Audience Preferences and Information Disclosure*

Boris Ginzburg[†]

April 16, 2012

Quite preliminary and rather incomplete

Abstract

Many situations involve an informed agent choosing when to reveal a payoff-relevant state of the world to an uninformed decision-maker. This paper differs from previous studies analysing such interactions by considering the case when the informed party commits to a particular disclosure policy before learning the state, and also by explicitly looking at how disclosure is affected by the distribution of decision-maker's preference parameter. It turns out that at an equilibrium, the informed party will pool some states into at most one set, fully separating the rest. The complexity of this set depends on the complexity of the distribution of decision-maker's preference parameter. The exact disclosure policy depends on that distribution as well, and while prior research has largely concluded that disclosure of all states is optimal under fairly general conditions, this paper finds that full revelation occurs only when the decision-maker is likely to be biased against the informed party. These results help shed some light on a number of phenomena, such as degree and effects of political censorship, impact of varying levels of central bank transparency, and firms' disclosure of information related to product quality, financial situation, or environmental performance.

Keywords: persuasion games, communication, disclosure, political censorship, product quality, transparency

JEL codes: C72, D82, D72, L15

1 Introduction

On February 17 1982, an escalator in the Aviamotornaya station of the Moscow Metro suffered a failure, causing an accident which claimed eight lives. Since

*I am grateful to V. Bhaskar for guidance and helpful advice. I also thank Roland Bénabou, Philippe Jehiel, Christian Krestel, John Quah, Francesco Squintani, Felix Várdy, Jörgen Weibull, Andriy Zapechelnyuk, and seminar audiences at UCL and BICEPS for valuable comments.

[†]Department of Economics, University College London

the station had to be closed, it was impossible to entirely conceal the news of the disaster. The government-controlled Soviet media, however, provided few details of the accident and did not mention the death toll. The aim of concealing the truth was probably to prevent the public from forming a negative opinion about the safety of the metro, and by extension, about the competence of the government. But rather than shoring up the support of the government, media silence had an opposite effect - within days, rumours began to circulate that hundreds of people had been killed. Clearly, the attempt at controlling public opinion has backfired; perhaps the government would have been better off if it had published the true story[1].

When is it optimal for an informed agent to disclose payoff-relevant information to an uninformed party which then makes a payoff-relevant decision, and what are the effects of various disclosure policies? These questions has received substantial attention in both theoretical and applied research. A number of studies in economic theory have focused on the so-called persuasion games, in which an agent chooses whether to reveal a state of the world to a decision maker, who then selects an action that affects payoffs of both sides. The initial analysis of such games (Grossman, 1981; Milgrom, 1981) suggested that at an equilibrium, Receiver would learn every state of the world. Seidmann and Winter (1997), Koessler (2003), and Mathis (2008) show that this result holds in various fairly general settings. Exceptions to the full disclosure result have largely been due to uncertainty over whether the informed agent has precise information (Shin, 1994), or due to informed agent's preferences being uncertain (Wolinsky, 2003) or non-monotonic in decision-maker's action (Giovannoni and Seidmann, 2007).

In contrast, there is no consensus in applied literature that that full revelation is the best policy for the informed party in various settings. For instance, researchers analysing political communication frequently conclude that censorship of unfavourable information can be successfully used to manipulate political outcomes in the interest of the censoring agent. Thus, various studies have shown that media bias has played a role in determining election outcomes in Peru and Brazil (Boas, 2005), Mexico (Lawson and McCann, 2005), and Russia (Enikolopov et al., 2011). On the other hand, some studies suggest that non-disclosure is not always the optimal strategy - Dyczok (2006) questions the effectiveness of censorship in supporting Kuchma's government in Ukraine; Kern and Hainmueller (2009) report that the East German government enjoyed greater public support in areas where the population could watch West German television; Goldstein (1989) shows that censorship of anti-government caricatures in nineteenth-century France has at times led to an increase of support for the message they contained.

Similarly, research on disclosure of information by firms and non-governmental organisations has led to ambiguous conclusions regarding the optimality of full revelation. For example, Jin and Leslie (2003) conclude that mandatory disclosure causes firms to divulge more information about the quality of their prod-

[1]The accident, its reporting, and subsequent rumours are mentioned in Zolotov (2002).

ucts, compared to the situation when disclosure is voluntary. Similarly, the rich literature on disclosure of financial information (see a review by Healy and Palepu 2001) does not provide clear-cut support for the full revelation result. Neither do firms always provide full information about their environmental performance, as Patten (2002) shows. Finally, Rothman et al. (2011) report that the vast majority of health advocacy organisations do not reveal the full extent of their relationships with the pharmaceutical industry.

This paper contributes to the existing literature in two ways. First, it examines what happens when an informed party has to commit to a particular disclosure strategy (i.e. to choose which states will be revealed and which are to be hidden) before learning the state. This makes the model applicable to a wide range of situations. For example, disclosure of politically sensitive information (e.g. censorship of anti-government news stories, or restrictions on hate speech or incitement of violence) is often regulated by either laws or by bureaucratic instructions, none of which can be quickly adjusted in reaction to a new story. Similarly, firms can commit to revealing or concealing particular facts (such as information about their product quality, financial status, etc.) if these facts have to be determined by an outside expert (e.g. an auditor assessing a company's financial situation, or a reviewer evaluating a theatre play) who then makes them public. Finally, commitment may arise as a credible equilibrium strategy in a repeated interaction - consider intelligence agencies that choose never to comment on whether a suspected spy has been working for them even when the suspect is innocent.

Second, this paper adds to the previous studies by explicitly modeling possible uncertainty about the uninformed party's preferences and its effect on disclosure policy. Thus, it can be used to analyse how the equilibrium level of political censorship differs in societies with different distributions of attitudes towards the government (e.g. how much censorship would be expected at an equilibrium in a society that largely supports the government, compared to one that is against it, or to one which is sharply divided), how the distribution of customers' ex ante preference for a firm's product affects the optimal disclosure policy by the firm, and so on. This would help explain the often-noted fact that the amount of disclosure often varies in different settings - for example, different authoritarian regimes allow different amounts of media freedom (Egorov et al., 2009), central banks differ in their level of transparency (Eijffinger and Geraats, 2006), commercial banks have been revealing varying amounts of financial information (Pérignon and Smith, 2010), while firms vary in the amount of information on their environmental record that they reveal (Cho and Patten, 2007).

In addition to the works mentioned above, two articles offer a theoretical approach similar to the one used in this paper. Rayo and Segal (2010) offer a model in which an informed agent can commit to a strategy of disclosing information to a decision-maker; they then characterise optimal revelation policy. In contrast to their work - which while considering uncertainty over decision-maker's preferences derives its main results for a uniform distribution of the relevant preference parameter - I look explicitly at the relationship between the distribution of the audience's preferences and the optimal disclosure strategy. Kamenica and Gentzkow (2009) also allow the informed party to commit and derive general conditions under which it can benefit from an imperfect disclosure strategy; unlike their work, this paper explicitly characterises optimal revelation policy and examine the effect of decision-maker's preferences on it.

To address these question, this paper proposes a model in which an informed party (such as a government or a firm) chooses a partition (representing disclosure strategy) of the space of possible news. The news are then realised, and a decision-maker (a voter, a customer, an investor, etc) finds out which element of the partition contains the news. She then chooses an action, on which the informed agent's payoff depends. Decision-maker's payoff, on the other hand, depends on her action, the news, and her preference parameter, which is her private information. Specifically, she benefits from taking an action favourable to the informed agent when the value of the news is higher then her preference parameter.

The results of the paper suggest that the equilibrium disclosure policy crucially depends on the decision-maker's preferences. In general, an optimal disclosure policy will consist of singletons plus at most one non-singleton set, suggesting that the news are either fully disclosed, or pooled into one set. The number of disjoint intervals comprising this set cannot be higher than the number of peaks in the density of decision-maker's preference parameter - in other words, diclosure policy is likely to be simple if the distribution of preferences is simple.

The actual disclosure policy will largely depend, at an equilibrium, on the shape of decision-maker's preferences, indicating, for example, that the effect of censorship crucially depends on the existing opinion of the population. Somewhat paradoxically, more information will be revealed at ane equilibrium if the decision-maker is skeptical about taking an action that the informed party prefers. Full revelation emerges as a special case, being an equilibrium if and only if the density of the preference parameter is an increasing function - which corresponds to a decision-maker that is biased against the informed agent. On the other hand, decreasing density makes it optimal for the informed agent to pool all states, disclosing nothing. Finally, unimodal density in general produces a partially separating equilibrium.

2 Model

There are two players: Sender (he) and Receiver (she). State of the world is a pair $(\tau, \omega) \in [0,1]^2$; ω is what Sender chooses whether to disclose (referred to as "news"), and τ is Receiver's preference parameter, which is her private information. Higher values of ω indicate "better" news, and higher values of τ suggest that Receiver is predisposed against Sender. The cdf of τ and ω are F and G, respectively, and the associated densities are f and g. Assume that τ and ω are independent. Further, assume that f is continuously differentiable, and that g is strictly positive everywhere on $[0,1]$.

Receiver chooses an action $c \in C$, where the action set C is a compact subset of $\mathbb{R}$. Sender's strategy is a partition $\mathcal{P}$ of the set of news $[0, 1]$. I assume that $\mathcal{P}$ contains a finite or countable number of non-singleton sets, and that each element of $\mathcal{P}$ has a finite or countable number of boundary points. Denote by $\mathfrak{P}$ the set of all partitions of the $[0, 1]$ interval that meets this assumption.

The timing of the game is as follows. First, Sender commits to a disclosure policy by choosing $\mathcal{P} \in \mathfrak{P}$, which is announced to Receiver. Then, Nature draws τ and ω from F and G, respectively; and Receiver learnsτ. Next, the element of $\mathcal{P}$ to which ω belongs is disclosed to Receiver. Receiver then chooses $c \in C$. Finally, payoffs are realised in the following way: Sender's payoff equals c, while Receiver's payoff is $c(\omega - \tau)$.

Let us pause for a moment to examine the intuition of the model. In the context of this model, Sender and news can refer to, respectively, a government and some politically-relevant information that it can hide, a firm and the quality of its product, a central bank and the degree to which the currency is stable, etc. Receiver's preference parameter τ can indicate median voter's preferences for supporting or opposing the government, a buyer's reservation value, or an investor's ex ante trust in the currency - all of which are unknown but come from a known distribution. Alternatively, we can see τ as representing the preferences of a unit population of agents, distributed with a known density.

By choosing a partition of the news space, Sender can select how much information he discloses. For example, full disclosure, in which Receiver always learns the exact value of news, corresponds to partitioning the news space into singletons. On the other extreme, a partition that contains only one set (the entire $[0, 1]$ interval) reveals no information to Receiver. Sender can also reveal the news with imperfect precision, for example, choosing a partition that contains several large sets - in the real world, this may correspond to, for example, announcing that the news are "good", "moderate", or "bad", without giving more detail. Note that the sets comprising $\mathcal{P}$ need not be connected - Sender can, for example, reveal whether the news are "moderate" or "exceptional", without specifying whether the latter means good or bad.

Sender always wants to choose $\mathcal{P}$ that would encourage Receiver to pick higher c - depending on the specific application, higher c may mean that Receiver supports a more pro-government party, buys more of the good, etc (note that C can be both discrete and continuous). Receiver, on the other hand, is better off with larger c if and only if the news are good enough to exceed a threshold given by her preferences (i.e. when $\omega \geq \tau$). Higher τ thus means a Receiver that is gains less choosing an action favourable for Sender - for example, this might be a citizen who is generally opposed to the government but may change her mind if the news are very good.

Several assumptions are implied when the model is applied to specific situations. First, it is assumed that Receiver's preferences are independent of the news. This may not hold in the long term - for example, a citizen may become less inclined to support the government if bad news keep coming - but in a one-shot interaction this should hold. Alternatively, we may think of τ as indicating Receiver's preferences caused by factors other than those captured in ω - for instance, if ω characterises how well the government is conducting its foreign policy, then τ can show a citizen's level of support for its economic programme.

Additionally, the model assumes that Sender can hide the news or reveal them up to a subset of the news space, but he cannot lie. In some situations, this is straightforward - when thinking of firms choosing whether to disclose their financial situation or produc quality, we may assume that they will be subjected to very severe punishment if caught lying. If Sender is a government, which may be not restricted by law, we may assume that revealing information means either providing hard evidence of the news (photos and videos, testimony by independent witnesses), or allowing to independent media to report the news; we can also suppose that failing to do so will lead Receiver to disbelief whatever the government is saying.

3 Analysis and Results

Let us say that Sender has chosen a partition $\mathcal{P}$. If the news picked by Nature fall into a set $S \in \mathcal{P}$, Receiver's expected payoff from taking an action c, given her preference parameter τ equals $\mathrm{E}\left[c(\omega-\tau) \mid \omega \in S\right] = c\left(\mathrm{E}\left[\omega \mid \omega \in S\right]-\tau\right)$. If $\tau < \mathrm{E}\left[\omega \mid \omega \in S\right]$, this expression is maximised at $c = \max\{C\}$, while if $\tau > \mathrm{E}\left[\omega \mid \omega \in S\right]$, it is maximised at $c = \min\{C\}$. This describes Receiver's best response. Since only $\max\{C\}$ and $\min\{C\}$ will ever be chosen by Receiver, we can normalise C to $\{0,1\}$ from now on.

Given that $\omega \in S \in \mathcal{P}$, Sender's expected payoff equals the probability that $\tau < \mathrm{E}\left[\omega \mid \omega \in S\right]$, which is $F\left(\mathrm{E}\left[\omega \mid \omega \in S\right]\right)$. Denote denote by μ_S the probability that ω falls in S (i.e. the measure of S associated with g). Then, $\mu_S \equiv \int\limits_{\omega \in S} g(\omega)\, d\omega$. Then Sender's overall expected payoff from choosing $\mathcal{P}$ equals

$$v(\mathcal{P}) = \int\limits_{S \in \mathcal{P}} F\left(\mathrm{E}\left[\omega \mid \omega \in S\right]\right) d\mu_S$$

which Sender maximises by choosing $\mathcal{P} \in \mathfrak{P}$.

For subsequent analysis, it is useful to make the following simplifications of Sender's strategy space. First, note that by assumption, the number of non-singleton sets in $\mathcal{P}$ is at most countable. Hence, if $\mathcal{P}$ contains any zero-measure sets other than singletons, then the overall measure of such sets is zero. But note that $v(\mathcal{P}) = v\left(\mathcal{P} \bigcup \mathcal{M}\right) = v(\mathcal{P} \setminus \mathcal{M})$ for any countable collection $\mathcal{M}$ of zero-measure sets. Thus, every $\mathcal{P} \in \mathfrak{P}$ that contains zero-measure non-singleton sets gives Sender the same payoffs as another partition in which these sets are further split into singletons. Because of this, without loss of generality we can restrict attention to partitions in which every non-singleton element has a positive measure.

Second, every $S \in \mathcal{P}$ has at most a countable number of boundary points, so the overall measure of all boundary points of S is zero. But $v(\mathcal{P}) =$

$v(\mathcal{P} \bigcup \{S \bigcup M\}) = v(\mathcal{P} \bigcup \{S \setminus M\})$, where M is a zero-measure set. So, without a change in Sender's payoff, a positive-measure set S can be split by transforming all boundary points that are not attached to intervals[2] into singletons, and this can be done to all positive-measure sets because there is a countable number of them. Thus, we can only look at partitions that contain either singletons or sets that are countable collections of intervals.

Finally, we can also restrict attention to partitions in which every ω that forms a singleton set in $\mathcal{P}$ has a neighborhood of such news. This is because every ω that does not have this property forms a boundary point of some positive-measure set. Therefore, the overall measure of these points is zero, so they can be pooled with some positive-measure set without a change in $v(\cdot)$.

Proposition 1. $\max\limits_{\mathcal{P} \in \mathfrak{P}} \{v(S)\}$ *exists.*

Proof: see Appendix.

This proposition ensures the existence of a pure-strategy equilibrium.

Note that if a set $S \in \mathcal{P}$ is a singleton $\{\omega\}$, then $d\mu_S = g(\omega)$, and $\mathrm{E}[\omega \mid \omega \in S] = \omega$. If S is not a singleton, i.e. if $\mu_S > 0$, let $t_S \equiv \mathrm{E}[\omega \mid \omega \in S] = \frac{1}{\mu_S} \int\limits_{\omega \in S} \omega g(\omega)\, d\omega$. The expression for Sender's expected payoff then becomes:

$$v(\mathcal{P}) = \sum_{S \in \mathcal{P}\,:\,\mu_s > 0} F(t_S)\,\mu_S + \int\limits_{\omega \in S \in \mathcal{P}\,:\,\mu_S = 0} F(\omega) g(\omega) d\omega$$

Clearly, the strategy that maximises it depends on the shapes of f and g. In order to characterise it, define for every positive-measure set S a function $z_S(\omega) \equiv \int_{t_S}^{\omega} f(t_S) - f(x)\, dx$. Now consider a partition $\mathcal{P}$ that is a candidate for an optimal partition. We can check for several kinds of deviations from $\mathcal{P}$. First, we can take some news ω that are pooled into a positive-measure set A and disclose them, i.e. turn them into singleton elements of the partition. We can also remove them from A and pool them with some other set B instead. Finally, we can also take some other news that under $\mathcal{P}$ are not pooled into any positive-measure set (i.e. that forms a singleton element of $\mathcal{P}$), and merge them with some $A \in \mathcal{P}$. If $\mathcal{P}$ is optimal, none of these deviations can be beneficial. This is captured in the following necessary condition for an optimum:

Proposition 2. *Suppose that $\mathcal{P}$ maximises $v(\cdot)$. Then the following must hold for every positive-measure set $A \in \mathcal{P}$:*

1. $z_A(\omega) \geq 0$ *for any* $\omega \in A$
2. $z_A(\omega) \geq z_B(\omega)$ *for any* $\omega \in A$ *and any positive-measure set* $B \in \mathcal{P}$
3. $z_A(\omega) \leq 0$ *for any* ω *such that* $\{\omega\}$ *forms a singleton element of* $\mathcal{P}$

[2] I.e. all $\omega \in S$ for which there exists a neighborhood T that contains no other elements of S besides ω.

Proof: see Appendix.

Using this condition, we can substantially narrow down the set of possible equilibrium strategies, in the following way:

Proposition 3. *For any g and as long as f is horizontal on no more than one interval, there always exists an optimal Sender's strategy in which the partition $\mathcal{P}$ contains at most one positive-measure set.*

Proof: see Appendix.

Proposition 3 ensures that at the optimum, almost all shapes of f induce a revelation strategy under which the news are pooled into at most one set. In other words, Sender will either disclose the news precisely, or suppress them altogether, revealing nothing except that the news are in of a kind that are not disclosed. Sender will never want to take the middle road of revealing the news imprecisely, i.e. disclosing them up to a subset of the news space.

Note that if f has a horizontal section, it is possible that there exists a partition with several positive-measure sets that brings the same expected payoff to Sender. In particular, if f is uniform on $[0, 1]$, any partition will give Sender the same payoff.

Additionally, note that any positive-measure set $S \in \mathcal{P}$ can be split into sets S_1 and S_2 such that $t_S = t_{S_1} = t_{S_2}$; the resulting partition will yield the same payoff to Sender as $\mathcal{P}$. From now on, however, we can assume that, when having several optimum strategies, Sender will always choose a partition with not more than one positive-measure set - perhaps because, all other things being equal, he has a preference for "simpler" strategies.

From now on, let us denote by S a positive-measure set that is a part of $\mathcal{P}$ at the equilibrium.

By eliminating all partitions with more than one positive-measure set, Proposition 3 substantially decreases the set of potentially optimal strategies and reduces the problem of determining the optimal partition to finding the optimal partition. Nevertheless, there are still many potential shapes of S. In particular, S can be considered more or less complex depending on the number of disjoint intervals it includes. The following proposition puts a restriction on the complexity of S:

Proposition 4. *If f has $m < \infty$ local weak maxima, then at the equilibrium, S includes no more than m disjoint intervals.*

Proof: see Appendix.

This proposition underscores the importance of f, the distribution of Receiver's preference parameter, in determining the optimal revelation strategy. It shows that in most cases, we should not expect a very complicated disclosure policy. Optimal disclosure strategy will only be "complex" - i.e. include a set of suppressed news consisting of many disjoint intervals - when the distribution f of Receiver's preference parameter is "complex" (i.e. has many peaks). For distributions with a small number of peaks, this again reduces the space of possible optimal strategies. This result suggest that for the most part, we are unlikely to see very complex diclosure and censorship policies in real-life situations, as long as Sender is optimising.

With these results in mind, we can look at the actual equilibrium disclosure policies for some specific shapes of f and g. We can start by examining the case that has often been the focus of research on disclosure - namely, the full disclosure case, which in this setting corresponds to a partition $\mathcal{P}$ consisting entirely of singletons.

Proposition 5. *Full disclosure is an equilibrium strategy if and only if f is weakly increasing. Furthermore, full disclosure is the unique equilibrium strategy if and only if f is strictly increasing.*

Proof: see Appendix.

This result show that full revelation can only emerge if fairly restrictive conditions are met. Recall that f is the density of τ, which can be interpreted as a minimum value of the news that still make Receiver better off from taking the action that Sender prefers. The higher the τ, the bigger is the payoff to Receiver from taking a pro-Sender action. Increasing f thus means that Receiver is predisposed against Sender - e.g. in the censorship application it may imply that the distribution of the society's attitudes is skewed towards opposing the government.

The idea that Sender is better off disclosing all the news when Receiver is biased against him may sound somewhat paradoxical. The intuitive explanation is that when τ is high, Receiver is better off not acting in Sender's interest if the news are hidden. Increasing f means that high τ is more likely, which encourages disclosure.

We can now look at the other possible extreme - the strategy of pooling all the news into one set, which is equivalent to communicating no information to Receiver.

Proposition 6. *Pooling all the news is the unique equilibrium strategy if f is strictly decreasing.*

Proof: see Appendix.

Similarly in spirit to the previous result, this suggests that Sender is better off hiding all states when the Receiver is more likely to be on his side. Note that this is a sufficient but not a necessary condition. Based on Proposition 2, a necessary condition for full pooling to be an equilibrium is that $z_S(\omega) = \int_{t_S}^{\omega} f(t_S) - f(x)\, dx \geq 0$ for all ω - meaning, in particular, that f must be weakly decreasing at t_S.

Finally, we can also look at a more general result, one that considers all unimodal densities f.

Proposition 7. *If f strictly increasing on $(0, k)$ and strictly decreasing on $(k, 1)$ for some $k \in (0, 1)$, then there is a unique equilibrium strategy $S = [a, 1]$, such that $0 \leq a < k$, and $t_S > k$.*

Proof: see Appendix.

The unimodal case thus gives rise to a partially separating equilibrium in which news are pooled over a certain interval. It is easy to see that full separation and full pooling emerge as special cases of this result when k is 1 and 0, respectively.

Note that as at the optimum, $t_S > k$, greater k - i.e. a peak further to the right - will in general lead to more states being disclosed. This, together with the result from Proposition 5, suggests that more news will be disclosed if τ is more likely to be large. Recalling that larger τ means that Receiver is more skeptical about taking an action favourable for Sender, we can again conclude that if Receiver is biased against the Sender, more information will be revealed, while if Receiver becomes pre-disposed towards Sender, a greater range of news become disclosed.

The idea that more extensive revelation is optimal is Receiver is more inclined to take an action favourable to Sender has several implications for real-world situations. On the normative side, suppose that Receiver represents a population of citizens who choose whether to adopt a racist model of behaviour, depending on their existing views and some incitement from an activist, which may be more or less effective. If Sender is a government that can choose whether to restrict hate speech, we can say that such a restriction will be effective if citizens are already pre-disposed against racism; if they are ex ante inclined to believe racist statements, restrictions on hate speech would be counterproductive. Similarly, we can say that a central bank is better off with more transparency if investors have a higher level of trust in the currency.

On the positive side, these results suggest that firms operating in markets with skeptical consumers are better off revealing more, or that less political censorship is optimal in societies where citizens are skeptical about the government, or . Note that this is not saying that more freedom of information will actually be observed in such societies, as there may be other factors (such as ideological constraints, legal requirements, disagreements within the ruling group, or the fact that the regime may be seeking support from only part of the public) that drive governments' behaviour. Rather, this suggests that when the population is not inclined to trust the government, political censorship is likely to harm rather than help it.

If we assume that authoritarian regimes do try to maxmise the support of the public, we can make certain conclusions about possible causes of certain political changes. A rise in the level of support for the government - for example, due to an outburst of patriotic feeling caused by a war - may encourage an optimising government to reduce media freedom. On the other hand, when facing a decline in popular support, a rational authoritarian ruler may choose to relax political censorship. Writing about the French Revolution, Alexis de Tocqueville has famously noted that "the regime which is destroyed by a revolution is almost always an improvement on its immediate predecessor, and experience teaches that the most critical moment for bad governments is the one which witnesses their first steps toward reform". In this case, perhaps the regime becomes less repressive precisely because it is in danger of being overthrown.

4 Conclusions

This paper has examined optimal information disclosure of a Sender who commits to a particular disclosure policy in advance. The main focus of the paper was the effect of Receiver's preference distribution f on the disclosure strategy.

In general, Sender's strategy space was quite large, allowing him to pick an arbitrary information partition subject to a few technical restrictions. Nevertheless, the set of possible optimal strategies has been found to be quite small. An optimal information partition generally has at most one positive-measure set, and the number of disjoint intervals of which this set consists is limited by the number of peaks of f.

Specific optimal disclosure policies in general depend on the shape of f. In particular, full revelation - the benchmark result that has often been the focus of other work in this area - has been found to only occur when f is increasing.

In general, more information is likely to be revealed when Receiver is biased against the Sender. Thus, for instance, a firm facing largely skeptical consumers is more likely to reveal information about its product than a firm whose customers are ex ante willing to trust it. Similarly, we can expect less political censorship at the equilibrium when the distribution of views in the population is skewed towards opposing the government. Alternatively, when censorhip is present under such circumstances, we can expect it to hurt the government. We can also expect an optimising government to change political censorship in response to a shift in public opinion, restricting the freedom of information when citizens become more willing to support it, and relaxing censorship when the views of the population turn against it.

References

Boas, Taylor C., "Television and Neopopulism in Latin America: Media Effects in Brazil and Peru," *Latin American Research Review*, 2005, *40* (2), 27–49.

Cho, Charles H. and Dennis M. Patten, "The role of environmental disclosures as tools of legitimacy: A research note," *Accounting, Organizations and Society*, 2007, *32* (7-8), 639–647.

Dyczok, Marta, "Was Kuchma's censorship effective? mass media in Ukraine before 2004," *Europe-Asia Studies*, 2006, *58* (2), 215–238.

Egorov, Georgy, Sergei Guriev, and Konstantin Sonin, "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data," *American Political Science Review*, 2009, *103* (04), 645–668.

Eijffinger, Sylvester C.W. and Petra M. Geraats, "How transparent are central banks?," *European Journal of Political Economy*, March 2006, *22* (1), 1–21.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya, "Media and Political Persuasion: Evidence from Russia," *American Economic Review*, December 2011, *101* (7), 3253–85.

Giovannoni, Francesco and Daniel J. Seidmann, "Secrecy, two-sided bias and the value of evidence," *Games and Economic Behavior*, May 2007, *59* (2), 296–315.

Goldstein, Robert Justin, "The Debate over Censorship of Caricature in Nineteenth-Century France," *Art Journal*, 1989, *48* (1), pp. 9–15.

Grossman, Sanford J, "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law & Economics*, December 1981, *24* (3), 461–83.

Healy, Paul M. and Krishna G. Palepu, "Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature," *Journal of Accounting and Economics*, September 2001, *31* (1-3), 405–440.

Jin, Ginger Zhe and Phillip Leslie, "The Effect Of Information On Product Quality: Evidence From Restaurant Hygiene Grade Cards," *The Quarterly Journal of Economics*, May 2003, *118* (2), 409–451.

Kamenica, Emir and Matthew Gentzkow, "Bayesian Persuasion," NBER Working Papers 15540, National Bureau of Economic Research, Inc November 2009.

Kern, Holger Lutz and Jens Hainmueller, "Opium for the Masses: How Foreign Media Can Stabilize Authoritarian Regimes," *Political Analysis*, 2009, *17* (4), 377–399.

Koessler, Frederic, "Persuasion games with higher-order uncertainty," *Journal of Economic Theory*, June 2003, *110* (2), 393–399.

Lawson, Chappell and James A. McCann, "Television News, Mexico's 2000 Elections and Media Effects in Emerging Democracies," *British Journal of Political Science*, 2005, *35* (01), 1–30.

Mathis, Jérôme, "Full revelation of information in Sender-Receiver games of persuasion," *Journal of Economic Theory*, November 2008, *143* (1), 571–584.

Milgrom, Paul R., "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, Autumn 1981, *12* (2), 380–391.

Patten, Dennis M., "The relation between environmental performance and environmental disclosure: a research note," *Accounting, Organizations and Society*, November 2002, *27* (8), 763–773.

Pérignon, Christophe and Daniel R. Smith, "The level and quality of Value-at-Risk disclosure by commercial banks," *Journal of Banking & Finance*, February 2010, *34* (2), 362–377.

Rayo, Luis and Ilya Segal, "Optimal Information Disclosure," *Journal of Political Economy*, 2010, *118* (5), 949 – 987.

Rothman, Sheila M., Victoria H. Raveis, Anne Friedman, and David J. Rothman, "Health Advocacy Organizations and the Pharmaceutical Industry: An Analysis of Disclosure Practices," *Am J Public Health*, 2011, *101* (4), 602–609.

Seidmann, Daniel J. and Eyal Winter, "Strategic Information Transmission with Verifiable Messages," *Econometrica*, January 1997, *65* (1), 163–170.

Shin, Hyun Song, "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, October 1994, *64* (1), 253–264.

Wolinsky, Asher, "Information transmission when the sender's preferences are uncertain," *Games and Economic Behavior*, February 2003, *42* (2), 319–326.

Zolotov, Andrei Jr., "The Dangers of Riding the Escalator," *The Moscow Times*, April 2002, (2417).

5 Appendix

5.1 Proof of Proposition 1

The proof that $\max_{\mathcal{P}\in\mathfrak{P}} \{v(S)\}$ exists follows several steps.

Step 1. A metric on $\mathfrak{P}$.

Take a partition $\mathcal{P} \in \mathfrak{P}$, and denote its positive-measure sets by $\left(S^i\right)_{i\in\mathbb{N}}$ such that $\inf\left(S^i\right) \leq \inf\left(S^{i+1}\right)$.[3] For every $S^i \in \mathbb{N}$, denote by $\left(s^i_j\right)_{j\in\mathbb{N}}$ an increasing sequence of the boundary points of S^i.[4] The sequence of such sequences, $\left(\left(s^i_j\right)_{j\in\mathbb{N}}\right)_{i\in\mathbb{N}}$, fully describes[5] a partition $\mathcal{P}$ (up to a closure[6]). To measure the distance between two partitions $\mathcal{P}$ and $\hat{\mathcal{P}}$, we can measure the Euclidean

[3] Recall that every $\mathcal{P} \in \mathfrak{P}$ must include at most a countable number of positive-measure sets, so it is possible to assign to each set a natural number.

[4] Again, recall that the number of boundary points for each $S^i \in \mathcal{P}$ is at most countable.

[5] In case $\mathcal{P}$ has a finite number N of positive-measure sets, let $\left(s^i_j\right)_{j\in\mathbb{N}} = (1, 1, \ldots)$ for every $i > N$. Similarly, if some set S^i has a finite number N of boundary points, let $\left(s^i_j\right)_{j\in\mathbb{N}} = \left(s^i_1, s^i_2, \ldots s^i_{N-1}, s^i_N, s^i_N, \ldots\right)$

[6] Recall that the overall measure of all boundary points of all positive-measure sets in $\mathcal{P}$ is zero, so $v(\mathcal{P}) = v(\mathrm{cl}(\mathcal{P}))$, where $\mathrm{cl}(\mathcal{P})$ is the closure of $\mathcal{P}$. Thus, we can treat the as the same partition.

distance between each pair of boundary points in each pair of positive-measure sets belonging to $\mathcal{P}$ and $\hat{\mathcal{P}}$, and then find the largest such distance. Formally, let the measure $d : \mathfrak{P} \times \mathfrak{P} \to \mathbb{R}$ be defined as:

$$d\left(\mathcal{P}, \hat{\mathcal{P}}\right) \equiv \sup_{i,j\in\mathbb{N}} \left\{\left|s_j^i - \hat{s}_j^i\right| \,:\, s_j^i \in S^i \in \mathcal{P}\,,\, \hat{s}_j^i \in \hat{S}^i \in \mathcal{P}\right\}$$

Step 2. Sequential compactness of $(\mathfrak{P}, d)$.

Let us prove that every sequence in $\mathfrak{P}$ has a convergent subsequence. Take an arbitrary sequence of partitions $(\mathcal{P}_v)_{v\in\mathbb{N}}$ lying in $\mathfrak{P}$. Define $h : \mathbb{N}^2 \to \mathbb{N}$ as a bijective function that gives a number to every pair (i, j).

First, for every partition $\mathcal{P}_v$ take the element $\left(s_j^i\right)_v$ for which $h(i,j) = 1$. The sequence of these elements along v, $\left(s_j^i\right)_{v\in\mathbb{N}}$ is bounded between 0 and 1; therefore, by Bolzano-Weierstrass Theorem, it must have a convergent subsequence. Denote this convergent subsequence by $\left(s_j^i\right)_{v\in V_1}$, and denote its limit by l_j^i. Thus, the sequence of partitions must have a subsequence $(\mathcal{P}_v)_{v\in V_1}$ in which the "first" boundary point converges to l_j^i.

Now suppose that there exists a subsequence of partitions $(\mathcal{P}_v)_{v\in V_n}$ in which every boundary point s_j^i converges as long as such that $h(i,j) \leq n$. Then $(\mathcal{P}_v)_{v\in V_n}$ must have a subsequence in which s_j^i for wich $h(i,j) = n+1$ converges as well.

By induction, therefore, there must exist a subsequence of partitions over which every point s_j^i converges to a limit l_j^i. Then that subsequence of partitions converges to a limit described by $\left(\left(l_j^i\right)_{j\in\mathbb{N}}\right)_{i\in\mathbb{N}}$.

Step 3. Compactness of $\mathfrak{P}$.

We we have shown that every sequence of partitions in $\mathfrak{P}$ converges, so $\mathfrak{P}$ is sequentially compact, which in metric spaces implies compactness. Thus, $v(\cdot)$ is a continuous function with compact support; by Weierstrass theorem it must then have a maximum.

5.2 Proof of Proposition 2

To prove (1), take a partition $\mathcal{P}$ containing a positive-measure set A. Now take a w belonging to the interior of A and suppose that $z_A(w) < 0$.[7] Consider a deviation from $\mathcal{P}$ to a partition $\hat{\mathcal{P}}$ that differs from $\mathcal{P}$ in that an interval $[w, r]$ is removed from A and instead all the news in $[w, r]$ are disclosed (i.e. turned into singleton elements of the partition). If $r = w$, then $v\left(\hat{\mathcal{P}}\right) = v(\mathcal{P})$. Recall

[7] The assumption that w is in the interior of A is without loss of generality, since for every w on the boundary of A such that $z_A(w) < 0$, there must (by continuity of $z_A(w)$) be another w' in the neighborhood of w for which this property also holds.

that $v(\mathcal{P}) = \sum\limits_{S\in\mathcal{P}\,:\,\mu_s>0} F(t_S)\mu_S + \int\limits_{\omega\in S\in\mathcal{P}\,:\,\mu_S=0} F(\omega)g(\omega)d\omega$. Then

$$v\left(\hat{\mathcal{P}}\right) - v(\mathcal{P}) = F\left(t_{A\setminus[w,r]}\right)\mu_{A\setminus[w,r]} + \int_w^r F(\omega)g(\omega)d\omega - F(t_A)\mu_A$$

Note that that

$$\mu_{A\setminus[w,r]} = \int\limits_{\omega\in A} g(\omega)\,d\omega - \int_w^r g(\omega)\,d\omega$$

and

$$t_{A\setminus[w,r]} = \frac{\int\limits_{\omega\in A} \omega g(\omega)\,d\omega - \int_w^r \omega g(\omega)\,d\omega}{\int\limits_{\omega\in A} g(\omega)\,d\omega - \int_w^r g(\omega)\,d\omega}$$

Taking the derivative of $v\left(\hat{\mathcal{P}}\right) - v(\mathcal{P})$ with respect to r yields

$$\frac{\partial\left[v\left(\hat{\mathcal{P}}\right) - v(\mathcal{P})\right]}{\partial r} = g(r)\left[f\left(t_{A\setminus[w,r]} - r\right) - F\left(t_{A\setminus[w,r]}\right) + F(r)\right] = -g(r)\,z_{A\setminus[w,r]}(r)$$

If $r = w$, then $A\setminus[w,r] = A$, and $v\left(\hat{\mathcal{P}}\right) = v(\mathcal{P})$, so the difference is zero. If $\mathcal{P}$ is an optimal strategy, the difference must be weakly decreasing in r at $r = w$. But if $z_A(w) < 0$, then

$$\left.\frac{\partial v\left(\hat{\mathcal{P}}\right)}{\partial r}\right|_{r=w} = -g(w)\,z_A(w) > 0$$

As g is assumed to be strictly positive everywhere. Therefore, increasing r is a profitable deviation for Sender, which means that $\hat{\mathcal{P}}$ is not optimal.

Parts (2) and (3) are proved analogously. To prove (2), suppose that for some positive-measure sets $A, B \in \mathcal{P}$, $z_A(w) < z_B(w)$ for some $w \in A$. Consider a deviation from $\mathcal{P}$ to $\hat{\mathcal{P}}$ in which an interval $[w,r]$ is removed from A and pooled with B. Again, $v\left(\hat{\mathcal{P}}\right) = v(\mathcal{P})$ for $r = w$. But then

$$\left.\frac{\partial\left[v\left(\hat{\mathcal{P}}\right) - v(\mathcal{P})\right]}{\partial r}\right|_{r=w} = -g(w)\,z_A(w) + g(w)\,z_B(w) > 0$$

Hence, Sender benefits from a deviation in which r is increased, and thus $\mathcal{P}$ cannot be an equilibrium strategy.

Finally, to prove (3), assume that $z_A(w) > 0$ for some positive-measure set A and for some w that is not part of any positive-measure set. Now take some interval $[w,r]$ such that every $\omega \in [w,r]$ is a singleton element of the partition, and consider a change from $\mathcal{P}$ to $\hat{\mathcal{P}}$ in which this interval is pooled with A. Then

$$\left.\frac{\partial\left[v\left(\hat{\mathcal{P}}\right)-v\left(\mathcal{P}\right)\right]}{\partial r}\right|_{r=w}=g\left(w\right)z_{A}\left(w\right)>0$$

so again there is a profitable deviation.

5.3 Proof of Proposition 3

Suppose that there exists a partition $\mathcal{P}$ which includes two positive-measure sets A and B, and assume without loss of generality that $t_A \leq t_B$. Denote by $\hat{\mathcal{P}}$ a partition that differs from $\mathcal{P}$ in that A and B are pooled together.

It is possible to prove in several steps that either $\hat{\mathcal{P}}$ gives Sender the same expected payoff as $\mathcal{P}$, or that there exists a partition that gives a higher payoff.

Step 1. If $t_A = t_B$, then $v\left(\hat{\mathcal{P}}\right) = v\left(\mathcal{P}\right)$.

Note that $\mu_{A\bigcup B} = \mu_A + \mu_B$, and

$$t_{A\bigcup B}=\frac{\int\limits_{\omega\in A}\omega g\left(\omega\right)d\omega+\int\limits_{\omega\in B}\omega g\left(\omega\right)d\omega}{\mu_{A}+\mu_{B}}=\frac{\mu_{A}}{\mu_{A}+\mu_{B}}t_{A}+\frac{\mu_{B}}{\mu_{A}+\mu_{B}}t_{B}=t_{A}$$

Hence,

$$\begin{aligned}&v\left(\hat{\mathcal{P}}\right)-v\left(\mathcal{P}\right)=F\left(t_{A\bigcup B}\right)\mu_{A\bigcup B}-F\left(t_{A}\right)\mu_{A}-F\left(t_{B}\right)\mu_{B}=\\&=F\left(t_{A}\right)\left(\mu_{A}+\mu_{B}\right)-F\left(t_{A}\right)\mu_{A}-F\left(t_{B}\right)\mu_{B}=0\end{aligned}$$

Step 2. If $t_A < t_B$, then for $\mathcal{P}$ to be an optimal partition, f must be increasing on $[t_A, t_B]$.

Suppose that $\mathcal{P}$ is an optimum partition, and consider the following deviation: take $C \subseteq A$ such that $t_C = t_A = t_{A\setminus C}$. Now remove it from A and pool with B; call the resulting partition $\mathcal{P}'$. Then

$$\begin{aligned}&v\left(\mathcal{P}\right)-v\left(\mathcal{P}'\right)=F\left(t_{A}\right)\mu_{A}+F\left(t_{B}\right)\mu_{B}-F\left(t_{A\setminus C}\right)\mu_{A\setminus C}-F\left(t_{B\bigcup C}\right)\mu_{B\bigcup C}=\\&=F\left(t_{A}\right)\mu_{A}+F\left(t_{B}\right)\mu_{B}-F\left(t_{A}\right)\left(\mu_{A}-\mu_{C}\right)-F\left(t_{B\bigcup C}\right)\left(\mu_{B}+\mu_{C}\right)=\\&=F\left(t_{A}\right)\mu_{C}+F\left(t_{B}\right)\mu_{B}-F\left(\tfrac{\mu_{B}}{\mu_{B}+\mu_{C}}t_{B}+\tfrac{\mu_{C}}{\mu_{B}+\mu_{C}}t_{A}\right)\left(\mu_{B}+\mu_{C}\right)\end{aligned}$$

The expression above must be non-negative for $\mathcal{P}$ to be the optimal partition. Denote $\gamma \equiv \frac{\mu_C}{\mu_B+\mu_C}$, and note that we can choose μ_C to be of any value between 0 and μ_A. Then

$$\gamma F\left(t_{A}\right)+\left(1-\gamma\right)F\left(t_{B}\right)\geq F\left(\gamma t_{A}+\left(1-\gamma\right)t_{B}\right)\ ,\ \forall\gamma\in\left[0,\frac{\mu_{A}}{\mu_{A}+\mu_{B}}\right]$$

Now consider a deviation from $\mathcal{P}$ to $\mathcal{P}''$ of the following form: take $D \subseteq B$ such that $t_D = t_B = t_{B \setminus D}$, remove D from B and pool it with A. Then

$$\begin{aligned}
&v(\mathcal{P}) - v(\mathcal{P}'') = F(t_A)\mu_A + F(t_B)\mu_B - F\left(t_{B\setminus D}\right)\mu_{B\setminus D} - F\left(t_{A\bigcup D}\right)\mu_{A\bigcup D} = \\
&= F(t_A)\mu_A + F(t_B)\mu_B - F(t_B)(\mu_B - \mu_D) - F\left(t_{A\bigcup D}\right)(\mu_A + \mu_D) = \\
&= F(t_A)\mu_A + F(t_B)\mu_D - F\left(\tfrac{\mu_A}{\mu_A+\mu_D}t_A + \tfrac{\mu_D}{\mu_A+\mu_D}t_D\right)(\mu_A + \mu_D)
\end{aligned}$$

Denote $\delta \equiv \frac{\mu_A}{\mu_A+\mu_D}$; note that μ_D can be chosen between 0 and μ_B. Then

$$\delta F(t_A) + (1-\delta)F(t_B) \geq F(\delta t_A + (1-\delta)t_B) \ , \forall \delta \in \left[\frac{\mu_A}{\mu_A+\mu_B}, 1\right]$$

This inequality and the previous one, taken together, imply that F must be convex on $[t_A, t_B]$, and thus f must be increasing.

Step 3. If $t_A < t_B$ and f is constant on $[t_A, t_B]$, then $v\left(\hat{\mathcal{P}}\right) = v(\mathcal{P})$.

If f is horizontal then F is linear, which implies

$$\begin{aligned}
&v\left(\hat{\mathcal{P}}\right) - v(\mathcal{P}) = F\left(t_{A\bigcup B}\right)\mu_{A\bigcup B} - F(t_A)\mu_A - F(t_B)\mu_B = \\
&= F\left(\tfrac{\mu_A}{\mu_A+\mu_B}t_A + \tfrac{\mu_B}{\mu_A+\mu_B}t_B\right)(\mu_A + \mu_B) - F(t_A)\mu_A - F(t_B)\mu_B = 0
\end{aligned}$$

Step 4. If $t_A < t_B$, and f is not constant on $[t_A, t_B]$, $\mathcal{P}$ cannot be optimal.

If f is not constant on $[t_A, t_B]$, then in the neighbourhood of either t_A or t_B it must be strictly increasing. Assume it is increasing in the neighbourhood of t_A.

Suppose that $\mathcal{P}$ is optimal. Then, as we have established, f must be increasing on (t_A, t_B), and thus $z_A(\omega) = \int_{t_A}^{\omega} f(t_A) - f(x)\,dx < 0$ for every $\omega \in (t_A, t_B)$ - thus, by Proposition 2, no news in the interval (t_A, t_B) belong to A. On the other had, increasing f implies that $f(t_B) > f(t_A)$. Then for every $\omega \geq t_B$, $z_A(\omega) = \int_{t_A}^{\omega} f(t_A) - f(x)\,dx < \int_{t_B}^{\omega} f(t_A) - f(x)\,dx < \int_{t_B}^{\omega} f(t_B) - f(x)\,dx = z_B(\omega)$. Hence, by Proposition 2 no $\omega \geq t_B$ can belong to A. Therefore, optimal $\mathcal{P}$ implies that $t_A \geq \max\{A\}$, which is not possible.

In a similar way we can show that $\mathcal{P}$ cannot be optimal if f is increasing in the neighbourhood of t_B.

To summarise, we have shown that unless f is constant between the expected values of all positive-measure elements of $\mathcal{P}$ (in which case they can all be pooled together without a loss in utility), $\mathcal{P}$ cannot be optimal. Yet Proposition 1 has established that optimal strategy must exist. Therefore, there is always an optimal strategy in which $\mathcal{P}$ has at most one positive-measure set.

5.4 Proof of Proposition 4

We have earlier established that without loss of generality S can be thought of as a countable union of disjoint intervals. Thus, we can write $S = \bigcup\limits_{i \in I} [a_i, b_i]$ such that $0 \le a_i \le b_i \le a_{i+1} \le 1$, $\forall i \in I$, where I is countable.

Let us start by taking a set S comprising $|I|$ disjoint intervals. What is the smallest number of local maxima that f needs to have in for S to be Sender's equilibrium strategy?

Observe that since we have assumed the number of weak local maxima to be finite, f cannot be constant at any interval. This means that $z_S(\cdot)$ cannot equal zero on any interval $[p,q] \subseteq [0,1]$, since if it was zero, this would mean that $f(\omega) = f(t_S)$, $\forall \omega \in [p,q]$, i.e. that f is horizontal. The fact that $z_S(\cdot)$ cannot be zero on an interval, together with Proposition 2, implies that $z_S(\omega)$ is decreasing at $\omega = a_i$ and increasing at $\omega = b_i$, $\forall i \in I$. This means - since $z_S(\omega)$ is continuously differentiable - that $\frac{dz_S}{d\omega}(a_i) > 0$ and $\frac{dz_S}{d\omega}(b_i) < 0$.

Hence, for every $i \in I$, $z_S(\omega)$ must have at least one local maximum $c_i \in (a_i, b_i)$ and at least one local minimum $d_i \in (b_i, a_{i+1})$. At a local maximum, $\frac{dz_S^2}{d\omega^2}(c_i) = -f'(c_i) < 0$, while at a local minimum, $\frac{dz_S^2}{d\omega^2}(d_i) = -f'(d_i) > 0$. But f is assumed to be continuously differentiable, and thus for every $i \in I$ there must be news $w_i \in (c_i, d_i)$ such that (i) $f'(w_i) = 0$, and (ii) in some neighbourhood of w_i, $f'(\omega) > 0$ for $\omega < w_i$ and $f'(\omega) < 0$ for $\omega > w_i$. This w_i is therefore a local maximum of f, and there must be such a point in every interval (a_i, a_{i+1}). There are $|I| - 1$ such intervals, which gives us $|I| - 1$ local maxima.

Additionally, note that $z_S(t_S) = 0$, and also, $\frac{dz_S}{d\omega}(t_S) = 0$. This gives several possibilities. If $t_S \in (b_i, a_{i+1})$ for some $i \in I$, then it is a local maximum, and there are not one but at least two local minima in (b_i, a_{i+1}) - one to the left of t_S and one to the right. So there is one more pair of a maximum and a minimum of $z_S(\omega)$, and by the above logic, there must be a local maximum of f in addition to the ones analysed in the previous paragraph. If $t_S \in (a_i, b_i)$ for some $i \in I$, then it is a local minimum, and there are two local maxima in (a_i, b_i) - again, f must have an extra local maximum. If $t_S = a_i$ for some $i \in I$, then the shape of $z_S(\omega)$ implies that in some neighbourhood of a_i, $\frac{dz_S^2}{d\omega^2}(\omega) = -f'(\omega) < 0$ for $\omega < a_i$ and $\frac{dz_S^2}{d\omega^2}(\omega) = -f'(\omega) > 0$ for $\omega > a_i$. Consequently, f must have an additional local maximum at a_i. Finally, if $t_S = b_i$ for some $i \in I$, then the shape of $z_S(\omega)$ implies that f must have not one but at least two local maxima in (c_i, d_i).

Hence, if S forms part of an equilibrium, f must have at least $|I| - 1$ local maxima plus one more. Therefore, $m \ge |I|$.

5.5 Proof of Proposition 5

First statement. To prove necessity, suppose that f is not weakly increasing - this implies that is is strictly increasing on some interval $[p,q]$. Let $\hat{\mathcal{P}} \equiv \left\{ [p,q], \{\omega\}_{\omega \in [0,1] \setminus [p,q]} \right\}$ be a partition consisting of the interval $[p,q]$ and singletons. Then the difference in Sender's payoff from $\hat{\mathcal{P}}$ and from the fully revealing partition $\left\{\{\omega\}_{\omega\in[0,1]}\right\}$ equals:

$$\begin{aligned}
&v\left(\hat{\mathcal{P}}\right)-v\left(\left\{\{\omega\}_{\omega\in[0,1]}\right\}\right)=F\left(t_{[p,q]}\right)\mu_{[p,q]}-\int_p^q F\left(\omega\right)g\left(\omega\right)d\omega=\\
&=\Pr\left(\omega\in[p,q]\right)F\left[\mathrm{E}\left(\omega\mid\omega\in[p,q]\right)\right]-\Pr\left(\omega\in[p,q]\right)\mathrm{E}\left(F\left[\omega\right]\mid\omega\in[p,q]\right)=\\
&=\Pr\left(\omega\in[p,q]\right)\left[F\left[\mathrm{E}\left(\omega\mid\omega\in[p,q]\right)\right]-\mathrm{E}\left(F\left[\omega\right]\mid\omega\in[p,q]\right)\right]>0
\end{aligned}$$

where the last inequality sign follows from Jensen's inequality and the fact that decreasing f implies concave F. Hence, if f is not weakly increasing, full disclosure cannot be optimal.

To prove sufficiency, consider a weakly increasing f. Pick an arbitrary partition $\tilde{\mathcal{P}}$ containing one[8] positive-measure set S. Now consider a partition $\mathcal{P}'$ that differs from $\tilde{\mathcal{P}}$ by having singletons instead of S - i.e. every $\omega\in S$ is a singleton element of $\tilde{\mathcal{P}}$. If $f\left(\omega\right)=f\left(t_S\right)$ for all $\omega\in S$, then

$$\begin{aligned}
&v\left(\tilde{\mathcal{P}}\right)-v\left(\mathcal{P}'\right)=F\left(t_S\right)\mu_S-\int\limits_{\omega\in S}F(\omega)g(\omega)d\omega=\\
&=\int\limits_{\omega\in S}\left[F\left(t_S\right)-F(\omega)\right]g(\omega)d\omega=\int\limits_{\omega\in S}f\left(t_S\right)\left[t_S-\omega\right]g(\omega)d\omega=\\
&=f\left(t_S\right)\int\limits_{\omega\in S}\left[t_S-\omega\right]g(\omega)d\omega=f\left(t_S\right)\left[t_S\mu_S-t_S\mu_S\right]=0
\end{aligned}$$

On the other hand, if $f\left(w\right)\neq f\left(t_S\right)$ for some $w\in S$, then either $w>t_S$ and $f\left(w\right)>f\left(t_S\right)$, or $w<t_S$ and $f\left(w\right)<f\left(t_S\right)$. In either case, $z_S\left(w\right)=\int_{t_S}^{w}f\left(t_S\right)-f\left(x\right)dx<0$. From Proposition 3 it follows that $\tilde{\mathcal{P}}$ containing S cannot be an optimal.

Hence, a partition containing a positive-measure set S can only be optimal if $f\left(\omega\right)=f\left(t_S\right)$ for all $\omega\in S$. But every strategy that fits this criterion yields the same expected payoff to Sender as the fully revealing strategy. Thus, full disclosure must be an equilibrium strategy.

Second statement. To prove necessity, suppose, that f is not strictly increasing. Then f is weakly decreasing on some interval $[p,q]\subseteq[0,1]$. Defining $\hat{\mathcal{P}}\equiv\left\{[p,q],\{\omega\}_{\omega\in[0,1]\setminus[p,q]}\right\}$ as above and using the same reasoning, we can prove that $v\left(\hat{\mathcal{P}}\right)-v\left(\left\{\{\omega\}_{\omega\in[0,1]}\right\}\right)\geq 0$, so full disclosure cannot be a unique equillibrium strategy.

To prove sufficiency, note that if f is strictly increasing, then for any partition $\mathcal{P}$ containing a positive-measure set S, we have $\omega>t_S\Leftrightarrow f\left(\omega\right)>f\left(t_S\right)$. Pick news $w\in S$ such that $w>t_S$ (such w exists as $t_S<\max\left(S\right)$), and observe that $z_S\left(w\right)=\int_{t_S}^{w}f\left(t_S\right)-f\left(x\right)dx<0$. Proposition 2 then ensures that $\mathcal{P}$ is not an equilibrium strategy.

[8]Proposition 3 has already established that, in general, a partition chosen by Sender will have at most one positive-measure set.

5.6 Proof of Proposition 6

Take a strictly decreasing f, and consider a partition $\mathcal{P}$. If $\mathcal{P}$ is fully revealing, it cannot be optimal by Proposition 5. Now suppose $\mathcal{P}$ is not fully revealing, i.e. it contains a positive-measure set S. If there exists $\omega \notin S$, then $z_S(\omega) = \int_{t_S}^{\omega} f(t_S) - f(x)\,dx > 0$ - so $\mathcal{P}$ cannot be optimal. But Proposition 1 states that an optimal partition must exist. Therefore, $\mathcal{P} = \{[0,1]\}$ is an equilibrium, and as such it is unique.

5.7 Proof of Proposition 7

From Propositions 4 and 5 it follows that under a unimodal density f with a peak on $(0,1)$, the set S will consist of exactly one interval. Therefore, t_S must be in the interior of S. If $t_S \leq k$, this would mean that f is increasing on some neighbourhood of t_S, which would mean that $z_S(\omega) < 0$ for the news in that neighborhood, which in the equilibrium cannot hold. Thus, $t_S \in (k, 1]$. Then $z_S(\omega) > 0$ for all $\omega > t_S$. Similarly, $z_S(k) > 0$, and $z_S(\omega) > 0$ for some $\omega < k$. Hence, $S = [a, 1]$.