Optimal Information Censorship

Ginzburg, Boris

Department of Economics, Universidad Carlos III de Madrid

9 May 2019

Optimal Information Censorship*

Boris Ginzburg†

9th May 2019

Abstract

This paper analyses Bayesian persuasion of a privately informed receiver in a linear framework. The sender is restricted to censorship, that is, to strategies in which each state is either perfectly revealed or hidden. I develop a new approach to finding optimal censorship strategies based on direct optimisation. I also show how this approach can be used to restrict the set of optimal censorship schemes, and to analyse optimal censorship under certain classes of distributions of the receiver's type.

Keywords: Bayesian persuasion; censorship

JEL codes: D82, D83

1 Introduction

This paper analyses a Bayesian persuasion game between a sender and a privately informed receiver. Both the sender and the receiver have linear utility functions. There is a continuum of states. The sender wants the

*I am grateful to V. Bhaskar for guidance and advice. I also thank the editor, three anonymous referees, Miguel Ballester, Antonio Cabrales, Mikhail Drugov, Alessandro Ispano, Philippe Jehiel, Anton Kolotilin, Christian Krestel, Johannes Münster, David Myatt, Stephen Morris, Andriy Zapechelnyuk, and various audiences for valuable comments. Support from the Ministerio de Economía y Competitividad (Spain) grant MDM 2014-0431, and Comunidad de Madrid grant MadEco-CM (S2015/HUM-3444), is gratefully acknowledged.

†Department of Economics, Universidad Carlos III de Madrid, Calle Madrid 126, 28903 Getafe, Spain. Email: boris.ginzburg@uc3m.es. Website: sites.google.com/site/borgin.

receiver to act; the receiver only wants to act if his type is lower than the state.

The paper differs from the existing literature in two respects: it focuses on a particular class of persuasion strategies; and it develops a novel approach to finding the optimal strategy.

Specifically, the paper analyses a model in which the sender is restricted to censorship strategies: every state of the world is either revealed perfectly, or not revealed at all. In other words, the sender pools some states into a single set, and for each of the other states she sends a unique message.

A restriction to censorship strategies is relevant to a number of situations in which the sender needs to choose whether to transmit information that originates from an exogenous source. Consider, for example, a firm that wants to persuade customers to buy its product. A customer's willingness to do so depends on his preference type, as well as on a state of the world, which reflects the product's quality. The firm cannot credibly commit to an experiment that maps states to messages. It can, however, submit its product for review by independent experts. Some reviewers tend to detect and report very high quality; others – very low quality, etc. By choosing a set of reviewers, the firm can choose which states are revealed.

As another example, consider an authoritarian government that seeks to maximise the number of citizens that take a certain action (such as voting for the government, joining a pro-government rally, or refusing to join an anti-government protest). A citizen's willingness to do so depends on her type (which measures the degree to which she supports the government's ideology), and on the news (which is a measure of how competent the government is)[1]. The government can restrict the set of available news by choosing which independent media outlets are allowed to operate in the country[2].

The key contribution of this paper is in developing a novel approach to finding the optimal censorship scheme when the sender is constrained to censorship strategies. Specifically, the paper takes a direct optimisation approach: for a given censorship strategy it checks whether profitable deviations

[1]Thus, the citizen wants to support the government when his belief in the government's competence outweighs his dislike of the government's ideology. This is in line with models of expressive voting (Brennan and Hamlin, 1998), in which voters, irrespective of the outcome of the vote, derive intrinsic utility from voting for an alternative that is correct from their point of view.

[2]See the working paper version of Kolotilin et al. (2017) for an application of Bayesian persuasion to government censorship of the media.

exist.

Intuitively, because the sender is restricted to either revealing the state perfectly or censoring it, her persuasion strategy is fully described by a set S of states that are censored. For that set to be optimal, the sender must be unwilling to reveal any state that belongs to S, or to censor any state that does not belong to S.[3] The change in the sender's expected payoff resulting from such deviations is driven by two factors. First, whether a state is censored or not affects the sender's expected payoff when Nature draws that state. This happens because if the state is revealed, the receiver acts if and only if his type is below the state; while if it is censored, the receiver acts if and only if his type is below the expected value of the state conditional on it being in S. Second, whether or not a state is censored changes the aforementioned expectation, and hence affects the sender's payoff in the event *any* state is censored. The magnitude of these two effects depends on the shape of the distribution of the receiver's type – hence, that distribution affects the optimal censorship strategy. This logic underlies Proposition 1, which establishes a necessary condition for a censorship strategy to be optimal.

While this condition is only necessary and not sufficient, it substantially reduces the set of potentially optimal strategies. The rest of the paper shows that this result is sufficiently powerful to gain several new insights.

First, I demonstrate how Proposition 1 can be used to restrict the set of optimal strategies. In general, S can be "complex", consisting of many disjoint intervals. However, Proposition 2 shows that at the optimum, this "complexity" of S is bounded by the number of peaks of the density of the receiver's type.

Second, I analyse the simple case when the density of the receiver's type is single-peaked. In this case, censorship is known to be an unconstrained optimal persuasion strategy[4]. The result of Proposition 1 can then be used to describe the effect of a sufficiently large shift in the location of the peak, and in the distribution of the state. Specifically, Propositions 3 and 4 show that the sender censors more states when the peak decreases, or when the expected state increases.

Third, I examine the case when the density of type is bimodal. In that case, censorship is not, in general, an optimal strategy[5]. But what if the

[3]More precisely, the paper analyses marginal changes in the sender's payoff that occur if she reveals or censors an infinitesimally small interval of states.

[4]See Alonso and Câmara (2016b) and Kolotilin (2018).

[5]See Proposition 3 in Kolotilin (2018).

sender is restricted to censorship strategies? In Propositions 5 and 6, I characterise optimal censorship policies for different classes of bimodal distributions. Specifically, I show that depending on the shape of the bimodal density, the sender either censors intermediate states while revealing high and low states, or reveals intermediate states while censoring high and/or low states.

This paper belongs to the growing literature on Bayesian persuasion with linear utilities (see, for example, Gentzkow and Kamenica, 2016; Kolotilin and Zapechelnyuk, 2019; Kolotilin and Li, 2019; Dworczak and Martini, forthcoming). In particular, Kolotilin et al. (2017) and Kolotilin (2018) analyse linear persuasion in a setting in which, as in this paper, the receiver has private information. Alonso and Câmara (2016b) also study the case when the receiver is privately informed. More generally, a number of papers have studied Bayesian persuasion of a group of heterogeneous receivers[6]. My paper differs from the rest of the literature by focusing on censorship strategies, and using an optimisation approach that checks for the existence of profitable deviations[7].

The rest of the paper is structured as follows. Section 2 outlines the model. Section 3 derives the direct optimisation approach to finding the optimal censorship policy. Section 4 shows how that approach can be used to gain new insights about optimal censorship. Section 5 concludes. All proofs are in the Appendix.

2 Model

A sender (she) is facing a receiver (he). The receiver has a type $t \in [0, 1]$. The type is drawn from a distribution G with a continuously differentiable density g. There is a state of the world $\omega \in [0, 1]$, drawn by Nature from a smooth distribution F with a strictly positive density f.

The receiver can choose action $a \in \{0, 1\}$. I will say that the receiver "acts" if he chooses action 1. The sender's payoff equals a – thus, the sender

[6]See Taneva (2018). In addition, Alonso and Câmara (2016a), Wang (2013), Bardhi and Guo (2018), and Chan et al. (2018) study Bayesian persuasion of a heterogeneous group of voters.

[7]In contrast, Kolotilin (2018) as well as Dworczak and Martini (forthcoming) use duality approach, while Gentzkow and Kamenica (2016), Alonso and Câmara (2016b), and Kolotilin et al. (2017) use concavification.

aims to maximise the probability that the receiver acts. If the receiver does not act, his payoff equals 0. If he acts, his payoff equals $\omega - t$. Thus, the receiver wants to act if and only if his type is lower than the state.

At the beginning of the game, Nature draws t from G; the receiver is informed about t. Next, the sender selects a set of states $S \subseteq [0,1]$ that are censored, i.e. not revealed to the receiver. I will refer to S as the sender's *censorship strategy*. For tractability, I will assume that S has a finite number of boundary points. Furthermore, I will assume that every boundary point of S is either an upper or a lower boundary point[8], but not both – this means that S does not contain any "unattached" points[9]. Thus, either $S = \emptyset$, or $S = \bigcup_{i=1}^{n} [p_i, q_i]$ such that $0 \leq p_i < q_i < p_{i+1} \leq 1, \forall i = 1, .., n$ for some integer n. After the sender has chosen S, Nature draws the state ω from F. Next, if $\omega \notin S$, the receiver learns it; otherwise, he updates his beliefs. He then chooses action $a \in \{0, 1\}$. Finally, payoffs are realised.

3 Optimisation Approach

This section will derive the optimisation approach to finding the optimal censorship strategy.

Suppose the sender has chosen some S. Then if $\omega \notin S$, the receiver learns the state. He then acts if $t < \omega$, and does not act if $t > \omega$. The sender's expected payoff thus equals $G(\omega)$. If S is nonempty and $\omega \in S$, the receiver's payoff from acting equals $\mathrm{E}_F[\omega - t \mid \omega \in S]$. Thus, the receiver acts if and only if $t < t_S$, where $t_S \equiv \mathrm{E}_F[\omega \mid \omega \in S]$. The sender's expected payoff then equals $G(t_S)$.

Given S, let $v(S)$ be the sender's expected payoff. It then equals

$$v(S) = \int_{\omega \notin S} G(\omega)\, dF(\omega) + \mu_S G(t_S)$$

[8] Formally, state w is a lower boundary point of S if there exists some $\varepsilon > 0$ such that all $\omega \in (w - \varepsilon, w)$ are outside S and all $\omega \in (w, w + \varepsilon)$ belong to S. A state w is an upper boundary point of S if there exists some $\varepsilon > 0$ such that all $\omega \in (w - \varepsilon, w)$ belong to S and all $\omega \in (w, w + \varepsilon)$ are outside S.

[9] This assumption is without loss of generality, because if there were such points, their total mass would be zero (as the number of boundary points of S is assumed to be finite), so a censorship strategy that contains such points is payoff-equivalent to another censorship strategy that does not contain them.

where $\mu_S \equiv \int_{\omega \in S} dF(\omega)$ is the probability that the state falls in S.[10]

The sender chooses S to maximise $v(S)$. We can consider the following deviations: first, the sender may deviate to censoring a small interval of states around some $\omega \notin S$; second, she can deviate to revealing a small interval of states around some $\omega \in S$. If S is optimal, the sender must not gain from such deviations. In particular, the change in her payoff from the deviation should be negative as the width of the interval converges to zero.

This logic underlies the key result of the paper: a necessary condition for S to constitute an optimal censorship strategy. It is summarised in the following proposition:

Proposition 1. *Suppose that S maximises $v(\cdot)$. Then*

- $z_S(\omega) \geq 0$ *for any* $\omega \in S$; *and*
- $z_S(\omega) \leq 0$ *for any* $\omega \notin S$,

where $z_S(\omega) \equiv G(t_S) - G(\omega) + (\omega - t_S)\, g(t_S)$.

To see the intuition, consider a state $\omega \notin S$. Suppose the sender deviates to censoring ω (that is, pools it with S). This will have two effects on her payoff. First, whenever ω is drawn, the sender will now receive $G(t_S)$ instead of $G(\omega)$. Second, pooling ω with S will move t_S towards ω, which will change the sender's payoff every time a state is censored. The magnitude of the shift in t_S is proportional to the distance between ω and t_S, while the marginal effect of shifting t_S on the sender's payoff equals $g(t_S)$, i.e. the slope of G at t_S. Hence, the marginal change in the sender's payoff from censoring ω instead of revealing it is proportional to

$$G(t_S) - G(\omega) + (\omega - t_S)\, g(t_S) = z_S(\omega)$$

If the initially chosen S is optimal, this must be weakly negative. By similar logic, $z_S(\omega)$ must be weakly positive at any $\omega \in S$.

Since $z_S(\omega)$ is continuous, Proposition 1 implies that at the equilibrium, any boundary point of S must be a state ω at which $z_S(\omega) = 0$, that is, at which $g(t_S) = \frac{G(\omega)-G(t_S)}{\omega - t_S}$.[11] Since $g(t_S)$ is the slope of G at t_S, any boundary

[10] Technically, t_S is only defined when S is nonempty. If $S = \emptyset$, then $v(S) = \int_0^1 G(\omega)\, dF(\omega)$.

[11] Technically, the condition $g(t_S) = \frac{G(\omega)-G(t_S)}{\omega - t_S}$ is only defined for $\omega \neq t_S$. It is also possible for t_S to be at the boundary of S.

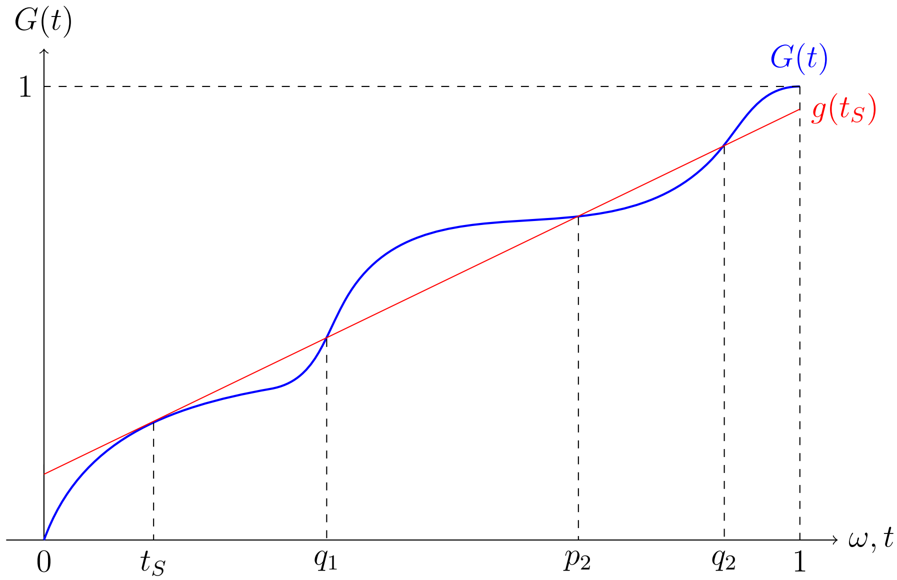

Figure 1: Optimal censorship. Here, $S = [0, q_1] \cup [p_2, q_2]$.

point of S must be a point at which a line that is tangent to G at t_S intersects G. At a given state ω, if G is below (above) that tangent line, then $g(t_S)$ is greater (smaller) than $\frac{G(\omega)-G(t_S)}{\omega - t_S}$, and hence ω is censored (revealed). Figure 1 illustrates this result.

Proposition 1 provides a necessary condition for a censorship strategy to be optimal. The condition links the existence of profitable deviations from S to the shape of G. Note that this condition is not sufficient – it may still be optimal for the sender to deviate at a positive-measure subset of states. Nevertheless, this result can be used to analyse optimal censorship strategies. This is shown in the next section.

4 Optimal Censorship Strategies

This section will show how the optimisation approach derived previously can be used to derive optimal censorship strategies for various distributions of the receiver's type. I will start by showing how Proposition 1 can be used to restrict the set of optimal censorship strategies for a generic G. Then I will show how it can be applied to characterising optimal censorship strategies when G is unimodal or bimodal.

For the subsequent analysis, the following result[12] will be useful:

Lemma 1. *$S = \emptyset$ is optimal if and only if G is convex.*

Hence, full revelation is optimal if and only if g is increasing on $[0, 1]$.

4.1 General distributions

Consider any distribution G. What can we say about the optimal censorship strategy? As described in the model, the set S of censored states is either $S = \emptyset$, or $S = \bigcup_{i=1}^{n} [p_i, q_i]$. Hence, S can include any finite number n of disjoint intervals. That n can be large – that is, the censorship strategy can be "complex". The next result will show, however, that the complexity of S is bounded by the complexity of the density of the receiver's type.

Referring to Figure 1, Proposition 1 implies that G must be flatter than $g(t_S)$ at each interior p_i, and steeper than $g(t_S)$ at each interior q_i. Hence, for all interior boundary points p_i and q_i, G needs to be convex on some interval within $[p_i, q_i]$, and concave on some interval within $[q_i, p_{i+1}]$. Thus, for optimal S to include many disjoint intervals $[p_i, q_i]$, G must contain a sufficiently large number of alternating convex and concave sections. A point at which G changes from being convex to being concave is a local maximum of g. Hence, for n to be large, g must have a large number of local maxima. This intuition underlies the following result:

Proposition 2. *If g has $m < \infty$ local weak maxima, then $S = \bigcup_{i=1}^{n} [p_i, q_i]$ with $n \leq m + 1$.*

Hence, the optimal S cannot consist of more disjoint intervals than the number of peaks of g plus 1. Relatively "simple" distributions of the receiver's types induce relatively "simple" censorship strategies.

Note that S is fully characterised by a vector of its n upper boundary points and n lower boundary points. Proposition 2 then implies that the optimal censorship strategy is a result of a $2(m + 1)$-variable optimisation problem. This result can be compared to Theorem 2 in Kolotilin et al. (2017), which (in the case when the sender is not restricted to censorship strategies)

[12] Kolotilin (2018) shows a similar result when the sender is not restricted to censorship strategies.

shows that the sender's choice of an optimal receiver's interim utility is (in the case when the sender is only interested in the receiver's action) a result of an m-variable optimisation problem.

4.2 Unimodal distributions

Suppose that g has a unique peak k. In that case, Kolotilin et al. (2017) and Kolotilin (2018) show that censorship is the optimal persuasion strategy when the sender is not restricted to censorship strategies. Specifically, the sender chooses upper-censorship (pooling together all states above a certain cutoff, and perfectly revealing all other states). Of course, upper-censorship is also optimal in my setup; the following lemma derives the result from Kolotilin et al. (2017) and Kolotilin (2018) using the approach developed in Section 3:

Lemma 2. *Suppose G is convex on $(0, k)$ and concave on $(k, 1)$ for some $k \in (0, 1)$. Then there exists a unique optimal censorship strategy $S = [p, 1]$, such that $0 \leq p < k$, and $k < t_{[p,1]} < 1$.*

The lower boundary of S is a point p at which $z_{[p,1]}(p) = 0$. If $p > 0$, then Proposition 1 implies that p is given by the condition $z_{[p,1]}(p) = 0$. Graphically, p it is the point at which the line that is tangent to G at t_S intersects G. It is also possible for the tangent line never to intersect G. This happens when $z_{[0,1]}(0) \geq 0$. In that case, we have a corner solution in which $p = 0$, and $S = [0, 1]$.

Since Proposition 1 describes the necessary condition for S to be optimal based on the curvature of G, Lemma 2 ensures that at the optimum, $p < k < t_S$. Hence, the optimisation approach of this paper implies that a change in G that shifts k, or a change in F that shifts t_S, induce a change of the optimal censorship strategy if the shift is sufficiently large.

In particular, if k moves far enough that it ends up below p or above t_S, the optimal S has to change as well, as the following result shows:

Proposition 3. *Take a unimodal distribution G with mode k that induces a censorship policy $S = [p, 1]$. Take another distribution $\hat{G}$ with mode $\hat{k}$. If $\hat{k} > t_S$, then $\hat{G}$ induces a censorship policy $\hat{S}$ such that $\hat{S} \subset S$. If $\hat{k} < p$, then $\hat{G}$ induces a censorship policy $\hat{S}$ such that $S \subset \hat{S}$.*

Hence, the sender censors more (less) states if the modal receiver becomes more (less) willing to act.

Similarly, suppose that F is replaced by another distribution $\hat{F}$ that puts a larger mass of states to the left. If the shift in F is sufficiently strong, it affects the optimal censorship strategy, as the following result shows.

Proposition 4. *Take a unimodal distribution G with mode k. Consider a distribution of states F, which induces a censorship policy $S = [p, 1]$. Take another distribution $\hat{F}$. If $E_{\hat{F}} [\omega \mid \omega > p] < k$, then $\hat{F}$ induces a censorship policy $\hat{S} \subset S$.*

Hence, when the state tends to be worse (better) from the sender's point of view, the optimal censorship strategy is less (more) restrictive.

4.3 Bimodal distributions

Suppose that G is bimodal. We can look at two classes of bimodal distributions.

First, suppose that for some $\underline{k}, \overline{k}$ such that $\underline{k} < \overline{k}$, g is increasing on $(0, \underline{k})$, decreasing on $\left(\underline{k}, \overline{k}\right)$, and increasing on $\left(\overline{k}, 1\right)$. Then we have the following result:

Proposition 5. *Suppose G is convex on $(0, \underline{k})$, concave on $\left(\underline{k}, \overline{k}\right)$, and convex on $\left(\overline{k}, 1\right)$ for some $\underline{k}, \overline{k}$ such that $0 < \underline{k} < \overline{k} < 1$. Then the optimal censorship strategy is $S = [p, q]$, where $0 \leq p < q \leq 1$ and $\underline{k} < t_{[p,q]} < \overline{k}$.*

In words, the sender censors states over some intermediate interval $[p, q] \subseteq [0, 1]$. This is illustrated in Figure 2.

Intuitively, if $t_S \leq \underline{k}$, then for all $\omega < t_S$, $G(\omega)$ lies above the line that is tangent to it at t_S. Then by Proposition 1 all of these states have to be revealed, which is impossible, since $t_S \equiv \mathrm{E}_F [\omega \mid \omega \in S]$. By similar reasoning, we cannot have $t_S \geq \overline{k}$. Hence, at the optimum, $t_S \in \left(\underline{k}, \overline{k}\right)$. The boundaries p and q of S are then the points at which the tangent line crosses G. Depending on the shapes of F and G, it is possible that the tangent line only crosses G once, or never – in that case, $p = 0$ and/or $q = 1$.

Next, consider a different class of bimodal distributions. Suppose that for some $\underline{k}, \overline{k}$ such that $\underline{k} < \overline{k}$, g is decreasing on $(0, \underline{k})$, increasing on $\left(\underline{k}, \overline{k}\right)$, and decreasing on $\left(\overline{k}, 1\right)$. For these distributions, Kolotilin (2018) shows that the optimal persuasion strategy is interval revelation: the sender sends one message for all states that are sufficiently low, another message for all states that are sufficiently high, and perfectly reveals all intermediate states.

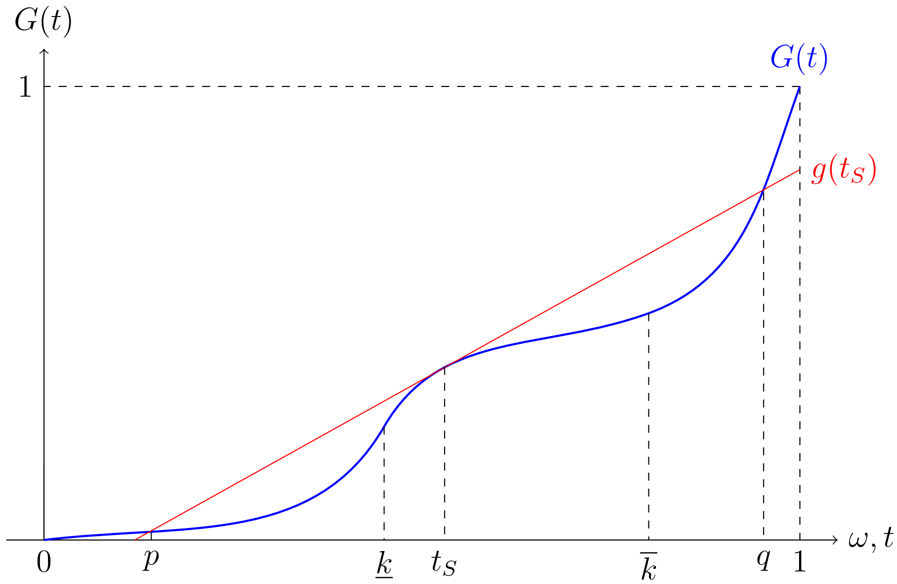

Figure 2: Optimal censorship under bimodal (convex-concave-convex) G.

That strategy, however, is not a censorship strategy. What if the sender is restricted to censorship strategies – that is, if she has to send the same message for all states that are not perfectly revealed? The next result characterises such a constrained optimal persuasion strategy.

Proposition 6. *Suppose G is concave on $(0, \underline{k})$, convex on $\left(\underline{k}, \overline{k}\right)$, and concave on $\left(\overline{k}, 1\right)$ for some $\underline{k}, \overline{k}$ such that $0 < \underline{k} < \overline{k} < 1$. Then the optimal censorship strategy is either (i) $S = [0, q]$; or (ii) $S = [p, 1]$; or (iii) $S = [0, q] \cup [p, 1]$; where $0 < q \leq p < 1$.*

In words, the optimal censorship strategy includes cutoffs p and q such that the sender reveals all states in the $[q, p]$ interval, and censors all states below q and above p. The location of these cutoffs depends on the shapes of F and G. In particular, there may be a corner solution in which either $p = 1$ or $q = 0$ (this corresponds to, respectively, cases (i) and (ii) in the proposition)[13]. It is also possible to have $p = q$ – this implies that $S = [0, 1]$, and thus no states are revealed.

Figure 3 illustrates Proposition 6. Intuitively, depending on the shapes of F and G, there are three possibilities. First, it is possible that $t_S < \underline{k}$.

[13]But note that Lemma 1 ensures that we cannot have both $q = 0$ and $p = 1$.

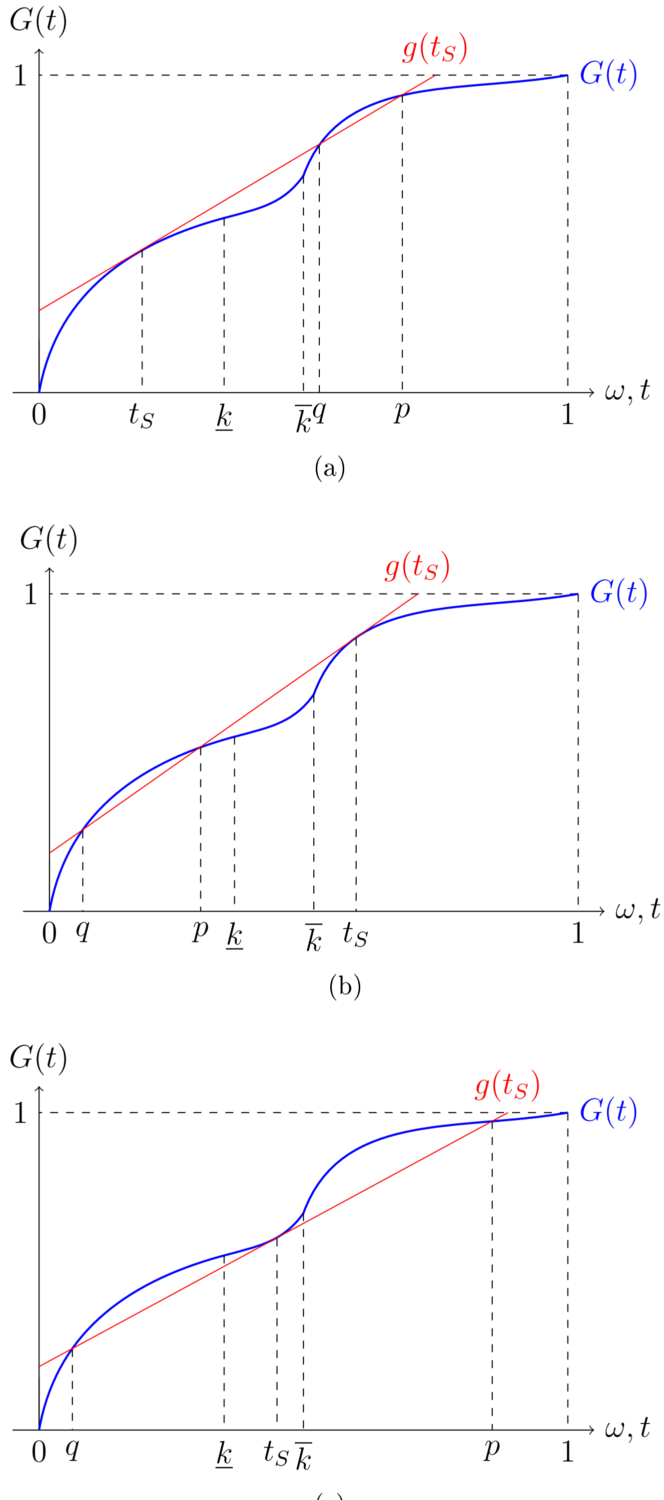

Figure 3: Optimal censorship under bimodal (concave-convex-concave) G.

Then all $\omega \in [0, \underline{k}]$ lie below the line that is tangent to G at t_S; hence, by Proposition 1 these states are censored. The tangent line can then intersect G at most twice. If it does intersect G twice, at states q and p, then all states $\omega \in (q, p)$ are revealed, and the rest are censored (this is shown in Figure 3a). If the tangent line intersects G once, at some point q, then $S = [0, q]$. Finally, if it never intersects G, then $S = [0, 1]$.

It is also possible to have $t_S > \overline{k}$. Then all $\omega \in [\overline{k}, 1]$ belong to S. For the rest of the state space, the tangent line can intersect G at most twice. If it does intersect G twice, at some states q and p, then all states $\omega \in (q, p)$ are revealed, and the rest are censored (this is shown in Figure 3b). If the tangent line intersects G once, then $S = [p, 1]$ for some p; and if it never intersects G, then $S = [0, 1]$.

Finally, if $t_S \in [\underline{k}, \overline{k}]$, then the tangent line can intersect G at most once on $[0, \underline{k}]$, and at most once on $[\overline{k}, 1]$. If the tangent line intersects G on one of these intervals only, then $t_S \notin S$, which is impossible. If the tangent line does not intersect G, then $S = \emptyset$, which for bimodal G is ruled out by Lemma 1. Hence, the tangent line must intersect G at some point $q \in [0, \underline{k}]$, and at some point $p \in [\overline{k}, 1]$. Then $S = [0, q] \cup [p, 1]$ (this is shown in Figure 3c).

5 Conclusions

In many persuasion settings, the sender is restricted to censorship strategies: she can either reveal a state of the world perfectly, or hide it. This paper has examined optimal censorship in a linear setting with a privately informed receiver. Its main contribution was in developing a simple optimisation approach, described in Proposition 1, to analysing optimal censorship.

The optimisation approach produces a condition for a censorship strategy to be optimal. While this condition is only a necessary and not a sufficient condition, the paper shows how it can be used to gain insights about optimal censorship strategies in a number of situations.

References

Alonso, Ricardo and Odilon Câmara, "Persuading voters," *The American Economic Review*, 2016a, *106* (11), 3590–3605.

— **and** — , "Political disagreement and information in elections," *Games and Economic Behavior*, 2016b, *100*, 390–412.

Bardhi, Arjada and Yingni Guo, "Modes of persuasion toward unanimous consent," *Theoretical Economics*, 2018, *13* (3), 1111–1150.

Brennan, Geoffrey and Alan Hamlin, "Expressive voting and electoral equilibrium," *Public Choice*, 1998, *95* (1-2), 149–175.

Chan, Jimmy, Seher Gupta, Fei Li, and Yun Wang, "Pivotal persuasion," 2018.

Dworczak, Piotr and Giorgio Martini, "The simple economics of optimal persuasion," *Journal of Political Economy*, forthcoming.

Gentzkow, Matthew and Emir Kamenica, "A Rothschild-Stiglitz approach to Bayesian persuasion," *American Economic Review*, 2016, *106* (5), 597–601.

Kolotilin, Anton, "Optimal information disclosure: A linear programming approach," *Theoretical Economics*, 2018, *13* (2), 607–635.

— **and Andriy Zapechelnyuk**, "Persuasion meets delegation," 2019.

— **and Hongyi Li**, "Relational Communication," 2019.

— **, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li**, "Persuasion of a Privately Informed Receiver," *Econometrica*, 2017, *85* (6), 1949–1964.

Taneva, Ina A, "Information design," 2018.

Wang, Yun, "Bayesian persuasion with multiple receivers," 2013.

6 Appendix

Proof of Proposition 1. To prove the first part, take a set S. Now take a state w belonging to the interior of S and suppose that $z_S(w) < 0$.[14]

[14] If w is on the boundary of S, take instead another state w' in the neighbourhood of w that lies in the interior of S such that $z_S(w') < 0$. Such a state must exist because $z_S(\cdot)$ is continuous.

Consider a deviation from S to $\hat{S} = S \setminus [w, r]$ for some $r > w$. Let $L(w, r) \equiv v\left(\hat{S}\right) - v(S)$. Then $L(w, r) = G\left(t_{S\setminus[w,r]}\right)\mu_{S\setminus[w,r]} + \int_w^r G(\omega)dF(\omega) - G(t_S)\mu_S$. If $r = w$, then $L(w, r) = 0$. For S to be optimal, $L(w, r)$ must be weakly decreasing in r at $r = w$. Differentiating yields:

$$\frac{\partial L(w,r)}{\partial r} = g\left(t_{S\setminus[w,r]}\right)\frac{\partial t_{S\setminus[w,r]}}{\partial r}\mu_{S\setminus[w,r]} + G\left(t_{S\setminus[w,r]}\right)\frac{\partial \mu_{S\setminus[w,r]}}{\partial r} + G(r) f(r)$$

Hence,

$$\left.\frac{\partial L(w,r)}{\partial r}\right|_{r=w} = g(t_S)\left.\frac{\partial t_{S\setminus[w,r]}}{\partial r}\right|_{r=w}\mu_S + G(t_S)\left.\frac{\partial \mu_{S\setminus[w,r]}}{\partial r}\right|_{r=w} + G(w) f(w)$$

Note that

$$t_{S\setminus[w,r]} = \frac{\int_{\omega\in S}\omega dF(\omega) - \int_w^r \omega dF(\omega)}{\int_{\omega\in S} dF(\omega) - \int_w^r dF(\omega)} = \frac{\int_{\omega\in S}\omega dF(\omega) - \int_w^r \omega dF(\omega)}{\mu_{S\setminus[w,r]}}$$

and thus

$$\left.\frac{\partial t_{S\setminus[w,r]}}{\partial r}\right|_{r=w} = \left.\frac{-\mu_{S\setminus[w,r]} r f(r) + t_{S\setminus[w,r]}\mu_{S\setminus[w,r]} f(r)}{\mu^2_{S\setminus[w,r]}}\right|_{r=w} = \frac{f(w)}{\mu_S}(t_S - w)$$

Also,

$$\mu_{S\setminus[w,r]} = \int_{\omega\in S} dF(\omega) - \int_w^r dF(\omega)$$

and thus

$$\left.\frac{\partial \mu_{S\setminus[w,r]}}{\partial r}\right|_{r=w} = -f(w)$$

Therefore,

$$\left.\frac{\partial L(w,r)}{\partial r}\right|_{r=w} = g(t_S) f(w)(t_S - w) - G(t_S) f(w) + G(w)f(w) = -f(w) z_S(w)$$

Since f is strictly positive everywhere, the derivative is strictly positive when $z_S(w) < 0$, so S is not optimal.

The second part is proved analogously. Suppose that $z_S(w) > 0$ for some $w \notin S$. Now take some interval $[w, r]$ such that $[w, r] \cap S = \emptyset$, and consider a deviation from S to $\hat{S} = S \cup [w, r]$. If S is optimal, then $L(w, r) = v\left(\hat{S}\right) - v(S)$ must be weakly decreasing at $r = w$. Differentiating yields

$$\left.\frac{\partial L(w,r)}{\partial r}\right|_{r=w} = f(w) z_S(w) > 0$$

so again S is not optimal. $\square$

Proof of Lemma 1. To show that full revelation is optimal when g is increasing, suppose that S is nonempty. Then $z_S(\omega) > 0$ for all ω, so S is not optimal. To prove the second part of the lemma, suppose that $S = \emptyset$, and that G is not convex. Then G must be concave on some interval $[p, q]$. Consider a deviation to to $S = [p, q]$. The change in the sender's payoff from such deviation equals

$$v([p,q]) - v(\emptyset) = \mu_{[p,q]} [G(\mathrm{E}_F [\omega \mid \omega \in [p,q]]) - \mathrm{E}_F (G(\omega) \mid \omega \in [p,q])]$$

This is positive by Jensen's inequality, and hence $S = \emptyset$ is not optimal. $\square$

Proof of Proposition 2. Proposition 1 implies that at every p_i for $i \in \{2, .., n\}$, G crosses the line is tangent to G at t_S from above. Furthermore, at every q_i for $i \in \{1, .., n-1\}$, G crosses the line is tangent to G at t_S from below (note that we exclude p_1 and q_n because there may be a corner solution with $p_1 = 0$ or $q_n = 1$). Hence, $g(p_i) < g(t_S)$ for all $i \in \{2, .., n\}$; and $g(q_i) > g(t_S)$ for all $i \in \{1, .., n-1\}$. Hence, for all $i \in \{2, .., n-1\}$, we have $g(p_i) < g(q_i)$ and $g(p_{i+1}) < g(q_i)$, with $p_i < q_i < p_{i+1}$. Thus, g must have a peak between p_i and p_{i+1} for all $i \in \{2, .., n-1\}$, which gives us $n-2$ peaks. In addition, since $g(q_1) > g(t_S) > g(p_2)$, there must be another peak of g between 0 and p_2. Hence, g must have at least $n-1$ peaks. $\square$

Proof of Lemma 2. Since the optimal S is nonempty by Lemma 1, t_S is well-defined. Suppose that $t_S \leq k$. Then, as g is increasing on $[0, t_S]$, we have $z_S(\omega) = \int_\omega^{t_S} [g(x) - g(t_S)]\, dx < 0$ for all $\omega < t_S$. Then if S is optimal, all $\omega < t_S$ do not belong to S. This cannot hold, as $t_S \equiv \mathrm{E}[\omega \mid \omega \in S]$. Hence, at the optimum, $t_S > k$.

Then $z_S(\omega) > 0$ for all $\omega \geq k$, so $[k, 1] \subseteq S$. Note that $\frac{dz_S(\omega)}{d\omega} = g(t_S) - g(\omega)$. This is negative at $\omega = k$, and since g is monotone increasing on $[0, k]$, $\frac{dz_S(\omega)}{d\omega}$ changes sign at most once on that interval. Hence, on that interval there is at most one state at which z_S crosses zero. If such a state p exists, it is the lower boundary of S. Otherwise, $z_S(\omega) \geq 0$ for all $\omega \in [0, k]$, so $S = [0, 1]$.

To show uniqueness, suppose on the contrary that there exist p and $\tilde{p}$ such that $\tilde{p} < p$, and both $[p, 1]$ and $[\tilde{p}, 1]$ are optimal censorship strategies. Then $z_{[\tilde{p},1]}(\tilde{p}) \geq 0$, and $z_{[p,1]}(p) = 0$, the latter because $p > 0$. Thus,

$$g\left(t_{[p,1]}\right) = \frac{G\left(t_{[p,1]}\right) - G(p)}{t_{[p,1]} - p}$$

The derivative of the left-hand side of this equation with respect to p equals $g'\left(t_{[p,1]}\right)\frac{dt_{[p,1]}}{dp}$. This is negative, since g is decreasing at $t_{[p,1]}$. The derivative of the right-hand side equals

$$\begin{aligned}
&\frac{\left[t_{[p,1]}-p\right]\left[g\left(t_{[p,1]}\right)\frac{dt_{[p,1]}}{dp}-g\left(p\right)\right]-\left[G\left(t_{[p,1]}\right)-G\left(p\right)\right]\left[\frac{dt_{[p,1]}}{dp}-1\right]}{\left[t_{[p,1]}-p\right]^2}\\
=&\frac{g\left(t_{[p,1]}\right)\frac{dt_{[p,1]}}{dp}-g\left(p\right)}{t_{[p,1]}-p}-g\left(t_{[p,1]}\right)\frac{\frac{dt_{[p,1]}}{dp}-1}{t_{[p,1]}-p}\\
=&\frac{g\left(t_{[p,1]}\right)-g\left(p\right)}{t_{[p,1]}-p}\\
>&0
\end{aligned}$$

where the inequality follows from the fact that $t_{[p,1]}>p$ and $g\left(t_{[p,1]}\right)>g\left(p\right)$ (the latter is because the line that is tangent to G at $t_{[p,1]}$ crosses G from below at p). Hence decreasing p increases the left-hand side while decreasing the right-hand side. Thus,

$$g\left(t_{[\tilde{p},1]}\right)>\frac{G\left(t_{[\tilde{p},1]}\right)-G\left(\tilde{p}\right)}{t_{[\tilde{p},1]}-\tilde{p}}$$

which implies that $z_{[\tilde{p},1]}\left(\tilde{p}\right)=G\left(t_{[\tilde{p},1]}\right)-G\left(\tilde{p}\right)-\left(t_{[\tilde{p},1]}-\tilde{p}\right)g\left(t_{[\tilde{p},1]}\right)<0$, so $[\tilde{p},1]$ cannot be an optimal strategy. $\square$

Proof of Proposition 3. By Lemma 2, $\hat{G}$ must induce a censorship policy $\hat{S}=[\hat{p},1]$ such that $t_{\hat{S}}>\hat{k}$ and $\hat{p}<\hat{k}$. If $\hat{k}>t_S$, this implies that $t_{\hat{S}}>t_S$. Hence, $\hat{p}>p$, so $\hat{S}\subset S$. On the other hand, if $\hat{k}<p$, then $\hat{p}<\hat{k}<p$, so $S\subset\hat{S}$. $\square$

Proof of Proposition 4. By Lemma 2, G and $\hat{F}$ must induce a censorship policy $\hat{S}=[\hat{p},1]$ such that $\mathrm{E}_{\hat{F}}\left[\omega\mid\omega>\hat{p}\right]>k$. Since $\mathrm{E}_{\hat{F}}\left[\omega\mid\omega>p\right]<k$, it must be that $\hat{p}>p$, so $\hat{S}=[\hat{p},1]\subset S$. $\square$

Proof of Proposition 5. Since the optimal S is nonempty by Lemma 1, t_S is well-defined. If $t_S\leq\underline{k}$, then, as g is increasing on $[0,\underline{k}]$, we have

$$z_S\left(\omega\right)=\int_{\omega}^{t_S}\left[g\left(x\right)-g\left(t_S\right)\right]dx<0$$

for all $\omega < t_S$. Then for S to be optimal, all $\omega < t_S$ must not belong to S, which cannot hold. Similarly, if $t_S \geq \overline{k}$, then, as g is increasing on $\left[\overline{k}, 1\right]$, we have $z_S(\omega) < 0$ for all $\omega > t_S$, which implies that S is not optimal. Hence, $t_S \in \left(\underline{k}, \overline{k}\right)$. Then $z_S(\omega) > 0$ for all $\omega \in \left(\underline{k}, \overline{k}\right)$, hence $\left(\underline{k}, \overline{k}\right) \subset S$.

Then $\frac{dz_S(\omega)}{d\omega} = g(t_S) - g(\omega)$ is negative at $\omega = \underline{k}$. Since g is monotone increasing on $[0, \underline{k}]$, $\frac{dz_S(\omega)}{d\omega}$ changes sign at most once on that interval. Hence, on that interval there is at most one state at which z_S crosses zero. If such a state p exists, it is the lower boundary of S. Otherwise, $z_S(\omega) \geq 0$ for all $\omega \in [0, \underline{k}]$, so $p = 0$.

Similarly, $\frac{dz_S(\omega)}{d\omega} = g(t_S) - g(\omega)$ is positive at $\omega = \overline{k}$. Since g is monotone increasing on $\left[\overline{k}, 1\right]$, $\frac{dz_S(\omega)}{d\omega}$ changes sign at most once on that interval. Hence, on that interval there is at most one state at which z_S crosses zero. If such a state q exists, it is the upper boundary of S. Otherwise, $z_S(\omega) \geq 0$ for all $\omega \in \left[\overline{k}, 1\right]$, so $q = 1$. $\square$

Proof of Proposition 6. Since the optimal S is nonempty by Lemma 1, t_S is well-defined. Take some S, and suppose it is optimal. There are three possibilities: $t_S < \underline{k}$; $t_S > \overline{k}$; and $t_S \in \left[\underline{k}, \overline{k}\right]$.

If $t_S < \underline{k}$, then, as g is decreasing on $[0, \underline{k}]$, we have

$$z_S(\omega) = \int_{\omega}^{t_S} [g(x) - g(t_S)]\, dx > 0$$

for all $\omega \leq \underline{k}$. Hence, $[0, \underline{k}] \subseteq S$. Note that $\frac{dz_S(\omega)}{d\omega} = g(t_S) - g(\omega)$ is positive at $\omega = \underline{k}$, hence $z_S(\omega)$ is increasing at $\underline{k}$. Since g is increasing on $\left[\underline{k}, \overline{k}\right]$ and decreasing on $\left[\overline{k}, 1\right]$, $\frac{dz_S(\omega)}{d\omega}$ changes sign at most twice on $[\underline{k}, 1]$. Hence, on that interval there are at most two states at which $z_S(\omega)$ crosses zero. If there are two such states, call them q and p, and then $S = [0, q] \cup [p, 1]$. If there is one such state, call it q, and then $S = [0, q]$. Finally, if there are no such states, then $S = [0, q] \cup [p, 1]$ with $q = p$.

If $t_S > \overline{k}$, then, as g is decreasing on $\left[\overline{k}, 1\right]$, we have $z_S(\omega) > 0$ for all $\omega \geq \overline{k}$. Hence, $\left[\overline{k}, 1\right] \subseteq S$. Note that $\frac{dz_S(\omega)}{d\omega} = g(t_S) - g(\omega)$ is negative at $\omega = \overline{k}$, hence $z_S(\omega)$ is decreasing at $\overline{k}$. Since g is decreasing on $[0, \underline{k}]$ and increasing on $\left[\underline{k}, \overline{k}\right]$, $\frac{dz_S(\omega)}{d\omega}$ changes sign at most twice on $\left[0, \overline{k}\right]$. Hence, on that interval there are at most two states at which $z_S(\omega)$ crosses zero. If there are two such states, call them q and p, and then $S = [0, q] \cup [p, 1]$. If there is one such state, call it p, and then $S = [p, 1]$. Finally, if there are no such states, then $S = [0, q] \cup [p, 1]$ with $q = p$.

If $t_S \in \left[\underline{k}, \overline{k}\right]$, then, as g is increasing on $\left[\underline{k}, \overline{k}\right]$, we have $z_S(\omega) < 0$ for all $\omega \in \left(\underline{k}, \overline{k}\right)$. Hence, the interval $\left(\underline{k}, \overline{k}\right)$ does not belong to S. Note that $\frac{dz_S(\omega)}{d\omega} = g(t_S) - g(\omega)$ is positive at $\omega = \underline{k}$, and increasing on $[0, \underline{k}]$. Also, $\frac{dz_S(\omega)}{d\omega} = g(t_S) - g(\omega)$ is negative at $\omega = \overline{k}$, and increasing on $\left[\overline{k}, 1\right]$. Hence, $z_S(\omega)$ crosses zero at most once on the $[0, \underline{k}]$ interval, and at most once on the $\left[\overline{k}, 1\right]$ interval. If $z_S(\omega)$ does not cross zero anywhere, then $S = \emptyset$, which cannot be the case by Lemma 1. If $z_S(\omega)$ crosses zero on $[0, \underline{k}]$ but not on $\left[\overline{k}, 1\right]$, then $S = [0, q]$ for some $q \in [0, \underline{k}]$ – but in this case $t_S < \underline{k}$, which is a contradiction. Similarly, if $z_S(\omega)$ crosses zero on $\left[\overline{k}, 1\right]$ but not on $[0, \underline{k}]$, then $S = [p, 1]$ for some $p \in \left[\overline{k}, 1\right]$ – but in this case $t_S > \overline{k}$, which is a contradiction. Hence, $z_S(\omega)$ crosses zero exactly once on $[0, \underline{k}]$, and exactly once on $\left[\overline{k}, 1\right]$. Hence, $S = [0, q] \cup [p, 1]$ for some $q \in (0, \underline{k})$ and $p \in \left(\overline{k}, 1\right)$. $\square$