics*. 2nd Ed. Basingstoke, Hampshire [England]; New York: Palgrave Macmillan.

[42] Stern, Rachel E., and Jonathan Hassid. 2012. "Amplifying Silence: Uncertainty and Control Parables in Contemporary China." *Comparative Political Studies* 45 (10): 1230-1254.

[43] Stockmann, Daniela, and Mary E. Gallagher. 2011. "Remote Control: How the Media Sustain Authoritarian Rule in China." *Comparative Political Studies* 44 (4): 436-467.

[44] Svolik, Milan W.. 2012. *The Politics of Authoritarian Rule.* Cambridge, UK; New York, NY: Cambridge University Press.

[45] Wintrobe, Ronald. 1998. *The Political Economy of Dictatorship.* Cambridge, UK; New York, NY: Cambridge University Press.

[46] Xu, Beina, and Eleanor Albert. 2017. *Media Censorship in China*. Available at: `https://www.cfr.org/backgrounder/media-censorship-china`.

Chapter 4

Reputation and Media Selection

"The sinister fact about literary censorship in England is that it is largely voluntary .... So far as the daily newspapers go, this is easy to understand. The British press is extremely centralized, and most of it is owned by wealthy men who have every motive to be dishonest on certain important topics."

— George Orwell, *Orwell's Proposed Preface to Animal Farm*

4.1 Introduction

The prevalence of low-quality news contents, particularly the so-called "fake news" has become a growing concern in democracies. This issue is considered to be severely problematic, to the extent that "it was alleged that fake news might have been pivotal in the election of President Trump" (Allcott and Gentzkow, 2017: 232). This phenomenon has been attributed to the relatively low barrier to enter the news market nowadays. Thus, an intuitive way to address this problem is to fact-check news and correct the rumors.

Fact-checking, however, is not without problem. The number of reports to be checked is high. People may doubt the neutrality of fact checkers. Research also shows that fact-checkers tend to disagree among themselves (Lim, 2018). Moreover,

corrections against rumors "frequently fail to reduce misperceptions" and can backfire (Nyhan and Reifler, 2010). Hence, alternatively, some scholars and public intellectuals suggest that providing the audience reputational information of the media outlets could help them correctly choose news sources and discipline the media. There are already organizations, such as the Trust Project, beginning to provide this kind of quality evaluation. In early 2019, Microsoft starts providing ratings on the quality of news outlets in its Edge mobile browser via an add-on named "NewsGuard."[1]

The logic of this suggested solution is quite straightforward. If an objective quality information is provided to the audience, it is believed that they can select news outlets wisely, and the news providers can also be disciplined to provide reports of good-quality. This follows directly from the reputation effect characterized by Kreps and Wilson (1982).

This logic, however, is incomplete. Provision of reputational information makes some media outlets more influential than the rest. Because, unlike the reputation of individuals, the reputation of firms is tradable (Tadelis, 1999 and 2002), providing reputational information to consumers create a new dimension of strategic consideration which can be exploited by interested parties. The latter may have incentives to own the news outlets with established reputation in order to manipulate public opinion quickly and effectively effectively. This paper aims at investigating the promised positive effects brought by provision of reputational information building on the framework originally proposed by Tadelis (1999 and 2002).

In this framework, interested parties can establish new media outlets or take over the existing ones to engage in political persuasion and information manipulation.

[1]See the following webpage for how Microsoft rates the news outlets: `https://www.newsguardtech.com/ratings/rating-process-criteria/`.

When consumers know nothing about the media outlets, all the media outlets are considered equal. When the consumers learn the reputations of the outlets, the outlets are differentiated, some are more popular, and hence more profitable and more influential, than the others. That is, reputation is important for the media outlets to be "powerful". While some new entrants have built up their own with struggle, it is common that interested parties choose to takeover the established and reputable news outlets and altered their editorial policies. For example, Hong Kong newspapers taken over by pro-Beijing corporations or sponsored by the Chinese government have changed their political stances significantly. According to some content analyses, those newspapers "lent overwhelm support to Beijing with almost no news items favorable the democrats" and the editors tend to work as gatekeepers by removing contents positive to the democrats and manipulating the headlines "to the extent that they did not match the stories" (Ngok, 2007: 963-967).

This kind of episodes also happens in consolidated democracies. In 2011, Swiss newspaper *Basler Zeitung* was purchased by Christoph Blocher, a prominent figure of the rightwing Swiss People's Party. This transaction was intentionally kept secret at the beginning through an indirect ownership of his daughter and he denied this transaction for months after it was eventually confirmed. This transaction was accused to be a political plot of the Party.[2] The *London Evening Standard* is said to begin favoring more conservative positions after purchased by Russian businessman and former KGB agent Alexander Lebedev. In a report by Media Reform Coalition and Goldsmiths, University of London, the newspaper was criticized for being biased against Labour Party's Sadiq Khan in coverage of the 2016 London Mayoral Election

[2]Newspaper takeover reveals rightwing strategy.
https://www.swissinfo.ch/eng/newspaper-takeover-reveals-rightwing-strategy/31752732

campaigns.[3] More recently, President Trump also accused the *Washington Post*, which is owned by Amazon, a lobbyist for the Internet company.[4] In a recent study, Martin and McCrain (2019) show that, after bought by the right-wing broadcasting group Sinclair, the news coverage of many local stations in the US have turned conservative. They further show that, while there exists a small decline in viewership at those stations, "the vast majority of viewers watching before the acquisition date continued to watch afterwards, despite the substantial changes in political content" (20).

This paper introduces a new theoretical framework to evaluate the impact of reputation on the media market and the effect of potential market regulations on social welfare. Apart from adding new insights to the discussion of media market regulation, this paper argues that exerting influence on consumers information by media ownership is not only a feasible strategy in autocracies. It is possible even in a competitive media market. Moreover, providing the audience reputational information is not likely to eliminate the existence of biased media outlets. Finally, the analysis shows that, while banning all the transactions of media ownerships may not be socially optimal, imposing partial restriction on the market is likely to be beneficial for social welfare.

4.2 Related Literature

This paper relates to the literature concerning about the behavior of media outlets with reputational concerns, regulations on media ownership, media capture and media power, and the impact of biased media.

[3] Londons monopoly newspaper biased in mayoral coverage, report shows. https://www.gold.ac.uk/news/mayoral-election-coverage-bias/

[4] Trump claims Washington Post is 'lobbyist' for Amazon. http://thehill.com/homenews/administration/381089-trump-claims-washington-post-is-lobbyist-for-amazon

In political economy literature, there are analyses on the behavior of media outlets with reputational concerns, for example, Gentzkow and Shapiro (2006), Ashworth and Shotts (2010), and Sheen (2019), but there lacks a formal analysis on impact of the availability of reputational information. In other words, the existing literature only regards reputation per se as a goal, instead of treating reputation as a tool or a valuable characteristic for the media to achieve their ultimate objectives, such as maximizing profit or influencing public opinion. In this paper I study a game theoretical model to investigate this important question.

I adopt the model originally proposed by Tadelis (1999, 2002). The key differences between this paper and Tadelis (1999, 2002) are: 1. this paper is the first application of his model in media political economy, and 2. this paper provides social welfare evaluations and policy implications that are not the focus of Tadelis (1999, 2002). Moreover, Tadelis (1999, 2002) focuses on non-existence of good-type-only equilibrium in the long run(three or more periods). In this paper, considering the nature of the media market, I focus on a two-period case, and show the conditions for the non-existence of good-type-only equilibrium. That is, I show that provision of reputation *per se* could not eliminate the low-quality media from the market even in the short run.

Media ownership and concentration is considered as a determining factor of the well-functioning of democracies. The debate on regulations of media market has a long tradition across social sciences (see Baker, 2007, Downing, 2011). In political economy literature, Anderson and McLaren (2012) analyze media mergers and contend that reasonable regulation on media merger is desirable in protection of the benefit from diversity of reports. This paper introduces a new theoretical framework

to analyze media mergers and acquisition motivated by political considerations, evaluates the impact of reputation on the media market, and assess the effect of potential market regulations on social welfare.

This paper also contributes to the literature on media capture and media power (Besley and Prat, 2006; Prat, 2017), and consumer behavior in a competitive news market (Sobbrio, 2014). This paper argues that exerting influence on consumers information by media ownership is not only a feasible strategy in autocracies (see Gehlbach and Sonin, 2007), it is possible even in a competitive media market through obscure ownership rule. Moreover, providing the audience reputational information is not likely to eliminate the existence of biased media outlets. This paper argues that, while banning all the transactions of media ownerships may not be socially optimal, imposing partial restriction on the market is likely to be beneficial for social welfare.

This paper also connects to recent research focusing on assessing the impact of biased media on public opinion. Martin and Yurukoglu (2017) show that the audience that are not very familiar with political biases of media outlets can be influenced by the slanted media. Wolton (2019) argues that the audience do not need to be unaware of the media bias to be influenced, and biased media are not necessarily bad for democracy. In psychological research, Pennycook and Rand (2018) finds that people are susceptible to fake news due to lazy thinking, this paper argues that even a perfect rational ant not lazy Bayesian could be persuaded by fake news that is spread from a reputable outlet. This is particularly true for those who are not very well informed about political information and media quality, and this type of citizens are essentially the majority in most societies.

4.3 Model Setup

I study the reputational effect by analysing an adverse-selection model proposed by Tadelis (1999). The model has 2 periods.[5] The actors are: a continuum of editorial teams and a continuum of short-lived citizens C in each period. It can be understood as citizens require information from the media to make decisions in different periods. The editorial teams operate newspapers and the citizens want to make decisions after learning information from the newspapers.[6]

In this overlapping generation model, there are three generations. Some editorial teams will retire from the economy at some point and some new players will want to enter the market. To ensure that the number of editorial teams is constant, I assume that there is a measure 1 of the teams (one generations) that will operate for two periods, and in each period, there is also a measure 1 of teams (two generations) that will only operate for one period, and thus the total measure of teams is 2 throughout the game (illustrated as Figure 4.1).

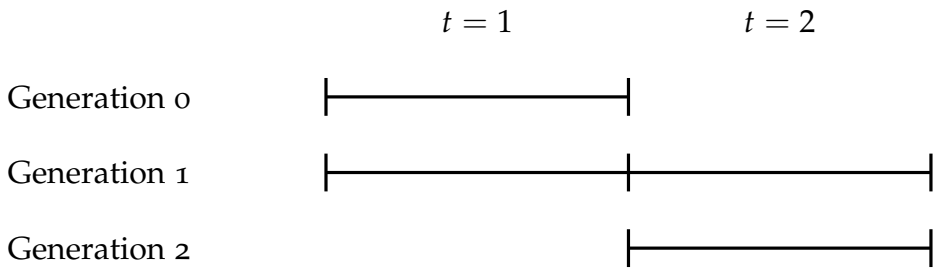


Figure 4.1: Generations of the Editorial Teams

There are two types of editorial teams, good or bad, $\tau \in \{G, B\}$. The proportion

[5]For the media market, transactions do not occur frequently, and hence a two-period model is rather suitable for our purpose. However, it may be still interesting to consider what will change if the game is of a longer time horizon. Tadelis (1999) discusses the possibility and shows that there is no equilibrium that only good-type buyers can successfully purchase reputable names. In other words, a rather pessimistic result arises in a longer timeframe. This is however not the focus of the present paper.

[6]I call the media outlets newspapers, but the analysis applies to other information sources such as TV channels.

for the editorial teams to be of good-type is γ. Only the newspapers run by the good-type teams are able to provide correct information to their readers. The bad-type exploit the trust from their readers. For example, they want to persuade their readers to make decisions that serve for their own political agenda. Suppose the state of the world is $\omega_t \in [0,1]$. The probability of the good-type providing correct information is $Pr(n_t^j = \omega_t) = P_G$ and that of the bad-type is P_B, and without loss of generality, I assume that $1 > P_G > P_B = 0$. It means that the bad-type newspapers cannot refrain from misleading the readers (maybe because they care a lot about the issue), and the good-type newspapers sometimes commit mistakes.[7] Suppose there are more readers than editorial teams (and thus newspapers), and hence the newspapers are operating in a competitive market. A market of newspaper ownerships is created as some editorial teams retire after the first period.

The number of readers is constant in each period. As in Tadelis (1999), it is assumed that the readers are on the long side of the market.

At the very beginning of the game, the editorial teams establish their own newspapers and report news. As in Baron (2006), Gentzkow and Shapiro (2006), Anderson and McLaren (2012) and many other papers, readers read newspapers and then make (private) decisions guided by the newspapers. The readers are Bayesians. However because of limited capability and willingness in information acquisition (as in Bernhardt et al. 2008, Duggan and Martinelli 2011, Gehlbach and Sonin 2014), they are not well-informed about politics nor very interested in it. Hence each reader only reads one newspaper. A reader enjoys a payoff normalized to 1 if a correct decision ($a_t^j = \omega_t$) is made, and if the decision is wrong, the payoff is 0. At the end of the

[7] In reality, the difference between the good- and bad-type editors should be less dramatic, and hence the public are even less likely to differentiate them from their behavior.

first period, either the readers or some organizations aggregate the information of the newspapers' reporting performances, and the reputational record of the media is publicized to every actor.

In the following period, the Generation-0 teams retire and can sell their established newspapers. The Generation-1 editorial teams can decide to rename the newspapers, purchase the brands from the Generation-0 teams, or simply continue operating the newspapers; the Generation-2 editorial teams can establish new newspapers or takeover the newspapers operated by the Generation-0 teams.

It is assumed that all of these activities are not observable by the readers in each period.[8] What the readers know is the previous reporting performance, denoted as h, i.e., whether a newspaper provides correct information at Period 1. A history consists of three possible records for each period denoted as $\{S, F, N\}$. S means the report was correct, F means the reports was wrong, and N means the newspaper is just established or renamed, and thus without performance record.

After the transactions, the newspapers report information. The readers again learn information from the chosen newspapers and make decisions. Then the payoffs are realized.

A reader R_t^i's problem is to select a newspaper j in order to maximize

$$U_{R_t^i} = \mathbf{1}\{a_t^i = \omega_t\},$$

which is the utility from making a correct decision.

Assume that the utility of each period is not discounted, the overall payoff to a newspaper j is:

[8]The results will hold as long as the observability is not perfect and free.

$$U_{N^j} = \begin{cases} f(W_1^j) + V, & \text{for Generation 0} \\ f(W_1^j) + f(W_2^j) - V, & \text{for Generation 1} \\ f(W_2^j) - V, & \text{for Generation 2.} \end{cases}$$

W_t^j is the reputation of the newspaper j at time t, which is the probability that the newspaper can provide correct information $Pr(n_t^j = \omega_t)$ to the reader. $f(W_t^j)$ is a function that maps reputation to profit, and V is the price of a reputable name. $f(W_t^j)$ can be any strictly increasing function; for simplicity, in the baseline analysis, I assume that $f(W_t^j) = W_t^j$. The general case is discussed in Section 4.7.

The timing of this game can be summarized as follows:

1. At the beginning of $t = 1$, the Generation-0 and Generation-1 editorial teams establish newspapers and report news

2. Each reader purchases a newspaper and makes a decision a_1^i

3. Newspapers records are publicised, Generation-0 teams retire, and $t = 1$ ends

4. At the beginning of $t = 2$, the Generation-1 editorial teams can decide to rename the newspapers, purchase the names from the Generation-0 teams, or simply continue operating the newspapers; the Generation-2 editorial teams can establish new newspapers or takeover the newspapers operated by the Generation-0 teams

5. The newspapers report news

6. Each reader selects a newspaper and makes a decision a_2^i

7. The payoffs are realized and the game ends.

The equilibrium concept is Perfect Bayesian Equilibrium (PBE). In addition to understanding the equilibrium strategies of the actors, we are interested in the composition of media market after transactions of ownership. The equilibrium allocation of the names in the society is characterized based on Rational-Expectations Equilibrium

of an overlapping generation model. An equilibrium allocation is supported by the following parameters and it is assumed that the actors can correctly expect them: ϕ denotes the proportion of good-type of Generation 1 that failed at $t = 1$ subsequently purchase S names at $t = 2$; ψ is the proportion of bad-type of Generation 1 that successfully buy the S names; δ is the proportion of the good-type new entrants that buy the S names, and ρ is that for the bad-type new entrants. In Section 4.6, I discuss the consequence when the public mistakenly expect the allocation of good names.

4.4 Analysis

At the beginning of $t = 1$, W_1^j is the same for all newspapers because all of them are just created, and thus $W_1^j = \gamma P_G$ for both good and bad types.

The readers choose a newspaper depending on the reputation of the newspapers, i.e. the probability that the newspaper can provide correct information $Pr(n_t^j = \omega_t)$ to the reader. Hence, at the first period, because all the newspaper seem to be the same to the readers, each of them selects a newspaper randomly. When the reputations of newspapers vary, a reader selects a newspaper ramdomly from those with highest reputation.

If no trade is allowed, only the good-type editorial teams can provide correct information to the readers, and therefore in all the following periods newspapers with a S record are of good quality, and the readers can make optimal purchasing decision by learning the reporting histories of the papers. To be competitive in the market, all newspapers have to build up their own reputation, and in the extreme case as shown in this model, only the good-type editorial teams are capable of doing so. Hence to address a purely adverse selection problem, it seems socially optimal to

prohibit all transactions of ownerships, but in most countries, it is usually not possible to ban all this kinds of transactions, and it is also not optimal once moral hazard is involved (discussed in Section 4.8). In this section I discuss the social implications of an entirely free market of reputation. Section 4.5 discusses what if the market is partially restricted.

When reputation is tradable, there is no equilibrium without transaction of the name, because otherwise a good reputation perfectly signals the quality of the editorial teams, and hence, the newspapers with good names can charge higher price. This creates an incentive for those without good records, such as interest groups that just enter the media market, to purchase the names from the retired operators. Moreover, as a newspaper with a better reputation is also with a larger readership, it is in a better position to manipulate the electorate (Kennedy and Prat, 2018).

Proposition 4.4.1. *The ownerships of newspapers with successful records will be traded in all equilibrium.*

When purchases by both existing teams and new entrants are allowed, we can solve for the rational expectations equilibrium by equating the supply and demand of the S names. The supply are those newspapers with S names operated by Generation 0 (γP_G) and the buyers of the names are those failed to obtain the S records on their own at $t = 1$ and those new entrants at $t = 2$, formally, $\phi^a\gamma(1 - P_G) + \psi^a(1 - \gamma) + \delta^a\gamma + \rho^a(1 - \gamma)$. Recall that ϕ^a denotes the proportion of good-type of Generation 1 that failed at $t = 1$ subsequently purchase S names at $t = 2$; ψ^a is the proportion of bad-type of Generation 1 that successfully buy the S names; δ^a is the proportion of the good-type new entrants that buy the S names, and ρ^a is that for the bad-type new entrants. The superscript "a" stands for "all" as every willing buyer can enter the

market.

The market-clearing condition is therefore:

$$\gamma P_G = \phi^a \gamma(1 - P_G) + \psi^a(1 - \gamma) + \delta^a \gamma + \rho^a(1 - \gamma). \tag{4.1}$$

In a competitive market, the premium from a good reputation equals the price of the name. That is, all the potential buyers bid for the names available on the market with the equilibrium price.

We first calculate the rational expected posterior beliefs for a newspaper being operated by a good-type editorial team when it is with S record and without past record[9] respectively as follows:

$$\begin{aligned} Pr^a(G|S) &= \frac{\gamma P_G + \phi^a \gamma(1 - P_G) + \delta^a \gamma}{2\gamma P_G} \\ &= \frac{P_G + \phi^a(1 - P_G) + \delta^a}{2P_G} \end{aligned}$$

and,

$$\begin{aligned} Pr^a(G|N) &= \frac{(1 - \phi^a)\gamma(1 - P_G) + (1 - \delta^a)\gamma}{(1 - \phi^a)\gamma(1 - P_G) + (1 - \phi^a)(1 - \gamma) + (1 - \delta^a)\gamma + (1 - \rho^a)(1 - \gamma)} \\ &= \frac{2\gamma - \gamma P_G - \phi^a \gamma(1 - P_G) - \delta^a \gamma}{2 - 2\gamma P_G}. \end{aligned}$$

At the beginning of $t = 2$, the price of the reputation of a newspaper with a S record is

$$\begin{aligned} V^a &= W_2(G|S) - W_2(G|N) \\ &= (Pr^a(G|S) - Pr^a(G|N)) * P_G \\ &= \frac{P_G + \phi^a + \delta^a - \phi^a P_G - 2\gamma P_G}{1 - \gamma P_G}. \end{aligned}$$

[9] All those with with a bad record in the first period will change their names to erase their failures if they are not able to purchase the names with good records because $Pr(G|N) = \gamma > Pr(G|F) = \frac{\gamma(1-P_G)}{\gamma(1-P_G)+(1-\gamma)}$ when no one changes its name.

When only the bad-type purchase the S names, once $\gamma > \frac{1}{2}$, the price of the good reputation is 0. That means when there are more likely to have high-quality editorial teams in each generation and the bad-type are very competitive in the market of names, the names become meaningless. However, as suggested in Proposition 1, in every equilibrium there should still be transactions of names, otherwise, a name with S record will again turn to have positive value as it is a perfect indicator of quality when there is no trade in expectation.

Proposition 4.4.2. *The equilibrium price of reputation is*

$$V^a = \max\{0, \frac{P_G + \phi^a + \delta^a - \phi^a P_G - 2\gamma P_G}{1 - \gamma P_G}\}.$$

From comparative statics of V^a, it is clear that when higher proportion of good-type buyers successfully purchase the names, the price is higher. On the other hand when the editorial teams are more likely to be of good-type, the price is lower, because of less scarcity of the good teams.

Corollary 4.4.1. *The equilibrium price V^a increases in ϕ^a and δ^a and decreases in γ.*

Following any reasonable allocation rule that is not explicitly modelled here, some of the buyers successfully acquire the reputable names. Those Generation-0 editorial teams that failed in the first period and fail to acquire the good names instead rename their old newspaper, mimicking the new entrants that also fail to acquire the good names.

There are multiple equilibria. In equilibria with $P_G > \frac{\phi^a + \delta^a}{1 + \phi^a}$, there are some reputable names are acquired by the bad types. A numerical example is when $\phi^a = \delta^a = 0.2$, and $P_G > \frac{1}{3}$. Hence we know provision of reputational information *per se* does not guarantee the elimination of low-quality editorial teams from the

market.

4.5 Market Restrictions

Considering the public goods nature of the media, countries worldwide impose various regulations on this industry. For example, in the UK, the Communications Act 2003 requires the government to maintain a sufficient media plurality and media mergers "may be subject to intervention on media public interest grounds."[10] In this section, I first motivate why providing reputational information could sometimes lead to problematic outcomes, and then analyse the potential impact of market regulations. The analysis shows that although many people think that the problem of low-quality reporting is at least partially caused by the low entry barrier of the media market, encouraging pluralism (allowing purchases of the established media outlet by new entrants) could be more beneficial for the society than allowing only mergers and acquisitions by the existing players.

Why ignorance could be bliss

As shown in the previous section, the public awareness of newspaper reputation motivates interest groups to purchase good names. We can thus evaluate whether the provision of reputational information is unconditionally beneficial by comparing the public welfare with and without the information.

Without knowing the reputational information, the readers choose newspapers as when at the first period, the probability for them to select a good newspaper is $Pr(G) = \gamma$. After learning the reputational information and the transactions of media ownership settled, the probability for a reader to select a good newspaper

[10]Current Legal and Regulatory Context of Media Plurality, the Parliament of the United Kingdom. https://publications.parliament.uk/pa/ld201314/ldselect/ldcomm/120/12013.htm

becomes $Pr^a(G|S) = \frac{P_G+\phi^a(1-P_G)+\delta^a}{2P_G}$. Comparing the two probabilities, we know that once $P_G(2\gamma-1) > \phi^a(1-P_G)+\delta^a$, the probability of selecting a good newspaper becomes lower after the public is provided with the reputational information. This means, if the media environment is already sufficiently good at the beginning ($\gamma > \frac{1}{2}$), provision of reputational information could lower social welfare unless a good proportion of the good names are purchased by good editorial teams. It is also worth noting that, after provided with the reputational information and the interest groups are motivated to purchase the good names, ignoring media reputation, i.e., selecting a newspaper randomly, is not going to be beneficial to the public, because by doing so, the probability for them to select a good newspaper will be further lower, decreasing from $\frac{P_G+\phi^a(1-P_G)+\delta^a}{2P_G}$ to $\frac{P_G+\phi^a(1-P_G)+\delta^a}{2}$.

Proposition 4.5.1. *Provision of the reputational information decreases the probability for a reader to select a good newspaper if* $P_G(2\gamma-1) > \phi^a(1-P_G)+\delta^a$.

This result suggests that provision of the reputational information about the newspapers could be detrimental to the public, and to avoid this negative outcome, the government needs to implement some regulations to ensure $\phi^a(1-P_G)+\delta^a > P_G(1-\gamma)$. That is, some regulations need to be applied in order to have sufficient good editorial teams successfully purchasing the good names. We first discuss the effects of some entry barriers.

Only purchases by old teams allowed

When only the old players in the market are allowed to purchase the retiring newspapers, the market-clearing condition is:

$$\gamma P_G = \phi^o\gamma(1-P_G)+\psi^o(1-\gamma). \tag{4.2}$$

The posterior beliefs are:

$$\begin{aligned} Pr^o(G|S) &= \frac{\gamma P_G + \phi^o \gamma (1 - P_G)}{2\gamma^o P_G} \\ &= \frac{P_G + \phi^o (1 - P_G)}{2P_G} \end{aligned}$$

and,

$$\begin{aligned} Pr^o(G|N) &= \frac{(1-\phi^o)\gamma(1-P_G) + \gamma}{(1-\phi^o)\gamma(1-P_G) + (1-\phi^o)(1-\gamma) + 1} \\ &= \frac{2\gamma - \gamma P_G - \phi^o \gamma + \phi^o \gamma P_G}{2 - 2\gamma P_G}. \end{aligned}$$

At the beginning of $t = 2$, the price of the reputation of a newspaper with S record is

$$V^o = (Pr^o(G|S) - Pr^o(G|N)) * P_G = \frac{P_G + \phi^o - \phi^o P_G - 2\gamma P_G}{1 - \gamma P_G}.$$

The superscript "o" stands for "old" as only old teams are eligible to buy the names.

Only purchases by new entrants allowed

On the other hand, when only new entrants to the market are allowed to purchase the retiring newspapers, the corresponding market-clearing condition is:

$$\gamma P_G = \delta^n \gamma + \rho^n (1 - \gamma). \tag{4.3}$$

As before, we can calculate the posterior beliefs as follows:

$$\begin{aligned} Pr^n(G|S) &= \frac{\gamma P_G + \delta^n \gamma}{2\gamma P_G} \\ &= \frac{P_G + \delta^n}{2P_G} \end{aligned}$$

and,

$$\begin{aligned} Pr^n(G|N) &= \frac{\gamma(1-P_G) + (1-\delta^n)\gamma}{\gamma(1-P_G) + (1-\gamma) + (1-\delta^n)\gamma + (1-\rho^n)(1-\gamma)} \\ &= \frac{2\gamma - \gamma P_G - \delta^n \gamma}{2 - 2\gamma P_G}. \end{aligned}$$

At the beginning of $t = 2$, the price of the reputation of a newspaper with S record is

$$V^n = (Pr^n(G|S) - Pr^n(G|N)) * P_G = \frac{P_G + \delta^n - 2\gamma P_G}{1 - \gamma P_G}.$$

The superscript "n" stands for "new" as only new entrants are eligible to buy the names.

Social welfare evaluation

Because of the multiplicity of equilibria, as Tadelis (1999: 556) argues, it is almost natural to find some way of selecting the equilibrium in which only good types bought the names. Similarly, it is also possible to select an equilibrium in favor of the bad types. Hence, I evaluate the social welfare by comparing the best and worst equilibria under the three scenarios. The benchmark for comparison is the proportion of the good-type among the newspapers with S record ($2\gamma P_G$). In other words, I consider whether it is possible to achieve market clearing by only the good-type editorial teams or by the bad-type ones in the three cases. When the proportion of the newspapers operated by the good-type editorial teams increases, the people are more likely to make correct decisions based on the information they acquire, and hence the social welfare increases subsequently.

Considering the case without market restriction, the market is cleared by only the good types when $\phi^a(1 - P_G) + \delta^a = P_G$, and by only the bad types when $\psi^a + \rho^a = \frac{\gamma P_G}{1-\gamma}$. As the former condition is feasible in all values of P_G, we know that the market of names can be cleared by only the good types. On the other hands, as both ψ^a and ρ^a are in $[0, 1]$, the bad types can clear the market when $\frac{P_G}{2} < \frac{1-\gamma}{\gamma}$.

Following the same process, we know that when only purchase by old teams is allowed, the market can be cleared by the good types when $P_G < \frac{1}{2}$ and the bad types

when $P_G < \frac{1-\gamma}{\gamma}$. When only purchase by new entrants is allowed the market can always be cleared by the good types and can be cleared only by the bad types when $P_G < \frac{1-\gamma}{\gamma}$.

This analysis shows, as summarized in Table 4.1, that when only purchases by old teams are allowed, there is an undersupply of the good-type editorial teams, so even when we want to apply some sort of selection mechanism to clear the market by only the good types, it is unachievable. This is possible for the other two cases, but a further advantage of the "only purchases by new entrants allowed" case is that there are fewer bad types in the market, so it is more unlikely that all the names are purchased by them. Hence, the worst equilibrium in the third case is better than that in the first case. Thus we can conclude that imposing restriction on mergers and acquisitions may increase the likelihood for achieving preferable outcomes.

Table 4.1: Comparing the market clearing conditions

	By only the good-type	By only the bad-type
No restriction	Always	if $\frac{P_G}{2} < \frac{1-\gamma}{\gamma}$
Only purchases by old teams allowed	if $P_G < \frac{1}{2}$	if $P_G < \frac{1-\gamma}{\gamma}$
Only purchases by new entrants allowed	Always	if $P_G < \frac{1-\gamma}{\gamma}$

Proposition 4.5.2. *If social welfare is evaluated by the the proportion of the good-type among the newspapers with S record ($Pr(G|S)$), allowing only purchases by new entrants yields the best potential outcome.*

4.6 Scrutiny of Ownership Transactions

An appropriate institution design reduces the probability of having an undesirable equilibrium, but a large part of equilibrium-selecting relies on the scrutiny process underlying the media ownership transaction process. If transactions of media owner-

ship are subject to careful scrutiny by an independent and benevolent agency, ideally, the public will be able to enjoy a high quality media environment. Many countries have set up this kind of institutions, such as Federal Communications Commission in the US, but we should realistically expect that it is not possible to avoid all potentially scandalous acquisitions. Overall, the review should focus on avoiding type 2 error: buyers with suspicious background should be barred.

Also the public may overestimate or underestimate the capacity and quality of governmental screening. I denote the former as $\tilde{\delta} > \delta$ or $\tilde{\phi} > \phi$, and the latter $\tilde{\delta} < \delta$ or $\tilde{\phi} < \phi$. When the public overestimate the quality of governmental scrutiny, there will be: 1. more bad-type newspapers than the public believe, and 2. the price of good names will be higher. While the former will be detrimental to social welfare for sure, the latter concerning about the resource available to and the origin of profits for the name buyers, and hence the effect is rather ambiguous. I discuss about the relationship between price and social welfare in Section 4.7.

4.7 Discussion on the Media's Payoff Structure

Price of the newspaper ownership determines the potential buyers of the reputable names. Higher price could select a wrong group of buyers for several reasons. If those with bad intent are more resourceful, higher price would only be affordable for them. The problem can be addressed if governmental scrutiny functions properly, but will be very problematic if it is not.

Moreover, as reputation maps to benefits for newspaper owners, it is important to understand where is the main source of the returns: whether economic or political? In the baseline analysis, it is assumed that $f(W_t^j) = W_t^j$ for simplicity. However, $f(W_t^j)$

can be any strictly increasing function in reputation W_t^j. For example, it can be that for the good-type the return from reputation is $f^G(W_t^j) = f_E^G(W_t^j)$, which is purely from economic benefit. For the bad-type, the return from reputation is $f^B(W_t^j) = f_E^B(W_t^j) + f_P^B(W_t^j)$, the first term being the economic return and the latter political. Economic returns include revenue from newspaper selling and advertising. Political returns can be material or immaterial rent that only owners with own political agenda can enjoy, such as income from governmental advertising, readership from partisan audience (as in Bernhardt et al., 2008), or because newspapers successfully influence readers' decisions. If the rent from reputation consists predominantly of political benefit, buying a name is an effective and essentially a cheap way of affecting the public opinion, and the readers will be likely to subject to political manipulation.[11] If this is the case, only those with access to non-economic rents have incentive to buy the names, and the names are no longer profitable for good-type editorial teams to acquire. Such a case is socially worrisome, and hence maintaining $f_E^G(W_t^j) > f_E^B(W_t^j) + f_P^B(W_t^j)$ is crucial, yet not always possible.

4.8 Considering Moral Hazard

A change in the behaviour of the bad-type editorial teams enables us to analyse the case with both adverse selection and moral hazard. In this extension, as in Tadelis (2002), the bad types can decide whether to exert effort in reporting. The accuracy of their reports is a function of the levels of effort. Formally, the accuracy is $P_B = eP_G$ and the cost of making effort is $c(e)$. It is further assumed that $c'(e) > 0$, $c''(\cdot) > 0$,

[11] For example in Hungary, the son of Mr. Orbans former finance minister purchased leading news website, Origo. According to a New York Times report titled "The Website That Shows How a Free Press Can Die", before the purchase, "Origo had struggled financially. Now it was flush with government advertising revenue, which more than tripled after the sale." https://www.nytimes.com/2018/11/24/world/europe/hungary-viktor-orban-media.html

and $c(0) = 0$.

With this setup, it is clear that the existence of market of reputation can incentivize the retiring bad-type editorial teams to report news with effort. When without the market (because of total market restriction or perfect observability of transactions), the bad types of Generation 0 and Generation 2 will not exert any effort as they have "no future" when making decisions. Hence, at $t = 1$, only the bad types of Generation 1 have incentive to exert effort, and the expected value of information is $W_1 = [\gamma + \frac{1}{2}(1-\gamma)e]P_G$. On the other hand, if there is a market of names, all of the bad types will exert effort and thus the expected value of information is $W_1 = [\gamma + (1-\gamma)e]P_G$.

When allowing trading of names, more editorial teams would exert effort than when there is no market, but the equilibrium level of effort may decrease as more teams are willing to exert effort, and hence it is now harder to differentiate types of teams by their names. Therefore, the social welfare evaluation depends on the specific cost function chosen (Tadelis, 2002: 869). Overall we know that, unlike in the adverse selection model, banning the transaction of names could be socially detrimental.

Proposition 4.8.1. *When considering the problem of moral hazard, banning all transactions in the market of reputation may not be socially optimal as when concerning only about the problem of adverse selection.*

4.9 Concluding Remarks

Nowadays, many argue that the problem of low-quality news reporting can be addressed by providing the public with the reputational information regarding the media. In this paper, I study a stylised model to evaluate the impact of reputation on

media power and discuss the social welfare implications of different types of market restrictions and scrutiny. This framework points out that, instead of being a cure to the problem of low quality reporting, media reputation itself can be a source of concern about political manipulation. The model is particularly suitable for analysing an economy consisting of many media outlets and a significant proportion of them is operated by new entrants. The analysis indicates that when the readers know the reputation of newspapers, a market of names (reputation) is created naturally. When the market faces only adverse selection problem, forbidding all transactions of names is the first-best option. If that is not feasible, allowing only mergers and acquisitions by new entrants could deliver most preferable social welfare outcomes than no restriction or allowing only mergers and acquisitions by old teams. If moral hazard is involved, banning all transactions may be no longer socially optimal as the premium from a name with good reputation can incentivise the newspapers to produce high-quality reports. Overall, I argue that reputation information can make some newspapers more influential than the others and this leads to a need for institutional and governmental interventions on transactions of media ownership.

There are other related issues not discussed in this paper worth future investigation. For example, we need to consider what means "high-quality" journalism. In the model of this paper, it is *ex post* clear that the bad-type media outlets will always report low-quality news, but in reality they can employ more subtle reporting strategies to manipulate news and can preserve their to reputation for a longer period. Such strategies include focusing on national or global news instead of local news (George and Waldfogel, 2006) or disguising biased reporting in the name of balance reporting (Sharpiro, 2016). These strategies lead to public ignorance on certain issues, and

further complicate the relationship between reputation and good journalism. Also, it is clear in the analysis that governmental scrutiny of media ownership transactions could be helpful in ruling out purchase proposals by the bad teams. However, if the interest groups apply a strategy short than direct ownership to compromise the editorial teams, for example the Chinese government pays internationally respected newspapers such as *Daily Telegraph* in the UK to include propaganda supplements in their publications,[12] then it is hard for democratic governments to intervene this practice due to the promise of free press. How to address this issue remains a critical question for future research.

[12]Inside China's audacious global propaganda campaign. https://www.theguardian.com/news/2018/dec/07/china-plan-for-global-media-dominance-propaganda-xi-jinping

Bibliography

[1] Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31 (2): 211-236.

[2] Anderson, Simon P., and John McLaren. 2012. "Media Mergers and Media Bias with Rational Consumers." *Journal of the European Economic Association* 10 (4): 831-859.

[3] Ashworth, Scott, and Kenneth W. Shotts. 2010. "Does Informative Media Commentary Reduce Politicians' Incentives to Pander?." *Journal of Public Economics* 94 (11-12): 838-847.

[4] Baker, C. Edwin. 2007. *Media Concentration and Democracy: Why Ownership Matters.* Cambridge: Cambridge University Press.

[5] Baron, David. 2006. "Persistent Media Bias." *Journal of Public Economics* 90 (1-2): 1-36.

[6] Bernhardt, Dan, Stefan Krasa, and Mattias Polborn. 2008. "Political Polarization and The Electoral Effects of Media Bias." *Journal of Public Economics* 92: 1092-1104.

[7] Besley, Timothy, and Andrea Prat. 2006. "Handcuffs for the Grabbing Hand?: Media Capture and Government Accountability." *American Economic Review* 96 (3): 720-736.

[8] Downing, John D. H.. 2011. "Media Ownership, Concentration, and Control: The Evolution of Debate." *The Handbook of Political Economy of Communications* Ed. Janet Wasko, Graham Murdock, and Helena Sousa. Oxford: Blackwell Publishing Ltd.

[9] Duggan, John, and Cesar Martinelli. 2011. "A Spatial Theory of Media Slant and Voter Choice." *The Review of Economic Studies* 78 (2): 640-666.

[10] 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.

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

[12] George, Lisa M., and Joel Waldfogel. 2006. "The "New York Times" and the Market for Local Newspapers." *The American Economic Review* 96(1): 435-447.

[13] Gentzkow, Matthew, and Jesse M. Shapiro. 2006. "Media Bias and Reputation." *Journal of Political Economy* 114 (2): 280-316.

[14] Kennedy, Patrick and Andrea Prat. 2018. "Where Do People Get Their News?." *Columbia Business School Research Paper* No. 17-65.

[15] Kreps, David M., and Robert Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27 (2): 253-279.

[16] Lim, Chloe. 2018. "Checking How Fact-checkers Check." *Research and Politics* July- September: 1-8.

[17] Martin, Gregory J., and Josh McCrain. 2019. "Local News and National Politics." *American Political Science Review* 113 (2): 372- 384.

[18] Ngok, Ma. 2007. "State–Press Relationship in Post-1997 Hong Kong: Constant Negotiation amidst Self-Restraint." *The China Quarterly* 192 (December): 949-970.

[19] Nyhan, Brendan, and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32: 303-330.

[20] Pennycook, Gordon, and David G. Rand. 2018. "Lazy, Not Biased: Susceptibility to Partisan Fake News Is Better Explained by Lack of Reasoning Than by Motivated Reasoning." *Cognition* forthcoming.

[21] Prat, Andrea. 2015. "Media Capture and Media Power." In Simon P. Anderson, Joel Waldfogel, and David Stromberg (Eds.). *Handbook of Media Economics Vol. 1b*. North Holland.

[22] Prat, Andrea. 2017. "Media Power." *Journal of Political Economy* forthcoming.

[23] Shapiro, Jesse M.. 2014. "Special Interests and the Media: Theory and An Application to Climate Change." *Journal of Public Economics* 144: 91-108.

[24] Sheen, Greg Chih-Hsin. 2019. "Media with Reputational Concerns: Yes Men or Watchdogs?." *Working Paper*

[25] Sobbrio, Francesco. 2014. "Citzen-Editors' Endogenous Information Acquisition and News Accuracy." *Journal of Public Economics* 113: 43-53.

[26] Tadelis, Steven. 1999. "What's in a Name? Reputation as a Tradeable Asset." *American Economic Review* 89 (3): 548-563.

[27] Tadelis, Steven. 2002. "The Market for Reputations as an Incentive Mechanism." *Journal of Political Economy* 110 (4): 854-882.

[28] Wolton, Stephane. 2019. "Are Biased Media Bad for Democracy?." *American Journal of Political Science* forthcoming.

Appendix A

Appendix for Chapter 2

A.1 Proofs

Proof of Lemma 2.4.1

The lemma is proved directly from Bayes' Rule. Given $\overline{\gamma} > \underline{\gamma}$ and the media outlet reports truthfully, $Pr(\tau^G = h|r = x) \geq p \geq Pr(\tau^G = h|r \neq x)$ always holds. $Pr(\tau^G = h|r = x) \geq p \geq Pr(\tau^G = h|r \neq x)$ only occurs, given $\overline{\gamma} > \underline{\gamma}$, when the media outlet reports truthfully. When solving for $(1) \geq p$ and $p \geq (2)$, $\overline{\gamma}$ is cancelled out everywhere. Thus $Pr(\tau^G = h|r = x) \geq p \geq Pr(\tau^G = h|r \neq x)$ is independent of the value of $\overline{\gamma}$. As weakly dominated strategies are ruled out, the only strategy that can provide informative report to the voter is the truthful reporting strategy.

Proof of Proposition 2.4.1

For the truthful equilibrium to sustain, the condition is that when $s_J^M = x$, $Pr(\tau^M = h|r = x) \geq Pr(\tau^M = h|r \neq x)$, and when $s_J^M \neq x$, $Pr(\tau^M = h|r \neq x) \geq Pr(\tau^M = h|r = x)$. That is, when the media receiving signal $s_J^M = x$, there is no profitable deviation for it to report $r \neq x$, and vice versa. Thus the necessary and sufficient condition for the existence of a truthful equilibrium is:

$$\frac{q(1-\overline{\mu}E(\gamma))}{q(1-\overline{\mu}E(\gamma))+(1-q)(1-\underline{\mu}E(\gamma))} = \frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}. \tag{A.1}$$

Simple algebra yields that this equation can hold only if $q = 0$, $q = 1$ or $\overline{\mu} = \underline{\mu}$.

When these three situations do not occur, it is always profitable for M to report $r = x$ as

$$\frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}} > \frac{q(1-\overline{\mu}E(\gamma))}{q(1-\overline{\mu}E(\gamma)) + (1-q)(1-\underline{\mu}E(\gamma))}.$$

This result means being "yes man" is a weakly dominant strategy for the media outlet. Hence even the competent media outlet does not report truthfully.

Proof of Proposition 2.4.2

To analyze the general case, we need to find the posteriors corresponding to different actions when $\rho = 1$ and when $\rho = 0$. We find the posteriors when $\rho = 1$ with the similar procedure as in the previous case. Recall that $s_J^M = x$ only when $x = \omega$ because the probability that the media outlet and the government both receive the same wrong signal is 0 (Caillaud and Tirole, 2002). In other words, the case $s_V = \omega \neq \{s_J^M = x\}$ is ruled out. That means $Pr(s_V = s_J^M = x = \omega | s_J^M = x) = 1$.

Suppose truthful reporting, then the posterior for V about M after observing $r = x$ is again

$$Pr(\tau^M = h | r = x) = \frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}}. \tag{A.2}$$

When the media outlet receives $s_J^M = x$, it has no incentive to deviate away from reporting $r = x$, so to check the existence of a truthful equilibrium, we only need to consider whether, when the media outlet receiving $s_J^M \neq x$, there is a profitable deviation from reporting $r \neq x$ to $r = x$.

Suppose the actors follow truthful equilibrium, and the probability ρ that the voter receives a private signal $s_V = \omega$ is 1, when $r \neq x$, there are three possible relations of $\{\omega, x, s_J^M, s_V\}$ could happen: $s_V = \omega = s_J^M \neq x$, $s_V = \omega = x \neq s_J^M$, and

$s_V = \omega \neq r \neq x$.

The posterior beliefs about the type of M after observing each of the above relations are as follows:

1. If the media report is correct while the government decision is wrong:

$$Pr(\tau^M = h | s_V = \omega = s_J^M \neq x) = \frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}};$$

2. If the government decision is correct and the media report is wrong:

$$Pr(\tau^M = h | s_V = \omega = x \neq s_J^M) = \frac{q(1-\overline{\mu})}{q(1-\overline{\mu}) + (1-q)(1-\underline{\mu})};$$

3. If both the government decision and the media report are wrong:

$$Pr(\tau^M = h | s_V = \omega \neq s_J^M \neq x) = \frac{q(1-\overline{\mu})}{q(1-\overline{\mu}) + (1-q)(1-\underline{\mu})}.$$

For a competent media outlet, the probability for the first relation to occur is $(1 - E(\gamma))\overline{\mu}$; for the second relation to occur is $E(\gamma)(1-\overline{\mu})$; for the third relation to occur is $(1 - E(\gamma))(1-\overline{\mu})$.

Thus the posterior belief $Pr(\tau^M = h | r \neq x)$ supposing truthful equilibrium is:

$$(1-E(\gamma))\overline{\mu}(\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}) + E(\gamma)(1-\overline{\mu})(\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}) + (1-E(\gamma))(1-\overline{\mu})(\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}). \quad \text{(A.3)}$$

Recall that if the media outlet is caught cheating, the voter believes it is incompetent ($Pr(\tau^M = h | s_V = \omega \neq (r = x)) = 0$), thus when the competent media deviates from truthful reporting, the posterior belief $Pr(\tau^M = h | r = x)$ is:

$$[(1-E(\gamma))\overline{\mu}]0 + E(\gamma)(1-\overline{\mu})(\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}) + [(1-E(\gamma))(1-\overline{\mu})]0. \quad \text{(A.4)}$$

Solving for $(A.3) \geq (A.4)$, we have the inequality (A.5) as the condition for the competent media to adopt the truth-reporting strategy:

$$(1-\overline{\mu})(\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}) \geq (E(\gamma)-\overline{\mu})(\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}). \tag{A.5}$$

Given $\overline{\mu} \geq E(\gamma)$, the LHS of (A.5) is positive; while the RHS of (A.5) is negative, the condition always holds.

Thus we can conclude that a competent media outlet has no incentive to deviate from truth-reporting when $\rho = 1$.

However an incompetent media outlet has incentive to deviate from truth-reporting when $\rho = 1$:

The probabilities for $s_V = \omega = s_J^M \neq x$, $s_V = \omega = x \neq s_J^M$, and $s_V = \omega \neq r \neq x$ to occur are $(1-E(\gamma))\underline{\mu}$; $E(\gamma)(1-\underline{\mu})$; $(1-E(\gamma))(1-\underline{\mu})$. Thus the parallel condition to (A.5) is:

$$(1-\underline{\mu})(\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}) \geq (E(\gamma)-\underline{\mu})(\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}). \tag{A.6}$$

Because the RHS of (A.6) is non-negative, the condition for the incompetent media to not deviate from the truthful equilibrium is more complicated than that for its competent counterpart.

Rearranging (A.6) and combining it with the condition $\overline{\mu} \geq E(\gamma) \geq \underline{\mu}$, we have the condition (A.7) for the incompetent media to follow the truthful equilibrium:

$$min\{\overline{\mu}, \frac{q(\overline{\mu}-\overline{\mu}^2)+(1-q)(\underline{\mu}-\underline{\mu}^2)}{\overline{\mu}(1-\underline{\mu})+q(\overline{\mu}(\underline{\mu}-\overline{\mu}))}\} \geq E(\gamma) \geq \underline{\mu}. \tag{A.7}$$

When $\rho \in (0,1)$, the condition for a competent media outlet to not deviate from the truthful equilibrium is:

$$\rho[(A.3)] + (1-\rho)\frac{q(1-\overline{\mu}E(\gamma))}{q(1-\overline{\mu}E(\gamma))+(1-q)(1-\underline{\mu}E(\gamma))} \geq \rho[(A.4)] + (1-\rho)\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}. \tag{A.8}$$

Rearranging (A.8), we have the condition (2.5) in the main text.

Following the same procedure, we have the condition (2.6) in the main text for the incompetent type.

Proof of Proposition 2.5.1

Two Media Outlets 1. When $\rho = 0$: Simple algebra shows that $\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$ is greater or equal to both Ψ_2 and $\frac{q(1-\overline{\mu})}{q(1-\overline{\mu})+(1-q)(1-\underline{\mu})}$ when $\overline{\mu} \geq \underline{\mu}$, and it is always true by definition, unless $q = 0$, $q = 1$ or $\overline{\mu} = \underline{\mu}$.

2. When $\rho = 1$: For a competent media outlet, suppose a truthful equilibrium, the condition for it to not deviate is $A + B \geq C$, where,

$$A = [(1 - E(\gamma))E(\mu)\overline{\mu} + E(\gamma)[1 - E(\mu)]\overline{\mu} + [1 - E(\gamma)][1 - E(\mu)]\overline{\mu}]\frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}}$$

$$B = [E(\gamma)E(\mu)(1 - \overline{\mu}) + [1 - E(\gamma)][1 - E(\mu)](1 - \overline{\mu}) + E(\gamma)[1 - E(\mu)](1 - \overline{\mu})$$
$$+ [1 - E(\gamma)]E(\mu)(1 - \overline{\mu})]\frac{q(1-\overline{\mu})}{q(1-\overline{\mu}) + (1-q)(1-\underline{\mu})}$$

$$C = E(\gamma)(1 - \overline{\mu})\frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}}.$$

With some algebra, the condition becomes

$$(1 - \overline{\mu})^2[q\overline{\mu} + (1-q)\underline{\mu}] \geq [E(\gamma)[1 - \overline{\mu} + E(\mu)\overline{\mu}] - \overline{\mu}]\overline{\mu}[q(1-\overline{\mu}) + (1-q)(1-\underline{\mu})],$$

because $E(\gamma)[1 - \overline{\mu} + E(\mu)\overline{\mu}] - \overline{\mu}$ on the RHS is negative whenever $\overline{\mu} \geq E(\gamma)$, the condition holds. Hence a competent media outlet has no incentive to deviate from truth-reporting when $\rho = 1$.

For an incompetent media outlet, the parallel condition for it to follow a truthful

equilibrium is

$$[\underline{\mu} - E(\gamma)E(\mu)\underline{\mu}]\frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}} + (1-\underline{\mu})\frac{q(1-\overline{\mu})}{q(1-\overline{\mu}) + (1-q)(1-\underline{\mu})} \geq E(\gamma)(1-\underline{\mu})\frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}}.$$

The condition can be simplified as

$$(1-\underline{\mu})\frac{q(1-\overline{\mu})}{q(1-\overline{\mu}) + (1-q)(1-\underline{\mu})} \geq [E(\gamma)[1-\underline{\mu} + E(\mu)\underline{\mu}] - \underline{\mu}]\frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}}.$$

3. The conditions for the truthful equilibrium when $\rho \in (0,1)$ are the convex combination of the above two cases, and thus we can conclude that a truthful equilibrium does not always exist.

3 or More Media Outlets 1. $\frac{q\overline{\mu}}{q\overline{\mu}+(1-q)\underline{\mu}}$ is always greater or equal to Ψ_n because $\overline{\mu} \geq \underline{\mu}$.

2. When $\rho = 0$: Consider $n \to \infty$, an incompetent media outlet will deviate from truthful reporting if

$$E(\gamma)(1-\underline{\mu})\left(\frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}} - \frac{q(1-\overline{\mu})}{q(1-\overline{\mu}) + (1-q)(1-\underline{\mu})}\right) - [1-E(\gamma)]\mu\frac{q\overline{\mu}}{q\overline{\mu} + (1-q)\underline{\mu}} - [1-E(\gamma)](1-\underline{\mu})\frac{q(1-\overline{\mu})}{q(1-\overline{\mu}) + (1-q)(1-\underline{\mu})} \geq 0. \tag{A.9}$$

3. Simplify (A.9), we have the condition

$$[E(\gamma) - \underline{\mu}]\overline{\mu} - (1-\overline{\mu})E(\mu) \geq 0$$

for an incompetent media outlet to deviate from truthful reporting.

4. When $\rho = 1$ and $n \to \infty$, the condition for an incompetent media outlet to deviate from truthful reporting is again (A.9) and thus the two cases are equivalent.

Appendix B

Appendix for Chapter 3

B.1 Proofs

All proofs that are not derived directly from the main text are presented here.

Proof of Corollary 3.4.1

Proof. The information value when the society is unstable is derived by comparing the expected payoffs with and without truthful reporting. That is,

$$Pr(m = b|m(s) = s)(-D_b) - Pr(m = g|m(s) = s)(1-\beta)\mu_g^A D_g - (-D_g) =$$

$$D_g - (1 - p - q + 2pq)D_b - (p + q - 2pq)(\beta\mu_g^A + (1-\beta))D_g$$

The difference in information values when the society is unstable and stable is $V_u - V =$

$$(1 - (p+q-2pq)(\beta\mu_g^A + (1-\beta)) - p)D_g - (1-p-q+2pq)(1 - (\mu_b^A + (1-\mu_b^A)\beta))D_b.$$

We first check whether $(1 - (p+q-2pq)(\beta\mu_g^A + (1-\beta)) - p) > 0$. Suppose

$$(1 - (p+q-2pq)(\beta\mu_g^A + (1-\beta)) - p) < 0 \Rightarrow$$

$$p > \frac{1+q+\beta q}{2-2q+\beta q}.$$

Because $\frac{1+q+\beta q}{2-2q+\beta q} > 1$ when $q > \frac{1}{3}$ and given $q > \frac{1}{2}$ by assumption, by contradiction, we always have $(1 - (p+q-2pq)(\beta\mu_g^A + (1-\beta)) - p) > 0$.

Thus $V_u - V > 0$ when

$$D_g > \frac{(1-p-q+2pq)(1-(\mu_b^A+(1-\mu_b^A)\beta))D_b}{(1-(p+q-2pq)(\beta\mu_g^A+(1-\beta))-p)} \Rightarrow$$
$$D_g > \frac{(1-q)(1-p)(1+\beta)D_b}{(1+q+\beta q)-(2-2q+\beta q)p}.$$

Note that because $q \in (\frac{1}{2}, 1]$, $p \in [0,1]$, $\beta \in [0,1)$

$$[(1+q+\beta q)-(2-2q+\beta q)p]-[(1-q)(1-p)(1+\beta)] =$$
$$(2q-p)+(2q\beta-\beta)+(\beta+q-2\beta q)p > 0 \Rightarrow$$
$$D_b > \frac{(1-q)(1-p)(1+\beta)D_b}{(1+q+\beta q)-(2-2q+\beta q)p}$$

and by definition $D_g > D_b$, the condition $D_g > \frac{(1-q)(1-p)(1+\beta)D_b}{(1+q+\beta q)-(2-2q+\beta q)p}$ always holds. As such we prove that the dictator is more willing to commit to no censorship when the society is unstable than when it is stable. □

B.2 Private Signal

Suppose apart from the news report, the dictator receives a private signal about the state of the world from his viziers or internal media (see Egorov and Sonin, 2011, Dimitrov, 2017). The dictator then decides whether to censor a negative news report. Similar to q for the media signal, we set the precision of the dictator's private signal $\phi \in (\frac{1}{2}, 1]$, meaning that $Pr(v = g|\omega = G) = Pr(v = b|\omega = B) = \phi$. For the public, the posteriors are still the same, but the dictator now has better knowledge about the state of the world even without the information from the media. We denote the posteriors for the dictator as $Pr(\omega = B|v, m) = \mu_{vm}^A, v \in \{b, g\}, m \in \{b, g\}$.

If the media outlet reports truthfully, the posteriors for the government, the posteriors for the public and the prior will have the following relations: 1). When the quality of the dictator's signal is higher than the media outlet's signal ($\phi > q$),

$\mu^A_{bb} > \mu^A_b > \mu^A_{bg} > \mu^A_{gb} > \mu^A_g > \mu^A_{gg}$; 2). when the quality of the dictator's signal is lower than the media outlet's signal ($\phi < q$), $\mu^A_{bb} > \mu^A_b > \mu^A_{gb} > \mu^A_{bg} > \mu^A_g > \mu^A_{gg}$. We denote the damage for the dictator depending on his information as $-D_{vm}$. As before, based on his posterior beliefs about the state of the world, the relation of different damage levels (and also comparing to $-D_g$ and $-D_b$) is as follows:

$$\begin{cases} -D_{gg} < -D_g < -D_{gb} < -D_{bg} < -D_b < -D_{bb} < 0 \text{ if } \phi > q, \\ -D_{gg} < -D_g < -D_{bg} < -D_{gb} < -D_b < -D_{bb} < 0 \text{ if } \phi < q. \end{cases}$$

The expected payoff matrix for the dictator is Table B.1.

Table B.1: U_A when $m = b$

	$c = 0$	$c = 1$
$v = g$	$-\mu^A_{gb} D_{gb} - (1 - \mu^A_{gb})\beta D_{gb}$	$-\mu^A_{gb} D_{gb} - k$
$v = b$	$-\mu^A_{bb} D_{bb} - (1 - \mu^A_{bb})\beta D_{bb}$	$-\mu^A_{bb} D_{bb} - k$

When $v = g$ and $m = b$, the commitment constraint for no censorship is:

$$(1 - \mu^A_{gb})\beta D_{gb} < k. \tag{B.1}$$

When $v = b$ and $m = b$, the commitment constraint for no censorship is:

$$(1 - \mu^A_{bb})\beta D_{bb} < k. \tag{B.2}$$

Given $\mu^A_b > \mu^A_{gb}$ and $D_{gb} > D_b$, Inequality (B.1) is more difficult to be satisfied than Inequality (3.2). Given $\mu^A_{bb} > \mu^A_b$ and $D_b > D_{bb}$, Inequality (B.2) is easier to be satisfied than Inequality (3.2). Given $v = g$, the probability that $m = b$ is $Pr(m = b|v = g) = \phi + pq - p\phi - q\phi$, and given $v = b$, the probability that $m = b$ is $Pr(m = b|v = b) = 1 - q - \phi - p + q\phi + pq + p\phi$. The *ex ante* commitment constraint is a weighted average of Inequalities (B.1) and (B.2) with the weight of $Pr(s = b|v = g)$ and $Pr(s = b|v = b)$ respectively.

Hence, when $1-q-p<0$, having private information can ease the dictator's commitment problem as the *ex ante* commitment constraint is more lenient. On the other hand, when $1-q-p>0$, having private signal deteriorates the commitment problem. This is because when the bad state of the world is likely to be realized, censoring the media report is not very helpful, and the truthful information from the media can help the dictator to prepare for the challenge. While if the media report is likely to be a false alarm, the dictator with his private information is more likely to censor the negative report. In other words, the dictator can choose to avoid the immediate damage from the negative report by censorship and hopes that the report turns out to be incorrect. When ϕ is approaching to 0.5, these two cases are almost equally likely to occur, and thus the strength of commitment problem faced by the dictator is similar to that of the baseline result.

Proposition B.2.1. *Having private information can ease the dictator's commitment problem when* $1-q-p<0$. *When* $1-q-p>0$, *having private signal deteriorates the commitment problem. When* $\phi \to 0.5$, *these two cases are equally likely to occur.*

Unless the dictator's private signal is perfect ($\phi=1$), information from the media is always valuable to the dictator.[1] However, dictators are always under pressure not only from outside but also from inside of their government. Maintaining powerful secret services is not only costly but also risky. For example, the South Korean dictator Park Chung-hee was assassinated at Korean Central Intelligence Agency by his intelligence chief, Kim Jae-kyu, in 1979. Thus according to Egorov and Sonin (2011), the

[1]Some scholars, e.g., Dimitrov (2017), contend that the dictator's possession of an effective internal information channel relegates information from public media to redundancy. This argument is only true when the internal channel provides perfect information. Formally, when $\phi=1$ and $v=b$, μ^A_{bb} and μ^A_{bg} are both equal to 1, and thus learning information from the media cannot further decrease damage, the dictator is indifferent from censoring the media or not. Whenever $\phi<1$, information from public media is valuable as the dictator's posterior belief is more accurate with one more piece of information, as long as it is truthfully reported.

dictators tend to employ mediocre viziers to avoid potential betrayal, and this could mean that the quality of the private signal for the dictator should be low (ϕ is close to 0.5). Moreover, even when the sources of private signals are loyal to the dictators, it is still possible that the dictator cannot receive truthful information. An extreme example is that the long-lived Portuguese dictator António de Oliveira Salazar was "died without knowing he had been replaced as Premier."[2] As such, once the actors have to make the strategic decisions as depicted in this paper, both qualitatively and quantitatively, the baseline result is robust when taking dictators' private information into account.

Another reason that a dictator cannot solely rely on his private information channel is that when he needs to convince the public that the state of the world is good (when the society is unstable), he has to rely on media outlets that the public deem as credible to convey this message. Even when the private signal can provide perfect information to the dictator, that signal cannot alter the posterior of the public because state-owned media cannot commit to report truthfully.[3] Thus there is a demand for truthful reporting from the dictator and he inevitably needs to face the commitment problem suggested by this paper.

Proof of Proposition 5. Inequalities (B.1) and (B.2) are derived in the main text. It is also shown in the main text that Inequality (B.1) is more difficult to hold than Inequality (3.2), and Inequality (B.2) is easier to hold than Inequality (3.2).

[2] "Salazar, Ruler of Portugal for 40 Years, Dies," New York Times, July 28, 1970. Retrieved from `http://www.nytimes.com/1970/07/28/archives/salazar-ruler-of-portugal-for-40-years-dies-salazar-dies-in-lisbon.html`.

[3] Assume that the preference of the state-owned media is aligned with the dictator. Once the state-owned media are believed to have reported information truthfully, they can exploit this trust to mislead the public in order to benefit the regime. Gehlbach and Sonin (2014) investigate how the dictator can optimally manipulate the public by an editorial policy that provides a certain amount of negative information to make the state-controlled media trustworthy, but they do not discuss how the dictator can adhere to this editorial policy.

Suppose the media outlet reporting truthfully,

$$\begin{aligned} Pr(m=b|v=b) &= Pr(\omega=B)Pr(v=b|\omega=B)Pr(m=b|\omega=B) \\ &\quad + Pr(\omega=G)Pr(v=b|\omega=G)Pr(m=b|\omega=G) \\ &= pq\phi + (1-p)(1-q)(1-\phi) \\ &= 1-q-\phi-p+q\phi+pq+p\phi; \end{aligned}$$

whereas,

$$\begin{aligned} Pr(m=b|v=g) &= Pr(\omega=B)Pr(v=g|\omega=B)Pr(m=b|\omega=B) \\ &\quad + Pr(\omega=G)Pr(v=g|\omega=G)Pr(m=b|\omega=G) \\ &= p(1-\phi)q + (1-p)\phi(1-q) \\ &= \phi + pq - p\phi - q\phi. \end{aligned}$$

Hence when $1-q-p-2\phi+2q\phi+2p\phi = (1-q-p)(1-2\phi) > 0$, the dictator's commitment constraint is more likely to be Inequality (B.2). Given $\phi > 0.5$, the commitment problem is less severe when $1-q-p < 0$. When $1-q-p > 0$, having private signal makes the commitment problem more severe. When $\phi \to 0.5$, $(1-q-p)(1-2\phi) \to 0$, and thus these two cases are equally likely to occur. Given that $(1-\mu^A_{gb})\beta D_{gb} > (1-\mu^A_b)\beta D_b > (1-\mu^A_{bb})\beta D_{bb}$, the realized commitment constraint, which is the average of Inequalities (B.1) and (B.2) with equal weight, is close to Inequality (3.2). □

B.3 Selection Bias

The potential sample selection bias thus created can be readily captured by the following statistical model (Davidson and MacKinnon, 2004). To be consistent with the

notations above, denote the government's censorship decision over message i by c_i and whether it is censored or not by m_i (an indicator variable).[4] They are related to two latent variables, c_i° and m_i° as follows:

$$c_i = \begin{cases} c_i^\circ & \textbf{if } m_i^\circ > 0; \\ \textbf{unobserved/self-censored} & \textbf{Otherwise}. \end{cases} \tag{B.3}$$

$$m_i = \begin{cases} 1 \textbf{ (truthful report)} & \textbf{if } m_i^\circ > 0; \\ 0 \textbf{ (self-censorship)} & \textbf{Otherwise}. \end{cases} \tag{B.4}$$

According to our theory, one's self-censorship decision hinges critically on the dictator's commitment ability to no censorship. As a result, c_i° and m_i° should be correlated, and therefore we specify a joint bivariate process for their generation:

$$\begin{bmatrix} c_i^\circ \\ m_i^\circ \end{bmatrix} = \begin{bmatrix} \boldsymbol{X_i\gamma} \\ \boldsymbol{W_i\psi} \end{bmatrix} + \begin{bmatrix} u_i \\ v_i \end{bmatrix}, \qquad \begin{bmatrix} u_i \\ v_i \end{bmatrix} \sim \mathcal{NID}\left(0, \begin{bmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{bmatrix}\right) \tag{B.5}$$

where c_i° measures the government's censorship decision when a full sample—i.e., including those self-censored messages—is available, and m_i° captures an individual citizen's (unobservable) decision to self-censor him/herself. The bivariate data-generating process specified above allows their random components to be correlated and normally and independently distributed ($\mathcal{NID}$) with standard deviation σ and correlation ρ.[5]

$\boldsymbol{X_i}$ is a vector of exogenous variables that determine the government's decisions over censorship. In King et al.'s (2013) language, the vector must include a major variable measuring whether a message has the collective action potential or not. However, as we have explained—and as King et al. (2013) also admit in the paragraph quoted above— since the sample is truncated by individuals' self-censorship (when $c_i = c_i^\circ$), estimating γ without taking m_i° into account will definitely lead to biased results.

[4] Let's suppose both are observable here for the exposition purposes. Moreover, c_i can either be an indicator variable like m_i, or take a more general form of a continuous variable measuring the strength of censorship (e.g., the duration of censorship, or the number of messages deleted).

[5] The variance of v_i is set to 1 since only the sign of m_i° can be observed.

Through this simple model of sample selectivity, our contributions can be easily explicated. First of all, the comparative statics derived from our theoretical model can make it relatively easy to specify $\boldsymbol{W_i}$ for the determinants of one's self-censorship behavior since the commitment constraint shown in our model can all be viewed as issue and media outlet specific. Second, from a Maximum Likelihood Estimation perspective, this simple statistical model also allows us to derive a full log-likelihood function.[6]

$$\sum_{m_i=0} \log \Phi(-\boldsymbol{W_i}\boldsymbol{\psi}) + \sum_{m_i=1} \log \left[\frac{1}{\sigma}\phi(\frac{c_i - \boldsymbol{X_i}\boldsymbol{\gamma}}{\sigma})\right] + \sum_{m_i=1} \log \Phi \left[\frac{\boldsymbol{W_i}\boldsymbol{\psi} + \frac{\rho(c_i - \boldsymbol{X_i}\boldsymbol{\gamma})}{\sigma}}{\sqrt{1-\rho^2}}\right] \quad \text{(B.7)}$$

For the first two terms, they are simply the log-likelihood functions for a probit (regarding the part choosing to self-censor themselves) and an OLS (for testing King et al.'s (2013) hypothesis) models respectively. When $\rho = 0$ (i.e., the media outlet's self-censorship decision is not correlated with the government's censorship decision), γ can be estimated independently of the estimation of $\boldsymbol{\psi}$, and the result in King et al. (2013) is therefore unbiased even if the information about self-censorship is entirely missing. By contrast, if $\rho \neq 0$ (according to our theory), the presence of the third term implies that γ can only be estimated without bias when we know the joint distribution of the two choices—i.e., the government censorship and self-censorship. Hence, since all the key parameters in our model such as β (correlating positively with the penetration of media outlets), q (the quality of media outlets), and α (the perceived probability that a report would be disliked by the dictator) are all observables, self-censorship decision m_i can be modeled as a linear combinations of them in empirical implementation. This step then lays the very foundation for estimating consistently

[6]Based on the following expression for how much each observation contributes to the likelihood function:

$$I(m_i = 0)P(m_i = 0) + I(m_i = 1)P(m_i = 1)f(c_i^{\circ}|m_i = 1) \quad \text{(B.6)}$$

both parameter vectors, γ and $\boldsymbol{\psi}$.

Moreover, when we relax the assumption that m_i is observable, more data-driven approaches to address the unobservability of m_i in estimating γ can be adopted. For example, applying the often-used EM algorithm (Little and Rubin, 2002) to such a typical missing data problem will require empirical researchers to evaluate the probability distribution of a relevant latent variable when taking the expectation of the complete-data log-likelihood (the E step), given the observed data and an initial guess of the estimand of interest, γ^{old}—i.e., $P(m_i^{\circ}|\mathbf{X}, \mathbf{W}, \gamma^{old})$. Specifying such a conditional probability is often a challenge in implementing an EM algorithm, but the parameters derived from the comparative statics in our model let empirical researchers of censorship solve it easily by leveraging the variations in $\mathbf{W}$. [7]

B.4 EM Algorithm

In reality, one's self-censorship decision, m_i, is unobservable, but this doesn't mean our theoretical predictions have no way to help empirical researchers better estimate how messages' collective action potential affects the dictator's censorship decision (i.e., γ). The unobservability of m_i constitutes a typical missing data problem for estimating γ, and our contribution can be readily highlighted by applying the EM algorithm to the problem.

Assume that after the n^{th} iteration the current estimate for γ is given by γ_n. Since the objective is to maximize the log-likelihood function, $L(\gamma)$, we need to compute an updated estimate of γ such that ,

$$L(\gamma) > L(\gamma_n) \tag{B.8}$$

[7] Please see appendix B.4 for the details about how the EM algorithm actually works to estimate γ when the data on self-censorship are missing.

Equivalently, we want to maximize the following difference,

$$L(\gamma) - L(\gamma_n) = \ln P(\mathbf{Y}|\gamma) - \ln P(\mathbf{Y}|\gamma_n) \tag{B.9}$$

where $\mathbf{Y}$ denotes a set of complete data. It comprises of a set of observables , $\mathbf{Y} = (\mathbf{X}, \mathbf{W})$ and a hidden random vector of self-censorship m_i°. The total probability $P(\mathbf{Y}|\gamma)$ may be written in terms of the hidden variable m_i° as,

$$P(\mathbf{Y}|\gamma) = \sum_m P(\mathbf{Y}|m_i^\circ, \gamma) P(m_i^\circ|\gamma) \tag{B.10}$$

Plugging (B.10) into (B.9) and applying Jensen's Inequality yield:

$$\begin{aligned} L(\gamma) - L(\gamma_n) &= \ln \left[\sum_m P(\mathbf{Y}|m_i^\circ, \gamma) P(m_i^\circ|\gamma) \right] - \ln P(\mathbf{Y}|\gamma_n) \\ &= \sum_m \left\{ P(m_i^\circ|Y, \gamma_n) \ln \left[\frac{P(\mathbf{Y}|m_i^\circ, \gamma) P(m_i^\circ|\gamma)}{P(m_i^\circ|Y, \gamma_n) P(\mathbf{Y}|\gamma_n)} \right] \right\} \\ &\triangleq \Delta(\gamma|\gamma_n) \end{aligned} \tag{B.11}$$

which lower-bounds the difference between the likelihood function before and after γ is updated.

$$L(\gamma) \geq L(\gamma_n) + \Delta(\gamma|\gamma_n) \tag{B.12}$$

Define

$$l(\gamma|\gamma_n) \triangleq L(\gamma_n) + \Delta(\gamma|\gamma_n) \tag{B.13}$$

so that the relationship in (B.11) can be made explicit as,

$$L(\gamma) \geq l(\gamma|\gamma_n) \tag{B.14}$$

When $\gamma = \gamma_n$, we observe that $\Delta(\gamma|\gamma_n) = 0$, and therefore $L(\gamma) - L(\gamma_n) = 0$ and $L(\gamma) = L(\gamma_n) = l(\gamma|\gamma_n)$. Given this equality, any γ that increases $l(\gamma|\gamma_n)$ will also increase $L(\gamma)$. In order to achieve the greatest possible increase in the value of $L(\gamma)$,

the EM Algorithm calls for choosing a value of γ such that $l(\gamma|\gamma_n)$ is maximized. The **updated** value can be denoted as γ_{n+1}. Formally, we have

$$\begin{aligned}
\gamma_{n+1} &= \arg\max_{\gamma} \{l(\gamma|\gamma_n)\} \\
&= \arg\max_{\gamma} \left\{ L(\gamma_n) + \sum_{m} \left\{ P(m_i^{\circ}|Y,\gamma_n) \ln \left[\frac{P(\mathbf{Y}|m_i^{\circ},\gamma)P(m_i^{\circ}|\gamma)}{P(m_i^{\circ}|Y,\gamma_n)P(\mathbf{Y}|\gamma_n)} \right] \right\} \right\} \\
&= \arg\max_{\gamma} \left\{ \sum_{m} \{P(m_i^{\circ}|\mathbf{Y},\gamma_n) \ln [P(\mathbf{Y},m_i^{\circ}|\gamma)]\} \right\} \\
&= \arg\max_{\gamma} \left\{ E_{\mathbf{m}|\mathbf{Y},\gamma_n} [\ln P(\mathbf{Y},m_i^{\circ}|\gamma)] \right\}
\end{aligned} \tag{B.15}$$

The final expression reached can be anatomized by the following steps;

1. *E-step*: Determine the **conditional expectation**, $E_{\mathbf{m}|\mathbf{Y},\gamma_n}\left[lnP(\mathbf{Y},m_i^{\circ}|\gamma)\right]$.

2. *M-step*: Maximize this expression with respect to γ.

As a consequence, even if the self-censored messages are missing in the sample, using EM algorithm along with the observables suggested by our theory can still help find an unbiased estimate for γ.