The Value of Information in Delegation

Maren Vairo†

June 7, 2023

Abstract

We study an optimal delegation problem in which the principal can jointly design a signal that may be used by the agent before making his decision (an information policy), together with the set of actions that the agent can choose from after observing the information disclosed by the principal (a delegation policy). Transfers are not allowed. We find that the optimal joint information and delegation policy features a double-censorship structure, in which the principal censors both the realizations of the state that she discloses and the actions available to the agent. We show that information and discretion serve as substitutes: the principal gives less discretion to the agent after realizations of the state that are fully disclosed, and vice versa. The intuition is that more discretion makes the agency problem more salient and thus lowers the principal’s gains from information disclosure. We apply these findings to a monopoly regulation setting where the regulator can jointly regulate the market segmentation (i.e., by limiting the extent to which the monopolist can engage in third-degree price discrimination), and the set of (segment-specific) prices that the monopolist can use.

*I am grateful to Eddie Dekel and Bruno Strulovici for their guidance and support; and to them, Piotr Dworczak, Yingni Guo, Alessandro Pavan, Udayan Vaidya, and Asher Wolinsky for helpful comments.
†Department of Economics, Northwestern University. Email: mvairo@u.northwestern.edu
1 Introduction

Consider a principal (she) and an agent (he) contracting over an action. The agent has private information that is relevant for determining the optimal course of action and therefore the principal benefits from delegating decision-making to him, but his preferences are misaligned with those of the principal, and monetary transfers are not feasible. In response to the agent’s bias, the principal may benefit from committing to limit the amount of discretion given to the agent, by delegating decision-making to him while optimally restricting the set of actions that the agent can choose from. Alternatively, the principal may be able to control the type of information that the agent can use when choosing an action, and in this way influence the agent’s decision-making. There is extensive work that studies the two respective problems—optimal delegation and information design—individually, but very little is known about the two of them in conjunction. Nonetheless, as we argue below, there are several practical situations (such as price regulation, or the legislation over college admissions) that feature a designer, who cannot resort to monetary transfers, jointly using these two tools.

This paper studies the principal’s problem of choosing the optimal joint information and delegation policies. Specifically, we consider a simple model in which there are two sources of payoff-relevant uncertainty. One of these sources, which we denote by $\omega$, is in principle unobserved by both players, but the principal can optimally reveal information about it by means of designing an experiment. The second source, denoted by $\theta$, is the agent’s private type, which is drawn from an exogenously given distribution and whose realization is only known by the agent. Both the principal’s and the agent’s payoff (and their respective optimal actions) depend on $\omega$ and $\theta$, and we assume the two to be independent random variables. Consequently, our model effectively captures scenarios in which: i) the agent has some specific expertise that the principal needs to elicit (e.g. his taste or horizontal type), and ii) the principal has control over the information that is made available about some other attribute of the environment that is independent of the agent’s type. We assume that the principal commits at the outset to an information policy, which maps realizations of $\omega$ into distributions over signals, and to a delegation policy which specifies, for every signal realization, the set of permissible actions that the agent can choose from. At the time of choosing an action, the agent knows his private type and the signal realization that was drawn from the principal’s information policy.

One application that fits this setting is that of a government that regulates college admis-
sions. In this application, the university (the agent) is better informed about the institutional
details that make an applicant well-suited to attend there. Because this information might
be valuable for the government, it justifies delegating admissions decisions to the university.
In many cases, the government intervenes by regulating the type of information about stu-
dents (e.g., gender and ethnicity) that colleges are allowed to use in their admissions process.
Furthermore, the government may limit the university’s discretion by placing restrictions on
the admission decisions that are allowed for every realization of this information (e.g., by im-
posing gender- or ethnicity-based quotas). Of course, the two policy interventions interact
with each other: Implementing a gender-based quota is only feasible if schools are allowed
to ask the student’s gender in their application form. Another motivating example, which
we will discuss in detail throughout the paper, is that of a monopoly regulator who jointly
controls the extent of price discrimination and the magnitude of prices that are allowed in
the market. Even though our model is very stylized and, to that extent, doesn’t capture the
full complexity of these environments, our results shed light on some of the important forces
that shape the optimal joint information and delegation structure.

When solving for the optimal policy, the principal is faced with the task of balancing
the following two key trade-offs. Firstly, for a fixed information structure, the principal who
designs a delegation policy faces a commitment vs. flexibility trade-off: Granting the agent
greater flexibility (i.e., a broader set of permissible actions) proves beneficial as it allows the
agent to condition his actions on $\theta$. However, it also gives rise to ex-post suboptimal actions
resulting from the agent’s bias. Secondly, for a fixed delegation set, the information designer
may benefit from using information about $\omega$ to steer the agent’s decision making. Nonethe-
less, these benefits are again limited by the extent of the agent’s misalignment. Individually,
each of these two trade-offs appear, respectively, in the standard (purely) delegation and
(purely) information design problems. There is a final third force that affects the principal’s
incentives for information provision in our environment, which is absent in each of these two
separate problems: The fact that the principal may condition the mechanism on the public
signals, renders information disclosure more valuable relative to the setting in which the
mechanism is fixed. In this paper, we study how these trade-offs interact with each other
when the principal uses information design and delegation in conjunction, and how these
interactions shape the optimal policy.

In our main result, we show that the optimal policy entails double-censorship, meaning
that both the delegation and the information policy can be described by a one-sided cen-
sorship cutoff. More specifically, under a regularity assumption and assuming for simplicity
that the agent is always biased towards higher actions, we show in Proposition 1 that the optimal delegation policy uses an action cap for every realized message about \( \omega \). To simplify the analysis, without loss of generality, we relabel the messages such that each message equals its induced posterior mean about \( \omega \), which we denote by \( x \). After this relabeling, the delegation policy is described by a cutoff on the \( \theta \)-space, \( \theta^*(x) \), such that all agents below the cutoff are able to choose their first-best actions, and all agents above it are pooled together at the same action. This cutoff, which is a measure of the amount of flexibility given to the agent, decreases with the expected size of the agent’s bias conditional on \( x \). Eventually, the agent’s bias may become sufficiently high that the principal chooses to give the agent no discretion in which case she ignores the private information possessed by the agent.

To understand this result note that, because the agent observes the signal before choosing an action, his incentive compatibility constraint has to hold for every realization of \( x \). This makes the principal’s delegation problem separable across values of \( x \), and as a consequence the optimal delegation policy can be obtained by solving a standard optimal delegation problem pointwise for each message realization. Subject to this simplification, it is well known (see Alonso and Matouschek (2008), and Amador and Bagwell (2013)) that, under the assumptions made in our model, the optimal delegation policy takes the form of a cap, with the cap becoming more restrictive when the agent’s bias is higher. This gives rise to the message-dependent cutoff function shown in Figure 1.1.

Applying the solution to the delegation part of the problem derived in Proposition 1, we then solve for the principal’s optimal information policy. In Proposition 2, we establish that the optimal information policy features a one-sided censorship, as depicted in Figure 1.1. This policy can be characterized by a cutoff on the expected bias of the agent as a function of the message \( x \) (which we denote by \( b^* \) in the figure) such that: all realizations of \( \omega \) that lead to an agent-bias below \( b^* \) are pooled together, while those that lead to an agent-bias above \( b^* \) are fully revealed.

An important implication of the joint delegation and information policy, which is illustrated in Figure 1.1, is that information and flexibility serve as substitutes in the principal’s problem: The principal gives less discretion to the agent after realizations of the state that are fully disclosed, and vice versa. The key intuition is that granting more discretion makes the agency problem at the information design stage more salient, and thus lowers the principal’s gains from information disclosure. This can be better understood in light of the two aforementioned trade-offs. The principal’s relative cost from granting the agent discretion is

\footnote{We can generalize our results to the case in which the agent’s bias is allowed to change sign. In that case, it can be shown that the optimal delegation and information policies will each consist of a two-sided censorship. Other than that, the main insights from our results remain virtually unchanged.}
Figure 1.1: Optimal information and delegation policy. The horizontal axis depicts the expected magnitude of the agent’s bias, conditional on the realized message $x$. On the vertical axis, we plot the optimal cutoff type $\theta^*$, which is the highest type $\theta$ that is allowed to choose his first-best action. At the optimal information policy, all realizations of $\omega$ that lead to a bias below $b^*$ are pooled, and those that lead to a bias above $b^*$ are fully revealed.

higher following realizations of $\omega$ at which the agent’s bias is relatively high. This implies that those realizations of $\omega$ will feature more restrictive delegation sets. As a result, the principal endogenously enjoys higher control over the action when the agent’s bias is high, and therefore finds it relatively less costly to disclose information about $\omega$ in those cases. In Section 4, we use comparative statics to further formalize the notion of substitutatibility between information provision and flexibility.

Throughout the paper, we illustrate our results within the context of a simple monopoly regulation model. Specifically, we consider the problem of a regulator who can jointly limit the extent of demand-based price discrimination carried out by the monopolist, and the set of prices that the monopolist can use in each segment of the market. The first aspect of the regulation problem is studied by Bergemann et al. (2015), while the second one is addressed by Alonso and Matouschek (2008), Bhaskar et al. (2020), and Amador and Bagwell (2022). We apply our results to describe the optimal joint policy of the regulator. Our findings imply that the regulator will use more stringent price caps and allow for more price discrimination in high-valuation segments of the market, while the opposite will be true in low-valuation segments. Consistent with our previous discussion, the intuition behind this result follows
from the fact that the monopolist is relatively more biased towards setting high prices when facing a high-demand segment of the market. This leads the principal to optimally give the monopolist little discretion to choose the magnitude of the price in that portion of the market, but greater power to finely discriminate across different consumer types. The reverse reasoning applies to the left tail of market valuations.

**Related literature.** By allowing the principal to jointly control information and the agent’s discretion, this paper belongs to the intersection between two distinct bodies of the literature: the one on information design pioneered by Kamenica and Gentzkow (2011), and the one on delegation initiated by Holmström (1980). Recent work by Kolotilin and Zapechelnyuk (2019) and Kleiner et al. (2021) has highlighted the connection between these two design problems, and proposed a unified framework that allows to solve them using common techniques. Our contribution consists of studying the two problems jointly. Relatedly, several papers have studied the role of communication in delegation: Dessein (2002) compares cheap talk with full delegation of authority; Szalay (2005) incorporates costly information acquisition by the agent; and Kartik (2009) considers verifiable disclosure by the agent. Unlike in our setting where the principal designs the information policy, in those papers it is the agent who is in charge of choosing the information structure.

Within papers that study communication in delegation, the most closely related to ours is Lou (2022), who studies optimal information disclosure about the agent’s private type in a delegation setting. In their model, the information that results from the signal structure chosen by the principal is privately observed by the agent and hence has to be elicited by the mechanism. In contrast, our model introduces two dimensions of uncertainty: one which is fully and privately observed by the agent, and another which is publicly observed and whose information is optimally garbled by the principal. Interestingly, the results derived from these two settings exhibit stark contrast. While Lou (2022) finds that the optimal delegation and information policies are finite, the optimal policy in our setting typically involves a continuum of actions and messages. Additionally, in contrast to our findings, they conclude that delegation and information provision act as complements in the principal’s problem. This outcome arises in their model because the principal resorts to obedient mechanisms, where each message corresponds to a distinct action, leading to a one-to-one relationship between information and the size of the delegation set. In our setting, due to public observability of the principal’s message and the residual uncertainty that stems from the agent’s private type, the principal will benefit from conditioning the delegation set on the realized message. As demonstrated, these message-contingent delegation sets tend to contain more actions when information is scarce, and vice versa. Given these differences, we view their
paper and ours as complementary, offering insights into the principal’s incentives to provide information in scenarios where information must be elicited (as in their paper) versus situations where information is publicly observed and utilized as an input in the mechanism (as in our paper).

Finally, within the information design literature, our paper borrows from the techniques by Dworczak and Martini (2019), which allows us to show the optimality of censorship information policies in our setting. To that extent, we also relate to Kolotilin et al. (2022) who provide conditions under which censorship information policies are optimal. Even though their result cannot be directly applied in our setting, the proof of our main result is closely related to theirs. We also conduct comparative statics analysis on the informativeness of the optimal policy under different delegation schemes in Section 4, which relates our work to recent papers providing conditions for comparative statics in information design problems (Curello and Sinander, 2022; Kolotilin et al., 2022; Whitmeyer, 2022).

2 Model

**General framework.** A principal and an agent, indexed respectively by \( i = P, A \), contract over an action, \( y \in Y \equiv [y, \bar{y}] \). Aside from \( y \), the principal’s and the agent’s payoffs depend on the agent’s type \( \theta \), and on the state variable \( \omega \). \( \theta \) is the agent’s private information and is drawn from the cumulative distribution function (cdf) \( F_\theta(\theta) \) on \( \Theta = [\theta, \bar{\theta}] \), with a continuous density that satisfies \( f_\theta(\theta) > 0 \) for all \( \theta \in \Theta \). \( \omega \) is not observed by the principal or the agent, but the principal can disclose information about it. Its prior distribution is given by the cdf \( F_\omega(\omega) \) on \( \Omega = [\omega, \bar{\omega}] \), with a continuous density that satisfies \( f_\omega(\omega) > 0 \) for all \( \omega \in \Omega \). We assume that \( \theta \) and \( \omega \) are independently distributed.\(^3\)

The principal’s and the agent’s payoffs as a function of the model variables are:

\[
u_i(y, \omega, \theta) = -\frac{y^2}{2} + (\gamma_i \omega + \alpha_i \theta + \beta_i) y, \quad i = P, A,
\]

Where \( \gamma_i, \alpha_i, \beta_i \) are commonly known constants. Without loss, we normalize \( \beta_P = 0 \) and set \( \alpha_A > 0 \). We assume that \( \gamma_P \neq 0 \).\(^4\)

The parametric assumptions on \( u_i \) imply that, for every realization of \( (\omega, \theta) \), both the principal and the agent’s preferences are single-peaked. We denote the resulting optimal

---

\(^3\)An analogous information structure is used in Kolotilin et al. (2017), who study persuasion of a receiver with private information about his payoffs.

\(^4\)The assumption that \( \gamma_P \neq 0 \) will ensure uniqueness of the optimal information policy. Except for uniqueness, all of the results in the paper will go through in the case in which the principal’s payoff is state-independent \( (\gamma_P = 0) \).
action for each of them by \( y_i(\omega, \theta) \equiv \gamma_i \omega + \alpha_i \theta + \beta_i, \ i \in \{A, P\} \). We will assume that the action space is large enough to include all possible principal and agent optimal actions: i.e., \( \{y : \exists i \in \{A, P\}, \omega \in \Omega, \theta \in \Theta \text{ such that } y_i(x, \theta) = y\} \subset Y \).

**Principal’s problem.** The principal can jointly choose the information about \( \omega \) that is available at the time of contracting with the agent and, conditional on this information, the set of actions that the agent can choose from. Formally, the principal’s problem is to choose an information policy \( \sigma : \Omega \rightarrow \Delta(S) \), where \( S \) is an arbitrary message space, which maps realizations of \( \omega \in \Omega \) into a distribution over messages in \( S \); and a delegation policy \( \Gamma : S \rightarrow 2^Y \) which specifies the set of actions that the agent can choose from for every realization of the public message \( s \in S \). The structure of the principal’s problem is captured by the following timing of events:

1. The principal commits to an information policy \( \sigma \) and a delegation policy \( \Gamma \)
2. \( \theta \) is drawn from \( F_\theta \) and \( \omega \) is drawn from \( F_\omega \), and a message \( s \) is drawn from \( \sigma(\omega) \)
3. The agent observes \( \theta \) and \( s \), and takes an action \( y \in \Gamma(s) \).

**Discussion of modeling assumptions.** Besides assuming that players’ payoffs are quadratic in the action, we assume that the optimal action is affine in \( \omega \) and \( \theta \), and that the marginal utility change from increasing \( \theta \) is independent of the value of \( \omega \). All of these assumptions help with tractability, and enable us to provide a sharp characterization of the optimal policy. Although very stylized, we believe that this structure allows us to capture the key interaction between delegation and information provision in the principal’s problem, in isolation from additional forces purely driven by complementarities in players’ payoff functions. They also allow us to draw new conclusions within the quadratic-payoffs settings, which is a workhorse model studied both in the delegation and in the communication literature (see respectively Holmström (1980) and Crawford and Sobel (1982), among many others). We will discuss the implications from relaxing some of these assumptions in Section 5.

One interpretation of the model is that \( \theta \) captures the specific expertise about the environment possessed only by the agent, whereas \( \omega \) represents broader institutional details that the principal can directly observe and disclose information about. Another interpretation is that, regardless of whether or not \( \omega \) is known by the agent, the principal is able to control the extent to which the agent uses information about \( \omega \) when choosing his action. With the second interpretation in mind, we now introduce our running application in the paper, which is that of a monopoly regulator who can jointly regulate the degree of demand-based price discrimination and the prices that the monopolist can use.
Example: Monopoly regulation. A monopolist (agent) sells to a market composed by a continuum of buyers with linear demand function \( q(y, \omega) = \omega - y \), where \( y \) is the price and \( \omega \in \Omega \) is the buyer-specific intercept of the demand curve.\(^5\) The population of buyers is distributed over \( \Omega \) according to the cdf \( F_\omega \). The monopolist chooses a profit-maximizing price, subject to his privately known marginal cost \( \theta \), and the information about \( \omega \) that is available to him. Under these assumptions, the monopolist’s utility function can be written as:

\[
u_A(y, \omega, \theta) = (\omega - y)(y - \theta) = -y^2 + (\omega + \theta)y - \omega \theta.\]

For simplicity, we will assume that \( \bar{\theta} \leq \omega \), which implies that, for every realization of \( (\omega, \theta) \), the monopolist always finds it optimal to serve the market when the set of prices is unrestricted.

The regulator (principal) can jointly restrict how much the monopolist can price-discriminate across buyer types \( \omega \) (information policy), and the prices that can be used in each segment of the market (delegation policy). More specifically, the regulator’s problem will be to partition the market \( \Omega \) into segments, and to determine what prices are permissible in each segment. Subject to this, the monopolist chooses a price for each segment of the market, which has to lie in the segment-specific permissible set.

The regulator’s objective function is a weighted average of consumer and monopolist’s surplus. Her utility function is thus given by:

\[
u_P(y, \omega, \theta) = \rho \int_y^\omega (\omega - s) \, ds + u_A(y, \omega, \theta) = - \left( 1 - \frac{\rho}{2} \right) y^2 + (\theta - (\rho - 1)\omega)y + \left( \frac{\rho}{2} \omega - \theta \right) \omega,
\]

Where \( \rho \) represents the redistribution concern of the regulator. We will assume that \( \rho \in (1, 2) \), which means that the regulator cares weakly more about the consumer’s welfare than she does about the monopolist’s \( (\rho > 1) \), but that the redistribution concern is not sufficiently high that the regulator’s incentives are fully aligned with those of the consumers \( (\rho < 2) \). Without this latter assumption, the regulator’s problem would be trivial since she would choose to restrict the monopolist to set the lowest possible price without any regards for what his marginal cost is. The assumption that \( \rho \) is strictly greater than 1 ensures that the optimal policy is unique \( (it is the analogue of assuming that \gamma_P \neq 0) \). Our characterization applies to cases in which the regulator’s objective is arbitrarily close to total surplus \( (i.e.,\)

\(^5\)The linear demand function can be microfounded by assuming that the buyer’s preferences can be represented by the utility function \( u(q, \omega) = \frac{q^2}{2} + \omega q \). In this sense, we can interpret \( \omega \) as the buyer’s marginal valuation for the good when she has zero units of it.
for $\rho$ in any small right-neighborhood of 1), and it also applies to the limiting case in which $\rho = 1$ although uniqueness of the optimal policy is no longer ensured in this case.

The version of this problem in which $\omega$ is commonly observed by the monopolist and the regulator, and the regulator can only restrict the set of prices is studied by Alonso and Matouschek (2008). Alternatively, the problem where the set of prices that the monopolist can use is unrestricted but the regulator can control the degree of price discrimination is the focus of Bergemann et al. (2015). These two settings will be natural benchmarks for our analysis of this example. Later, we will apply our results to discuss how the regulator can improve upon them, by jointly controlling the two features of the monopoly pricing problem.

3 Main Results

In this section, we characterize the optimal policy. We show that, in our framework, the optimal information policy features one-sided censorship (Proposition 2), and that, following every message realization about $\omega$, the optimal delegation policy is implementable by means of an action cap (Proposition 1). Further, we establish the substitutes property, which states that the agent is granted less flexibility following realizations of $\omega$ that are fully disclosed (Corollary 1). All proofs are collected in the Appendix.

3.1 Preliminaries: Simplifying the Principal’s Problem

We begin by formally setting up the principal’s problem. Let $x = \mathbb{E}[\omega|s] \in \Omega$ denote the posterior mean about $\omega$ after having observed a message $s$. Because both the principal’s and the agent’s payoffs are linear in $\omega$, we can restrict attention to unbiased information policies in which the principal’s message is the posterior mean about $\omega$ itself. Hence, we redefine the message space to be $S = \Omega$. Moreover, using standard results (Blackwell, 1953; Gentzkow and Kamenica, 2016), we can recast the problem of choosing an information policy as one of optimizing over distributions over posterior means of $\omega$ which are mean-preserving contractions of the prior $F_\omega$.

In addition, we apply the revelation principle and write the problem of choosing a delegation policy as a mechanism design problem, whereby the principal commits to an outcome
function $y : \Omega \times \Theta \rightarrow Y$ subject to truthful reporting of $\theta$ (for every $x$) by the agent.\footnote{As shown by Kovác and Mylovanov (2009), the restriction to deterministic mechanisms is without loss of optimality in our environment (i.e., under Assumption 1).}

Applying this, we write the principal’s problem as choosing a distribution over posterior means $G_\omega \in \Delta(\Omega)$ and an outcome function $y : \Omega \times \Theta \rightarrow Y$ to solve:

$$\max_{G_\omega, y(x, \theta)} \int_\Omega \int_\Theta u_F(y(x, \theta), x, \theta) \, dF_\theta(\theta) \, dG_\omega(x)$$  \hspace{1cm} \text{ (Opt)}

subject to $u_A(y(x, \theta), x, \theta) \geq u_A(y(x, \theta'), x, \theta)$ \hspace{1cm} \forall \theta, \theta' \in \Theta, \forall x \in \Omega,$  \hspace{1cm} \text{ (IC)}

$F_\omega$ is a mean-preserving spread of $G_\omega,$  \hspace{1cm} \text{ (BP)}

Where we have used the linearity of the principal’s payoff in $\omega$ and the law of iterated expectations to write $E[u_F(y(x, \theta), \omega, \theta)] = E[E[u_F(y(x, \theta), \omega, \theta) \mid x, \theta]] = E[u_F(y(x, \theta), x, \theta)].$

Henceforth, we refer to $y : \Omega \times \Theta \rightarrow Y$ as a delegation policy and to $G_\omega \in \Delta(\Omega)$ as an information policy, and to a pair $(G_\omega, y)$ plainly as a policy. Also, to avoid expanding the notation, for a given cdf $F$ with real-valued support, we refer to the probability measure associated to $F$ simply as $F.$

We now argue that the principal’s problem can be solved through the following two-step procedure. The first step is to solve for the optimal outcome function $y(x, \cdot)$ for every realization of the posterior belief $x \in \Omega.$ Indeed, any delegation policy that, together with some $G_\omega,$ is part of a solution to (Opt) must be optimal almost everywhere in $G_\omega.$ This is because both the principal and the agent observe the realization of $x$ and can condition their actions (the delegation policy in the case of the principal and the report in the case of the agents) on it. Further, since the principal’s payoff is separable in $x,$ her optimal delegation policy is obtained by solving the problem pointwise for every $x.$\footnote{Requiring optimality for realizations of $x$ outside of the support of $G_\omega$ is without loss, and simplifies the problem since it makes the delegation part of the problem independent of the support of the information policy.} In other words, the principal doesn’t benefit from committing to a mechanism before the realization of $x.$

This allows us to write the problem of choosing an optimal delegation policy as:

$$V(x) \equiv \max_{y(x, \cdot)} \int_\Theta u_F(y(x, \theta), x, \theta) \, dF_\theta(\theta)$$  \hspace{1cm} \text{ (Opt Del)}

subject to (IC).

The solution to (Opt Del) recovers, not only the optimal delegation rule, but also the
principal’s value from inducing the posterior mean \( x \in \Omega \). Armed with this value function, the second step consists of finding an information policy that maximizes the principal’s ex-ante utility. The resulting program is a standard one-dimensional information design problem, which can be stated as

\[
\max_{G_\omega} \int_\Omega V(x) \, dG_\omega(x) \tag{Opt Info}
\]

s.to \((BP)\).

We now proceed to solve the principal’s problem in two steps. We begin by finding the optimal delegation policy in Section 3.2, and then proceed to solve for the outer information design problem in Section 3.3.

### 3.2 Optimal Delegation

Having simplified the principal’s problem, we now turn to solving the inward problem of finding the optimal delegation policy for every realization of \( x \). To do so, we introduce the following assumption.

**Assumption 1.**

i) The function \( \alpha_A F_\theta(\theta) + (y_A(x, \theta) - y_P(x, \theta)) f_\theta(\theta) \) is strictly increasing in \( \theta \), for all \( x \in \Omega \), and

ii) \( y_A(x, \theta) - y_P(x, \theta) \geq 0 \), for all \( x \in \Omega \) and \( \theta \in \Theta \)

The first part of the assumption is a standard regularity condition used in the bulk of the delegation literature. As shown by Alonso and Matouschek (2008), Amador and Bagwell (2013) and Amador et al. (2018), the weak version of it (i.e., assuming weak instead of strong monotonicity) ensures that, for every \( x \), an interval delegation policy —i.e., one that is implementable by means of a (possibly degenerate) interval delegation set— is optimal. We use strong monotonicity because it allows us to show that the optimal policy is unique.\(^9\) Uniqueness of the optimal policy will prove useful in Section 4 when we conduct comparative statics. Other than that, except for uniqueness in Proposition 1, all the results in the paper remain valid if we assume weak monotonicity instead.

The second part of the assumption implies that the agent’s bias is always non-negative, and hence it suffices to look at delegation policies that take the form of a *cap*, meaning that, for every realization of \( x \), the principal censors the action space from above. This assumption is introduced mostly for exposition purposes. The setting where the agent’s bias is negative is analogous to this one, with the difference that the optimal delegation policy entails a floor

\(^9\)Throughout the paper, we refer to uniqueness in the almost-everywhere sense, that is, unique up to probability-zero events.
instead of a cap. We discuss the case in which the sign of the bias is allowed to change in the Online Appendix and show that, the main difference in that case is that both the delegation and the information policy may feature two-sided (instead of one-sided) censorship.

Proposition 1 states the solution to the inner problem of finding an optimal delegation policy for every $x$. To state the result, we have to consider two separate regions of values of $x$. The first region is one where the principal and the agent are sufficiently aligned that the principal benefits from granting the agent discretion. In such a region, the optimal policy features a cutoff on the space of agent types, denoted by $\theta^*(x)$, such that all agent types below the cutoff get to choose their first-best action, whereas all agent types above it are pooled together at a constant action which coincides with the bliss point of the cutoff type, $y_A(x, \theta^*(x))$. Otherwise, if $x$ is such that the magnitude of the agent’s bias is sufficiently large, the optimal policy features no discretion at all, and the principal chooses an action that is constant in $\theta$. This is summarized in Proposition 1.

**Proposition 1 (Optimal delegation policy).** The unique optimal delegation policy is:

$$y(x, \theta) = \begin{cases} 
\min\{y_A(x, \theta), y_A(x, \theta^*(x))\}, & \text{if } \mathbb{E}_\theta[y_P(x, \theta)] \geq y_A(x, \theta) \\
\mathbb{E}_\theta[y_P(x, \theta)], & \text{if } \mathbb{E}_\theta[y_P(x, \theta)] < y_A(x, \theta)
\end{cases}$$

where $\theta^*(x) : \mathbb{E}_\theta[y_P(x, \theta)|\theta \geq \theta^*(x)] = y_A(x, \theta^*(x))$.

For notational convenience, we will extend the function $\theta^*(x)$ to be such that $\theta^*(x) = \emptyset$ whenever $\mathbb{E}_\theta[y_P(x, \theta)] < y_A(x, \theta)$. Under this extension, the function $\theta^*(x)$ is continuous on $\Omega$.

In the language of delegation sets, the above mechanism can be implemented by offering the agent a delegation set which takes the form of an interval, following every realization of $x \in \Omega$. The smallest delegation set that implements the allocation rule in Proposition 1 is equal to

$$D^*(x) = \begin{cases} 
[y_A(x, \theta), y_A(x, \theta^*(x))], & \text{if } \mathbb{E}_\theta[y_P(x, \theta)] \geq y_A(x, \theta) \\
\{\mathbb{E}_\theta[y_P(x, \theta)]\}, & \text{if } \mathbb{E}_\theta[y_P(x, \theta)] < y_A(x, \theta)
\end{cases}$$

One convenient observation that arises in our environment is that, because the agent’s bias $y_A(x, \theta) - y_P(x, \theta) = (\gamma_A - \gamma_P)x + (\alpha_A - \alpha_P)\theta + \beta_A$ is monotone in $x$, there is a sense in which the ‘amount of flexibility’ given to the agent at the optimum is also monotone in $x$. We illustrate this in the case in which $\gamma_A - \gamma_P > 0$, which implies that the agent’s (upward) bias increases with $x$. In this setting, the optimal delegation policy will feature a cutoff on $x$, above which the agent is given no flexibility (the delegation set is a singleton), and below
which the agent is given partial flexibility (the delegation set is a non-degenerate interval with a cap).

Moreover, in the region of low enough values of \( x \) where the optimal delegation set is non-degenerate, we may capture changes in flexibility at the intensive margin through the cutoff type \( \theta^*(x) \). Since \( \theta^*(x) \) is the highest agent type that gets to choose his first-best action at the optimal delegation policy, we can associate a higher value of \( \theta^*(x) \) to a higher degree of flexibility. As we show in Lemma 2, the function \( \theta^*(x) \) is decreasing when \( \gamma_A - \gamma_P > 0 \), thus establishing monotonicity at the intensive margin as well. The reasoning is analogous, but goes in the opposite direction, when \( \gamma_A - \gamma_P < 0 \). On the other hand, one can readily check from the expressions in Proposition 1 that the optimal delegation policy is constant in \( x \) whenever \( \gamma_A - \gamma_P = 0 \).

In summary, the amount of flexibility given to the agent, as captured by the cutoff \( \theta^*(x) \), is weakly decreasing in the magnitude of the agent’s bias. This intuitive property will be key for characterizing the shape of the principal’s value function \( V(x) \) and the resulting optimal information policy, which is the focus of the next section.

### 3.3 Optimal Information Provision

Having derived the optimal mechanism and the resulting value function \( V(x) \), the next step is to solve the outer problem of finding a distribution over posterior means about \( \omega \) that maximizes the expected value of \( V(x) \). Before stating the main result, we first introduce a simple class of information policies that will prove to be optimal.

**Definition 1 (Censorship information policy).** A (one-sided) censorship information policy is described by a cutoff \( \omega^* \in \Omega \) such that either:

i) All realizations \( \omega < \omega^* \) are fully revealed, and all realizations \( \omega \geq \omega^* \) are pooled together. We call this policy an *upper censorship*, or

ii) All realizations \( \omega > \omega^* \) are fully revealed, and all realizations \( \omega \leq \omega^* \) are pooled together. We call this policy a *lower censorship*.

The extreme cases of full information disclosure and no information disclosure are (up to measure-zero events) a special case of an upper censorship with cutoffs given, respectively,

---

\footnote{It is not necessarily the case that the optimal action cap, \( y_A(x, \theta^*(x)) \), is decreasing \( x \). This is because, even though \( \theta^*(x) \) decreases with \( x \), the optimal action of all agent types shifts upward whenever \( x \) increases, so the overall effect of a change in \( x \) on the action cap is ambiguous. Despite this, we argue that the relevant measure of flexibility is \( \theta^*(x) \) which is always monotone in \( x \). One way to see this is by noting that the diameter of the delegation set \( D^*(x) \) is given by \( y_A(x, \theta^*(x)) - y_A(x, \underline{\theta}) = \alpha_A(\theta^*(x) - \underline{\theta}) \), which moves in the same direction as \( \theta^*(x) \).}
by $\omega$ and $\bar{\omega}$. We say that a censorship information policy features *strict censorship* if it doesn’t involve full information disclosure.

One-sided censorship information policies have the desirable feature of being fully described by a scalar. Being able to restrict attention to this kind of information policies simplifies the information design problem in a very similar manner as restricting attention to one-sided interval delegation simplifies the delegation problem.\footnote{Kolotilin and Zapechelnyuk (2019) formalize the connection between the two problems and their solution.} In particular, the one-dimensionality of the solution to the problem will allow us to conduct comparative statics on the amount of information and the amount of flexibility given to the agent at the optimum.

In order to establish optimality of censorship information policies, we introduce the following assumption.

**Assumption 2.** If $\alpha_P > 0$, then the function

$$M(\theta^*) \equiv \frac{(1 - F_0(\theta^*))^2}{\alpha_A(1 - F_0(\theta^*)) - \alpha_P (E[\theta|\theta \geq \theta^*] - \theta^*) f_0(\theta^*)}$$

is strictly decreasing in $\theta^* \in \Theta$.

Henceforth, it is assumed that Assumption 2 holds. We will discuss the role that this assumption plays in our results after stating Proposition 2. The assumption is trivially satisfied when $\alpha_P \leq 0$.\footnote{The reason why we only require the assumption when $\alpha_P > 0$ is that, if $\alpha_P \leq 0$, then it follows from Proposition 1 that the optimal delegation policy doesn’t give the agent any discretion, which will in turn imply that the optimal information policy is fully informative (one of the extreme forms of censorship).} If $\alpha_P > 0$, then a sufficient condition for Assumption 2, that only depends on the distribution of $\theta$, is that the function

$$\frac{(E[\theta|\theta \geq \theta^*] - \theta^*) f_0(\theta^*)}{1 - F_0(\theta^*)}$$

is strictly decreasing in $\theta^*$, which is satisfied, for example, by the uniform, the exponential, and the Pareto distributions.

Proposition 2 describes the optimal information policy. We show that, paired with the optimal mechanism from the previous section, censorship information policies are optimal in our environment. We also characterize the type (upper or lower) of censorship, and provide conditions under which the censorship cutoff is interior.

To formalize this, let $y_P(x, \theta^*) = E[\theta|y_P(x, \theta)|\theta \geq \theta^*]$, which is the principal’s optimal
action given \( x \in \Omega \) and conditional on pooling all agent types above \( \theta^* \), and let

\[
\hat{x} \equiv \begin{cases} 
\max \{ x \in \Omega : \bar{y}_P(x, \theta) \geq y_A(x, \theta) \} \cup \{ \omega \}, & \text{if } \gamma_A - \gamma_P \geq 0 \\
\min \{ x \in \Omega : \bar{y}_P(x, \theta) \geq y_A(x, \theta) \} \cup \{ \omega \}, & \text{if } \gamma_A - \gamma_P < 0,
\end{cases}
\]  

(3.1)

By Proposition 1, if \( \gamma_A - \gamma_P \geq (\leq)0 \), \( \hat{x} \) represents the cutoff on \( x \) above (below) which the optimal delegation policy gives no discretion to the agent. In the complementary region, the principal will give full discretion up to a cutoff \( \theta^*(x) \), which weakly decreases (increases) with \( x \), and satisfies \( y_A(x, \theta^*(x)) = \bar{y}_P(x, \theta^*(x)) \).

As will be clear in Proposition 2, the extent to which the principal optimally chooses to disclose information hinges on the cutoff \( \hat{x} \) and on the ratio

\[
\Delta \equiv \begin{cases} 
\frac{\gamma_A^2}{(\gamma_A - \gamma_P)^2}, & \text{if } \gamma_A - \gamma_P \neq 0 \\
+\infty, & \text{if } \gamma_A - \gamma_P = 0,
\end{cases}
\]

Which, for reasons that will become clear later, we term the benefit-to-cost ratio of information disclosure.

Proposition 2 establishes that the optimal information policy uses censorship, and that the direction of the censorship hinges exclusively on whether the agent’s bias increases or decreases with \( \omega \). It also provides necessary and sufficient conditions for strict censorship to be optimal, and sufficient conditions on \( \Delta \) under which the optimal censorship cutoff is interior. A full characterization of the optimal policy can be found in the proof of Proposition 2 in the Appendix.

Proposition 2 (Optimal information policy). If \( \gamma_A - \gamma_P \geq 0 \) (respectively, \( < 0 \)), a lower (respectively, upper) censorship information policy with cutoff \( \omega^* \in \Omega \) is uniquely optimal. Moreover,

- There exists \( \bar{\Delta} > 0 \) such that, if either \( \hat{x} = \omega \) (respectively, \( = \bar{\omega} \)) or if \( \Delta > \bar{\Delta} \), then the optimal policy is full disclosure.

- If \( \hat{x} > \omega \) (respectively, \( < \bar{\omega} \)) and \( \Delta < \bar{\Delta} \), then the optimal policy features strict censorship. Moreover, there exists \( \bar{\Delta} < \bar{\Delta} \), such that if \( \Delta \in (\bar{\Delta}, \bar{\Delta}) \), the optimal censorship cutoff is interior.

Figure 3.1 depicts the optimal policy when \( \gamma_A - \gamma_P > 0 \) and \( \Delta \in (\Delta, \bar{\Delta}) \). Because the agent’s bias increases with \( \omega \), the optimal delegation policy features a cutoff \( \theta^*(x) \) which is strictly decreasing up to \( \hat{x} \), and becomes flat on \( x > \hat{x} \), which is the region in which the agent’s bias is high enough that the principal prefers choose the action herself. Moreover, as
we show in Lemma 4, the principal's value function $V(x)$ is \textit{S-shaped}: it is strictly concave to the left of $x^{*} \in (\omega, \hat{x})$ and strictly convex to the right of it. This feature allows us to show that there is an optimal lower censorship cutoff, given by $\omega^{*} \in (x^{*}, \overline{x})$, above which the principal fully discloses the state. The figure also shows the principal's value under the optimal policy, which is equal to $E_{\omega}[V(\omega)]$, where

$$
\overline{V}(\omega) \equiv \begin{cases} 
V(\omega^{*}) + V'(E[\omega|\omega \leq \omega^{*}])(\omega^{*} - \omega), & \text{if } \omega \leq \omega^{*} \\
V(\omega), & \text{if } \omega > \omega^{*},
\end{cases}
$$

is an upper convex envelope of $V(x)$, which coincides with $V(x)$ on the support of the optimal information policy.

![Diagram](a) Optimal delegation policy  (b) Optimal information policy

Figure 3.1: Optimal policy

An important implication of Proposition 2 is that, at the optimum, the realizations of $\omega$ which are fully disclosed by the principal will feature \textit{less} flexibility, compared to realizations of $\omega$ that are pooled. To see this, consider the case in which $\gamma_{A} - \gamma_{P} > 0$. In this setting, the optimal policy is a lower censorship and thus only high realizations of the state are disclosed. Moreover, because $\theta^{*}(x)$ is non-increasing, these realizations also coincide with smaller delegation sets.

In order to state this result formally, let $\omega^{*} \in \Omega$ be the optimal censorship cutoff. If the optimal information policy is a lower (respectively, upper) censorship, we say that $\omega$ belongs
to the pooling region if $\omega \in [\omega, \omega^\ast]$ (respectively, $\omega \in [\omega^\ast, \omega^\ast]$), and we say that $\omega$ belongs to the revealing region if $\omega \in [\omega^\ast, \omega]$ (respectively, $\omega \in [\omega, \omega^\ast]$).

Observe further that any censorship information policy can be implemented by means of the following simple direct signal structure $\sigma(\omega)$: if $\omega$ belongs to the revealing region, $\sigma(\omega)$ places all the mass on the message "$\omega$"; and if $\omega$ belongs to the pooling region, say $[\omega_1, \omega_2]$, $\sigma(\omega)$ is degenerate at the message "$E[\omega|\omega \in [\omega_1, \omega_2]]$". Because $\sigma(\omega)$ is degenerate for all $\omega \in \Omega$, the mapping from realizations of $\omega$ to delegation sets (or mechanisms) is deterministic. This allows us to say that, at the optimum, the principal commits at the outset to $D: \Omega \rightarrow 2^Y$, which specifies the delegation set\textsuperscript{13} for each realization of $\omega$.

Given this simplification, let $d(\omega) \equiv \text{diam } D(\omega)$ be the size of the delegation set, as measured by its diameter (i.e., the Lebesgue measure of the interval). The corollary below states that realizations of $\omega$ that are fully revealed are associated to smaller delegation sets relative to realizations that are censored.

**Corollary 1.** For any $\omega$ that belongs to the pooling region and $\omega'$ that belongs to the revealing region, it holds that $d(\omega) > d(\omega')$.

One interpretation of Corollary 1 is that information provision and flexibility serve as substitutes in the principal’s problem: she optimally provides less flexibility under realizations of the state that are fully revealed and vice versa. This result holds ex-post, in the sense that it compares delegation sets across different realizations of $\omega$. In section 4, we provide an ($\omega$-)ex-ante notion of substitutability between information and flexibility.

To understand the intuition behind this result, we use Proposition 1 to write the principal’s objective function in the information design problem as:

$$V(x) = \int_{\theta^\ast(x)}^{\theta(x)} u_P(y_A(x, \theta), x, \theta) dF_\theta(\theta) + \int_{\theta^\ast(x)}^{\bar{\theta}} u_P(y_P(x, \theta^\ast(x)), x, \theta) dF_\theta(\theta). \quad (3.2)$$

To fix ideas, suppose that $\gamma_A - \gamma_P \geq 0$. Consider first the case in which $x$ is such that $\theta^\ast(x) = \bar{\theta}$, so that the optimal delegation set is a singleton. In this case, the principal is effectively taking the action herself without contracting with the agent. Because the agency problem is muted, the principal strictly benefits from releasing information about $\omega$, since she can then use this information to optimally choose the action $y_P(x, \theta^\ast(x))$. Formally, this is equivalent to $V(x)$ being strictly convex whenever $\theta^\ast(x) = \bar{\theta}$. By extending this argument, we can show that, for a constant $\theta^\ast(x)$, the second term on the right-hand-side of (3.2) is convex for any $x$. Intuitively, for all $\theta \geq \theta^\ast(x)$, the outcome implemented by the optimal mechanism is $E[y_P(x, \theta)|\theta \geq \theta^\ast(x)]$, which coincides with what the principal would choose

\textsuperscript{13}As before, we let $D(\omega)$ be the smallest delegation set implementing the principal’s optimal mechanism.
subject to knowing that $\theta \geq \theta^*(x)$. Therefore, information about $\omega$ is beneficial in that it allows the principal to fine-tune her choice $E[y_P(x, \theta)|\theta \geq \theta^*(x)]$.

Alternately, the first term in (3.2) fully captures the agency problem faced by the principal: if the agent’s type is below $\theta^*(x)$, the agent has full discretion to act and will choose an action that differs from the principal’s first-best. Because of this, information disclosure need not be beneficial. In fact, if the principal and the agent are sufficiently misaligned, this first term in the principal’s value function will be concave, which would potentially offset the benefits from information provision described in the previous paragraph. Furthermore, the first term becomes more relevant in the principal’s objective whenever the agent has more discretion (i.e., when $\theta^*(x)$ is larger), which implies that the principal will benefit less from providing information about $\omega$ following realizations that lead to a more flexible optimal delegation set. Because, under the assumption that $\gamma_A - \gamma_P \geq 0$, the degree of flexibility given to the agent decreases with the posterior mean about $\omega$, this explains why the optimal information policy would use censoring from below.

The above arguments fully describe the principal’s incentives to provide information for a constant $\theta^*(x)$. However, at the optimal delegation policy, $\theta^*(x)$ varies with $x$ in the manner described in Proposition 1. Therefore, the shape of the optimal delegation-cutoff $\theta^*(x)$ will also play a role in determining the curvature of $V(x)$, and in that way the nature of the optimal information policy. This is where Assumption 2 comes in: It ensures that the curvature of $\theta^*(x)$ is not too pronounced so as to overturn the “agency vs. information exploitation” trade-off described in the previous paragraphs.

In the formal proof of Proposition 2, we exploit this intuition by showing that the function $V(x)$ switches concavity at most once, and that it goes from being strictly concave to strictly convex when $\gamma_A - \gamma_P \geq 0$, and from strictly convex to strictly concave when $\gamma_A - \gamma_P < 0$. Using this, we follow the duality approach in Dworczak and Martini (2019) in order to extend the proof of the main result in Kolotilin et al. (2022) to our setting to show unique optimality of censorship persuasion.\textsuperscript{14}

We argue that the parameter $\Delta$ captures the benefit-to-cost ratio that the principal faces when releasing information about $\omega$. $|\gamma_P|$ represents the extent to which the principal cares about matching the action to the value of $\omega$ and thus is a measure of the benefit from information, while $|\gamma_A - \gamma_P|$ governs the misalignment between the principal and the agent in terms of how their most-preferred action responds to information about $\omega$ thereby capturing the resulting agency costs. According to Proposition 2, if the benefit-to-cost ratio

\textsuperscript{14}Kolotilin et al. (2022) assume that the sender’s value function is twice continuously differentiable. This is not true in our setting, since the second derivative of $V$ may feature a discontinuity at $\hat{x}$. In the proof, we extend their result to piecewise twice-continuously differentiable value functions such as ours.
of information disclosure is sufficiently high, then the principal finds it optimal to fully disclose $\omega$. In particular, full informativeness is always optimal when $\gamma_A - \gamma_P = 0$. On the other hand, if $\Delta$ is low enough, the optimal policy will necessarily feature some withholding of information.

Proposition 2 also states that there is an open set of values of $\Delta$, namely $(\Delta, \bar{\Delta})$, under which the optimal censorship cutoff is interior. In general, the cutoff $\Delta$ cannot be derived in closed form. In fact, the principal’s objective function when choosing the optimal censorship cutoff, i.e.

$$W(\omega^*) \equiv \int_{\omega}^{\omega^*} V(\mathbb{E}[\omega| \omega \leq \omega^*]) dF_\omega(x) + \int_{\omega^*}^{\infty} V(x) dF_\omega(x),$$

in the case when $\gamma_A - \gamma_P \geq 0$, does not exhibit complementarities between $\omega^*$ and $\Delta$, and hence the optimal censorship cutoff $\omega^*$ need not be monotone in $\Delta$. In particular, it is typically not possible to assert that the optimal policy will be completely uninformative if and only if $\Delta < \Delta$. Nonetheless, in the proof of Proposition 2, we provide necessary and sufficient conditions (which do not depend exclusively on $\Delta$) for the optimal cutoff to be interior.

We now apply the results in Propositions 1 and 2 to the monopoly regulation example.

**Example: Monopoly regulation (contd.)** Given the monopolist’s marginal cost, $\theta$, and a posterior belief about the market’s valuation, $x$, the monopolist’s and the regulator’s optimal action are respectively given by:

$$y_A(x, \theta) = \frac{x + \theta}{2}, \quad y_P(x, \theta) = \frac{\theta - (\rho - 1)x}{2 - \rho}.$$ 

The monopolist’s optimal price, for a given market segment, is increasing in the segment’s average willingness to pay, $x$. If $\rho = 1$, the regulator’s objective is market surplus, and therefore her optimal price is constant in $x$ and equal to the monopolist’s marginal cost. Otherwise, when she has a redistribution concern ($\rho > 1$), the gain from reducing the price increases with the size of the market and therefore her optimal price becomes decreasing in $x$. Overall, the bias is given by $y_A(x, \theta) - y_P(x, \theta) = \frac{\rho(x - \theta)}{2(2 - \rho)}$, which is strictly positive and increasing in $x$. To be consistent with the problem studied in the rest of the paper and following Alonso and Matouschek (2008), we disregard the participation constraint of the monopolist. This means that, whenever the price cap imposed by the regulator binds, the monopolist might in some cases be selling at a loss (i.e., if his cost $\theta$ is above the price cap). Allowing for participation constraints is beyond the scope of this example and, moreover,
should not affect the main insights derived from the analysis.\footnote{The monopoly regulation problem with a participation constraint is studied in Kolotilin and Zapechelnyuk (2019), and Amador and Bagwell (2022). They show that the optimal policy continues to be a price cap, but the price cap is higher in the presence of a participation constraint.}

Suppose that Assumptions 1 and 2 hold. Applying our result in Proposition 1, the optimal price cap is given by:

$$
\bar{y}(x) = \begin{cases} 
\frac{y + \theta^*(x)}{2}, & \text{if } \frac{px}{2(2-\rho)} < \frac{\theta}{2-\rho} - \frac{\theta}{2} \\
\frac{E\theta - (\rho - 1)x}{2-\rho}, & \text{otherwise},
\end{cases}
$$

(3.3)

Where $\theta^*(x)$ is defined as in Proposition 1. Since this is an instance in which the size of the agent’s bias is increasing in $x$ (i.e., $\gamma_A - \gamma_P > 0$), the cutoff $\theta^*(x)$ decreases with $x$. For high valuation segments of the market, the monopolist’s overpricing bias is sufficiently high that the regulator chooses to set the price herself without eliciting any information regarding the monopolist’s marginal cost.

Turning to the optimal information policy, which dictates the extent to which the regulator allows the monopolist to charge different prices across different segments of the market, Proposition 2 implies that the optimal market segmentation will entail a lower censorship, whereby only high-valuation segments may face full price discrimination. Corollary 2 summarizes these results.

**Corollary 2.** In the monopoly regulation problem, the optimal segment-specific price cap is given by (3.3). The size of the set of permissible prices weakly decreases with the market’s average valuation, $x$.

The optimal market segmentation is described by a cutoff $\omega^* \in \Omega$, such that: the monopolist is restricted to use a uniform price in the market segment $[\omega, \omega^*]$, and the monopolist is allowed to fully price discriminate within the market segment $[\omega^*, \varpi]$. Moreover, there is a cutoff $\bar{p} \in (1, 2)$ such that the optimal policy allows full price discrimination if and only if $\rho > \bar{p}$.

The optimal policy can be explained as follows. The monopolist is restricted to charge a uniform price within the low-value segment of the market composed by all buyers whose valuation is below $\omega^*$. In that segment, the regulator imposes a price cap equal to $\frac{E|\omega| |\omega \leq \omega^*| + \theta^*(E|\omega| |\omega \leq \omega^*|)}{2}$. In the high-value segment of the market ($\omega > \omega^*$), the regulator allows for full price discrimination across consumer types. However, in this region, she exerts higher control over the magnitude of the price by using more restrictive price caps (relative to the uniform-price region below $\omega^*$). Furthermore, for sufficiently high-demand segments ($\omega > \hat{x}$), she mandates the monopolist to charge a cost-independent price equal to $\frac{E\theta - (\rho - 1)x}{2-\rho}$, which is strictly
decreasing in the buyer’s valuation.

Interestingly, the regulator optimally allows the monopolist to fully price discriminate if and only if her *redistribution concern is sufficiently high*. This result can be understood by writing the benefit-to-cost ratio of information disclosure, which is in this example is given by

\[ \Delta_{\text{mon}} = \frac{4(\rho - 1)^2}{\rho^2}, \]

which strictly increases with \( \rho \). To elaborate, a higher value of \( \rho \) affects the incentives to allow for price discrimination in two ways. On the one hand, it increases the bias term \( \gamma_A - \gamma_P \), thus making the regulator and the monopolist more misaligned in terms of the *slope* of the optimal price as a function of the market’s average valuation \( x \). This makes enabling price discrimination less attractive for the regulator. On the other hand, a higher value of \( \rho \) makes the regulator’s optimal pricing rule more sensitive to the value of \( \omega \), which makes a finer segmentation of the market more attractive. Our result implies that, at least for sufficiently high values of \( \rho \), the latter effect overturns the former, and thus full price discrimination is optimal if and only if \( \rho \) is high enough. We further extend this comparative statics result on \( \rho \) in Section 4 by showing that in fact the informativeness of the optimal market segmentation is increasing in \( \rho \).

Finally, we discuss how our findings relate to the results in Bergemann et al. (2015). Their paper derives the set of surplus divisions between the consumer and the monopolist that are attainable through some segmentation of the market, in a setting in which the monopolist can choose from an unrestricted set of prices. Here, we extend the scope of the regulator’s power by allowing her to also choose segment-specific price caps. A consequence of Propositions 1 and 2 is that the regulator strictly benefits from having the ability to use price caps, and that this in turn affects how she optimally segments the market.

Another important difference between the setting studied in this example and the model in Bergemann et al. (2015) is that, in their model, full market segmentation is equivalent to *first-degree* price discrimination. This is because of their unit demand assumption which implies that, in expectation, the monopolist is able to extract all the market surplus by charging a price equal to the average willingness-to-pay in each market segment. In particular, if the regulator’s objective were to maximize welfare, the unit demand assumption would imply that the first-best outcome could be achieved without any price caps, by simply allowing the monopolist to engage in complete price discrimination and extract all the surplus.

This is no longer true under our linear-demand environment: in each segment of the market, the monopolist conducts *third-degree* price discrimination, meaning that he is restricted to charging a fixed price for every unit consumed by buyers. Consequently, the
monopolist can only extract the full surplus from the last marginal unit sold. An im-
portant consequence of this is that achieving first-best welfare is no longer possible through full
market segmentation. In fact, it is well-known that, in the linear-demand environment and
absent the regulator’s ability to use price caps, allowing the monopolist to engage in full
third-degree price discrimination strictly reduces welfare (see Varian (1989)). This implies
that in our setting, the solution to the optimal regulation problem with $\rho = 1$ is non-trivial
and will typically require the regulator using price caps, as well as limiting the extent of
demand-based price discrimination that is allowed.

We conclude this section by discussing the intuition behind the shape of the optimal joint
delegation and information policy in light of the trade-offs faced by the principal when she
faces these two problems separately. The trade-off faced by the principal in the delegation
problem (with a fixed information policy) stems from the fact that, on the one hand, she
benefits from granting the agent the flexibility to tailor his action to his private information
about $\theta$, but on the other hand, this flexibility will give rise to ex-post suboptimal actions
due to the agent’s bias. A similar trade-off is faced by an information designer who faces a
fixed delegation set: she may benefit from providing information about $\omega$ since this allows
her to steer the agent’s actions in a direction that is favorable to her. However, the principal’s
ability to do so is limited by the degree of the agent’s misalignment, and hence the extent to
which she optimally provides information will depend on the severity of this agency problem.

In our setting, the principal is faced with the task of jointly resolving these two trade-offs.
The optimal policy is then shaped by the fact that the agency cost associated to information
disclosure is smaller whenever the agent’s actions are more restricted (i.e., when he faces
a tighter delegation set). It is precisely the interaction between these two trade-offs that
drives the results in Proposition 2 and Corollary 1. Finally, we note that the principal
experiences an additional benefit from information disclosure which is not present in the
standard fixed-action-space information design problem. This extra benefit stems from the
fact that the principal is able to tailor the delegation set to different realizations of $\omega$, which
makes information disclosure overall more attractive, relative to the case in which she is
restricted to use a fixed delegation set. In Section 5, we provide an informal discussion on
how to disentangle these different forces.

4 Information vs. Flexibility

In this section, we carry out comparative statics on the principal’s optimal policy, with a focus
on formalizing the notion of substitutability between flexibility and information provision
that appears in Corollary 1.
In general, establishing the substitutes property formally proves challenging. The reason is that the two “inputs” in the principal’s “production function” are very different objects lying in different spaces: one input (the delegation policy) is a mapping from \( \Omega \) to subsets of \( Y \), while the other (the information policy) is a distribution over posterior means about \( \omega \) and thus belongs to \( \Delta(\Omega) \).

Nevertheless, we can take advantage of the fact that, in our simple environment, both the optimal information policy and the optimal delegation set (for every realization of \( x \in \Omega \)) can be described by a real number. This enables us to carry out comparative statics on the informativeness and the amount of flexibility given to the agent, and in particular on the interaction between these two along the principal’s optimal frontier. Specifically, in our setting, it is possible to rank optimal information policies under different model primitives according to Blackwell informativeness based on the magnitude of the censorship cutoff \( \omega^* \). Similarly it will be enough to order the size of delegation sets in terms of the highest type \( \theta \) whose first-best action belongs to the delegation set. As mentioned above, this cutoff type \( \theta^*(x) \) fully determines the width of the optimal interval delegation set.

With this in mind, we now formally define the concepts of flexibility of a delegation policy, and of informativeness of an information policy. Our definition of informativeness specializes Blackwell informativeness to the class of censorship information policies. Similarly, flexibility of a delegation policy \( D : \Omega \rightarrow 2^Y \), will be described by the associated cutoff function

\[
\theta^*_D(x) \equiv \sup(\{\theta \in \Theta : y_A(x, \theta) \in D(x)\} \cup \{\theta\}),
\]

which is the highest type \( \theta \) that is able to choose his first-best action under \( D(x) \).

**Definition 2** (Flexibility and informativeness).

- Let \( D : \Omega \rightarrow 2^Y \) and \( D' : \Omega \rightarrow 2^Y \) be two delegation policies, and let \( \theta^*_D : \Omega \rightarrow \Theta \) and \( \theta^*_{D'} : \Omega \rightarrow \Theta \) be the associated cutoff functions. We say that \( D \) is more flexible than \( D' \) if \( \theta^*_D(x) \geq \theta^*_D(x) \), and that it is strictly more flexible if moreover \( \theta^*_D(x) > \theta^*_D(x) \) for some \( x \in \Omega \).

- For two lower (respectively, upper) censorship information policies \( G_\omega \in \Delta(\Omega) \) and \( G'_\omega \in \Delta(\Omega) \) with censorship cutoffs \( \omega^* \) and \( \omega'^* \), we say that \( G_\omega \) is more informative than \( G'_\omega \) if \( \omega^* \leq \omega'^* \) (respectively, \( \omega^* \geq \omega'^* \)), and we say that it is strictly more informative if the inequality is strict.

The first part of the definition defines a partial order on the space of delegation policies. We emphasize that this yields a reasonable order so long as one restricts attention to convex-valued delegation policies (i.e, those where \( D(x) \) is convex for all \( x \in \Omega \), which by
Proposition 1 holds true under the principal’s optimal delegation policy.\footnote{To clarify this, consider the following non-convex valued delegation policies: $D(x) = \{y_A(x, \theta_3)\}$, and $D'(x) = \{y_A(x, \theta_1), y_A(x, \theta_2)\}$, with $\theta_1 < \theta_2 < \theta_3$. According to Definition 2, $D$ is strictly more flexible than $D'$, even though $D'$ induces more actions.} This condition also holds under the natural (suboptimal) delegation policies of no discretion and full discretion, which we study as benchmarks in Proposition 3. The second part of the definition coincides with Blackwell informativeness and thus is standard.

We begin by showing that informativeness in the optimal joint delegation/information design problem lies in between the two extreme cases of no discretion and full discretion. To do so, let $D^O(x)$ be the optimal interval delegation policy described in Proposition 1. Consider also the optimal no discretion delegation policy $D^{ND}(x) \equiv \{\gamma P x + \alpha P \mathbb{E} \theta\}$, and the full discretion one $D^{FD}(x) \equiv Y$. The former is the allocation that would arise if the principal refrained from contracting with the agent altogether and chose the action herself. The latter arises if the principal doesn’t have the power to restrict the set of actions that the agent may choose from. Observe that, by Definition 2, $D^{FD}(x)$ is strictly more flexible than $D^O(x)$, which in turn is more flexible than $D^{ND}(x)$ (strictly so unless $D^O(x) = D^{ND}(x)$ for all $x \in \Omega$).

For each of these delegation policies, let $y^j(x, \theta) = \arg\max_{y \in D^j(x)} u_A(x, \theta)$, $j \in \{O, ND, FD\}$, be the action chosen by the agent of type $\theta$ given his belief $x \in \Omega$. Consider the principal’s information design problem when the delegation policy is set equal to $D^j$: \footnote{We can show that, under $j = FD$, the principal is indifferent across all information policies if $\Delta = 1$.}

\[
\max_{G_\omega} \int_\Omega \int_\Theta u_P(y^j(x, \theta), x, \theta) dF_\theta(\theta) dG_\omega(x), \quad j \in \{O, ND, FD\} \quad (\text{OptInfo}_j)
\]

s.to \ (BP)

When $j = O$, the problem amounts to the outer information design problem discussed in Section 3.3. In all three cases, the problem has a unique solution, which we denote by $G^j_\omega \in \Delta(\Omega)$, for $j \in \{O, ND, FD\}$. Our next result says that, as we move from the most flexible to the least flexible delegation scheme, the informativeness of the optimal signal increases in the Blackwell sense. In order to guarantee uniqueness, we focus on the (generic) case in which $\Delta \neq 1$.\footnote{\textit{Proposition 3.} Suppose that $\Delta \neq 1$. For each $j \in \{O, ND, FD\}$, $(\text{OptInfo}_j)$ has a unique solution, denoted by $G^j_\omega$. Moreover, $G^{ND}_\omega$ is more informative than $G^O_\omega$ (strictly so if and only if $\Delta < \overline{\Delta}$), and $G^O_\omega$ is more informative than $G^{FD}_\omega$ (strictly so if and only if $\Delta < 1$ and $G^O_\omega$ is non-degenerate).}

The comparison between the optimal information policy under no discretion relative to...
the one under optimal delegation is straightforward. As argued in Section 3.3, the principal’s value function is strictly convex whenever she doesn’t give the agent any discretion and chooses the action herself (disregarding the information contained in $\theta$). Intuitively, she strictly benefits from disclosing information about $\omega$ because this allows her to make a more informed decision. Therefore, the optimal information policy $G_{\omega}^{ND}$ entails full disclosure, which trivially implies that $G_{\omega}^{O}$ must be less informative. Moreover, the comparison is strict whenever the $G_{\omega}^{O}$ doesn’t entail full disclosure as well, which by Proposition 2 is equivalent to $\Delta < \bar{\Delta}$.

The intuition behind the comparison between $G_{\omega}^{O}$ and $G_{\omega}^{FD}$ is similar to the one that appeared in Corollary 1. More precisely, this part of the result stems from two forces. The first one is that the principal’s “agency cost” from disclosing information is lower when the agent is given little discretion to act. The second force is reminiscent to the one driving the first part of the comparison in Proposition 3: under optimal interval delegation, the principal is able to tailor the delegation set to different realizations of $x$, whereas under full discretion the delegation set is fixed throughout. This renders information disclosure relatively more beneficial when the principal has the ability to optimally delegate.

A natural question to ask is whether or not a similar comparative statics obtains for any pair of interval-delegation policies that can be ordered in terms of their flexibility. The answer to this is elusive. The reason is that the shape of the principal’s value function as a function of the posterior mean about $\omega$ for a given (possibly suboptimal) interval-delegation policy will depend heavily on how the delegation policy changes with $x$. If the delegation policy is arbitrary, it will generally not be true that the resulting value function will satisfy the properties needed in order to carry out comparative statics on the informativeness of the optimal information policy (see Curello and Sinander (2022)).

Next, we ask how a change in the model’s primitives that leads the principal to optimally choose a more flexible delegation policy affects the informativeness of the of the optimal $G_{\omega}$. Specifically, we study the effects of a decrease in the agent’s additive bias from $\beta_A^\prime$ to $\beta_A < \beta_A^\prime$. As expected, this change leads the principal to offer a more flexible delegation policy. In Proposition 4, we show that it also leads the principal to optimally choose a less informative information policy. To state the result, we focus on the case in which, both under $\beta_A$ and $\beta_A^\prime$, the principal finds it optimal to give the agent some degree of discretion for any realization of $x \in \mathbb{R}$. Formally, letting $\hat{x}(\beta)$ be defined as in (3.1) with $\beta_A = \beta$, we require that $\hat{x}(\beta_A^\prime) = \bar{\omega}$ if $\gamma_A - \gamma_P \geq 0$, and that $\hat{x}(\beta_A^\prime) = \omega$ if $\gamma_A - \gamma_P < 0$.$^{18}$

$^{18}$Our proof for the informativeness comparison in Proposition 4 uses twice differentiability of the value function $V(x)$ which is ensured by this assumption. The flexibility comparison does not rely on this assumption.
Proposition 4. Let \( \beta_A' > \beta_A \), and \((G_\omega^{\beta_A}, D^{\beta_A})\) be the optimal joint delegation and information policy under \( \beta \in \{\beta_A', \beta_A\} \). If \( \gamma_A - \gamma_P \geq 0 \) (respectively, < 0) and \( \hat{x}(\beta_A') = \omega \) (respectively, \( \hat{x}(\beta_A) = \omega \)) then: \( D^{\beta_A} \) is strictly more flexible than \( D^{\beta_A'} \), and \( G_\omega^{\beta_A} \) is more informative than \( G_\omega^{\beta_A'} \) (strictly so if and only if \( G_\omega^{\beta_A} \) is non-degenerate and \( G_\omega^{\beta_A} \neq F_\omega \)).

A similar version of the comparative statics result in Proposition 4 regarding flexibility can be found in Alonso and Matouschek (2008). The main insight in the proposition pertains to the interplay between the optimal flexibility and informativeness granted to the agent. More specifically, Proposition 4 underscores that environments where the principal benefits from giving the agent more flexibility to choose an action (i.e., when the size of the agent’s bias is pointwise lower) must necessarily feature less information provision.

The result may appear to be counterintuitive, given that one may expect the principal to provide less information when the agent’s bias is higher. This in principle would be true if the action set that the agent chooses from was fixed. However, the principal’s ability to design the delegation policy allows her to hedge against the agent’s greater misalignment by optimally restricting the degree of discretion. As a result, the agency aspect of the information design problem will endogenously be less severe when the agent is more misaligned.\(^\text{19}\) It then follows from an intuition very similar to the one outlined in Section 3.3 that the principal benefits more from providing information when the agent’s bias is higher (and thus the delegation sets are tighter).

We conclude this section by applying these results to the monopoly regulation example, and by providing an independent comparative statics on \( \rho \) which does not follow directly from Propositions 3 and 4.

Example: Monopoly regulation (contd.) We can apply Proposition 3 to provide a comparison with respect to the full discretion benchmark studied in Bergemann et al. (2015). Under full discretion, we can show that the optimal policy features a degenerate market segmentation for all \( \rho \in (1, 2) \), which is obviously less informative than the market segmentation that arises under optimal price regulation. Still, we can apply Corollary 2 and Proposition 3, to argue that the segmentation under optimal delegation is strictly more informative than in the full discretion benchmark whenever the former involves some price discrimination, which is true when \( \rho \) is sufficiently high.

Another natural comparative statics question that arises in this setting is that of how the optimal policy changes with the regulator’s preference for redistribution, which in our example is governed by \( \rho \). We now derive an analog comparative statics result for \( \rho \) as the

\(^{19}\text{Whitmeyer (2022) provides conditions under which adding an (exogenously given) action to the receiver’s action set leads the information designer to choose a more informative policy.}\)
one in Proposition 4, stating that an increase in $\rho$ leads to a less flexible (price-) delegation policy, and a more informative market segmentation. This result doesn’t directly follow from Proposition 4, since, unlike $\beta_A$, $\rho$ is not a constant additive term in the agent’s bias. We provide a separate proof for it in the Appendix. As in Proposition 4, we restrict attention to the case in which the regulator finds it optimal to give the monopolist some discretion to choose the price for any realization of the market’s demand $\omega$. Namely, we assume that $\hat{x}(\rho') = \overline{\omega}$, where $\hat{x}(\hat{\rho})$ is as in (3.1) with $\rho = \hat{\rho}$.

Claim 1. Consider two monopoly regulation environments with $\rho' > \rho$, and suppose that $\hat{x}(\rho') = \overline{\omega}$. The respective optimal regulations policies are such that: i) in each segment of the market, the regulator imposes a lower price cap under $\rho'$, and ii) more price discrimination is allowed under $\rho'$.

Claim 1 states that a regulator who is more concerned with distribution of surplus will allow for more price discrimination. To understand this result, first note that a higher value of $\rho$ means that the regulator’s optimal price is steeper in $\omega$. The fact that she cares more about adjusting the price to the buyer’s valuation $\omega$ implies that her benefit from allowing price discrimination is higher. In the language of Proposition 2, the benefit-to-cost ratio of information provision, $\Delta$, increases. Interestingly, this is true even though the monopolist’s optimal price is strictly increasing in $\omega$, while the regulator’s is strictly decreasing—i.e., the monopolist’s pricing rule responds to information about consumer’s valuation in the opposite direction compared to the regulator’s first best. This mechanism is present regardless of the use of price caps.\footnote{To see this, consider the full discretion benchmark. In the region of parameters that we focus on, namely $\rho \in (1, 2)$, we saw that the regulator’s optimal market segmentation is constant in $\rho$ and equal to the uniform segmentation. However, it is easy to show that the optimal segmentation is fully informative if and only if $\rho \geq 2$. Thus, if we allow for a larger region of parameters, the mechanism described in this paragraph leads the informativeness of the optimal policy to increase with $\rho$ in the full discretion benchmark as well.}

Second, when $\rho$ increases, the gap between the monopolist’s and the regulator’s optimal price, $y_A(x, \theta) - y_P(x, \theta)$ increases pointwise. As a result, and as stated in Claim 1, the regulator will resort to more stringent price caps. This gives rise to an endogenous increase in the regulator’s benefits from allowing price discrimination, which we described in the discussion of Proposition 4. Namely, under the regulator’s optimal price cap rule, an increase in $\rho$ leads the monopolist to enjoy less pricing discretion in every segment of the market. This turns the monopolist’s tendency to overprice into a less severe problem for the regulator, thereby reducing the agency cost from allowing price discrimination.
5 Concluding Remarks

We study a joint information/mechanism design problem without transfers, in which a principal, who faces a privately informed agent, optimally chooses: (i) an information policy that determines the extent to which the agent’s chosen action can be conditioned on payoff-relevant information (which is orthogonal to the agent’s private type), and (ii) a delegation policy that imposes restrictions on the agent’s feasible actions based on the realized information. We derive the optimal policy and demonstrate that it takes a simple ‘double-censorship’ form. Firstly, the optimal information policy involves setting a cutoff on the state space, such that realizations on one side of the cutoff are perfectly revealed, while all remaining realizations are pooled together. Secondly, given every possible message under the optimal information policy, the optimal delegation policy uses an action cap, which induces full separation of agent types on one side of the cutoff, and pooling at a constant action on the other side. Additionally, we establish that the optimal policy provides the agent with more flexibility following realizations of the state in which the agent has less information, and vice versa.

To maintain tractability, our model relies on a set of stylized assumptions regarding players’ utility functions and the information structure. These assumptions enable us to offer a precise description of the optimal policy. In concluding, we discuss the extent to which our results remain robust when relaxing these assumptions, as well as considering alternative timings in the contracting problem. Additionally, we provide further intuition for our main result.

Decomposing the incentives to provide information. As discussed in Section 3.3, in our setting, information provision has two impacts on the principal’s payoffs. Firstly, it influences the agent’s choice of action, potentially leading to benefits from providing information to optimally influence the agent’s decisions. Secondly, the principal benefits from the ability to condition the mechanism on the disclosed information. While the former effect arises in any information design setting, the latter appears because in our setting the principal has the ability to design a message-contingent action set. In this section, we offer an informal discussion on how to evaluate the significance of these two effects.

Without loss of insight, suppose that $\gamma_A - \gamma_P \geq 0$. We focus on realizations of beliefs $x \in \Omega$ under which the principal benefits from giving the agent some discretion (i.e., $x < \hat{x}$, with $\hat{x}$ defined as in (3.1)), given that this is precisely the region under which the two effects mentioned in the previous paragraph are at play. Consider the principal’s payoff from using
an action cap \( \tilde{y} \in (y_A(x, \theta), y_A(x, \overline{\theta})) \) when the posterior belief about \( \omega \) is \( x \):

\[
v(x, \tilde{y}) \equiv \int_{\theta}^{\theta(x, \tilde{y})} u_P(y_A(x, \theta), x, \theta) \, dF_\theta(\theta) + \int_{\theta(x, \tilde{y})}^{\tilde{y}} u_P(\tilde{y}, x, \theta) \, dF_\theta(\theta),
\]

Where \( \theta(x, \tilde{y}) = \{ \theta \in \Theta : y_A(x, \theta) = \tilde{y} \} \). Observe that the principal’s value in the delegation problem is \( V(x) = \max_{\tilde{y} \in Y} v(x, \tilde{y}) \), and that the optimal cap is \( \overline{y}(x) \equiv \arg \max_{\tilde{y} \in Y} \overline{v}(x, \tilde{y}) = y_A(x, \theta^*(x)) \). By the envelope theorem, the second derivative of \( V \) on \( x < \hat{x} \) can be written as

\[
V''(x) = \underbrace{v_{11}(x, \overline{y}(x))}_{\text{Agent’s response to information}} + \underbrace{ \overline{y}'(x) v_{12}(x, \overline{y}(x)) }_{\text{Information used in the mechanism}}.
\]

We argue that (5.1) decomposes the principal’s incentives to provide information into the two effects mentioned above.

Intuitively, a higher second derivative of \( V \) is associated to a higher benefit from providing information. This is made explicit in the proof of Proposition 2: There, a pointwise increase in \( V''(x) \) leads the principal to choose a smaller (lower-)censorship cutoff, and therefore to use a more Blackwell-informative policy.\(^{21}\) In the extreme case in which \( V''(x) > 0 \) for all \( x \in \Omega \) then the optimal policy is full disclosure, and the opposite is true when \( V''(x) < 0 \) for all \( x \in \Omega \). So we can informally assess the importance of the two effects by evaluating how they affect the magnitude of the second derivative in (5.1).

The first term in (5.1) captures the curvature of the principal’s value function if the action cap is fixed (evaluated at its optimal value \( \overline{y}(x) \)), and therefore isolates the effect through how the agent changes his action in response to information about \( \omega \). The sign of this effect is ambiguous, and hinges on the degree of misalignment between the principal and the agent. When we explicitly compute this term, we find that it is strictly positive when the parameter that we identified as the benefit-to-cost ratio of information provision (\( \Delta \)) is high enough, and that it is strictly negative whenever \( \Delta \) is sufficiently close to zero.

The second term is always strictly positive and captures the principal’s benefit from using information to optimally choose the action cap.

In summary, depending on the severity of the agent’s misalignment, the two effects in (5.1) may act in opposite directions. As we showed in Proposition 2, if the disagreement between the principal and the agent regarding how to optimally respond to information about \( \omega \) is strong enough, then a negative effect through the ‘agent’s response to information’ may

\(^{21}\)A general version of this comparative statics, that goes beyond censorship information policies, can be found in Curello and Sinander (2022).
fully offset the second positive effect. This would lead the principal to optimally withhold information, and in that way to relinquish the benefits from conditioning the mechanism on \( \omega \). If, on the other hand, the disagreement is small, these two effects go in the same direction and reinforce one another.

**Parametric assumptions on players’ utilities.** Naturally, one may ask to what extent our results are driven by the functional form assumptions on the principal’s and the agent’s utility functions. In this section, we argue that the key tension highlighted in the paper, leading to information disclosure being relatively less beneficial for the principal when the agent has more flexibility, continues to play a role more generally. However, the actual shape of the optimal policy might change, and in particular the presence of new opposing forces may result in ambiguity with respect to whether information and discretion are complements or substitutes for the principal.

To illustrate this, we discuss the consequence of relaxing the assumption that players’ optimal action, \( y_i(\omega, \theta) \), features no cross-effects between \( \omega \) and \( \theta \).\(^{22}\) It is possible to verify that, when \( (\partial^2 y_i(\omega, \theta)) / (\partial \omega \partial \theta) \neq 0 \) for some \( i \in \{A, P\} \), then the principal’s value function need no longer be S-shaped, and therefore the optimal information policy may not be a censorship policy. This further complicates the comparison in terms of how flexibility and information provision relate to each other at the principal’s optimum. Still, without fully solving for this version of the model, we can argue that (keeping other features of the model fixed) the sign of \( (\partial^2 y_i(\omega, \theta)) / (\partial \omega \partial \theta) \) will determine whether the substitutes property in Proposition 2 is reinforced or offset in the presence of these cross-effects. To fix ideas, suppose that the agent’s bias \( y_A(\omega, \theta) - y_P(\omega, \theta) \) is increasing in \( \omega \), which implies that the optimal amount of flexibility decreases with \( x \), and that \( (\partial^2 y_i(\omega, \theta)) / (\partial \omega \partial \theta) = \lambda_i \), where \( \lambda_i \) is a constant.

If the agent’s optimal action is supermodular in \( (\omega, \theta) \), then the result in Proposition 2 remains true: the optimal information policy features one-sided censorship, with realizations of the state that get fully revealed corresponding to tighter delegation sets. Informally, supermodularity in the agent’s optimal action makes the substitutes property between information provision and flexibility **stronger**. The intuition is that if \( \lambda_A > 0 \), an increase in the cutoff type \( \theta^{*} \) makes the agent’s marginal response to information about \( \omega \), as measured by \( \gamma_A + \lambda_A \theta^{*} \), higher. The assumption that \( y_A(\omega, \theta) - y_P(\omega, \theta) \) is increasing in \( \omega \) implies that the slope of the agent’s optimal action is already too high relative to the principal’s. Hence, when \( \lambda_A > 0 \), providing information becomes even costlier for the principal under realiza-

\(^{22}\)Similar conclusions can be obtained if we relax the assumption of quadratic payoffs, or if we dispense with Assumption 2.
tions of $x$ that feature a higher delegation cutoff $\theta^*(x)$. The opposite is true when $\lambda_A < 0$. In fact, when $\lambda_A < 0$, it is possible to construct an example where the optimal information policy may feature the opposite type of censorship: Namely, one where realizations of $\omega$ that get fully disclosed coincide with larger delegation sets.\footnote{In the general case, when $\lambda_A < 0$, the optimal information policy need not be a censorship.}

While it is possible for the latter type of complementarity in players’ optimal actions to arise in practice, resulting in a different form of the optimal policy compared to what we describe in Proposition 2, we consider these forces to be more mechanical in nature. Our analysis deliberately abstracts from any direct interactions between the two sources of uncertainty in players’ payoffs. By doing so, we identify a substitutes property between (information about) $\omega$ and $\theta$ that exclusively stems from the agency relationship.

**Alternative timings: inscrutability and private communication.** We have focused on a setting in which communication between the principal and the agent takes place after the information about $\omega$ is publicly revealed. In this environment, the optimal mechanism has an indirect implementation by means of ‘delegation sets’: for every realization of the posterior belief about $\omega$, the principal may simply restrict the agent’s action space, after which the agent will choose an action in an incentive compatible manner.

An alternative contracting problem would involve the principal eliciting a report from the agent prior to releasing information about $\omega$. This ex-ante communication has the potential to benefit the principal in two ways. Firstly, by asking the agent to report their type before learning about $\omega$, the principal faces a weaker form of incentive compatibility that requires truthful reporting to be optimal for the agent in expectation, rather than holding for every message realization as required in our setting. This is the well-known inscrutability principle (Myerson, 1983), which states that the designer does not benefit from disclosing information to the agent before eliciting a report. Secondly, under this protocol, it becomes feasible for the principal to engage in private communication by offering a menu of information policies $(G(x|\theta))_{\theta \in \Theta}$ that are conditioned on the agent’s reports. Due to these two features, our two-stage solution approach, which relied on the separability of the principal’s objective with respect to the posterior belief about $\omega$, is no longer applicable. Hence, extending our results to this environment is not straightforward, leaving open the possibility that the principal can achieve strictly better outcomes by eliciting the agent’s type prior to providing information about $\omega$.

As mentioned earlier, the timing that we study in this paper involves a straightforward implementation where the principal does not directly communicate with the agent but instead delegates decision-making while imposing constraints on the action space. It can be
argued that in many practical scenarios, the principal may be limited to using such simple delegation mechanisms, even if it results in some loss of optimality. Our results provide insights into the optimal policy when these constraints are present. However, analyzing the optimal information/mechanism design with ex-ante communication is an interesting question that we leave for future research.

References


BHASKAR, D., A. McCLELLAN, AND E. SADLER (2020): “Regulation design in insurance markets,” Available at SSRN 3501883.


Appendix

A Proofs

Proof of Proposition 1

Assumption 1 implies that part one in Proposition 2 of Alonso and Matouschek (2008) applies, and thus the optimal delegation set is connected. One can then solve for the one-dimensional problem of finding the optimal delegation policy within the class that are implementable by an interval delegation set. For completeness, we provide a direct proof (tailored to our environment), which also allows to establish uniqueness of the optimal mechanism. Certain steps in the proof will be used several times when showing Proposition 2.

We begin by reformulating the principal’s problem, using a standard monotonicity and envelope characterization of incentive compatibility.

Lemma 1. Given \( x \in \Omega \), the optimal delegation policy solves

\[
\max_{y(\cdot)} \quad U + \int_{\hat{\theta}} \left[ \alpha_A(1 - F_\theta(\theta)) - (y_A(x, \theta) - y_P(x, \theta)) f_\theta(\theta) \right] y(\theta) d\theta \\
\text{s.to} \quad U = u_A(y(\theta), x, \theta) \\
y(\cdot) \text{ non-decreasing}
\]

\[
u_A(y(\theta), x, \theta) = U + \alpha_A \int_\theta^\theta y(s) ds. \tag{ENV}
\]

Proof. The fact that monotonicity of \( y(\cdot) \) and (ENV) are necessary and sufficient for (IC) follows from standard arguments.\(^{24}\) Using this, we can write the principal’s objective in (Opt Del) as:

\[
\int_{\hat{\theta}} \left( -\frac{y(\theta)^2}{2} + y(\theta)y_P(x, \theta) \right) f_\theta(\theta) d\theta = \int_{\hat{\theta}} \left[ u_A(y(\theta), x, \theta) + y(\theta)(y_P(x, \theta) - y_A(x, \theta)) \right] f_\theta(\theta) d\theta = \\
= U + \int_{\hat{\theta}} \left[ \alpha_A(1 - F_\theta(\theta)) - (y_A(x, \theta) - y_P(x, \theta)) f_\theta(\theta) \right] y(\theta) d\theta
\]

Where the last equality follows from (ENV) and integration by parts. \( \square \)

We now apply Lemma 1 to show that the mechanism described in Proposition 1 is optimal.\(^{24}\) See Milgrom and Segal (2002).
For a fixed $x \in \Omega$, define the function $h_x : \Theta \to \mathbb{R}$ as

$$h_x(\theta^*) \equiv \int_{\theta^*}^{\theta} \frac{\partial u_p(y_A(x, \theta^*), x, \theta)}{\partial y} dF_{\theta}(\theta) = (1 - F_{\theta}(\theta^*))[\mathbb{E}_{\theta}[y_P(x, \theta)|\theta \geq \theta^*] - y_A(x, \theta^*)].$$

We distinguish two cases depending of whether or not the condition $y_A(x, \theta) < \mathbb{E}_{\theta}[y_P(x, \theta)]$ holds.

**Case 1 :** $y_A(x, \theta) < \mathbb{E}_{\theta}[y_P(x, \theta)]$. Suppose first that $x$ is such that $y_A(x, \theta) < \mathbb{E}_{\theta}[y_P(x, \theta)]$, and thus $h_x(\theta) > 0$. Since $h_x(\theta) < 0$ in a left-neighborhood of $\theta$, by continuity, there is $\theta^*(x) \in (\theta, \bar{\theta})$ such that $h_x(\theta^*(x)) = 0$, and $h_x(\cdot)$ crosses zero from above at $\theta^*(x)$—i.e. $\theta^*(x)$ satisfies the following second-order condition

$$h_x'(\theta^*(x)) = [y_A(x, \theta^*(x)) - y_P(x, \theta^*(x))]f_{\theta}(\theta^*(x)) - \alpha_A(1 - F_{\theta}(\theta^*(x))) < 0. \quad (A.1)$$

Moreover, by Assumption 1, $h_x(\cdot)$ is convex, which implies that $\theta^*(x)$ is the unique down-crossing of $h_x(\theta)$.

For a fixed value of $x$, let $y^*(\theta)$ be the mechanism described in Proposition 1, with $\theta^*(x)$ being the unique down-crossing of $h_x(\theta)$. Let $y(\theta)$ be any other incentive compatible mechanism, and let $U(\theta) = u_A(y(\theta), x, \theta)$. Letting $W(y)$ denote the value of the principal under mechanism $y$ and applying Lemma 1, we can write

$$W(y^*) - W(y) = u_A(y_A(x, \theta), x, \theta) - U(\theta) - \int_{\theta}^{\theta^*} h_x'(\theta)(y^*(\theta) - y(\theta)) d\theta =$$

$$u_A(y_A(x, \theta), x, \theta) - U(\theta) - \int_{\theta}^{\theta^*(x)} h_x'(\theta)(y_A(x, \theta) - y(\theta)) d\theta - \int_{\theta^*(x)}^{\theta^*} h_x'(\theta)(y_A(x, \theta^*(x)) - y(\theta)) d\theta. \quad (A.2)$$

We can rewrite the first term as

$$A \overset{(1)}{=} u_A(y_A(x, \theta), x, \theta) - U(\theta) + \int_{\theta}^{\theta^*(x)} h_x'(\theta) \left( \frac{U'(\theta)}{\alpha_A} - y_A(x, \theta) \right) d\theta \overset{(2)}{=}$$

$$(u_A(y_A(x, \theta), x, \theta) - U(\theta)) \left( 1 + \frac{h_x'(\theta)}{\alpha_A} \right) + \frac{h_x'(\theta^*(x))}{\alpha_A} (U(\theta^*(x)) - u_A(y_A(x, \theta^*(x)), \theta^*(x), x)) +$$

$$\frac{1}{\alpha_A} \int_{\theta}^{\theta^*(x)} \left( \frac{y_A(x, \theta)^2}{2} - U(\theta) \right) d\theta \overset{(3)}{\geq} 0,$$

Where (1) follows from (ENV), and (2) from integration by parts for the Riemann-Stieltjes integral. To show (3), we claim that each term on the left-hand-side of it is positive. This is because $U(\theta) \leq u_A(y_A(x, \theta), x, \theta)$ for all $\theta \in \Theta$, by definition of $y_A(x, \theta)$ being type-$\theta$'s
first-best action. Moreover, \(1 + \frac{h_x' \left( \theta \right)}{\alpha_\alpha} = \frac{\left( y_A(x, \theta) - y_p(x, \theta) \right) f_\alpha \left( \theta \right)}{\alpha_\alpha} \geq 0\). Thus, the first term is non-negative. The result for the second term follows from (A.1) and using again optimality of \(y_A(x, \theta^* (x))\). Finally, the integrand in the third term is equal to \((u_A(y_A(x, \theta), x, \theta) - U(\theta)) dh_x' (\theta)\), which is non-negative due to optimality of \(y_A(x, \theta)\) and convexity of \(h_x(\theta)\).

The second term in \(W(y^*) - W(y)\) can be written as

\[
B = \int_{\theta^*(x)}^{\bar{\theta}} h_x' (\theta) y(\theta) d\theta \geq 0,
\]

Where (1) follows from using the fact that \(h_x(\bar{\theta}) = h_x(\theta^*(x)) = 0\), and (2) obtains from using this fact again and Riemann-Stieltjes integration by parts. (3) holds because \(h_x(\theta) \leq 0\) for all \(\theta \geq \theta^*(x)\), and \(dy(\theta) \geq 0\) for all \(\theta \in \Theta\) by incentive compatibility of \(y\). Therefore, \(W(y^*) \geq W(y)\) for any incentive compatible \(y\), and thus the mechanism in Proposition 1 is optimal.

Finally, to establish uniqueness, we observe that strict convexity of \(h_x(\theta)\) implies that at least one of the inequalities in (A.2) and (A.3) will be strict unless: \(i) U'(\theta) = u_A(y_A(x, \theta), x, \theta)\) almost-everywhere in \(\theta < \theta^*(x)\), \(ii) U'(\theta^*(x)) = u_A(y_A(x, \theta^*(x)), x, \theta^*(x))\), and \(iii) dy(\theta) = 0\) almost-everywhere in \(\theta > \theta^*(x)\). Indeed, a violation of \(i)\) or \(ii)\) leads inequality (A.2) to hold strictly, whereas the inequality in (A.3) becomes strict if \(iii)\) doesn’t hold. These three conditions uniquely select the mechanism in Proposition 1.

Case 2: \(y_A(x, \theta) \geq \mathbb{E}_\theta \left[ y_p(x, \theta) \right]\). In this case, \(h_x(\theta) \leq 0\). By convexity of \(h_x(\cdot)\) and the fact that \(h_x(\bar{\theta} - \eta) < 0\) when \(\eta > 0\) is sufficiently small, it follows that \(h_x(\theta) < 0\) for all \(\theta \in (\bar{\theta}, \overline{\theta})\).

We argue that any optimal mechanism has to be constant almost everywhere in \(\theta\). To see this, consider an incentive compatible mechanism \(y(\cdot)\). We set \(y(\cdot)\) to be right-continuous, which is without loss by monotonicity. We can write the principal’s payoff under this mechanism as

\[
W(y) = U(\bar{\theta}) - \int_{\theta}^{\bar{\theta}} h_x'(\theta) y(\theta) d\theta = U(\bar{\theta}) + h_x'(\bar{\theta}) y(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} h_x(\theta) dy(\theta) \leq U(\bar{\theta}) + h_x'(\bar{\theta}) y(\bar{\theta}),
\]

Where the inequality follows from monotonicity of \(y(\theta)\) and the fact that \(h_x(\theta) < 0\) for all \(\theta \in (\bar{\theta}, \overline{\theta})\). It is a strict inequality whenever \(y(\theta) > y(\bar{\theta})\) for a positive measure of \(\theta\). The right-hand-side of the inequality is equal to the principal’s payoff under the constant mechanism where \(y(\theta) = y(\overline{\theta})\) for all \(\theta \in \Theta\). Thus, the optimal mechanism must be (almost everywhere) constant. Since the mechanism in Proposition 1 is uniquely optimal within the class of constant mechanisms, it is overall the unique optimal mechanism.
Proof of Proposition 2

Throughout, we focus on the case in which \( \gamma_A - \gamma_P \geq 0 \). The proof for \( \gamma_A - \gamma_P < 0 \) is analogous to the case with \( \gamma_A - \gamma_P > 0 \) which we analyze below.

Preliminaries: Properties of the solution to the delegation problem.

We start by deriving some useful properties of the optimal cutoff \( \theta^*(x) \) and the value function \( V(x) \). Let \( \theta'^*(x) \) be the fist derivative of \( \theta^*(x) \), and \( V'(x) \) and \( V''(x) \) be, respectively, the first and second derivatives of \( V(x) \). A key step of the proof is Lemma 4, where we show that \( V(x) \) has at most one inflection point. This will allow us to establish optimality of a censorship information policy in the next part of the proof.

Lemma 2.

i) \( \theta^*(x) \) is weakly decreasing (increasing) if \( \gamma_A - \gamma_P \geq (\leq)0 \), strictly so if \( \gamma_A - \gamma_P \neq 0 \).

ii) \( \theta'^*(x) \) exists and is continuous everywhere on \( \Omega \), except for at most one point.

Proof. Recall from Proposition 1 that \( \theta^*(x) \) is defined by

\[
\theta^*(x) = \begin{cases} 
\{ \theta^* \in \Theta : \bar{u}_P(x, \theta^*) = y_A(x, \theta^*) \}, & \text{if } x < \hat{x}, \\
\theta^*, & \text{if } x \geq \hat{x}. 
\end{cases}
\]

Which is continuous on \( \Omega \). If \( \hat{x} > \underline{\omega} \), it follows from (A.1) and interiority of \( \theta^*(x) \), that \( \theta^*(x) \) is differentiable on \((\underline{\omega}, \hat{x})\) with derivative

\[
\theta'^*(x) = -\frac{\partial h_x(\theta^*(x))}{\partial x} \frac{\partial}{\partial \theta^*} \int_{\theta^*(x)}^{\bar{\theta}} \left( \frac{\partial u_A(x, \theta^*(x))}{\partial x} - \frac{\partial u_P(x, \theta)}{\partial x} \right) dF_\theta(\theta) \frac{(\gamma_A - \gamma_P)(1 - F_\theta(\theta^*(x)))}{h_x'(\theta^*(x))}.
\]

This is continuous under our assumptions since it’s the composition of continuous functions. Moreover, \( \gamma_A - \gamma_P \geq 0 \) and (A.1) imply that \( \theta'^*(x) \leq 0 \) on \( x \in (\underline{\omega}, \hat{x}) \) (strictly so if \( \gamma_A - \gamma_P > 0 \)). On the other hand, \( \theta'^*(x) = 0 \) on \( x \in (\hat{x}, \bar{\omega}) \). Thus, \( \theta'^*(x) \) is well defined and continuous for all \( \Omega \) except possibly at \( \hat{x} \). In fact, if \( \gamma_A - \gamma_P > 0 \) and \( \hat{x} > \underline{\omega} \), then \( \theta^*(x) \) has a kink at \( \hat{x} \).

Lemma 3. \( V(x) \) is continuously differentiable on \( \Omega \).

Proof. Applying Proposition 1, for \( x \geq \hat{x} \), we can write

\[
V(x) = V_{ND}(x) \equiv \max_{y \in Y} \int_{\theta}^{\bar{\theta}} u_P(y, x, \theta) dF_\theta(\theta),
\]
Where we define \( V_{ND}(x) \) to be the principal’s value when she gives the agent no discretion. By our assumption on \( Y \), the maximizer is interior and equal to \( y_P(x, \theta) \). Thus, we can write

\[
V'(x) = V'_{ND}(x) = \int_{\theta}^{\theta^*} \frac{\partial u_P(y_P(x, \theta), x, \theta)}{\partial x} dF_\theta(\theta) = \gamma_P y_P(x, \theta),
\]

Which is continuous on \( x > \hat{x} \). If \( \hat{x} = \omega \), this establishes continuity of \( V'(x) \) on \( \Omega \).

If \( \hat{x} > \omega \), for \( x < \hat{x} \), we have:

\[
V(x) = V_{OID}(x) = \int_{\theta}^{\theta^*} u_P(y_A(x, \theta), x, \theta) dF_\theta(\theta) + \int_{\theta^*}^{\theta} u_P(y_A(x, \theta^*(x)), x, \theta) dF_\theta(\theta),
\]

Where we define \( V_{OID}(x) \) to be the principal’s value under optimal interval delegation. Because the optimal cutoff \( \theta^*(x) \) is interior, the envelope theorem implies that

\[
V'(x) = V'_{OID}(x) = \int_{\theta}^{\theta^*} \left( \gamma_A \frac{\partial u_P(y_A(x, \theta), x, \theta)}{\partial y} + \frac{\partial u_P(y_A(x, \theta), x, \theta)}{\partial x} \right) dF_\theta(\theta) + \int_{\theta^*}^{\theta} \frac{\partial u_P(y_A(x, \theta^*(x)), x, \theta)}{\partial x} dF_\theta(\theta),
\]

Where we are using the optimality condition \( h_x(\theta^*(x)) = 0 \). This is continuous in \( x \) due to continuity of \( \theta^*(x) \), \( u_P(y, \cdot, \theta) \) and \( y_A(\cdot, \theta) \).

Finally, we verify that \( V'(x) \) is continuous at \( \hat{x} \). To do so, we note that \( \lim_{x \uparrow \hat{x}} \theta^*(x) = \underline{\theta} \), so we can write

\[
\lim_{x \uparrow \hat{x}} V'(x) = \lim_{x \uparrow \hat{x}} V'_{OID}(x) = \int_{\theta}^{\theta^*} \frac{\partial u_P(y_A(\hat{x}, \theta), \hat{x}, \theta)}{\partial x} dF_\theta(\theta) = \int_{\theta}^{\theta^*} \frac{\partial u_P(y_P(\hat{x}, \theta), \hat{x}, \theta)}{\partial x} dF_\theta(\theta) = \lim_{x \downarrow \hat{x}} V'_{ND}(x) = \lim_{x \downarrow \hat{x}} V'(x) = V'(\hat{x}),
\]

Where the third equality follows from the fact that \( y_A(\hat{x}, \theta) = y_P(\hat{x}, \theta) \) by definition of \( \hat{x} \).

\[\square\]

\textbf{Lemma 4.}

i) If \( \gamma_A - \gamma_P > 0 \), there exists \( x^* \in [\omega, \hat{x}] \) such that \( V(x) \) is strictly concave on \( x \leq x^* \), and \( V(x) \) is strictly convex on \( x \geq x^* \).

ii) If \( \gamma_A - \gamma_P < 0 \), there exists \( x^* \in [\hat{x}, \omega] \) such that \( V(x) \) is strictly concave on \( x \geq x^* \), and \( V(x) \) is strictly convex on \( x \leq x^* \).

iii) If \( \gamma_A - \gamma_P = 0 \), then \( V(x) \) is strictly convex on \( \Omega \).
Proof. Write \( V(x) = \max\{V_{ND}(x), V_{OID}(x)\} \). \( V_{ND}(x) \) is twice continuously differentiable on \( \Omega \), with second derivative

\[
V''_{ND}(x) = \gamma_P^2.
\]

On the other hand, by Lemma 2, \( V_{OID}(x) \) is twice continuously differentiable on \( (\omega, \hat{x}) \cup (\hat{x}, \varpi) \), with right-second-derivative:

\[
V''_{OID}(x) = \begin{cases} 
-(\gamma_A - \gamma_P)^2 + \gamma_P^2 - \frac{\alpha_A(\gamma_A - \gamma_P)^2(1 - E_\theta(\theta^*(x)))^2}{h_\theta'(\theta^*(x))}, & \text{if } x < \hat{x} \\
-(\gamma_A - \gamma_P)^2 + \gamma_P^2, & \text{if } x \geq \hat{x}
\end{cases}
\]

If \( \gamma_A - \gamma_P = 0 \), \( V''_{OID}(x) \) is twice continuously differentiable on \( \Omega \), and \( V''_{OID}(x) = V''_{ND}(x) = \gamma_P^2 > 0 \), and thus \( V(x) \) is strictly convex since it is the pointwise maximum of two convex functions. This establishes point \( iii) \). Similarly, if \( \hat{x} = \omega \), then \( V(x) = V_{ND}(x) \) for all \( x \in \Omega \) which is strictly convex.

Suppose now that \( \gamma_A - \gamma_P > 0 \) and \( \hat{x} > \omega \). Let \( x^* \equiv \inf(\{x \in [\omega, \hat{x}] : V''_{OID}(x) \geq 0\} \cup \{\varpi\}) \). Observe that, if \( x < \omega \), then \( x^* < \hat{x} \), given that

\[
\lim_{x \uparrow \hat{x}} V''_{OID}(x) = \gamma_P^2 - \frac{(\gamma_A - \gamma_P)^2(y_A(\hat{x}, \theta) - y_P(\hat{x}, \theta))f_\theta(\theta)}{h_\theta'(\theta)} > 0.
\]

By definition of \( x^* \), \( V(x) \) is strictly concave on \([\omega, x^*]\). The result in Lemma 4 would follow immediately if \( x^* = \omega \). So suppose now that \( x^* < \omega \). We want to show that \( V'(x) \) is strictly increasing for all \( x \in (x^*, \omega] \). Take \( x \in (x^*, \hat{x}] \), in which case we can write

\[
V''(x) = V''_{OID}(x) = -(\gamma_A - \gamma_P)^2 + \gamma_P^2 - \frac{\alpha_A(\gamma_A - \gamma_P)^2(1 - E_\theta(\theta^*(x)))^2}{h_\theta'(\theta^*(x))} \overset{(1)}{=} -\frac{(\gamma_A - \gamma_P)^2}{h_\theta'(\theta^*(x))} + \gamma_P^2 + \alpha_A(\gamma_A - \gamma_P)^2 M(\theta^*(x)) > -\frac{(\gamma_A - \gamma_P)^2}{h_\theta'(\theta^*(x))} + \gamma_P^2 + \alpha_A(\gamma_A - \gamma_P)^2 M(\theta^*(x^*)) \overset{(3)}{=} V''(x^*) \overset{(4)}{=} 0,
\]

Where (1) and (3) follow from applying Proposition 1 which implies that, for all \( x' \in (\omega, \hat{x}) \),

\[
y_A(x, \theta^*(x')) - y_P(x, \theta^*(x')) = \alpha_P(E_\theta[\theta \geq \theta^*(x')] - \theta^*(x')).
\]

(2) follows from Assumption 2 and the fact that \( \theta^*(x) < \theta^*(x^*) \) by Lemma 2, and (4) follows from the definition of \( x^* \).

Thus, \( V(x) \) is strictly convex on \([x^*, \hat{x}] \). Furthermore, it is strictly convex on \([x^*, \varpi] \), since \( V'(x) \) is continuous at \( \hat{x} \) by Lemma 2 (and thus it doesn’t discontinuously drop at \( \hat{x} \)), and strictly increasing on \([\hat{x}, \varpi] \) given that \( V_{ND}(x) \) is strictly convex.

Optimality of censorship information policies.
We are now ready to show Proposition 2. We consider three different cases, which depend on the value of the cutoff $x^*$, defined as in Lemma 4, and show that under all cases the optimal information policy takes the form of a censorship. Our proof is closely related to the proof of Theorem 1 in Kolotilin et al. (2022), who show that censorship policies are uniquely optimal when the value function is strictly S-shaped, meaning that it is strictly concave to one side of a cutoff and strictly convex to the other side of it. This result doesn’t immediately carry over to our setting, since some of the assumptions made in that paper (such as twice differentiability of the sender’s function) do not hold in our setting. We extend their proof to account for these differences.

**Case 1: Optimality of full disclosure.**

If $x^* = \omega$, then the value function is strictly convex in $\Omega$ and therefore full disclosure (i.e., a lower censorship with cutoff $\omega^* = \omega$) is uniquely optimal. We summarize this finding in the following Lemma.

**Lemma 5.** If $\gamma_A - \gamma_P \geq 0$ (respectively, $\gamma_A - \gamma_P < 0$) and $x^* = \omega$ (respectively, $x^* = \overline{\omega}$), then full disclosure is uniquely optimal.

**Case 2: Optimality of interior censorship.**

We describe the conditions under which the optimal information policy features an interior censorship cutoff. We proceed through a series of auxiliary Lemmas (Lemmas 6 and 7), and conclude by stating the sufficient conditions for optimality of an interior censorship cutoff in Lemmas 8 and 9. As before, to avoid repetition, we provide arguments only for the case in which $\gamma_A - \gamma_P \geq 0$, and formally state the results for the case in which $\gamma_A - \gamma_P < 0$ without providing a proof.

Suppose that $x^* > \omega$ and that $\gamma_A - \gamma_P \geq 0$. By Lemma 4, $x^* > \omega$ implies that $\gamma_A - \gamma_P \neq 0$. Let $m(x) \equiv \mathbb{E}[\omega|\omega \leq x]$, and define the function $H : \Omega \to \mathbb{R}$ as

$$H(x) \equiv V(m(x)) - V(x) + V'(m(x))(x - m(x)).$$

(A.2)

Intuitively, an optimal lower censorship cutoff has to satisfy the first-order condition $H(\omega^*) = 0$. Lemma 6 gives conditions under which such an optimal cutoff exists and is unique.

**Lemma 6.** If $\gamma_A - \gamma_P \geq 0$, $x^* > \omega$ and $H(\overline{\omega}) < 0$, there exists a unique $\omega^* \in (x^*, \overline{\omega})$ such that: i) $H(\omega^*) = 0$, and ii) $H(x)$ is strictly decreasing on a neighborhood of $\omega^*$. Moreover, $m(\omega^*) \leq x^*$.

**Proof.** Because $V(x)$ is strictly concave on $(m(x^*), x^*)$, it holds that $H(x^*) > 0$. It then follows by the assumption that $H(\overline{\omega}) < 0$ and continuity of $V'(x)$ in Lemma 3 that there

---

25See also Dworczak and Martini (2019) for a closely related approach.
exists $\omega^* \in (x^*, \overline{\omega})$ such that $H(\omega^*) = 0$, and that $H(x)$ crosses zero strictly from above at $\omega^*$.

Next, we show that $m(\omega^*) \leq x^*$. Suppose by contradiction that $m(\omega^*) > x^*$. By continuity of $m(x)$, we can pick an $x' < \omega^*$ sufficiently close to $\omega^*$ that satisfies $m(x') > x^*$. For such an $x'$ we can write

$$H(x') = -\int_{m(x')}^{x'} [V'(s) - V'(m(x'))] \, ds \leq 0,$$

where the inequality follows from the fact that $V(x)$ is weakly increasing on $(m(x'), x') \subset (x^*, \overline{\omega})$. This contradicts $H(x') > 0$ for all $x'$ belonging to a left-neighborhood of $\omega^*$.

To show uniqueness, suppose by contradiction that there exist $\omega^*$ and $\omega^{*'}$ that satisfy conditions $i)$ and $ii)$ in Lemma 6. Without loss of generality, suppose that $\omega^* < \omega^{*'}$. We can then write

$$H(\omega^{*'}) = -\int_{m(\omega^{*})}^{\omega^{*'}} [V'(s) - V'(m(\omega^{*'}))] \, ds = -\int_{m(\omega^{*})}^{\omega^{*}} [V'(s) - V'(m(\omega^*))] \, ds =$$

$$H(\omega^*) + \int_{m(\omega^*)}^{m(\omega^{*})} [V'(s) - V'(m(\omega^*))] \, ds - \int_{m(\omega^{*})}^{\omega^{*}} [V'(s) - V'(m(\omega^*))] \, ds \leq -\int_{\omega^*}^{\omega^{*'}} [V'(s) - V'(\omega^*)] \, ds \leq 0,$$

where (1) follows from the fact that $m(\omega^*) < m(\omega^{*'}) \leq x^*$ implies that $V'(m(\omega^*)) > V'(m(\omega^{*'}))$. (2) follows from $H(\omega^*) = 0$ and $V'(s) < V'(m(\omega^*))$ for all $s \in (m(\omega^*), m(\omega^{*'}))$ since $V$ is concave on that interval. (3) follows from the fact that, by definition of $\omega^*$,

$$(\omega^*-m(\omega^*))V'(m(\omega^*)) = \int_{m(\omega^*)}^{x^*} V'(s) \, ds + \int_{x^*}^{\omega^*} V'(s) \, ds \leq (x^*-m(\omega^*))V'(m(\omega^*))+x^* V'(\omega^*),$$

which implies that $V'(m(\omega^*)) \leq V'(\omega^*)$. (4) follows from the fact that $V$ is convex on $(\omega^*, \omega^{*'})$. We thus have $H(\omega^{*'}) < 0$, a contradiction. \hfill \Box

We now show that a lower censorship policy with cutoff $\omega^*$ as given by Lemma 6 is uniquely optimal. Formally, this policy can be described by the cdf

$$G_{\omega^*}(x) \equiv \begin{cases} 
0, & \text{if } x \in [\omega, m(\omega^*)) \\
F_\omega(\omega^*), & \text{if } x \in [m(\omega^*), \omega^*) \\
F_\omega(x), & \text{if } x \in [\omega^*, \overline{\omega}]. 
\end{cases}$$
Proof. Part i) is immediate, given that supp($G_{\omega^*}$) = $\{m(\omega^*)\} \cup [\omega^*, \overline{\omega}]$.

For part ii), fix $x \in [\omega, \omega^*) \setminus \{m(\omega^*)\}$. Suppose first that $x \leq x^*$, in which case we have
\[
p(x) = V(\omega^*) + V'(m(\omega^*))(x - \omega^*) = V(m(\omega^*)) + V'(m(\omega^*))(x - m(\omega^*)) + V'(m(\omega^*))(m(\omega^*) - \omega^*) + V(\omega^*) - V(m(\omega^*)) > V(x) + V'(m(\omega^*))(m(\omega^*) - \omega^*) + V(\omega^*) - V(m(\omega^*)) = V(x),
\]

Where the inequality follows from the fact that $V$ is strictly concave on $(x, m(\omega^*))$ (respectively, on $(m(\omega^*), x)$ if $x < m(\omega^*)$ (respectively, $x > m(\omega^*)$), and thus $V(m(\omega^*)) + V'(m(\omega^*))(x - m(\omega^*)) > V(x)$. The final equality follows from the fact that $H(\omega^*) = 0$.

If $x \in [x^*, \omega^*)$, letting $\alpha \in (0, 1]$ be such that $x = \alpha x^* + (1 - \alpha)\omega^*$ (i.e., $\alpha = \frac{\omega^* - x}{\omega^* - x^*}$), we can write
\[
p(x) = \alpha p(x^*) + (1 - \alpha)p(\omega^*) \geq \alpha V(x^*) + (1 - \alpha)V(\omega^*) > V(x),
\]

Where the first inequality follows from the fact that $p(x^*) \geq V(x^*)$ by the previous step, and $p(\omega^*) = V(\omega^*)$; and the second inequality follows from strict convexity of $V$ on $(x^*, \omega^*)$.

Next, we show part iii). First, $p(x)$ is piecewise convex: it is affine to the left of $\omega^*$, and it is strictly convex to the right of $\omega^*$ given that $\omega^* > x^*$. Second, it has a kink at $\omega^*$, and the slope increases at this point. This follows because
\[
\lim_{x \uparrow \omega^*} p'(x) = V'(m(\omega^*)) \leq V'(\omega^*) = \lim_{x \downarrow \omega^*} p'(x).
\]
Thus, \( p(x) \) is convex on \( \Omega \).

Finally, for part \( iv \), applying the definitions of \( G_{\omega^*} \) and \( p \), we have

\[
\int_{\omega^*} p(x) dG_{\omega^*}(x) = F_{\omega}(\omega^*)V(m(\omega^*)) + \int_{\omega^*} V(x) dF_{\omega}(x) = \int_{\omega^*} p(x) dF_{\omega}(x).
\]

At this point, we could directly apply Theorem 1 in Dworczak and Martini (2019) which states that, if a pair \((p, G_{\omega^*})\) satisfies the conditions in Lemma 7, then \( G_{\omega^*} \) is optimal. However, because we are interested in establishing uniqueness of the solution as well, we provide a short direct proof.

**Lemma 8.** If \( \gamma_A - \gamma_P \geq 0 \), \( x^* > \omega \) and \( H(\omega) < 0 \), then a lower censorship policy with cutoff \( \omega^* \in (x^*, \omega) \) is uniquely optimal.

**Proof.** Let \( G \in \Delta(\Omega) \) be any mean-preserving contraction of \( F_{\omega} \). The principal’s value under \( G \) is

\[
\int_{\omega^*} V(x) dG(x) \quad \leq \quad \int_{\omega^*} p(x) dG(x) \quad \leq \quad \int_{\omega^*} p(x) dF_{\omega}(x) = \quad \int_{\omega^*} p(x) dG_{\omega^*}(x) \quad \leq \quad \int_{\omega^*} V(x) dG_{\omega^*}(x).
\]

Where (1) follows from part \( ii \) of Lemma 7; (2) follows from part \( iii \) and \( F_{\omega} \) being a mean-preserving spread of \( G \); (3) follows from part \( iv \); and (4) follows from part \( i \). Therefore, \( G_{\omega^*} \) is optimal.

To show that it is uniquely optimal, let \( G \) be any other optimal information policy, in which case inequalities (1) and (2) must hold with equality. To evaluate (2), write

\[
\int_{\omega^*} p(x)[F_{\omega}(x) - G(x)] = \int_{\omega^*} p(x)[F_{\omega}(x) - G(x)] = \int_{\omega^*} V(x) [F_{\omega}(x) - G(x)].
\]

Since \( V(x) \) is strictly convex on \([\omega^*, \omega]\), (2) can hold with equality only if all realizations of \( \omega \) above \( \omega^* \) are revealed.

On the other hand, to evaluate (1), we write

\[
\int_{\omega} [p(x) - V(x)] dG(x) = \int_{\omega^*} [p(x) - V(x)] dG(x).
\]

By part \( ii \) of Lemma 7, the integrand is strictly negative for all \( x \neq m(\omega^*) \). Thus, (1) holds with equality if and only if \( G \) assigns all the remaining probability mass to \( m(\omega^*) \). These two conditions uniquely select \( G_{\omega^*} \) as the optimal information policy.

\[\square\]
Finally, we state without proof the counterpart to Lemma 8 for the case where $\gamma_A - \gamma_P < 0$. Let $\hat{m}(x) = \mathbb{E}[\omega|\omega \geq x]$ and

$$\hat{H}(x) = V(x) - V(\hat{m}(x)) + V'(\hat{m}(x))(x - \hat{m}(x)).$$

If $\hat{H}(\omega) > 0$, let $\omega^{**} \in (\omega, x^*)$ be the unique down-crossing of $\hat{H}(x)$, which exists by analogous arguments as in Lemma 6.

**Lemma 9.** If $\gamma_A - \gamma_P < 0$, $x^* < \overline{\omega}$ and $\hat{H}(\omega) > 0$, then an upper censorship policy with cutoff $\omega^{**} \in (\omega, x^*)$ is uniquely optimal.

**Case 3: Optimality of no disclosure.**

We now show that, in all remaining cases, the optimal policy is no disclosure.

**Lemma 10.** If $\gamma_A - \gamma_P \geq 0$ (respectively, $\gamma_A - \gamma_P < 0$) and $H(\overline{\omega}) \geq 0$ (respectively, $\hat{H}(\omega) \leq 0$), then no disclosure is uniquely optimal.

**Proof.** Suppose that $\gamma_A - \gamma_P \geq 0$ and that $H(\overline{\omega}) \geq 0$. Observe that $H(\overline{\omega}) \geq 0$ implies that $\mathbb{E}\omega < x^*$. To see this, note that $\mathbb{E}\omega \geq x^*$ implies that $V$ is strictly convex on $(\mathbb{E}\omega, \overline{\omega})$, which in turn implies that $H(\overline{\omega}) < 0$.

Consider the affine dual function

$$\hat{p}(x) = V(\mathbb{E}\omega) + V'(\mathbb{E}\omega)(x - \mathbb{E}\omega),$$

and the uninformative information policy $G_{\overline{\omega}}$. Note that $\hat{p}(\mathbb{E}\omega) = V(\mathbb{E}\omega)$.

Moreover, we can show that $V(x) < \hat{p}(x)$ for all $x \in \Omega \setminus \{\mathbb{E}\omega\}$. To see this, consider first $x \in [\omega, x^*] \setminus \{\mathbb{E}\omega\}$. In this case, $V(x) < V(\mathbb{E}\omega) + V'(\mathbb{E}\omega)(x - \mathbb{E}\omega) = \hat{p}(x)$, by strict concavity of $V(x)$ on $(\omega, \mathbb{E}\omega) \subset (\omega, x^*)$. For $x \in (x^*, \overline{\omega}]$, writing $x = \alpha x^* + (1 - \alpha)\overline{\omega}$ with $\alpha \in (0, 1]$, we have

$$\hat{p}(x) = \alpha \hat{p}(x^*) + (1 - \alpha)\hat{p}(\overline{\omega}) > \alpha V(x^*) + (1 - \alpha)V(\overline{\omega}) = V(x),$$

where the first inequality follows from $\hat{p}(x^*) > V(x^*)$ by and $H(\overline{\omega}) \geq 0$, and the last inequality follows from convexity of $V$ on $(x^*, \overline{\omega})$.

Now, let $G \in \Delta(\Omega)$ be a mean-preserving contraction of $F_\omega$. Similar to the proof of Lemma 8, we have

$$\int_\omega V(x) dG(x) \overset{(1)}{=} \int_\omega \hat{p}(x) dG(x) \overset{(2)}{=} \int_\omega \hat{p}(x) dF_\omega(x) \overset{(3)}{=} \int_\omega \hat{p}(x) dG_{\overline{\omega}}(x) \overset{(4)}{=} \int_\omega V(x) dG_{\overline{\omega}}(x),$$

Where (1) follows from $V(x) \geq \hat{p}(x)$ for all $x \in \Omega$, and the remaining relations are obvious.
Given that $V(x) < \tilde{p}(x)$ for all $x \neq E\omega$, inequality (1) will be strict unless $G(x)$ assigns all the probability mass to $E\omega$. Therefore, $G(x)$ is uniquely optimal.

Combining the above three cases (which exhaust all possibilities), we have established that a censorship policy is uniquely optimal. Moreover, these cases provide necessary and sufficient conditions for the optimal censorship cutoff to be interior.

Conditions on the optimal censorship cutoff. To complete the proof, we now consider the scenarios i)-ii) described in Proposition 2 when $\gamma_A - \gamma_P \geq 0$.

First, let us derive a necessary and sufficient condition for optimality of full disclosure. By the above lemmas, it suffices to check that $V''(x) > 0$ in a neighborhood of $\omega$. This is obvious when $\hat{x} = \omega$. If $\hat{x} > \omega$ and letting $V''(\omega)$ be the right-limit of $V''(x)$, we can write

$$V''(\omega) = -(\gamma_A - \gamma_P)^2 + \gamma_P^2 - \alpha_A(1 - F_\theta(\theta^*(\omega)))^2 h_x'(\theta^*(\omega)).$$

Thus, if $\hat{x} > \omega$, full disclosure is optimal if and only if

$$\Delta > \bar{\Delta} \equiv 1 + \frac{\alpha_A(1 - F_\theta(\theta^*(\omega)))^2}{h_x'(\theta^*(\omega))}.$$  \hspace{1cm} (A.3)

Finally, by continuity, we know from the above paragraph that $\omega^* \to \omega$ as $\Delta \uparrow \bar{\Delta}$, and that $\omega^* > \omega$ for all $\Delta < \bar{\Delta}$. Therefore, there is $\Delta < \bar{\Delta}$ such that $\omega^* \in (\omega, \bar{\omega})$ for all $\Delta \in (\Delta, \bar{\Delta})$.

**Proof of Proposition 3**

Consider first the information design problem under no discretion, i.e. $j = ND$. In this case, using the definitions in the proof of Proposition 2, we can write the principal’s problem as

$$\max_{G_\omega} \int_\Omega V_{ND}(x) dG_\omega(x)$$

$$s.to \hspace{1cm} (BP).$$

We showed in Proposition 2 that $V_{ND}(x)$ is strictly convex, and therefore full disclosure is uniquely optimal. Therefore $G_{\omega}^{ND} = F_\omega$. It then follows immediately from the fact that $G_\omega^O$ satisfies (BP) that $G_{\omega}^{ND}$ is a mean-preserving spread of $G_\omega^O$, and that the relation is strict when $G_\omega^O \neq F_\omega$, which in turn is equivalent (by Proposition 2) to $\Delta < \bar{\Delta}$.

Next, consider the principal’s problem under full discretion. Define the principal’s value
from inducing the posterior mean $x \in \Omega$ as

$$V_{FD}(x) = \int_{\Theta} u_P(y_A(x, \theta), x, \theta) dF_\theta(\theta).$$

As before, the shape of the optimal information policy will depend on the curvature of the function $V_{FD}(x)$. One can check that this function is twice-differentiable, with second derivative

$$V''_{FD}(x) = - (\gamma_A - \gamma_P)^2 + \gamma_P^2,$$

Which is either strictly positive (when $\Delta > 1$) or strictly negative (when $\Delta < 1$). In either case, the optimal policy is unique, and will involve a censorship with a cutoff belonging to one of the endpoints of $\Omega$.

Suppose first that $\Delta > 1$, and thus $G_{FD}^\omega = F^\omega$. Recall from (A.3) that $\Delta < 1$, and thus by Proposition 2, $G_O^\omega = F^\omega$, so the optimal information policy in the two problems coincides. Next, suppose that $\Delta > 1$, in which case $G_{FD}^\omega$ is the Dirac delta function on $E^\omega$. Then, $G_O^\omega$ is more informative than $G_{FD}^\omega$, in the strict sense if and only if $G_O^\omega \neq G_{FD}^\omega$—i.e., $G_O^\omega$ is non-degenerate.

**Proof of Proposition 4**

As in the rest of the Appendix, we focus on the case in which $\gamma_A - \gamma_P \geq 0$. The arguments for the case with $\gamma_A - \gamma_P < 0$ are analogous. Let $\beta'_A > \beta_A$.

**Effects on flexibility.** For $\beta \in \{\beta_A, \beta'_A\}$, define

$$\hat{x}(\beta) = \left( \max \{ x \in \Omega : (\gamma_A - \gamma_P)x + \beta + \alpha_A(\theta - \mathbb{E}\theta) \leq 0 \} \cup \{ \omega \} \right).$$

Recall from Proposition 1 that $\hat{x}(\beta)$ is the cutoff on $x$ above which the agent is given no discretion, i.e., $\theta_D^\beta(x) = \theta$ for all $x \geq \hat{x}(\beta)$. Clearly, $\hat{x}(\beta_A) \geq \hat{x}(\beta'_A)$, and the inequality is strict unless $\hat{x}(\beta_A) = \omega$. Thus, $\theta_{D^\beta_A}^*(x) \geq \theta_{D^\beta'_A}(x) = \theta$ for all $x \geq \hat{x}(\beta'_A)$.

Next, suppose that $\hat{x}(\beta_A) > \omega$ (if not, the proof is completed). It follows from Proposition 1 that, for each $x < \hat{x}(\beta_A)$, the optimal cutoff $\theta_{D^\beta_A}^*(x)$ strictly decreases with $\beta_A$, and thus $\theta_{D^\beta'_A}^*(x) < \theta_{D^\beta_A}^*(x)$ for all $x < \hat{x}(\beta_A)$. This establishes the result (without using the assumption that $\hat{x}(\beta'_A) = \omega$).

**Effects on informativeness.** Henceforth, we assume that $\hat{x}(\beta'_A) = \bar{\omega}$, which implies (by the preceding part) that $\hat{x}(\beta_A) = \bar{\omega}$. It then follows that, for $\beta \in \{\beta_A, \beta'_A\}$, the value function $V(x; \beta)$ is twice differentiable. We now show that $V''(x; \beta_A) < V''(x; \beta'_A)$ for all $x \in \Omega$. To
show this, we write for any \( x \in \Omega \),

\[
V''(x; \beta'_A) = -(\gamma_A - \gamma_P)^2 + \gamma_P^2 + \alpha_A (\gamma_A - \gamma_P)^2 M(\theta_{D\beta_A}^*(x)) > 0 \quad (\ast)
\]

\[-(\gamma_A - \gamma_P)^2 + \gamma_P^2 + \alpha_A (\gamma_A - \gamma_P)^2 M(\theta_{D\beta_A}^*(x)) = V''(x; \beta_A),
\]

Where the inequality follows from the fact that \( \theta_{D\beta_A}^*(x) < \theta_{D\beta_A}^*(x) \) for all \( x \in \Omega \) (by the first part of this proof) and Assumption 2.

We now use this result to derive the informativeness comparison between \( G_{\omega}^{\beta_A} \) and \( G_{\omega}^{\beta'_A} \). For \( \beta \in \{ \beta_A, \beta'_A \} \), let \( \omega^*(\beta) \) be the optimal censorship cutoff under \( \beta \). Let \( H(x; \beta) = V(m(x); \beta) - V(x; \beta) + V'(m(x); \beta)(x - m(x)) \), which incorporates \( \beta \) as an argument in the function defined in (A.2).

If \( V''(\omega; \beta'_A) > 0 \) the result is immediate, since then \( V(\cdot; \beta'_A) \) is strictly convex on \( \Omega \) and the optimal policy under \( \beta'_A \) is fully informative. So suppose now that \( V''(\omega; \beta'_A) \leq 0 \), which implies by (\ast) that \( V''(\omega; \beta_A) \leq 0 \). By Lemmas 8 and 10, we have that for all \( \beta \in \{ \beta_A, \beta'_A \} \),

\[
\omega^*(\beta) = \inf(\{ x \in \Omega : H(x; \beta) \leq 0 \} \cup \{ \Omega \}.
\]

So, to show that \( \omega^*(\beta'_A) \leq \omega^*(\beta_A) \) it is enough to show that \( H(x; \beta'_A) \leq H(x; \beta_A) \) for all \( x \in \Omega \). We can show this by writing

\[
H(x, \beta'_A) = -\int_{m(x)}^{x} [V'(s; \beta'_A) - V'(m(x); \beta'_A)] ds - \int_{m(x)}^{x} \int_{m(x)}^{s} V''(t; \beta'_A) dt ds \quad (1)
\]

\[-\int_{m(x)}^{x} (x - s)V''(s; \beta'_A) ds < -\int_{m(x)}^{x} (x - s)V''(s; \beta_A) ds = H(x; \beta_A),
\]

Where (1) follows from integration by parts, and (2) follows from (\ast).

Combining the above arguments, we have that \( \omega^*(\beta'_A) \leq \omega^*(\beta_A) \), and the inequality is strict unless either \( \omega^*(\beta_A) = \omega \) (the optimal policy under \( \beta_A \) and \( \beta'_A \) is fully informative), or \( \omega^*(\beta_A) = \omega \) (the optimal policy under \( \beta_A \) and \( \beta'_A \) is fully uninformative).

**Proof of Claim 1**

Fix \( \rho' \) and \( \rho \) such that \( \rho' > \rho \). Recall that the bias in the monopoly regulation example is given by \( y_A(x, \theta) - y_P(x, \theta) = \frac{\rho(x-\theta)}{2(2-\rho)} \), which is strictly increasing in \( \rho \) for any \( (x, \theta) \). Therefore, an increase in \( \rho \) leads to a pointwise increase in the agent’s bias. We can then apply an argument analogous to the first part of the proof of Proposition 4 to show that the delegation policy under \( \rho' \) is less flexible than that under \( \rho \). Moreover, the comparison is strict, unless the optimal policy under \( \rho' \) features no discretion.

47
Next, let $\omega^*(\hat{\rho})$ be the optimal censorship cutoff for $\hat{\rho} \in \{\rho, \rho'\}$, and let $V(x; \hat{\rho})$ and $
abla^*_\hat{\rho}(x)$ be the respective value function and optimal cutoff from the delegation problem. The assumption that $\hat{x}(\rho') = \overline{\omega}$ implies that $V(x; \hat{\rho})$ is twice continuously differentiable for all $\hat{\rho} \in \{\rho, \rho'\}$. The comparison is immediate if $V''(\omega, \rho') > 0$, since then the optimal policy under $\rho'$ is full disclosure.

So suppose now that $V''(\omega, \rho') \leq 0$. Define

$$\tilde{H}(x; \hat{\rho}) \equiv -(2 - \hat{\rho}) \int_{m(x)}^x [V'(s; \hat{\rho}) - V'(m(x); \hat{\rho})] ds = -(2 - \hat{\rho}) \int_{m(x)}^x (x - s)V''(s; \hat{\rho}) ds,$$

i.e., $\tilde{H}$ is a rescaling (by a factor $2 - \hat{\rho} > 0$) of the function defined in (A.2). We rewrite it on the right-hand-side using integration by parts as in the proof of Proposition 4.

By Lemmas 8 and 10, we can write

$$\omega^*(\rho') = \inf(\{x \in \Omega : \tilde{H}(x; \rho') \leq 0\} \cup \{\overline{\omega}\}).$$

As in the proof of Proposition 4, we will show that $\tilde{H}(x; \rho') < \tilde{H}(x; \rho)$ for all $x \in \Omega$. To show this, we write

$$(2 - \rho')V''(x; \rho') = \frac{3\rho' - 2}{4} + \frac{\rho'^2}{8(2 - \rho')} M(\nabla^*_\rho(x)) > \frac{3\rho' - 2}{4} + \frac{\rho'^2}{8(2 - \rho')} M(\nabla^*_\rho(x)) = (2 - \rho)V''(x; \rho),$$

Where (1) follows from $\nabla^*_\rho(x) < \nabla^*_\rho(x)$ by the previous step and Assumption 2, and (2) follows from (A.1) which implies that $M(\nabla^*_\rho(x)) > 0$.

($\dagger$) implies that $V''(\omega, \rho) < 0$ whenever $V''(\omega, \rho') \leq 0$. Thus, we can write $\omega^*(\rho) = \inf(\{x \in \Omega : \tilde{H}(x; \rho) \leq 0\} \cup \{\overline{\omega}\})$. Moreover, ($\dagger$) implies that $\tilde{H}(x; \rho) > \tilde{H}(x; \rho')$, and therefore $\omega^*(\rho') \leq \omega^*(\rho)$. The inequality is strict unless either $\omega^*(\rho) = \omega$, or $\omega^*(\rho') = \overline{\omega}$. 

48