Optimal Delegation with Information Manipulation

Yichuan Lou*

Feb 28, 2021

Abstract

A principal delegates a decision to an agent, who has the capacity to process the relevant information. The principal cannot process information herself, but can jointly control the actions and the information available to the agent. I provide sufficient conditions under which the optimal mechanism (i) attains a perfect alignment of incentives—subject to the constraints on discretion and information, the agent plays exactly as the principal would want him to; and (ii) belongs to a simple class of delegation and disclosure rules, called monotone partitional rules, which specify a finite set of allowable actions from which the agent chooses his preferred one, and moreover, partition the state space into finite intervals so that the agent only learns to which interval the true state belongs. I then turn to the uniform-quadratic case which permits an explicit characterization of optimal mechanisms and clear comparative statics. Finally, I discuss two applications: the regulation of a monopolist, and the self-control of a dynamically inconsistent individual.

*Department of Economics, The University of Tokyo, 7-3-1 Hongo, Tokyo, 113-0033, Japan. E-mail: yichuanlou@e.u-tokyo.ac.jp. For continual guidance, I am grateful to my advisor Alessandro Lizzeri and my committee members Erik Madsen and Ennio Stachetti. For helpful comments and suggestions, I thank Dilip Abreu, Sylvain Chassang, Boyan Jovanovic, Sam Kapon, Xiaoye Liao, Laurent Mathevet, David Pearce, Jacopo Perego, Debraj Ray, and Ariel Rubinstein.

1 Introduction

Consider a principal who grants decision-making authority to an agent who is potentially better informed but has different objectives. Previous work has explored two ways for the principal to influence the agent's behavior in order to improve decision-making in this environment. The first way is by specifying a permissible set of actions (a delegation set) from which the agent may choose, which has been analyzed as a delegation problem (Holmström, 1977, 1984). The second way is by limiting the access to information to the agent, which has been analyzed as a persuasion problem (Kamenica and Gentzkow, 2011). In this paper, I combine these two ways of influencing behavior and ask how to simultaneously constrain discretion and information of the agent. The main result identifies a strong form of incentive compatibility: the principal's and the agent's incentives are *fully aligned* under the optimal joint restriction on discretion and information. In addition, the optimal mechanism belongs to a simple class of delegation and disclosure rules, called *monotone partitional rules*, that feature a finite set of actions and a monotone partitional information structure. Consequently, the agent only learns which partition element the true state lies in and then chooses from the delegation set an action that best serves his interests.

There are many environments that share the modeling features described above. For instance, a monopolist may collect and use consumer behavior data to segment the market and to make targeted offers. To prevent excessive discrimination and to protect consumer welfare, a regulator can adopt policy tools such as price control which limits the prices that can be charged, and data protection which limits the use of consumer data for discrimination. Similarly, in jury trial procedures, judges enforce detailed rules on what evidence can and cannot be presented to jurors, and additionally jurors are delegated a particular set of decisions they can make (e.g., the set of charges they can decide on). Finally, in the context of congressional oversight, congressional committees have the power only to inquire into matters and to make decisions within the scope of the authority delegated to them by their parent bodies. In all these settings, a principal delegates a permissible set of actions to an agent and additionally designs what information the agent has access to. The purpose of this paper is to provide an in-depth analysis for the joint control of actions and information.

I develop a stylized model in which a principal (she) relies on an agent (he) to make an informed decision and cannot use contingent transfers.[1] I begin my analysis by assuming that the principal can jointly commit to a delegation set and an information structure. The delegation set is a set of permissible actions from which the agent can freely choose. The information structure is modeled as a Blackwell experiment similar to that of Kamenica and Gentzkow (2011): a probability distribution over signal realizations as a function of the state of nature. After privately observing a signal realization, the agent then takes an action from the delegation set. I focus on the following question: what is an optimal mechanism as a combination of delegation and disclosure rules from the principal's perspective?

I begin by establishing two important results that greatly simplify the analysis. First, I show that the optimal mechanism is characterized by an *alignment principle*—subject to the constraints on discretion and information, the agent plays as if he shared the principal's preferences. More specifically, if the signal realization were publicly observed, the two parties would choose the same action from the delegation set. In fact, they may disagree about their most preferred actions, but given the well-adjusted noise of the signal and the moderate sparsity of the delegation set, both agree upon how to play optimally. An immediate implication is that the optimal delegation set must be finite, as the alignment property requires the implemented actions to be isolated far away from each other (relative to the preference divergence of the two parties).

Second, I show that the optimal mechanism features a finite delegation set and a monotone partitional information structure. Although the principal's choice of signals is not restricted a priori, in the optimal mechanism she partitions the state space into a finite disjoint union of intervals and pools each interval into a single signal realization. The principal,

[1]There are many situations in practice, like those described above, that transfers are infeasible.

in effect, introduces noise into the signal by letting the agent observe only which partition element the true state actually lies in. This feature greatly simplifies the characterization of the optimal mechanism by allowing me to focus on a finite-dimensional optimization problem over the cutoffs of intervals and the associated delegated actions.

Then, I explicitly characterize optimal mechanisms in a leading case of the model—the so-called uniform-quadratic case. This allows me to explicitly solve for the optimal mechanism and in turn to perform a comparative analysis of the optimal discretion and information bestowed on the agent. I show that the principal would always provide more discretion and more information to a less-biased agent. Also, the principal necessarily obtains higher welfare if the agent is less biased. In Section 6, I further investigate to what extent the comparative statics results that hold in the leading case can be generalized beyond the uniform-quadratic specification.

I next apply my results to two specific economic settings. First, I study the regulation of a monopolist who engages in both quality and price discrimination. In this context, a welfare-maximizing regulator jointly restricts prices (i.e., price control) and consumer information (i.e., data protection) available to a profit-maximizing monopolist. The information about consumer characteristics can be used to offer different quality levels and prices to different segments of the market. I provide conditions under which the optimal joint regulatory policy takes the form of a finite set of prices and a finite market segmentation. Specifically, the aggregate market is divided into finite segments in such a way that in each segment, the monopolist freely sets a quality level and chooses a price (from the delegated set of prices) so as to maximize his profits. The optimality of finite market segmentations implies that consumer privacy should be moderately preserved.

Next, I apply my results to a self-control problem. In the primary environment I study, the principal and the agent represent distinct individuals with conflict of interests. However, my model also captures a behavioral setup in which the principal and the agent represent the same individual at two points in time, who overweighs instantaneous rewards relative to future payoffs (e.g., hyperbolic discounting). As is well known, sophisticated, yet time-inconsistent individuals value commitment power (Strotz, 1995). Within a consumption-externality environment (Carrillo and Mariotti, 2000), I suppose that the current self possesses not only *hard commitment power* in that she can restrict the consumption choices available to the future self, but also *soft commitment power* in that she can restrict information (about the extent of the externality) available to the future self. I establish that people may choose to tie their hands by reducing the set of available consumption choices and to remain partially ignorant about the magnitude of the externality.

2 Related Literature

Since the paper studies the joint design of delegation and disclosure rules, it brings together the literatures on optimal delegation and Bayesian persuasion. The delegation literature begins with Holmström (1977, 1984), who considers a setting in which a principal faces a privately informed but biased agent and in which contingent transfers are infeasible. The agent is then given discretion over actions (in the form of a set of permissible actions) because his private information is valuable to the principal; the principal limits the degree of discretion because of preference misalignment. A key goal in this literature has been to identify sufficient conditions under which the principal optimally defines the permissible set as an interval. Melumad and Shibano (1991) solve the optimal delegation problem with quadratic preferences and a uniform distribution for the state. Following that, Martimort and Semenov (2006), Alonso and Matouschek (2008), and Amador and Bagwell (2013) extend the analysis by allowing for more general preferences and distributions, and provide optimal conditions for interval delegation. Most of these studies consider the agent's information as exogenous, whereas my work contributes by considering the information structure as an endogenous component. I am able to explore how the qualitative features of the optimal delegation set are altered in the presence of information manipulation.

Also, my work is related to the literature on Bayesian persuasion (Rayo and Segal 2010; Kamenica and Gentzkow 2011), where a sender chooses a signal and a receiver takes an action after observing a realization from this signal. The choice of a signal is isomorphic to my modeling of a disclosure rule, although the interpretation is different: my principal controls the information but does not observe the realization. Within this literature, several papers have also identified sufficient conditions for the optimality of monotone partitional signals in various settings; see, for example, Kolotilin (2018), Dworczak and Martini (2019), Ivanov (2021), and Mensch (2021).

The behavioral formulation captured by my model relates this paper to the literature on optimal provision of commitment devices for individuals with low self-control due to time-inconsistent preferences. Amador, Werning, and Angeletos (2006) study the problem of commitment through the restrictions on actions. In the context of a consumption-savings problem, they find that imposing a minimum level of savings (i.e., interval delegation) is always a feature of the optimal solution. On the other hand, Carrillo and Mariotti (2000) and Bénabou and Tirole (2002) study the problem of commitment through the manipulation of information. Both papers show that, in the presence of time inconsistency, Blackwell garbling of future self's information may increase the current self's payoff. My paper thus provides the first study of optimal joint provision of hard and soft commitment devices in self-control problems.

3 Model

3.1 Environment

There are two parties in the model on which I focus: a principal (she) and an agent (he). The principal has a payoff function $u_P(y, \theta) = \theta y - C_P(y)$, while the agent has a payoff function $u_A(y,\theta) = \theta y - C_A(y)$.[2] The value of y represents an action, and the value of θ represents a state.

I assume that θ has an absolutely continuous distribution μ_0 with bounded support $\Theta \equiv [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}$. The action y is chosen from a large compact interval $Y \subset \mathbb{R}$. For the remaining of the paper, I impose the following conditions on the primitives:

Assumption 1. *The following hold:*

(i) Both the functions $C_P : Y \to \mathbb{R}$ and $C_A : Y \to \mathbb{R}$ are continuously differentiable and strictly convex in y;

(ii) The functions $c_P \equiv C_P'$ and $c_A \equiv C_A'$ satisfy $c_P > c_A$.

Condition (i) in Assumption 1 ensures that the principal's preferred action $y_P(\theta) \equiv c_P^{-1}(\theta)$ and the agent's preferred action $y_A(\theta) \equiv c_A^{-1}(\theta)$ are both continuous and strictly increasing in $\theta \in \Theta$. Condition (ii) in Assumption 1 guarantees that the agent is always biased towards higher actions relative to the principal. Note that for each $\theta \in \Theta$, the principal's and the agent's payoff functions, $u_P(\cdot,\theta)$ and $u_A(\cdot,\theta)$, are single-peaked, respectively, around $y_P(\theta)$ and $y_A(\theta)$.

As I focus on decision-making with incomplete information, it is useful to identify the players' preferences over actions when the state is unknown. Let μ be a posterior belief and $m = \mathbb{E}_\mu[\theta]$ be the associated posterior mean. Since u_P and u_A are *linear* in the state θ, both players' expected payoffs depend on μ only through m. With slight abuse of notation, I will write $u_P(y,m) = my - C_P(y)$ and $u_A(y,m) = my - C_A(y)$ as the principal's and the agent's expected payoffs given y and m, respectively. Thus, given the posterior mean m, the principal's preferred action is $y_P(m) = c_P^{-1}(m)$ and the agent's is $y_A(m) = c_A^{-1}(m)$.

An interpretation. The model can be interpreted as follows. The two parties equally

[2] In fact, one can extend the analysis to more general payoff functions. Suppose that $u_P(y,\theta) = B(\theta)y - C_P(y) + D_P(\theta)$ and $u_A(y,\theta) = B(\theta)y - C_A(y) + D_A(\theta)$, where $B(\cdot)$ is strictly increasing. First, I can discard $D_P(\theta)$ and $D_A(\theta)$ without loss of generality because these two terms do not enter into the players' strategic considerations. Moreover, I can simply redefine the state as $\omega \equiv B(\theta)$.

share the benefits generated by an action y for any given state θ, which is equal to θy. The action y, however, renders different levels of cost to the principal and the agent: $C_P(y)$ and $C_A(y)$, respectively. Moreover, the agent incurs a lower marginal cost from higher actions, i.e., $c_A(y) < c_P(y)$, and is thus biased toward actions that are too high from the principal's perspective.

Two specifications. The set of preferences described above entails two particular specifications. The first specification is the case of quadratic preferences with a constant bias, which is widely-studied in the literature on cheap talk and delegation (for example, Crawford and Sobel 1982; Melumad and Shibano 1991; Alonso and Matouschek 2008). In this specification, the principal's payoff is $-\frac{(y-\theta)^2}{2}$ and the agent's payoff is $-\frac{(y-\theta-b)^2}{2}$, where y represents the action taken by the agent and $b > 0$ represents the agent's positive bias. This formulation is equivalent to letting $u_P(y,\theta) = \theta y - \frac{1}{2}y^2$ with $y_P(\theta) = \theta$, and $u_A(y,\theta) = \theta y - (\frac{1}{2}y^2 - by)$ with $y_A(\theta) = \theta + b$, and is therefore a special case. I explicitly solve for optimal mechanisms in this specification in Section 5.

The second specification considers the temptation and self-control problem of consumption with time-inconsistent preferences due to hyperbolic discounting, analyzed by Carrillo and Mariotti (2000). There is an individual with a horizon of three periods $t = 0, 1, 2$. At period $t = 1$, before taking a consumption decision x which generates an instantaneous utility $u(x)$, *self*-1 (the agent) learns about the probability ω of a negative externality $e(x)$ on the welfare of *self*-2. The welfare for *self*-1 from periods $t = 1, 2$ with externality probability ω is then $u(x) - \beta\omega e(x)$, where $0 < \beta \leq 1$. In contrast, the welfare for *self*-0 (the principal) from periods $t = 1, 2$ is $u(x) - \omega e(x)$. The discounting parameter β reflects the salience of the present.[3] The formulation is equivalent to $u_P(y,\theta) = \theta y - e(u^{-1}(y))$ and $u_A(y,\theta) = \theta y - \beta e(u^{-1}(y))$ by setting $\theta = \frac{1}{\omega}$ and $y = u(x)$, and hence is a special case as

[3]This setup represents a three-period version of quasi-hyperbolic discounting where benefits precede the cost, generating an intrapersonal conflict between selves. Examples include smoking, drinking, spending freely, etc., in which individuals tend to excessively satiate immediate gratification at the expense of future welfare.

well. I consider this specification in greater detail in Section 8.

3.2 The game

Prior to the realization of $\theta \in \Theta$, the principal can jointly commit herself to a delegation set and an information structure. The delegation set is denoted by D, which can be either countable, finite or infinite, or uncountable. The information structure σ consists of a signal realization space S and a measurable mapping $\sigma : \theta \to \Delta(S)$ that assigns to each state θ a probability distribution $d\sigma(\cdot|\theta)$ over S. The cost of every information structure is identical and set equal to zero.

The timeline of the game is as follows:

(i) The principal publicly selects a delegation set $D \subset Y$ and an information structure σ.

(ii) A pair (θ, s) is drawn according to the distributions μ_0 and σ, respectively.

(iii) The agent privately observes s and chooses an action $y \in D$.

3.3 Obedient recommendation mechanisms

Fix a delegation set D. In general, the principal could follow any rule for generating the agent's private information. However, a revelation principle-type argument establishes that I can restrict attention to information structures in which signal realizations are action recommendations that will be obeyed by the agent. It is therefore immediate that, instead of choosing a delegation D and an information structure σ separately, the principal can without loss of generality choose an *obedient recommendation mechanism* $\mathcal{M}$, which consists of a delegation set $D \subset Y$, and an action recommendation function (also an information structure) $\sigma : \Theta \to \Delta(D)$, and moreover, satisfies the obedience constraint (OB). In particular, OB is the requirement that when the agent is recommended to take action $y \in D$ according to σ, he would follow his recommendation y rather than choose any other available action $\hat{y} \in D$.[4] I describe the condition formally as follows.

Given a recommendation mechanism $\mathcal{M} = (D, \sigma)$ and the prior belief μ_0, every action recommendation $y \in D$ induces a posterior belief μ_y^σ and the associated posterior mean $m_y^\sigma = \mathbb{E}_{\mu_y^\sigma}[\theta]$. A recommendation mechanism $\mathcal{M} = (D, \sigma)$ is said to be obedient if

$$u_A(y, m_y^\sigma) \geq u_A(\hat{y}, m_y^\sigma) \quad \forall y, \hat{y} \in D.$$

The principal's problem is thus

$$\text{Maximize} \quad \int_\Theta \int_Y u_P(y, \theta) \, \mathrm{d}\sigma(y|\theta) \, \mathrm{d}\mu_0(\theta)$$

among all obedient recommendation mechanisms. $\mathcal{M}$ is said to be *optimal* if it solves the above problem.

All proofs are relegated to the Appendix.

3.4 Remarks

A key feature of my model is that the principal fully controls the quality of the agent's private information without learning its content. One might wonder why the principal relies on a biased agent to process the information and to make a biased choice subsequently. Why can't the principal simply assigns herself an information structure that informs herself perfectly and then chooses her first-best action? After all, there is no cost to adopt a particular information structure. To address this question, I offer three scenarios in which a principal may not be able to become informed herself. She therefore relies on some agent to gather, process, and use the information.

In the first scenario, the principal indeed has unlimited access to information but lacks

[4] The notion of obedience here is weaker than that in Bergemann and Morris (2016), which imposes no action restrictions, and therefore the agent is free to choose any action.

the time and resources to acquire information herself. Such a scenario fits in well with the application to organizations in which top management has formal authority over data access and usage, but is often too busy to have the time to process every piece of information. Top management thus relies on lower management to analyze data and to make informed decisions.

In the second scenario, the principal has unlimited access to information as well but faces technological constraints to observe the realization of information. In the earlier temptation and self-control problem of consumption, the current self can self-restrict information gathering about the unknown consumption externality, but cannot observe it herself because it is realized in the future.

In the third scenario, the principal not only has no access to information but also has no control over what information may be available to the agent. Instead, the principal has the ability to limit the degree of information the agent's action can effectively condition on. In the example of regulating a monopolist, considered in Section 7, even the regulator does not directly possess consumer data, she can impose limits, by law, on the monopolist's use of consumer data for discrimination. As a result, the monopolist might obtain excessive information on consumers himself, but by law, he is prohibited from using it for discriminatory offers.

4 Characterizing Optimal Mechanisms

In this section, I derive a general structure of optimal mechanisms and use it to simplify the principal's problem into a finite variable optimization problem. In particular, Section 4.1 states that local OB is sufficient to imply full OB. Section 4.2 provides a sufficient condition under which the principal finds it optimal to fully align the agent's incentives. I then use this insight to prove the optimality of a monotone partition in Section 4.3.

For the clarity of analysis, I assume for now that the delegation set D is finite and consists of N actions. I label it by $\{y_1, \ldots, y_N\} \subseteq Y$ with $y_i < y_{i+1}$. Later I will show that this finiteness restriction can be dropped without loss of generality. A mechanism $\mathcal{M} = (D, \sigma)$ is called *finite* if D is finite.

4.1 Preliminary observations

Recall that the agent's interim preference over actions depends solely on the posterior mean of the state. Therefore, the single-peakedness property of $u_A(y, \theta)$ with respect to y is inherited by the interim preference. Taking advantage of this observation, I can show (Lemma 1 below) that *local* OB is sufficient to establish full OB.

Lemma 1. *Let $\mathcal{M} = (D, \sigma)$ be a recommendation mechanism that implements N actions. Then local OB is sufficient to imply full OB.*

Lemma 2 (optimality of compromises)**.** *Let $\mathcal{M} = (D, \sigma)$ be an optimal obedient recommendation mechanism that implements N actions. Then it must satisfy*

$$y_P(m_i) \leq y_i \leq y_A(m_i) \quad \forall i = 1, \ldots, N.$$

Lemma 2 says that the implemented action in the optimal mechanism is always a *compromise*—it lies between the principal's preferred action and that of the agent. The logic is roughly as follows. Suppose that an action y is implemented outside the compromise region $[y_P(m), y_A(m)]$ for a given posterior mean m. Then a shift of action y in the direction closer to $[y_P(m), y_A(m)]$ will strictly benefit the principal without violating OB. An implication of Lemma 2 is that the agent's local downward OB must be *slack*. This is because y_{i-1} is farther to the left of the peak $y_A(m_i)$ than y_i. Then single-peakedness of u_A with respect to y implies that action y_i is strictly preferred by type y_i to action y_{i-1}. As a consequence, the principal does not need to consider such downward deviation by the agent.

This set of results from Lemma 1 and Lemma 2 allows me to focus on the relaxed program with only the local upward OB:

$$\max_{(D,\sigma)} \int_{\Theta} \left[\sum_{i=1}^{N} u_P(y_i, \theta)\sigma(y_i|\theta) \right] \mathrm{d}\mu_0(\theta)$$

subject to

$$u_A(y_i, m_i) \geq u_A(y_{i+1}, m_i) \quad \forall i = 1, \ldots, (N-1). \tag{OB$_{i,i+1}$}$$

An equivalent, yet convenient, way of writing (OB$_{i,i+1}$) is

$$m_i \leq \frac{C_A(y_{i+1}) - C_A(y_i)}{y_{i+1} - y_i} \quad \forall i = 1, \ldots, (N-1),$$

which says that the agent's marginal benefit of higher actions is smaller than his average rate of change of cost from implementing action y_i to action y_{i+1}. Hence a type y_i agent has no incentive to deviate to take action y_{i+1}.

4.2 Alignment principle

In this subsection, I introduce an incentive compatibility notion of aligned mechanisms. In an aligned mechanism, all relevant agent types play exactly as if they were maximizing the principal's payoff. An immediate implication is that the principal would also find it optimal to be obedient if she observed the mechanism's recommendation. The formal definition of aligned mechanisms is as follows.

Definition 1. An obedient recommendation mechanism $\mathcal{M} = (D, \sigma)$ is called aligned if, for any type $y \in D$,

$$y \in \arg\max_{\hat{y} \in D} u_P(\hat{y}, m_y) \cap \arg\max_{\hat{y} \in D} u_A(\hat{y}, m_y).[5]$$

In an aligned mechanism, the players' incentives are fully aligned. Given the optimal action restriction together with the information restriction, there is no way to improve the decision-making procedure—a reallocation of decision authority from the agent back to the

[5]The notion of aligned mechanism in this paper is similar to that in Frankel (2014), who considers the scenario where a principal delegates multiple decisions to an agent.

principal would not affect actions that are implemented.

Due to the single-crossing property of preferences and the agent's upward bias, the following result ensures that only local downward OB is effectively needed for the principal to be obedient.

Lemma 3. *Let $\mathcal{M} = (D, \sigma)$ be an obedient recommendation mechanism that implements N actions. Then it is aligned if and only if*

$$u_P(y_i, m_i) \geq u_P(y_{i-1}, m_i) \quad \forall i = 2, \ldots, N. \tag{4.1}$$

Likewise, (4.1) can be equivalently written as

$$m_i \geq \frac{C_P(y_i) - C_P(y_{i-1})}{y_i - y_{i-1}} \quad \forall i = 2, \ldots, N. \tag{4.2}$$

Let $\Delta_{C_P}(y, \hat{y}) \equiv \frac{C_P(y) - C_P(\hat{y})}{y - \hat{y}}$ denote the principal's average rate of change of cost from action $\hat{y}$ to y. From the principal's perspective, (4.2) thus requires that the marginal decrease in benefit is greater than the marginal decrease in cost from action y_i to y_{i-1}.

Let $C_{Diff} \equiv C_P - C_A$ denote the difference in costs. The next assumption provide sufficient conditions under which optimal mechanism is aligned.

Assumption 2. *The following hold:*

(i) The agent's cost function C_A satisfies $C_A''' \leq 0$;

(ii) The cost difference function C_{Diff} satisfies $C_{Diff}'' \geq 0$.

Condition (i) in Assumption 2 ensures that the agent's preferred action $y_A(\theta) = c_A^{-1}(\theta)$ is *convex* for all $\theta \in [\underline{\theta}, \overline{\theta}]$. Condition (ii) in Assumption 2 together with $C_{Diff}' = C_P' - C_A' > 0$ implies that the cost difference between Principal and Agent increases faster for higher realizations of the state.

The following is a sufficient result:

Proposition 1. *Let $\mathcal{M} = (D, \sigma)$ be an optimal obedient recommendation mechanism that recommends N actions. $\mathcal{M}$ is aligned if Assumption 2 holds.*

Proposition 1 indicates that it is always optimal for the principal to design a mechanism that completely align the agent's incentives with her own and hence avoid any interim conflict. The players may disagree about their preferred actions, but given the informational and discretionary constraints of the mechanism they agree on how to play optimally.

The optimality of aligned mechanisms turns out to be depending on the curvature of players' preferred actions. To see the intuition, suppose by contradiction that the mechanism is not aligned. By (4.2) there must exist an action recommendation y_i conditional on which the principal strictly prefers action y_{i-1} over y_i. In this case, the principal may wish to coarsen the original mechanism by collapsing the two recommendations y_{i-1} and y_i into a single recommendation $\hat{y}$ which takes a moderately small value (hence desirable to the principal). In turn, the modification reduces the informativeness of the mechanism and consequently the actions implemented become less responsive to the state, but at the same time, it mitigates the agent's incentive to take undesirable actions (specifically in this case, y_i) and hence improves efficiency in the allocation of actions from the principal's perspective. By requiring $C_A''' \leq 0$ and $C_{Diff}'' \geq 0$, I ensure that the agent's preferred action $y_A(\theta) = c_A^{-1}(\theta)$ is *convex* in θ, and that Agent is less biased for lower states. As an implication, the agent's incentive to deviate to higher actions (specifically, y_{i+1}) will be greatly mitigated when the posterior mean drops from m_i down to $\hat{m}$. By combining the two recommendations y_{i-1} and y_i, the principal is thus able to implement an action $\hat{y}$ small enough without violating OB, which delivers a welfare improvement for the principal.

Assumption 2 will be maintained for the remainder of the paper.

4.3 Monotone partitional mechanism

In practice, information disclosure often takes the form of monotone partitional signals—pooling, if present, is only between adjacent states. For example, schools adopt grading policies that contain only letter grades. Financial rating agencies also classify assets into different categories according to their riskiness.

Formally, a recommendation mechanism $\mathcal{M} = (D, \sigma)$ is called monotone partitional if the state space $\Theta = [\underline{\theta}, \overline{\theta}]$ is partitioned into a finite number N of intervals $\{(\theta_{i-1}, \theta_i]\}_{i=1}^N$ with $\theta_0 = \underline{\theta}$ and $\theta_N = \overline{\theta}$, and the agent observes only the interval that contains the realized state θ.

One apparent benefit of a monotone partitional mechanism is that it can be identified by two simple components: (i) the cutoffs $0 = \theta_0 < \theta_1 < \cdots < \theta_{N-1} < \theta_N = 1$; (ii) the delegated actions $y_1 < \cdots < y_N$. For a given realized state θ, a monotone partitional mechanism $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ recommends the agent to take action y_i with probability 1 if $\theta \in (\theta_{i-1}, \theta_i]$.

The following two theorems play an important role in my analysis. They establish that it is without loss of generality in focusing on mechanisms that partition the state space into a finite set of intervals and recommend different actions in different intervals. These features greatly simplify the characterization of the optimal mechanism by reducing it to a finite dimensional optimization problem over the cutoffs of intervals and the corresponding actions.

Theorem 1. *For any obedient recommendation mechanism that implements N actions, there exists another obedient recommendation mechanism $\mathcal{M} = (D, \sigma)$, such that $\mathcal{M}$ is a monotone partition and entails a weakly higher payoff for the principal.*

The result that the optimal mechanism (in the class of finite mechanisms) is a monotone partition stems from the fact that both parties in my setting prefers to take higher actions for higher realizations of the state of nature. Since the two parties' incentives are fully aligned in the optimal mechanism, intuitively there is no reason for the principal to induce "mismatched pairs" of states and actions.

I have so far restricted the mechanism to be finite. Below, I augment the space of mechanisms to allow for information structures with an arbitrary number of realizations, and show that an optimal mechanism exists and it is finite. Using this together with Theorem 1, I can now restrict attention to the class of finite monotone parititional mechanisms.

Theorem 2. *An optimal obedient recommendation mechanism exists and it is finite.*

The main idea of the proof follows from the approximation argument in Proposition 1 of Bergemann and Pesendorfer (2007), that the principal's payoff from any given mechanism with an arbitrary number of delegated actions can be approximated arbitrary well by a finite mechanism. In fact, it can be shown that the principal's set of payoffs generated by finite mechanisms is *dense* in the set of payoffs generated by all arbitrary mechanisms.

5 Uniform-Quadratic Case

The results in the previous section tell us a great deal about optimal mechanisms but do not fully describe them. In this section I first provide an explicit characterization of optimal mechanisms for the case of quadratic preferences mentioned in Section 3.1, under a further assumption that μ_0 is uniform on $\Theta = [0, 1]$. I then use this result, together with the bias parameter b, to discuss comparative statics on the agent's discretion and information in Section 5.2. Finally I conclude this section by analyzing the welfare effects of information manipulation on this delegation problem in Section 5.3.

5.1 Optimal partitional structure

I now turn to the study of optimal mechanisms for the uniform-quadratic case. As noted in Section 3.1, the quadratic formulation can be equivalently expressed as $u_P(y, \theta) = \theta y - \frac{1}{2}y^2$ with $y_P(\theta) = \theta$, and $u_A(y, \theta) = \theta y - (\frac{1}{2}y^2 - by)$ with $y_A(\theta) = \theta + b$. Moreover, since $C_A(y) = \frac{1}{2}y^2 - by$ and $C_{diff}(y) = by$ ensure that $C_A''' = 0$ and $C_{Diff}''' = 0$, Assumption 2 holds here and therefore Proposition 1 and Theorem 1 apply.

The following property will play an important role in the analysis.

Definition 2. A monotone partitional mechanism $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ features *alternating symmetry* if

$$\theta_i - \theta_{i-1} = \theta_j - \theta_{j-1} \quad \text{and} \quad y_i - \theta_{i-1} = y_j - \theta_{j-1}$$

for both odd (even) $i, j \in \{1, 2, \ldots, N\}$.

An illustration of this property is illustrated in Figure 1. It can be seen that all odd-numbered (even-numbered) elements are essentially "symmetric", in the sense that they have the same interval length, and the respective actions are located at the same positions relative to the associated intervals. The next lemma shows that the property indeed holds in the uniform-quadratic case.

Figure 1: Alternating symmetry.

Lemma 4. *In the uniform quadratic specification, the optimal mechanism satisfies the property of alternating symmetry.*

Lemma 4 significantly simplifies the characterization of the optimal mechanism, because it allows me to restrict attention to the first two elements; the rest of elements are merely "replications" of the first two. Moreover, the alternating symmetry structure also helps identify a finite upper bound, denoted $\overline{N}(b)$, on the size of an aligned monotone partitional mechanism. To see it, combining (4.1) and (4.2), and rearranging terms yields

$$y_{i+1} - y_{i-1} \geq 2b \quad \forall i = 2, \ldots, (N-1). \tag{5.1}$$

Together, (5.1) and the property of alternating symmetry imply $\overline{N}(b) \equiv \left\lfloor \frac{1}{2b} \right\rfloor + \left\lceil \frac{1}{2b} \right\rceil$.[6]

I am now ready to give an explicit characterization of optimal mechanisms in two steps. First, building on the previous observation that one only needs to focus on the first two elements, I obtain the suboptimal mechanism $\{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ for each size from one through $N \leq \overline{N}(b)$ by solving a simple linear program.[7] Second, the size of the optimal mechanism, denoted $N^*(b)$, can be determined by simply comparing the principal's expected payoffs under the resulting suboptimal mechanisms for each size $N \leq \overline{N}(b)$. The explicit expression of the optimal mechanism is summarized by the following proposition.

Proposition 2. *Suppose that $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ is an optimal obedient recommendation mechanism.*

[6] $\lfloor x \rfloor$ ($\lceil x \rceil$) is the largest (smallest) integer that is smaller (greater) than or equal to x.
[7] See Appendix.

(1) *If N is odd, then*

$$\theta_i = \begin{cases} \frac{(i-1)+(N-i)\theta_1}{N-1}, & \text{for } i \text{ odd}, \\ \frac{i(1-\theta_1)}{N-1}, & \text{for } i \text{ even}, \end{cases} \qquad y_i = \begin{cases} b + \frac{(i-2)+(N-i+1)\theta_1}{N-1}, & \text{for } i \text{ odd}, \\ b + \frac{(i-1)-(i-1)\theta_1}{N-1}, & \text{for } i \text{ even}, \end{cases}$$

with $\theta_1 = \frac{1-(N^2-1)b+\sqrt{(N+1)^2(N-1)^2b^2+1}}{N+1}$.

(2) *If N is even, then*

$$\theta_i = \frac{i}{N}, \quad y_i = b + \frac{i-1}{N} \quad \forall i = 1, \ldots, N.$$

Figure 2 displays the sizes $N^*(b)$ of optimal mechanisms together with the corresponding upper bounds $\overline{N}(b)$ for different values of $b > 0$. The figure indicates that the principal does not always prefer the partition with the largest possible number of elements, which, as previously noted in the Introduction, reflects the principal's trade-off between amount of information and efficiency of action allocations. The reason behind is that the mechanism with the highest number of elements, in spite of the most amount of information it generates in the first place, can result in the distortion of decision-making at the later stage. Indeed, the principal has to compromise between the two conflicting objectives of providing the agent with precise information and implementing actions sufficiently close to her optimal ones. It turns out that this trade-off, as illustrated in Figure 2, is resolved in favor of using mechanisms with *intermediate* sizes of partitions.

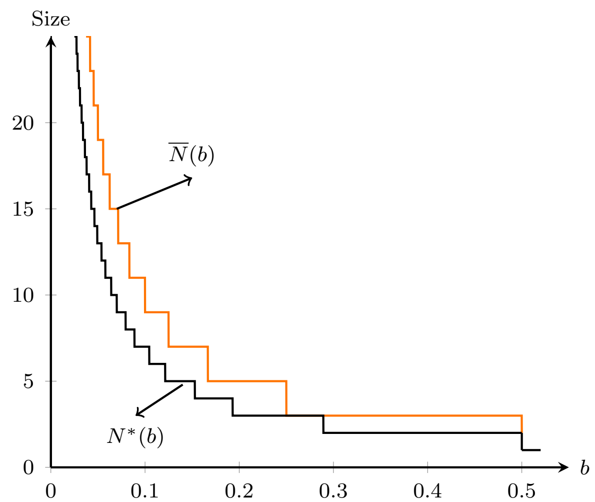

Figure 2: Optimal size of partitions and upper bounds.

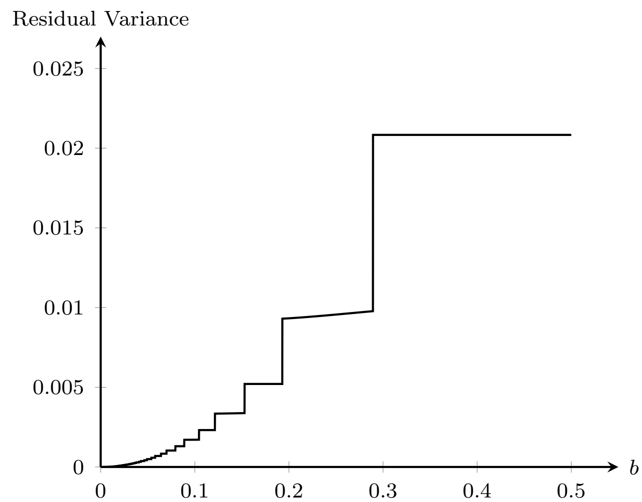

Figure 3: Informativeness of optimal mechanisms.

5.2 Comparative statics

In search for the optimal mechanism, the principal's problem is twofold: how much information the agent should possess, and how much discretion the agent should have. In this subsection, I examine comparative statics on these two features with respect to the divergence of the principal's and the sender's preferences.

Comparative statics on information. I measure informativeness as the expected residual variance of the state: $\mathbb{E}[variance(\theta|\mu)]$; a lower value of $\mathbb{E}[variance(\theta|\mu)]$ implies more precise posterior beliefs in expectation. Formally, a mechanism $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ is said to be more informative than another $\hat{\mathcal{M}} = \{(\hat{\theta}_i)_{i=1}^{\hat{N}}, (\hat{y}_i)_{i=1}^{\hat{N}}\}$ if

$$\sum_{i=1}^{N} \int_{\theta_{i-1}}^{\theta_i} (\theta - m_i)^2 \, \mathrm{d}\theta \leq \sum_{i=1}^{\hat{N}} \int_{\hat{\theta}_{i-1}}^{\hat{\theta}_i} (\theta - \hat{m}_i)^2 \, \mathrm{d}\theta.$$

This measure of informativeness is a natural one given the partitional information structure and the uniform prior assumption, because the posterior distribution of the state now is fully characterized by its mean and variance.

Figure 3 depicts a monotonic relationship between the agent's constant bias b and the informativeness of the corresponding optimal mechanism. In particular, it shows that a principal prefers to provide a greater amount of information if the agent's bias parameter b is smaller. This result might seem obvious in the uniform-quadratic specification. However, the impact of preference divergence on amount of information can be ambiguous in the general environment. On the one hand, the expected residual variance might not be an appropriate measure of informativeness when the prior distribution is not uniform. On the other hand, when preferences are less divergent the principal might also provide less information and still be able to steer the agent's incentives toward her own favor.

Comparative statics on discretion. I measure discretion by the *cardinality* of delegation sets. Formally, a mechanism $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ is said to grant more discretion than another mechanism $\hat{\mathcal{M}} = \{(\hat{\theta}_i)_{i=1}^{\hat{N}}, (\hat{y}_i)_{i=1}^{\hat{N}}\}$ if $N > \hat{N}$. Figure 2 illustrates the comparative statics on the agent's discretion: there is less discretion when the bias parameter b is bigger. This result confirms the standard intuition that an agent gets more discretion the more congruent he is with the principal.

5.3 Value of Information Manipulation

What is the value of limiting the agent's information in delegation problems? In this subsection I address this question and compare the players' payoffs in optimal delegation with and without information manipulation. When there is no information manipulation, the agent is assumed to have complete information about the state of nature. In this case, the principal simply faces a perfectly-informed agent.

Figure 4 displays the payoffs to the principal under optimal delegation with information manipulation compared to that without information manipulation. It is obvious that the payoffs are higher with information manipulation, because a full information structure can be viewed as a special informational policy under information manipulation. The value of keeping the agent partially informed simply reflects the aforementioned trade-off between quality of information and efficiency of action allocations.

It is interesting to note that the principal's gain from information manipulation is non-monotonic in the degree of bias. Clearly, for b small, the principal's payoff is close to her first-best level and the gains from information manipulation are small. As b increases, distorted action allocations become more severe and there is more scope for information manipulation to fix this incentive problem by aligning preferences. However, as b becomes sufficiently large, resorting to information manipulation improves this situation, but the cost of aligning incentives through limited information also increases. When $b \geq \frac{1}{2}$, the ability of limiting information is of no value whatsoever as the principal then optimally takes an uninformed action.

Figure 5, on the other hand, compares the payoffs to the agent in the two cases. It illustrates that in general the agent does not always achieve more information rents by possessing more private information. In fact, there are some regions of sufficiently large bias b in which an imperfectly-informed agent (possessing principal-optimal partitional information) is better off than a perfectly-informed agent.[8] As an implication, ignorance may generate strategic benefits for the agent even though private information guarantees the agent a positive rent under the optimal delegation scheme. Intuitively, when the agent's bias becomes sufficiently large (as long as $b \leq \frac{1}{2}$), by restricting his information in a manner that increases the distance between posterior means, it is possible to reduce the agent's incentive to take higher actions for intermediate types, and thus increase the principal's willingness to delegate more high actions.

To illustrate, consider the case where $b \in (\frac{2}{5}, \frac{1}{2})$. It follows from Proposition 2 that, with information manipulation, the principal optimally selects a binary partitional information structure with cutoffs $\{\frac{1}{2}, 1\}$, and a delegation set of $\{b, b + \frac{1}{2}\}$. The reader can easily verify that the agent's expected payoff is $\frac{1}{2}(b^2 + b + \frac{1}{4})$. In contrast, without information

[8] This phenomenon is also confirmed by Roesler and Szentes (2017) in the setting of static monopoly pricing.

manipulation, the principal's optimal delegation set takes the form of an interval: $[b, 1-b]$,[9] which delivers expected payoff $\frac{1}{2}(-\frac{8}{3}b^3 + b^2 + b + \frac{1}{3})$ to the agent. These imply, as Figure 5 shows, that the agent strictly prefers the binary partitional information structure under which the principal offers a delegation set including a sufficiently high action $b + \frac{1}{2}$, which allows the agent to respond to higher states.

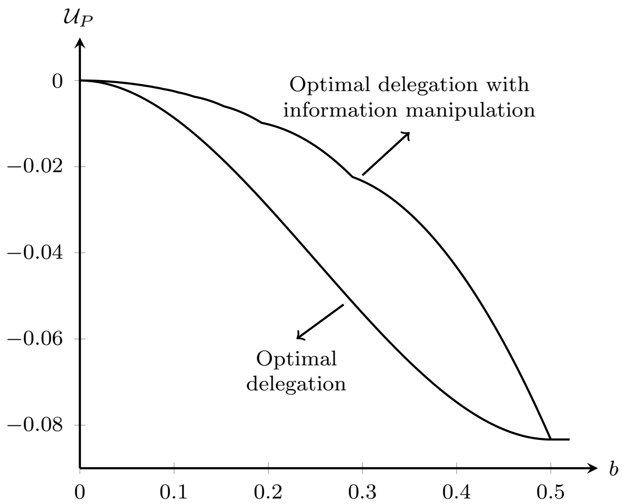

Figure 4: Principal's payoffs.

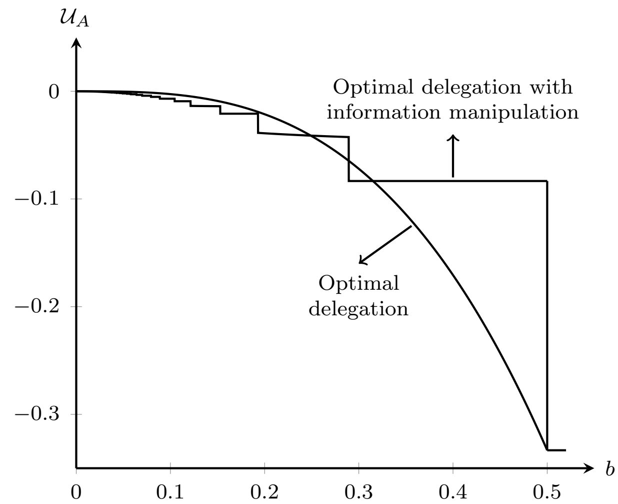

Figure 5: Agent's payoffs.

[9]In the uniform-quadratic case with perfectly-informed agent, the optimal delegation set takes the form $D^* = [b, 1-b]$ if $b \leq \frac{1}{2}$ and $D^* = \{\frac{1}{2}\}$ if $b > \frac{1}{2}$. This characterization was first given by Melumad and Shibano (1991).

6 General Comparative Statics

It is natural to ask at this point to what extent the comparative statics results that hold in the uniform-quadratic case can be generalized. This section establishes that the comparative statics on discretion and on the principal's payoffs still carry over to the general model.

To do so, I introduce the concept of an agent A with C_A being *uniformly less-biased* than another agent $\tilde{A}$ with $C_{\tilde{A}}$ if $c_A(y) \geq c_{\tilde{A}}(y)$ for each $y \in Y$. Recall that $y_A(\theta) = c_A^{-1}(\theta)$ and $y_{\tilde{A}}(\theta) = c_{\tilde{A}}^{-1}(\theta)$. It is straightforward to see that $y_{\tilde{A}}(\theta) \geq y_A(\theta) > y_P(\theta)$; namely, $y_A(\cdot)$ lies between $y_{\tilde{A}}(\cdot)$ and $y_P(\cdot)$. Note that this notion of preference divergence encompasses that of the uniform-quadratic case, which is captured by a single parameter b.

I am now ready to generalize the comparative statics results on the agent's discretion and the principal's payoffs.

Theorem 3. *The principal always grants more discretion to the agent who is uniformly less-biased than another agent.*

Theorem 3 extends the observation in Section 5.2 that the principal optimally offers finer partitions to a more congruent agent, and as a result, offers more discretion. (I use "finer" informally here.)

Next, it is easy to establish that the principal necessarily obtains higher payoffs if the agent is less biased. Consider moving from $C_{\tilde{A}}$ to a uniformly less-biased C_A. Letting $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ denote an recommendation mechanism, it is immediate that

$$m_i \leq \frac{C_{\tilde{A}}(y_{i+1}) - C_{\tilde{A}}(y_i)}{y_{i+1} - y_i} \Rightarrow m_i \leq \frac{C_A(y_{i+1}) - C_A(y_i)}{y_{i+1} - y_i} \quad \forall i = 1, \ldots, (N-1).$$

This means the feasible set of obedient mechanisms with C_A is a superset of that with $C_{\tilde{A}}$, so the principal's expected optimal payoff is weakly higher facing agent C_A than agent $C_{\tilde{A}}$. This result is stated formally in the next theorem.

Theorem 4. *The principal always prefers the agent who is uniformly less-biased.*

7 Application: Monopoly Regulation

In this section, I apply my results to the classic problem of regulating a monopolist who is privately informed (Baron and Myerson 1982). In contrast to the standard literature, I suppose that (i) contingent transfers between the regulator and the monopolist are not feasible; (ii) in addition to restricting prices (price control policy) as in Alonso and Matouschek (2008), the regulator can restrict the monopolist's access to private information about consumers (data protection policy).

I solve for the optimal joint regulatory policy using the results of Theorem 1 and 2. I show that (i) the optimal price control policy takes the form of a finite set of prices, which is strikingly different from the price cap regulation, as is commonly observed in practice and in theory (see, e.g., Alonso and Matouschek 2008; Amador and Bagwell 2020); (ii) the optimal data protection policy takes the form of monotone partitions: adjacent consumer types are pooled together in the same market segments.

I then compare the qualitative features of optimal joint regulatory policy with two benchmarks: one where there is only price control, and one where there is only data protection.

7.1 Joint regulation

A monopolist sells a good to a consumer.[10] The monopolist can choose a quality level a and a unit price y to the consumer, who, in turn, chooses a quantity level q to maximize a quadratic utility function:

$$u(y, a, q, \theta) = (\theta + \lambda a)q - yq - \frac{1}{2}q^2,$$

which yields consumer demand $q(y, a, \theta) = \theta + \lambda a - y$. One can interpret $\theta \in [0, 1]$ as the baseline willingness to pay for the first unit of the product and henceforth refer to θ as the

[10]This framework of price-quality discrimination is also adopted in Argenziano and Bonatti (2020), who provide a microfoundation for consumers' privacy preferences in a dynamic environment.

consumer's type. The parameter λ represents the monopolist's marginal value of the quality of the product.

The monopolist has a constant marginal cost of producing quantity q that I normalize to zero and a fixed per-consumer cost of producing quality a. The monopolist seeks to maximize his profit $u_A(y, a, \theta) = yq - \frac{1}{2}a^2$. Fix a price y, the monopolist's optimal quality level can be written as $a^*(y) = \lambda y$. Substituting a^* into u_A yields the monopolist's payoff

$$u_A(y, \theta) = \theta y - \frac{1}{2}(2 - \lambda^2)y^2,$$

where $y_A(\theta) = \frac{1}{2-\lambda^2}\theta$ is the profit-maximizing price.

The regulator's payoff, then, is the sum of the profit and the consumer surplus

$$u_P(y, \theta) = \lambda^2\theta y - \frac{1}{2}(1 + \lambda^2 - \lambda^4)y^2 + \frac{1}{2}\theta^2,$$

where $y_P(\theta) = \frac{\lambda^2}{1+\lambda^2-\lambda^4}\theta$ is the welfare-maximizing price. To ensure that $y_A(\cdot) > y_P(\cdot)$ I assume that $\lambda \in (0, 1)$. Observe that a monopolist with λ is uniformly less-biased than another monopolist with λ' if $\lambda \geq \lambda'$. Therefore, the parameter λ fully measures the monopolist's upward bias.

The regulator and the monopolist share a prior μ_0 about θ. I suppose the regulator can employ two regulatory policies at the same time: one is price control, which specifies a set of prices D available to the monopolist; the other is data protection, which is formally an information structure (equivalently, a randomized price recommendation function) $\sigma : [0, 1] \to \Delta(D)$ that provides the monopolist with information about the consumer's private type. The monopolist privately observes a realization from σ, and then he chooses a quality level a, and sets a price y from D to maximize profit $u_A(y, \theta)$.

By Theorem 1 and 2, the optimal joint regulatory policy (D, σ) features a monotone partition. That is, the optimal price control D is a finite set of prices; and the optimal data protection policy partitions the state space $[0, 1]$ into finite intervals and the monopolist only observes which interval the consumer's type actually lies in.

Notice that the monopolist's pricing problem is essentially a problem of *third-degree price discrimination*, where different realizations of the information play the role of market segments. Thus, the single-consumer setup introduced above is equivalent to a setup where there is a continuum of consumers, each of whom has a private type θ, which is drawn independently according to μ_0. With this interpretation, an information structure σ then corresponds to a *market segmentation* which is a division of the aggregate market into different markets. To see why, observe that any information structure σ must lead to a distribution τ over posteriors $\mu \in \Delta([0,1])$ satisfying $\int \mu \, \mathrm{d}\tau(\mu) = \mu_0$. Thus, a probability measure τ can be viewed as a market segmentation, with the interpretation that $\tau(\mu)$ is the proportion of the population in market segment μ.

As a result, with a continuum of consumers, I can reexpress the regulator's data protection policy as a market segmentation. Accordingly, a monotone partition corresponds to a "finite monotone segmentation", which splits consumers into finite groups and each group only consists of consumers with similar types. The monopolist then offers different quality levels and different prices to different segments. One implication of the optimality of finite market segmentations is that the consumers' privacy is protected moderately, which is between no protection (i.e., perfect segmentation with each segment containing consumers of a single type) and full protection (i.e., no segmentation with a uniform quality level and a uniform price).[11]

7.2 Comparing Regulatory Policies

I compare the key features of optimal joint regulatory policies with two benchmarks: one with only price control and one with only data protection.

Pure price control (full information). In this benchmark, the regulator imposes no

[11] For the privacy implications of consumer data collection, I refer the reader to the survey by Acquisti, Taylor, and Wagman (2016).

restrictions on the monopolist's access to consumer privacy, and thus the monopolist is perfectly informed about each consumer's type. The regulator's search for the optimal price control policy can then be formulated as an optimal delegation problem in which the agent (monopolist) has perfect information. Most of the delegation literature has focused on characterizing sufficient conditions for interval delegation to be optimal. For instance, in the present regulatory environment, it follows immediately from Proposition 4 in Alonso and Matouschek (2008) that there exists a $\lambda^* \in (0,1)$ such that for all $\lambda \in (\lambda^*, 1)$, i.e., the monopolist's bias is sufficiently small, the optimal price control policy is a *price cap*, below which the monopolist can set any prices.

Pure data protection (full discretion). In this benchmark, the regulator imposes no restrictions on prices so the monopolist has full discretion. The regulator's seaerch for the optimal data protection policy can then be formulated as a Bayesian persuasion problem. Following Kamenica and Gentzkow (2011), I can then easily solve for the optimal information structure by representing the regulator's (sender's) payoff as a function of the monopolist's (receiver's) beliefs. It can be shown that when $\lambda \in (\sqrt{\frac{3-\sqrt{5}}{2}}, 1)$, full disclosure is optimal, i.e., there is quality and price discrimination against each consumer type and hence no protection of consumer privacy. When $\lambda \in (0, \sqrt{\frac{3-\sqrt{5}}{2}})$, no disclosure is optimal, i.e., there is no quality and price discrimination and hence full protection of privacy.[12]

Qualitative differences. By comparing the optimal joint regulatory policy with the two benchmarks, I obtain two striking qualitative insights. First, with additional data protection, significantly less pricing flexibility is delegated to the monopolist. Specifically, the regulator optimally grants only a finite set of prices in the presence of data protection, rather than an interval of prices (i.e., a price cap) in the full information benchmark. Second, when the marginal value of the quality λ is sufficiently large, i.e., the monopolist's bias is sufficiently

[12]For the welfare effects of third-degree price discrimination, see, for instance, Pigou (1920), Robinson (1933), Schmalensee (1981), Varian (1985), Aguirre, Cowan, and Vickers (2010), Cowan (2016).

small, the social welfare is enhanced when the regulator moderately protects consumer privacy, but that cannot happen without controlling prices at the same time.

8 Application: Temptation, Information, and Discretion

In this section, I examine the consumption decision of a dynamically inconsistent individual for goods that provide an immediate benefit and a delayed cost, and ask how self-restricting information and discretion can serve as complementary commitment devices and alleviate the effects of temptation.

Consider an individual who lives for three periods $t \in \{0, 1, 2\}$. In period 1 he decides the amount x of good he is going to consume. Consumption generates an instantaneous utility equal to $u(x)$. However, it also exerts a negative externality in period 2 by an amount $e(x)$ with magnitude ω. I assume that the individual's preferences exhibit time-inconsistency, due to quasi-hyperbolic discounting. The intertemporal payoffs from the perspective of *self*-0 and *self*-1 from periods $t = 1, 2$ with externality shock ω are respectively given by

$$U_0(x, \omega) = \beta[\delta u(x) - \delta^2 \omega e(x)] \quad \text{and} \quad U_1(x, \omega) = u(x) - \beta\delta\omega e(x),$$

where u is a concave utility function, e is a convex externality function, and both are strictly increasing and continuously differentiable. The value of δ represents the standard discount factor. Since the main focus here is on the impacts of time inconsistency, I assume that $\delta = 1$. The value of β is the hyperbolic adjustment that reflects the momentary salience of the present. When β is lower, the temptation for current consumption is higher. To highlight the demand for commitment devices, I assume $\beta \in (0, 1]$. As a result, *self*-1's excessive preference for the present leads to overconsuming from *self*-0's perspective. For analytical tractability I restrict attention to a specific class of functional forms.

Assumption 3. *The following hold:*

(i) The utility of consumption satisfies $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$ *for* $0 \le \gamma \le \frac{1}{2}$*;*

(ii) The externality of consumption is linear, i.e., $e(x) = x$*.*

Letting $y = u(x)$, *self*-0's and *self*-1's payoffs can be re-written as

$$U_0(y,\theta) = \theta y - (1-\gamma)^{\frac{1}{1-\gamma}} y^{\frac{1}{1-\gamma}} \quad \text{and} \quad U_1(x,\theta) = \theta y - \beta(1-\gamma)^{\frac{1}{1-\gamma}} y^{\frac{1}{1-\gamma}}, \tag{8.1}$$

where θ is defined to be the inverse of ω.[13]

As is well known, a sophisticated and dynamically inconsistent individual values commitment power (Strotz, 1995). I consider two commitment devices that are commonly studied in the literature. A hard commitment device ties the hands of *self*-1 by reducing the set of available consumption choices; a soft commitment device keeps *self*-1 ignorant about the magnitude of the externality through selective exposure to information.

I analyze the optimal combination of the hard and the soft commitment devices. One can easily examine that the utility specification in (8.1) satisfies Assumption 1 and 2, and hence can be formally captured by the general analysis in Section 4. In particular, the alignment principle from Proposition 1 implies that *self*-0 would optimally eliminate the disagreement with *self*-1 under self-restricted consumption and information. Moreover, I can easily use Theorem 2 to show that the optimal joint commitment features a finite consumption set and a monotone partitional information structure.

9 Conclusion

I have studied an optimal delegation problem in which the principal has full control over the agent's information environment. The principal's problem is to jointly select a delegation set and an information structure that maximize her expected payoffs.

[13] I have removed β in the expression of $U_0(y,\theta)$ without loss of generality.

Notable features of the optimal mechanism include: (i) the players act as if they shared identical preferences given the chosen information and discretion; (ii) the optimal information structure takes the form of monotone partitions.

For future research it would be interesting to extend my model by allowing for the possibility of contingent monetary transfers. In the design of static decision-making processes, there are roughly three incentive schemes available to align players' interests: delegation, persuasion, and monetary transfers. By employing all the three incentive instruments, I would be able to analyze the three-way interaction among allocation of authority, provision of information, and commitment to transfers in optimal mechanisms, and moreover, to study the extent to which players' conflicts can be reduced.

A Appendix

A.1 Proof of Lemma 2

Let $\mathcal{M} = (D, \sigma)$ be an optimal obedient mechanism that implements N actions. I first show that $y_P(m_i) \le y_i$ for all $i = 1, \dots, N$. Suppose instead that there exists an action $y_i \in D$ such that $y_i < y_P(m_i)$. I split the analysis into two cases, depending on whether type y_{i+1}'s local downward OB is binding.

Case 1. Suppose $u_A(y_{i+1}, m_{i+1}) > u_A(y_i, m_{i+1})$. It then follows that $m_{i+1} > \frac{C_A(y_{i+1}) - C_A(y_i)}{y_{i+1} - y_i} = \Delta_{C_A}(y_{i+1}, y_i)$. Consider now a perturbation in which the principal replaces y_i with a slightly bigger action $y_i^{\uparrow} = y_i + \varepsilon$ with $\varepsilon > 0$, while keeping all other delegated actions fixed. I argue that the principal must be strictly better off and hence a contradiction to optimality. To see this, let $p_i = \int_0^1 \sigma(y_i|\theta)\, \mathrm{d}\mu_0(\theta)$ denote the absolute probability of realization y_i. The perturbation only changes the principal's payoff for type y_i. The change in payoff is

$$\Delta \mathcal{U}_P^{\sigma} = -p_i[m_i y_i - C_P(y_i)] + p_i\Big[m_i y_i^{\uparrow} - C_P(y_i^{\uparrow})\Big] = p_i(y_i^{\uparrow} - y_i)\Big[m_i - \Delta_{C_P}(y_i^{\uparrow}, y_i)\Big] > 0$$

given that $\varepsilon > 0$ is small and $y_i < y_P(m_i)$. The only thing left to verify is that OB is preserved. By Lemma 1 I only need to check type y_{i-1}'s $(\mathrm{OB}_{i-1,i})$, type y_i's $(\mathrm{OB}_{i,i-1})$ and $(\mathrm{OB}_{i,i+1})$, and type y_{i+1}'s $(\mathrm{OB}_{i+1,i})$, that is,

$$m_{i-1} \le \Delta_{C_A}(y_i^{\uparrow}, y_{i-1}), \tag{$\mathrm{OB}_{i-1,i}$}$$

$$\Delta_{C_A}(y_i^{\uparrow}, y_{i-1}) \le m_i \le \Delta_{C_A}(y_{i+1}, y_i^{\uparrow}), \tag{$\mathrm{OB}_{i,i-1} + \mathrm{OB}_{i,i+1}$}$$

$$\Delta_{C_A}(y_{i+1}, y_i^{\uparrow}) \le m_{i+1}. \tag{$\mathrm{OB}_{i+1,i}$}$$

Since $y_i^{\uparrow} > y_i$, both $(\mathrm{OB}_{i-1,i})$ and $(\mathrm{OB}_{i,i+1})$ are automatically satisfied. As for $(\mathrm{OB}_{i,i-1})$ and $(\mathrm{OB}_{i+1,i})$, observe that they must be slack due to the fact that $\Delta_{C_A}(y_i^{\uparrow}, y_{i-1}) < C_A'(y_i) < C_P'(y_i) < m_i$, and $\Delta_{C_A}(y_{i+1}, y_i) < m_{i+1}$, respectively.

Case 2. Suppose $u_A(y_{i+1}, m_{i+1}) = u_A(y_i, m_{i+1})$. It then follows that $m_{i+1} = \Delta_{C_A}(y_{i+1}, y_i)$. Consider now a small increase in y_i and a small decrease in y_{i+1} by setting $y_i^{\uparrow} = y_i + \varepsilon$ and $y_{i+1}^{\downarrow} = y_{i+1} - \delta$ with $\varepsilon, \delta > 0$ and $\Delta_{C_A}(y_{i+1}^{\downarrow}, y_i^{\uparrow}) = \Delta_{C_A}(y_{i+1}, y_i)$. The change in the principal's payoff is then

$$\begin{aligned}\Delta\mathcal{U}_P^{\sigma} =& - p_i[m_i y_i - C_P(y_i)] + p_i\Big[m_i y_i^{\uparrow} - C_P(y_i^{\uparrow})\Big] \\ & - p_{i+1}[m_{i+1} y_{i+1} - C_P(y_{i+1})] + p_{i+1}\Big[m_{i+1} y_{i+1}^{\uparrow} - C_P(y_{i+1}^{\uparrow})\Big] \\ =& p_i(y_i^{\uparrow} - y_i)\Big[m_i - \Delta_{C_P}(y_i^{\uparrow}, y_i)\Big] + p_{i+1}(y_{i+1}^{\downarrow} - y_{i+1})\Big[m_{i+1} - \Delta_{C_P}(y_{i+1}^{\downarrow}, y_{i+1})\Big] > 0\end{aligned}$$

given that $\varepsilon > 0$ is small, $m_i > C_P'(y_i)$, and $m_{i+1} = \Delta_{C_A}(y_{i+1}, y_i) < \Delta_{C_P}(y_{i+1}^{\downarrow}, y_{i+1})$. I am left to verify type y_{i-1}'s ($\text{OB}_{i-1,i}$), type y_i's ($\text{OB}_{i,i-1}$) and ($\text{OB}_{i,i+1}$), type y_{i+1}'s ($\text{OB}_{i+1,i}$) and ($\text{OB}_{i+1,i+2}$), and type y_{i+2}'s ($\text{OB}_{i+2,i+1}$):

$$m_{i-1} \leq \Delta_{C_A}(y_i^{\uparrow}, y_{i-1}), \qquad (\text{OB}_{i-1,i})$$

$$\Delta_{C_A}(y_i^{\uparrow}, y_{i-1}) \leq m_i \leq \Delta_{C_A}(y_{i+1}^{\downarrow}, y_i^{\uparrow}), \qquad (\text{OB}_{i,i-1} + \text{OB}_{i,i+1})$$

$$\Delta_{C_A}(y_{i+1}^{\downarrow}, y_i^{\uparrow}) \leq m_{i+1} \leq \Delta_{C_A}(y_{i+2}, y_{i+1}^{\downarrow}), \qquad (\text{OB}_{i+1,i} + \text{OB}_{i+1,i+2})$$

$$\Delta_{C_A}(y_{i+1}^{\downarrow}, y_{i+2}) \leq m_{i+2}. \qquad (\text{OB}_{i+2,i+1})$$

Observe that ($\text{OB}_{i-1,i}$), ($\text{OB}_{i,i+1}$), ($\text{OB}_{i+1,i}$), and ($\text{OB}_{i+2,i+1}$) still hold by the construction of $y_i^{\uparrow}$ and $y_{i+1}^{\downarrow}$. Moreover, ($\text{OB}_{i,i-1}$) and ($\text{OB}_{i+1,i+2}$) are slack because $\Delta_{C_A}(y_i^{\uparrow}, y_{i-1}) < \Delta_{C_A}(y_i^{\uparrow}) < \Delta_{C_P}(y_i^{\uparrow}) < m_i$, and $m_{i+1} = \Delta_{C_A}(y_{i+1}, y_i) < \Delta_{C_A}(y_{i+2}, y_{i+1}^{\downarrow})$.

Proceeding in a similar fashion as above, I can also show that $y_i \leq y_A(m_i)$ for all $i = 1, \ldots, N$.

A.2 Proof of Proposition 1

Let $\mathcal{M} = (D, \sigma)$ be an optimal obedient mechanism that implements N actions. Assume the contrary, i.e., that there exists a type y_i with $2 \leq i \leq N$ such that $u_P(y_i, m_i) < u_P(y_{i-1}, m_i)$, or equivalently, $m_i < \Delta_{C_P}(y_i, y_{i-1})$.

First, I must have $i \neq N$. Otherwise the principal can derive a net benefit from joining the type y_{i-1} and type y_i into a single type associated with action y_{i-1}.

Second, I claim that ($\text{OB}_{i-1,i}$) must be binding, that is,

$$m_{i-1} = \Delta_{C_A}(y_i, y_{i-1}). \tag{A.1}$$

Suppose not. Then I can perturb the information structure σ by subtracting a small portion from realization y_i, and adding it to realization y_{i-1}. More formally, the perturbed information structure σ' is the same as σ except that

$$\sigma'(y_i|\theta) = (1-\varepsilon)\sigma(y_i|\theta),$$
$$\sigma(y_{i-1}|\theta) = \sigma'(y_{i-1}|\theta) + \varepsilon\sigma(y_i|\theta).$$

By doing so, the principal is strictly better off since she prefers action y_{i-1} over y_i, conditional on realization y_i. Moreover, ($\text{OB}_{i-1,i}$) is still satisfied given the hypothesis and that $\varepsilon > 0$ is small. A contradiction.

Before proceeding to the contradiction argument, note that type y_i's ($\text{OB}_{i,i+1}$) can be written as

$$m_i + \delta_i = \Delta_{C_A}(y_{i+1}, y_i) \tag{A.2}$$

with $\delta_i \geq 0$. With this notation, the contradiction assumption $m_i < \Delta_{C_P}(y_i, y_{i-1})$ can be rewritten as $\Delta_{C_P}(y_i, y_{i-1}) > \Delta_{C_A}(y_{i+1}, y_i) - \delta_i$, or equivalently,

$$\Delta_{C_{Diff}}(y_i, y_{i-1}) > \Delta_{C_A}(y_{i+1}, y_i) - \Delta_{C_A}(y_i, y_{i-1}) - \delta_i. \tag{A.3}$$

Moreover, define action $\bar{y} \in \mathbb{Y}$ as the solution to

$$\rho m_{i-1} + (1-\rho)m_i + (1-\rho)\delta_i = \Delta_{C_A}(y_{i+1}, \bar{y}),$$

where $\rho = \frac{p_{i-1}}{p_{i-1}+p_i}$.

Now consider a modification $\hat{\mathcal{M}} = (\hat{\sigma}, \hat{D})$ of the mechanism $\mathcal{M}$ as follows:

(i) The realization space (delegation set) $\hat{D}$ is given by

$$\hat{D} = \{y_1, \ldots, y_{i-1}, \hat{y}, y_{i+2}, \ldots, y_N\},$$

where the value of $\hat{y}$ will be specified later.

(ii) $\hat{\sigma}$ is described by

$$\hat{\sigma}(y|\theta) = \begin{cases} \sigma(y_{i-1}|\theta) + \sigma(y_i|\theta), & \text{if} \quad y = \hat{y}, \\ \sigma(y|\theta), & \text{if} \quad y \in \hat{D} \setminus \{\hat{y}\}. \end{cases}$$

There are two exhaustive cases to consider:

Case 1: $\bar{y} \leq y_{i-1}$. In this case, let $\hat{y} = y_{i-1}$. To check the obedience of $\hat{\mathcal{M}}$, it suffices to verify that type $\hat{y}$'s local upward OB constraint holds,

$$\hat{m} = \rho m_{i-1} + (1-\rho) m_i = \Delta_{C_A}(y_{i+1}, \bar{y}) - (1-\rho)\delta_i \leq \Delta_{C_A}(y_{i+1}, y_{i-1}).$$

It follows that the principal is strictly better off under $\hat{\mathcal{M}}$ by the hypothesis that she strictly prefers action y_{i-1} over y_i on the support of realization y_i. A contradiction.

Case 2: $\bar{y} > y_{i-1}$. In this case, let $\hat{y} = \bar{y}$. To verify obedience, it is sufficient to verify that both type y_{i-2} and type $\hat{y}$ are upwardly obedient. For type y_{i-2}, I have

$$m_{i-2} \leq \Delta_{C_A}(y_{i-1}, y_{i-2}) < \Delta_{C_A}(\hat{y}, y_{i-2}).$$

For type $\hat{y}$, I have

$$\hat{m} + (1-\rho)\delta_i = \rho m_{i-1} + (1-\rho) m_i + (1-\rho)\delta_i = \Delta_{C_A}(y_{i+1}, \bar{y}) = \Delta_{C_A}(y_{i+1}, \hat{y}). \tag{A.4}$$

Additionally, replacing m_{i-1} by $\Delta_{C_A}(y_i, y_{i-1})$ and m_i by $\Delta_{C_A}(y_{i+1}, y_i) - \delta_i$ in (A.4), I obtain

$$\Delta_{C_A}(y_{i+1}, \hat{y}) = \rho\Delta_{C_A}(y_i, y_{i-1}) + (1-\rho)\Delta_{C_A}(y_{i+1}, y_i) \tag{A.5}$$

Since I have assumed $\bar{y} > y_{i-1}$, it follows that $\rho \in (0, \frac{y_i - y_{i-1}}{y_{i+1}-y_{i-1}})$. And I must have $\hat{y} \to y_i$ from below as $\rho \to 0$, and $\hat{y} \to y_{i-1}$ from above as $\rho \to \frac{y_i - y_{i-1}}{y_{i+1}-y_{i-1}}$.

To prove that the modification makes the principal better off, I begin by expressing the difference in the principal's payoff between the original mechanism $\mathcal{M}$ and the modified mechanism $\hat{\mathcal{M}}$ (normalized by $p_{i-1} + p_i$) as

$$\begin{aligned}
\frac{\mathcal{U}_P(\hat{\mathcal{M}}) - \mathcal{U}_P(\mathcal{M})}{p_{i-1} + p_i} =& [\hat{m}\hat{y} - C_P(\hat{y})] - \rho[m_{i-1}y_{i-1} - C_P(y_{i-1})] - (1-\rho)[m_i y_i - C_P(y_i)] \\
=& \hat{m}\hat{y} - \rho m_{i-1}y_{i-1} - (1-\rho)m_i y_i \\
& - \rho[C_P(y_i) - C_P(y_{i-1})] - [C_P(y_{i+1}) - C_P(y_i)] + [C_P(y_{i+1}) - C_P(\hat{y})] \\
=& \hat{m}\hat{y} - \rho m_{i-1}y_{i-1} - (1-\rho)m_i y_i \\
& - \rho\Delta_{C_P}(y_i, y_{i-1})(y_i - y_{i-1}) - \Delta_{C_P}(y_{i+1}, y_i)(y_{i+1} - y_i) + \Delta_{C_P}(y_{i+1}, \hat{y})(y_{i+1} - \hat{y}).
\end{aligned} \tag{A.6}$$

Using the OB constraints (A.1), (A.2), (A.4) to substitute in for $m_{i-1}, m_i, \hat{m}$, and further replacing Δ_{C_P} with $\Delta_{C_A} + \Delta_{C_{Diff}}$, (A.6) simplifies to

$$\begin{aligned}
\frac{\mathcal{U}_P(\hat{\mathcal{M}}) - \mathcal{U}_P(\mathcal{M})}{p_{i-1} + p_i} =& [\Delta_{C_A}(y_i, y_{i-1}) - \Delta_{C_A}(y_{i+1}, y_i)]\rho(y_{i+1} - y_i) \\
& - \Delta_{C_{Diff}}(y_i, y_{i-1})\rho(y_i - y_{i-1}) \\
& - \Delta_{C_{Diff}}(y_{i+1}, y_i)(y_{i+1} - y_i) \\
& + \Delta_{C_{Diff}}(y_{i+1}, \hat{y})(y_{i+1} - \hat{y}) \\
& - (1-\rho)\delta_i\hat{y} + (1-\rho)\delta_i y_i.
\end{aligned}$$

Substituting with the contradiction assumption (A.3) yields

$$\begin{aligned}\frac{\mathcal{U}_P(\hat{\mathcal{M}})-\mathcal{U}_P(\mathcal{M})}{p_{i-1}+p_i} >& -\Delta_{C_{Diff}}(y_i,y_{i-1})\rho[y_{i+1}-y_{i-1}]+C_{Diff}(y_i)-C_{Diff}(\hat{y}) \\ & -(1-\rho)\delta_i\hat{y}+(1-\rho)\delta_i y_i-\rho\delta_i(y_{i+1}-y_i)\end{aligned} \tag{A.7}$$

I next show that the RHS of (A.7) is weakly positive. Note that by the previous steps, $\lim_{\rho\to 0^+}\hat{y}=y_i$ and $\lim_{\rho\to[(y_i-y_{i-1})/(y_{i+1}-y_{i-1})]^-}\hat{y}=y_{i-1}$. Using this and (A.7), I obtain

$$\lim_{\substack{\rho\to 0^+ \text{ or} \\ [(y_i-y_{i-1})/(y_{i+1}-y_{i-1})]^-}} \frac{\mathcal{U}_P(\hat{\mathcal{M}})-\mathcal{U}_P(\mathcal{M})}{p_{i-1}+p_i}=0.$$

Implicitly differentiating (A.5) with respect to ρ, I get that

$$\begin{aligned}\hat{y}'(\rho) &= -\frac{\Delta_{C_A}(y_{i+1},y_i)-\Delta_{C_A}(y_i,y_{i-1})}{\frac{\Delta_{C_A}(y_{i+1},\hat{y})}{\partial\hat{y}}}, \text{ and} \\ \hat{y}''(\rho) &= \frac{\frac{\partial^2\Delta_{C_A}(y_{i+1},\hat{y})}{\partial\hat{y}^2}\hat{y}'(\rho)[\Delta_{C_A}(y_{i+1},y_i)-\Delta_{C_A}(y_i,y_{i-1})]}{(\Delta_{C_A}(y_{i+1},\hat{y}))^2}.\end{aligned} \tag{A.8}$$

Recall that by Condition (i) in Assumption 1, $\hat{y}'(\rho)<0$, and by Condition (i) in Assumption 2, $\hat{y}''(\rho)\geq 0$. Consider the second derivative of (A.7) with respect to ρ:

$$-C''_{Diff}(\hat{y}')^2-C'_{Diff}\hat{y}''+\delta_i[\hat{y}'-(1-\rho)\hat{y}''].$$

This second derivative is weakly negative given the fact that $\hat{y}'(\rho)<0$, $\hat{y}''(\rho)\geq 0$, $C'_{Diff}>0$, $C''_{Diff}\geq 0$ (Condition (ii) in Assumption 2). Hence, the RHS of (A.7) is concave, and therefore, together with (A.8), is weakly positive. This result implies that $\frac{\mathcal{U}_P(\hat{\mathcal{M}})-\mathcal{U}_P(\mathcal{M})}{p_{i-1}+p_i}>0$ and thus yields a contradiction.

Notice that the finiteness restriction does not play any role in the above argument. So the alignment principle holds generally in the class of discrete mechanisms (i.e., D is discrete).

A.3 Proof of Theorem 1

Let $\mathcal{M} = (D, \sigma)$ be an optimal obedient mechanism that implements N actions. Suppose by contradiction that $\mathcal{M}$ is not a monotone partition. It follows that there must be two actions $y_j < y_k$ in D and a cutoff $x \in \theta$ such that the conditional probabilities of the lower and upper segment have strictly positive probabilities, i.e.,

$$\int_{\underline{\theta}}^{x} \mathrm{d}\mu_j^\sigma > 0, \int_{x}^{\overline{\theta}} \mathrm{d}\mu_j^\sigma > 0, \int_{\underline{\theta}}^{x} \mathrm{d}\mu_k^\sigma > 0, \int_{x}^{\overline{\theta}} \mathrm{d}\mu_k^\sigma > 0.$$

I consider two exhaustive cases separately.

Case 1: Type y_k's ($\mathrm{OB}_{k,k+1}$) is binding for any $2 \leq k \leq (N-1)$. In this case, consider a local perturbation in the information structure σ:

(a) On the interval $[\underline{\theta}, x)$, subtract an absolute weight ε proportionally from recommendation y_k, and distribute the weight ε proportionally to recommendation y_j. Specifically, for each state $\theta \in [\underline{\theta}, x)$, the conditional probability switching from recommendation y_k to y_j is given by

$$\frac{\sigma(y_k|\theta)\varepsilon}{\int_{\underline{\theta}}^{x} \sigma(y_k|\theta)\,\mathrm{d}\mu_0}.$$

(b) On the interval $(x, \overline{\theta}]$, subtract an absolute weight ε proportionally from recommendation y_j, and distribute the weight ε proportionally to recommendation y_k. Specifically, for each state $\theta \in (x, \overline{\theta}]$, the conditional probability switching from recommendation y_j to y_k is given by

$$\frac{\sigma(y_j|\theta)\varepsilon}{\int_{x}^{\overline{\theta}} \sigma(y_j|\theta)\,\mathrm{d}\mu_0}.$$

(c) On the interval $[\underline{\theta}, \overline{\theta}]$, subtract an absolute weight ν proportionally from recommendation y_{k-1}, and distribute the weight ν proportionally to recommendation y_k. Specifically, for each state $\theta \in [\underline{\theta}, \overline{\theta}]$, the conditional probability switching from recommendation y_{k-1} to y_k is given by

$$\frac{\sigma(y_{k-1}|\theta)\nu}{\int_{\underline{\theta}}^{\overline{\theta}} \sigma(y_{k-1}|\theta)\,\mathrm{d}\mu_0},$$

where the value of ν is determined by restoring the binding ($\mathrm{OB}_{k,k+1}$) under the perturbed mechanism.

To simplify, the perturbed mechanism $\hat{\mathcal{M}} = (\hat{\sigma}, \hat{D})$ can be described as

(i) the delegation set $\hat{D}$ is the same as D.

(ii) $\hat{\sigma}$ satisfies

$$\hat{\sigma}(y|\theta) = \begin{cases} \sigma(y|\theta), & \text{if} \quad y \in \hat{D} \setminus \{y_j, y_k\}, \\ \hat{\sigma}(y_j|\theta), & \text{if} \quad y = y_j, \\ \hat{\sigma}(y_k|\theta), & \text{if} \quad y = y_k, \end{cases}$$

where

$$\hat{\sigma}(y_j|\theta) = \begin{cases} \sigma(y_j|\theta) + \frac{\sigma(y_k|\theta)\varepsilon}{\int_{\underline{\theta}}^{x} \sigma(y_k|\theta)\,\mathrm{d}F(\theta)} - \frac{\sigma(y_j|\theta)\nu}{\int_{\underline{\theta}}^{\overline{\theta}} \sigma(y_j|\theta)\,\mathrm{d}F(\theta)}, & \text{if} \quad \theta \in [\underline{\theta}, x), \\ \sigma(y_j|\theta) - \frac{\sigma(y_j|\theta)\varepsilon}{\int_{x}^{\overline{\theta}} \sigma(y_j|\theta)\,\mathrm{d}F(\theta)} - \frac{\sigma(y_j|\theta)\nu}{\int_{\underline{\theta}}^{\overline{\theta}} \sigma(y_j|\theta)\,\mathrm{d}F(\theta)}, & \text{if} \quad \theta \in (x, \overline{\theta}], \end{cases}$$

and

$$\hat{\sigma}(y_k|\theta) = \begin{cases} \sigma(y_k|\theta) - \frac{\sigma(y_k|\theta)\varepsilon}{\int_{\underline{\theta}}^{x} \sigma(y_k|\theta)\,\mathrm{d}F(\theta)} + \frac{\sigma(y_j|\theta)\nu}{\int_{\underline{\theta}}^{\overline{\theta}} \sigma(y_j|\theta)\,\mathrm{d}F(\theta)}, & \text{if} \quad \theta \in [\underline{\theta}, x), \\ \sigma(y_k|\theta) + \frac{\sigma(y_j|\theta)\varepsilon}{\int_{x}^{\overline{\theta}} \sigma(y_j|\theta)\,\mathrm{d}F(\theta)} + \frac{\sigma(y_j|\theta)\nu}{\int_{\underline{\theta}}^{\overline{\theta}} \sigma(y_j|\theta)\,\mathrm{d}F(\theta)}, & \text{if} \quad \theta \in (x, \overline{\theta}]. \end{cases}$$

(iii) the value of ν is set to be

$$\nu = \frac{m_j^H - m_k^L}{m_k - m_j}\varepsilon,$$

where $m_k^L = \int_0^x \theta\,\mathrm{d}\mu_k^\sigma$ denotes the posterior mean conditional on the lower segment of signal y_k, and $m_j^H = \int_x^1 \theta\,\mathrm{d}\mu_j^\sigma$ denotes the posterior mean conditional on the upper segment of signal y_j.

First, OB remains to hold under $\hat{\mathcal{M}}$. To see this, notice that for $i \neq j, k$, type y_i's ($\mathrm{OB}_{i,i+1}$) is automatically satisfied, which is directly inherited from that of $\mathcal{M}$. For type y_j, the conditional posterior mean

$$\hat{m}_j = \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta \hat{\sigma}(y_j|\theta)\,\mathrm{d}\mu_0}{\int_{\underline{\theta}}^{\overline{\theta}} \hat{\sigma}(y_j|\theta)\,\mathrm{d}\mu_0} = \frac{(p_j - \nu)m_j + \varepsilon m_k^L - \varepsilon m_j^H}{p_j - \nu} = m_j + \frac{m_k^L - m_j^H}{p_j - \nu}\varepsilon$$

is smaller than yet close to m_j for $\varepsilon > 0$ small enough. Hence, $(\mathrm{OB}_{j,j+1})$ is satisfied. For type y_k, the conditional posterior mean

$$\hat{m}_k = \frac{\int_{\underline{\theta}}^{\overline{\theta}} \theta \hat{\sigma}(y_k|\theta)\,\mathrm{d}\mu_0}{\int_{\underline{\theta}}^{\overline{\theta}} \hat{\sigma}(y_k|\theta)\,\mathrm{d}\mu_0} = \frac{p_k m_k - \varepsilon m_k^L + \varepsilon m_j^H + \nu m_j}{p_k + \nu} = m_k$$

stays the same by construction. An immediate implication of the above equality is that type y_k's $(\mathrm{OB}_{k,k+1})$ is still binding under $\hat{\mathcal{M}}$.

Second, I argue that the perturbed mechanism $\hat{\mathcal{M}}$ strictly improves the principal's payoff. In order to see this, I can represent the principal's change in payoff by

$$\begin{aligned} \mathcal{U}_P(\hat{\mathcal{M}}) - \mathcal{U}_{\mathcal{P}}(\mathcal{M}) = & \int_{\underline{\theta}}^{x} [u_P(y_j,\theta) - u_P(y_k,\theta)] \frac{\sigma(y_k|\theta)\varepsilon}{\int_{\underline{\theta}}^{x} \sigma(y_k|\theta)\,\mathrm{d}\mu_0}\,\mathrm{d}\mu_0 \\ & + \int_{x}^{\overline{\theta}} [u_P(y_k,\theta) - u_P(y_j,\theta)] \frac{\sigma(y_j|\theta)\varepsilon}{\int_{x}^{\overline{\theta}} \sigma(y_j|\theta)\,\mathrm{d}\mu_0}\,\mathrm{d}\mu_0 \\ & + \int_{\underline{\theta}}^{\overline{\theta}} [u_P(y_k,\theta) - u_P(y_j,\theta)] \frac{\sigma(y_j|\theta)\nu}{\int_{\underline{\theta}}^{\overline{\theta}} \sigma(y_j|\theta)\,\mathrm{d}\mu_0}\,\mathrm{d}\mu_0 \\ = & \varepsilon \frac{m_j^H - m_k^L}{m_k - m_j} \{m_k(y_k - y_j) - [C_P(y_k) - C_P(y_j)]\} \\ = & \varepsilon \frac{m_j^H - m_k^L}{m_k - m_j} [u_P(y_k, m_k) - u_P(y_j, m_k)] \\ \geq & 0, \end{aligned} \tag{A.9}$$

where the inequality follows from the fact that the optimal mechanism must be aligned from Proposition 1. If $u_P(y_k, m_k) - u_P(y_j, m_k) > 0$, (A.9) yields $\mathcal{U}_P(\hat{\mathcal{M}}) - \mathcal{U}_{\mathcal{P}}(\mathcal{M}) > 0$, which contradicts the optimality of $\mathcal{M}$. If $u_P(y_k, m_k) - u_P(y_j, m_k) = 0$, then the value of j must be equal to $(k-1)$ due to the alignment principle. In this case, consider an alternative perturbation $\tilde{\mathcal{M}}$ by pooling type y_j $(= y_{k-1})$ and type y_k into a single type associated with some action $\hat{y}$. Analogous arguments to those used in the proof of Proposition 1, the perturbation would weakly increase Principal's payoff.

Case 2: Either type y_k's ($\text{OB}_{k,k+1}$) is not binding for $2 \leq k \leq N$, or $k = N$. This case can be treated by solely applying the "swapping" procedure (a)–(b) from Case 1. It is straightforward to see that OB is satisfied in the perturbed mechanism $\hat{\mathcal{M}} = (D, \hat{\sigma})$ given that $\varepsilon > 0$ is small enough. Doing so, the principal's marginal payoff change can be expressed as

$$\begin{aligned}\lim_{\varepsilon \to 0} \frac{\mathcal{U}_P(\hat{\mathcal{M}}) - \mathcal{U}_P(\mathcal{M})}{\varepsilon} =& [u_P(y_j, m_k^L) - u_P(y_k, m_k^L)] - [u_P(y_k, m_j^H) - u_P(y_j, m_j^H)] \\ =& (y_k - y_j)(m_j^H - m_k^L) > 0,\end{aligned}$$

which, again contradicts the hypothesized optimality of $\mathcal{M}$.

Combining **Case 1** and **Case 2**, I obtain that the principal can without loss focus on the class of monotone partitions.

A.4 Proof of Theorem 2

I prove the theorem in two steps. In the first step, I show that in the class of discrete mechanisms (i.e., D is discrete), the optimal mechanism exists and it is a finite monotone partition. In the second step, I prove that the finite monotone partitional mechanism is also maximal in the class of mechanisms with an arbitrary number of delegated actions.

First, using an argument analogous to the one used by Theorem 1, I obtain that the monotone partitional mechanism (not necessarily finite) is also maximal within the discrete class. To prove existence, invoking Theorem 1 and Lemma 5 below allows me to focus on the space of finite monotone partitional mechanisms and denote by $z = (x_1, \ldots, x_{\overline{N}}, y_1, \ldots, y_{\overline{N}}) \in [\underline{\theta}, \overline{\theta}]^{\overline{N}} \times [y_P(\underline{\theta}), y_A(\overline{\theta})]^{\overline{N}}$ a typical element of this space. Here $\overline{N}$ is the upper bound established in Lemma 5. It is easy to see that the space is compact under the Euclidean metric. By continuity of $u_P(y, \theta)$ and Weierstrass' theorem, an optimal mechanism exists.

Second, to prove the optimality of finite monotone partitional mechanisms in the class of mechanisms with an arbitrary delegation space, it suffices to show that for any payoff generated by an arbitrary obedient mechanism, I can construct a sequence of finite mechanisms such that the associated payoff sequences converges to the given one (proof postponed in Lemma 6). To see the sufficiency, suppose instead no finite mechanism obtains the payoff supremum. Then by Lemma

6, there exists a sequence of finite obedient mechanisms $\{(D_n, \sigma_n)\}_{n=1}^{\infty}$ such that $|D_n| < |D_{n+1}|$, $\mathcal{U}_P(D_n, \sigma_n) < \mathcal{U}_P(D_{n+1}, \sigma_{n+1})$, and $\lim_{n\to\infty} \mathcal{U}_P(D_n, \sigma_n)$ attains the supremum. But by the first step, there must exist $n^* < \infty$ and $|D_{n^*}| \leq \overline{N}$ such that for all $n > n^*$, $\mathcal{U}_P(D_n, \sigma_n) \leq \mathcal{U}_P(D_{n^*}, \sigma_{n^*})$, a contradiction.

Lemma 5. *Let $\mathcal{M} = (D, \sigma)$ be an optimal obedient mechanism that is discrete. Then, it must be finite.*

Proof of Lemma 5. Let y_{i-1}, y_i, y_{i+1} be three adjacent actions induced in the mechanism. By obedience condition, it follows that $m_i \leq \Delta_{C_A}(y_{i+1}, y_i)$. By the alignment principle, it follows that $m_i \geq \Delta_{C_P}(y_i, y_{i-1})$. Combining the two inequalities and using the fact that $c_P > c_A$ yields

$$\Delta_{C_P}(y_{i+1}, y_i) - \delta \geq \Delta_{C_A}(y_{i+1}, y_i) \geq m_i \geq \Delta_{C_P}(y_i, y_{i-1}),$$

where $\delta \equiv \inf_{y\in[y_P(\underline{\theta}), y_A(\overline{\theta})]} c_P(y) - c_A(y) > 0$. Then by the mean value theorem, there exist $y_l \in (y_{i-1}, y_i)$ and $y_h \in (y_i, y_{i+1})$ such that

$$c_P(y_h) - c_P(y_l) = \Delta_{C_P}(y_{i+1}, y_i) - \Delta_{C_P}(y_i, y_{i-1}) \geq \delta.$$

Since $c'_P > 0$ is continuous, this gives an upper bound K to the value of c'_P attainable on $[y_P(\underline{\theta}), y_A(\overline{\theta})]$. Thus, the above inequality yields

$$K \geq \frac{c_P(y_h) - c_P(y_l)}{y_h - y_l} \geq \frac{\delta}{y_h - y_l},$$

which then implies

$$y_{i+1} - y_{i-1} > y_h - y_l > \frac{\delta}{K} \equiv \varepsilon > 0.$$

Using the fact that the delegation set is bounded by $y_P(\underline{\theta})$ and $y_A(\overline{\theta})$ from Lemma 2, this inequality implies that the size of any aligned and discrete mechanism must be bounded above by $2\left\lceil \frac{y_A(\overline{\theta}) - y_P(\underline{\theta})}{\varepsilon} \right\rceil + 1$, which completes the proof. □

Lemma 6. *Suppose that $\mathcal{M} = (D, \sigma)$ is an obedient mechanism with an arbitrary D. Then, there* *exists a sequence of finite mechanisms such that*

$$\lim_{n\to\infty} \mathcal{U}_P(D_n,\sigma_n) = \mathcal{U}_P(D,\sigma).$$

Proof of Lemma 6. Because the principal's payoff depends on posterior beliefs only through the posterior mean, her expected payoff from an arbitrary mechanism $\mathcal{M} = (D,\sigma)$ can be written as

$$\mathcal{U}_P(D,\sigma) = \int_{\underline{\theta}}^{\overline{\theta}} u_P(y(m),m)\,\mathrm{d}G(m),$$

where G is the cumulative distribution over posterior means, and $y(m) \in D$ is the corresponding action recommendation that induces m. Thus, every pair (D,σ) can be equivalently described by the corresponding pair (G,y), which I use for the rest of the proof.

Take an arbitrary positive integer n. Let $\{B_1^n, B_2^n \ldots, B_n^n\}$ be a partition of the posterior mean space $[\underline{\theta},\overline{\theta}]$ such that

$$B_k^n = (b_{k-1}^n, b_k^n] \quad \forall k = 1,\ldots,n,$$

with $b_k^n = \frac{n-k}{n}\underline{\theta} + \frac{k}{n}\overline{\theta}$ and $b_0^n = \underline{\theta}$. Let M denote the support of distribution G and define a sequence of posterior means $\{m_j^n\}_{j=1}^{J^n}$ by

$$m_j^n = \frac{\int_{B_{k(j)}^n} m\,\mathrm{d}G(m)}{\int_{B_{k(j)}^n} \mathrm{d}G(m)}$$

such that $k(j)$ is the j^{th} partition element $B_{k(j)}^n$ that satisfies $B_{k(j)}^n \cap M \neq \emptyset$. By construction, I can define a discrete distribution G^n by

$$G^n(m) = \begin{cases} 0 & \text{for } m \in [\underline{\theta}, m_1^n), \\ G(b_{k(j)-1}^n) & \text{for } m \in [m_{j-1}^n, m_j^n), \\ 1 & \text{for } m \in [m_{J^n}^n, \overline{\theta}]. \end{cases}$$

Let $\underline{m}_j^n$ and $\overline{m}_j^n$ denote respectively the lower bound and the upper bound of $B_{k(j)}^n \cap M$. Define

$y^n : M \to Y$ as

$$y^n(m) = \begin{cases} y(\underline{m}_j^n) & \text{if } m \in [m_{j-1}^n, m_j^n) \text{ or } [m_{J^n}^n, \overline{\theta}] \text{ and } U_A(y(\underline{m}_j^n)|m_j^n) \geq U_A(y(\overline{m}_j^n)|m_j^n), \\ y(\overline{m}_j^n) & \text{if } m \in [m_{j-1}^n, m_j^n) \text{ or } [m_{J^n}^n, \overline{\theta}] \text{ and } U_A(y(\underline{m}_j^n)|m_j^n) < U_A(y(\overline{m}_j^n)|m_j^n). \end{cases}$$

The above construction allows me to obtain a sequence of finite mechanisms $\{(G^n, y^n)\}_{n=1}^{\infty}$. Notice that (G^n, y^n) satisfies the obedience condition. Also, both G^n and $y^n(\cdot)$ are *simple functions* which uniformly converge to G and $y(\cdot)$, respectively.

Now I am ready to show that the sequence of principal's expected payoffs from $\{(G^n, y^n)\}_{n=1}^{\infty}$ converges to that from (G, m). To see this, notice that both $u_P(y^n(m), m)$ and $G^n(m)$ are uniformly bounded, and converge uniformly to $u_P(y(m), m)$ and $G(m)$, respectively. By bounded convergence theorem, it then follows that

$$\lim_{n\to\infty} \int_{\underline{\theta}}^{\overline{\theta}} u_P(y^n(m), m) \, \mathrm{d}G^n(m) = \int_{\underline{\theta}}^{\overline{\theta}} u_P(y(m), m) \, \mathrm{d}G(m),$$

which completes the proof.

□

A.5 Proof of Lemma 4

Proof of Lemma 4. This is an immediate consequence of Lemma 7 and 8 below. □

Lemma 7. *Let $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ be an optimal mechanism in the uniform-quadratic case. Then, the following must hold*

$$y_i - \theta_i = y_j - \theta_j \quad \forall i, j = 1, \ldots, N. \tag{A.10}$$

Proof. Notice that I only need show that (A.10) holds for any two consecutive intervals. To see this, consider the following modified mechanism $\mathcal{M}(\varepsilon)$:

(a) If $\varepsilon > 0$, then replace action recommendation y_{i+1} with y_i throughout states $\theta \in [\theta_i, \theta_i + \varepsilon]$; if $\varepsilon < 0$, replace action recommendation y_i with y_{i+1} throughout states $\theta \in [\theta_i + \varepsilon, \theta_i]$.

(b) Replace action y_{i+1} from D with $\hat{y}_{i+1} = y_{i+1} + \varepsilon$.

It follows that obedience is satisfied by construction. Moreover, the principal's incremental change in payoff can be written as

$$\begin{aligned}\mathcal{U}_P(\mathcal{M}(\varepsilon)) - \mathcal{U}_P(\mathcal{M}) =& \int_{\theta_i}^{\theta_i+\varepsilon} \left[(\theta y_i - \frac{1}{2}y_i^2) - (\theta y_{i+1} - \frac{1}{2}y_{i+1}^2)\right] \mathrm{d}\theta \\ &+ \int_{\theta_i+\varepsilon}^{\theta_{i+1}} \left[(\theta \hat{y}_{i+1} - \frac{1}{2}\hat{y}_{i+1}^2) - (\theta y_{i+1} - \frac{1}{2}y_{i+1}^2)\right] \mathrm{d}\theta \\ =& \frac{\varepsilon}{2}\left[(y_{i+1} - \theta_{i+1})^2 - (y_i - \theta_i)^2\right] + \frac{\varepsilon^2}{2}[(y_i - \theta_i) + (y_{i+1} - \theta_{i+1})].\end{aligned} \tag{A.11}$$

Provided that $\mathcal{M}$ is optimal, I must have

$$(y_{i+1} - \theta_{i+1})^2 = (y_i - \theta_i)^2 \quad \forall i = 1, \dots, (N-1). \tag{A.12}$$

Otherwise, I can take the limit of ε to zero, and obtain

$$\lim_{\varepsilon \to 0^+} \frac{\mathcal{U}_P(\mathcal{M}(\varepsilon)) - \mathcal{U}_P(\mathcal{M})}{\varepsilon} = (y_{i+1} - \theta_{i+1})^2 - (y_i - \theta_i)^2,$$

or

$$\lim_{\varepsilon \to 0^-} \frac{\mathcal{U}_P(\mathcal{M}(\varepsilon)) - \mathcal{U}_P(\mathcal{M})}{\varepsilon} = (y_i - \theta_i)^2 - (y_{i+1} - \theta_{i+1})^2,$$

contradicting the optimality of $\mathcal{M}$.

Also, observe that the second term of (A.11) must be weakly negative, i.e., $(y_i - \theta_i) + (y_{i+1} - \theta_{i+1}) \leq 0$ due to the optimality of $\mathcal{M}$. This observation, combined with (A.12), leads to

$$y_i - \theta_i = y_{i+1} - \theta_{i+1} \quad \forall i = 1, \dots, (N-1).$$

□

An immediate implication of the above proof is that any optimal mechanism $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ must satisfy $y_i \leq \theta_i$ for all $i = 1, \dots, N$, which I will use later. Notice also that, by Lemma 7, an optimal mechanism $\{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ (with an abuse of notation) can be equivalently expressed as $\{(\theta_i)_{i=1}^N, d\}$, where $d = \theta_i - y_i$ for all $i = 1, \dots, N$. Fix $\mathcal{M} = \{(\theta_i)_{i=1}^N, d\}$ and denote by

$\Delta\theta_i = \theta_i - \theta_{i-1}$ the length of the i^{th} interval. The principal's ex-ante payoff from $\mathcal{M}$ is

$$\begin{aligned}
\mathcal{U}_P(\mathcal{M}) &= \sum_{i=1}^{N} \int_{\theta_{i-1}}^{\theta_i} (\theta y_i - \frac{1}{2} y_i^2)\, \mathrm{d}\theta \\
&= \sum_{i=1}^{N} \int_{\theta_{i-1}}^{\theta_i} (\theta y_i - \frac{1}{2} y_i^2 - \frac{1}{2}\theta^2)\, \mathrm{d}\theta + \sum_{i=1}^{N} \int_{\theta_{i-1}}^{\theta_i} \frac{1}{2}\theta^2\, \mathrm{d}\theta \\
&= \sum_{i=1}^{N} \frac{1}{2}(\theta_i - \theta_{i-1}) \left[d(\theta_i - \theta_{i-1}) - d^2 - \frac{2}{3}(\theta_i - \theta_{i-1})^2 \right] + \frac{1}{6} \\
&= \sum_{i=1}^{N} \phi(\Delta\theta_i, d) + \frac{1}{6},
\end{aligned} \tag{A.13}$$

where $\phi(x, d) = -\frac{1}{6}x^3 + \frac{1}{2}dx^2 - \frac{1}{2}d^2 x$ is strictly concave in x for $x > 0$.

Before proceeding to Lemma 8, let me first introduce the majorization relation and the majorization inequality. A sequence $\{x_i\}_{i=1}^N$ *majorizes* another sequence $\{\hat{x}_i\}_{i=1}^N$ if

$$x_1 + \cdots + x_i \geq \hat{x}_1 + \cdots + \hat{x}_i \quad \forall i = 1, \ldots, N-1.$$

and

$$x_1 + \cdots + x_N = \hat{x}_1 + \cdots + \hat{x}_N.$$

Karamata's inequality. Let sequences $\{x_i\}_{i=1}^N$ and $\{\hat{x}_i\}_{i=1}^N$ be non-increasing and $\phi(\cdot)$ denote a real-valued, concave function. If $\{\hat{x}_i\}_{i=1}^N$ majorizes $\{x_i\}_{i=1}^N$, then

$$\sum_{i=1}^{N} \phi(\hat{x}_i) \leq \sum_{i=1}^{N} \phi(x_i).$$

I am now ready to state and prove Lemma 8.

Lemma 8. *Let $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ be an optimal mechanism in the uniform-quadratic case. Then, the local upward OB is binding.*

Proof. There are two distinct cases: either N is odd, or N is even. I now show that the binding condition is necessary for optimality in both cases.

Case 1. N is odd. Let $\hat{\mathcal{M}} = \{(\hat{\theta}_i)_{i=1}^N, (\hat{y}_i)_{i=1}^N\}$ be an obedient mechanism that satisfies $d = \hat{\theta}_i - \hat{y}_i$ for some $d \geq 0$. Then I can construct another mechanism $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ which satisfies $d = \theta_i - y_i$, and moreover, features binding local upward OB. It can be shown that $\theta_i = 1-(b+d)(N-i)$ for odd i, and $\theta_i = (b+d)i$ for even i, which implies that $\Delta\theta_i = 1-(b+d)(N-1)$ for odd i, and $\Delta\theta_i = (b+d)(N+1)-1$ for even i. Denote by $\{x_i\}_{i=1}^N$ the non-increasing permutation of $\{\Delta\theta_i\}_{i=1}^N$.

Consider next the mechanism $\hat{\mathcal{M}}$. For even i, OB condition implies $\hat{\theta}_i \geq \theta_{i-2}+2(b+d) \geq \cdots \geq (b+d)i = \theta_i$. For odd i, OB condition implies $\hat{\theta}_i \leq \hat{\theta}_{i+2}+2(b+d) \leq \cdots \leq 1-(b+d)(N-i) = \theta_i$. I thus obtain $\Delta\hat{\theta}_i = \hat{\theta}_i - \hat{\theta}_{i-1} \geq \theta_i - \theta_{i-1} = \Delta\theta_i$ for even i and $\Delta\hat{\theta}_i \leq \Delta\theta_i$ for odd i. Likewise, I denote by $\{\hat{x}_i\}_{i=1}^N$ the non-increasing permutation of $\{\Delta\hat{\theta}_i\}_{i=1}^N$.

It can be shown that $\{\hat{x}_i\}_{i=1}^N$ majorizes $\{x_i\}_{i=1}^N$, and hence, by Karamata's inequality and (A.13), I obtain $\mathcal{U}_P(\mathcal{M}) = \sum_{i=1}^N \phi(x_i,d) + \frac{1}{6} \geq \sum_{i=1}^N \phi(\hat{x}_i,d) + \frac{1}{6} = \mathcal{U}_P(\hat{\mathcal{M}})$.

Case 2. N is even. In this case, consider a mechanism $\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\}$ which satisfies $\theta_i = \frac{1}{N}, y_i = b + \frac{i-1}{N}$ for all $i = 1,\ldots,N$. One can easily verify that $\mathcal{M}$ features binding local upward OB. The next step is to show that the mechanism $\mathcal{M}$ is indeed payoff superior to all obedient mechanisms of the same size N, and hence the claim in Lemma 8 follows.

To see this, consider an obedient mechanism $\hat{\mathcal{M}} = \{(\hat{\theta}_i)_{i=1}^N, (\hat{y}_i)_{i=1}^N\}$ that satisfies $\hat{d} = \hat{\theta}_i - \hat{y}_i$ for some $\hat{d} \geq 0$. The OB condition is equivalent to requiring $\hat{\theta}_{i+1} - \hat{\theta}_{i-1} \geq 2(b+\hat{d})$, which, together with the fact that $\sum_{i=1}^{N-1}(\hat{\theta}_{i+1} - \hat{\theta}_{i-1}) = 1$, implies $\hat{d} \leq \frac{1}{N} - b$. Now consider another mechanism $\hat{\mathcal{M}}' = \{\{(\hat{\theta}_i{}')_{i=1}^N, (\hat{y}_i')_{i=1}^N\}$ with $\hat{\theta}_i = \frac{1}{N}$ and $\hat{y}_i' = \frac{1}{N} - \hat{d}$. It follows that $\hat{\mathcal{M}}'$ is obedient. Then an analogous majorization argument to that used to prove **Case 1** can be applied to show that $\mathcal{U}_P(\hat{\mathcal{M}}') \geq \mathcal{U}_P(\hat{\mathcal{M}})$. The argument is similar and hence omitted.

So all it needs to show is then that $\mathcal{M}$ yields a higher payoff than $\hat{\mathcal{M}}'$. Note that by (A.13), $\mathcal{U}_P(\mathcal{M}) = \sum_{i=1}^N \phi(\frac{1}{N}, \frac{1}{N} - b) + \frac{1}{6} \geq \sum_{i=1}^N \phi(\frac{1}{N}, \hat{d}) + \frac{1}{6} = \mathcal{U}_P(\hat{\mathcal{M}}')$ due to the fact that $\hat{d} \leq \frac{1}{N} - b$ and $\phi(x,d)$ is increasing in d.

□

A.6 Proof of Proposition 2

It follows from Lemma 4 that one only needs to focus on the first two partition elements. The problem of finding the suboptimal mechanism for each size $N \leq \overline{N}(b)$ reduces to one of the following two trivial programs:

(1) If N is an odd number, then

$$\max_{\theta_1,\theta_2,y_1,y_2} \frac{N+1}{2}\int_0^{\theta_1} u_P(y_1,\theta)\,\mathrm{d}\theta + \frac{N-1}{2}\int_{\theta_1}^{\theta_2} u_P(y_2,\theta)\,\mathrm{d}\theta \tag{A.14}$$

subject to (i) $y_1 - \theta_1 = y_2 - \theta_2$, and (ii) $\frac{N+1}{2}\theta_1 + \frac{N-1}{2}\theta_2 = 1$.

(2) If N is an even number, then

$$\max_{\theta_1,\theta_2,y_1,y_2} \frac{N}{2}\int_0^{\theta_1} u_P(y_1,\theta)\,\mathrm{d}\theta + \frac{N}{2}\int_{\theta_1}^{\theta_2} u_P(y_2,\theta)\,\mathrm{d}\theta \tag{A.15}$$

subject to (i) $y_1 - \theta_1 = y_2 - \theta_2$, and (ii) $\frac{N}{2}\theta_1 + \frac{N}{2}\theta_2 = 1$.

Therefore, it is straightforward to verify that (2) and (2) are solutions to programs (A.14) and (A.15), respectively.

A.7 Proof of Theorem 3

Fix an arbitrary agent $\tilde{A}$ with $c_{\tilde{A}}$. Define the associated set of uniformly less-biased agents by $\mathbb{c}_{\tilde{A}} = \{c_A : c_A(y) \geq c_{\tilde{A}}(y), \forall y \in Y\}$, where $\mathbb{c}_{\tilde{A}}$ is endowed with the sup norm $\|\cdot\|_\infty$. The induced metric is then $d(c_A, c_{A'}) = \|c_A - c_{A'}\|_\infty$ for any $c_A, c_{A'} \in \mathbb{c}_{\tilde{A}}$.

Let $\mathcal{M}^*_{\tilde{A}}$ of size $N^*_{\tilde{A}}$ and $\mathcal{M}^*_A$ of size N^*_A be the optimal mechanisms with $c_{\tilde{A}}$ and c_A, respectively. It follows from Theorem 4 that $\mathcal{M}^*_A$ achieves a weakly higher expected payoff than $\mathcal{M}^*_{\tilde{A}}$, i.e., $\mathcal{U}_P(\mathcal{M}^*_A) \geq \mathcal{U}_P(\mathcal{M}^*_{\tilde{A}})$.

To prove $N^*_A \geq N^*_{\tilde{A}}$ for any $c_A \in \mathbb{c}_{\tilde{A}}$, note that it suffices to check "local monotonicity", i.e., $N^*_A \geq N^*_{\tilde{A}}$ for $\|c_A - c_{\tilde{A}}\|_\infty$ sufficiently small. Suppose by contradiction that $N^*_A < N^*_{\tilde{A}}$. Take $\mathcal{M}^*_{\tilde{A},N^*_A}$ as the optimal mechanism of size N^*_A with agent $c_{\tilde{A}}$ and it is obvious that $\mathcal{U}_P(\mathcal{M}^*_{\tilde{A}}) > \mathcal{U}_P(\mathcal{M}^*_{\tilde{A},N^*_A})$.

Let $Feas^N(c_A)$ be the set of all obedient mechanisms of arbitrary size N with agent c_A:

$$\left\{\mathcal{M} = \{(\theta_i)_{i=1}^N, (y_i)_{i=1}^N\} : \frac{\theta_{i-1}+\theta_i}{2} \leq \frac{C_A(y_{i+1}) - C_A(y_i)}{y_{i+1}-y_i}\right\}.$$

By definition, it follows that $Feas^N(c_A)$ is compact and continuous in c_A. Since $\mathcal{U}_P : [\underline{\theta}, \overline{\theta}]^N \times [\underline{y}, \overline{y}]^N \to \mathbb{R}$ is continuous, Berge's Maximum Theorem implies that the optimal expected payoff function

$$\mathcal{V}_P^N(c_A) \equiv \max_{\mathcal{M} \in Feas^N(c_A)} \mathcal{U}_P(\mathcal{M}) \tag{A.16}$$

is continuous. Take $\mathcal{M}^*_{A,N}$ to be an optimal solution for (A.16) by restricting the partition size to be N. Thus, for sufficiently small $\varepsilon > 0$ with $d(c_A, c_{\tilde{A}}) < \varepsilon$, there exists $\delta > 0$ small enough such that $\mathcal{U}_P(\mathcal{M}^*_{A,N}) = \mathcal{V}_P^N(c_A) \leq \mathcal{V}_P^N(c_{\tilde{A}}) + \delta = \mathcal{U}_P(\mathcal{M}^*_{\tilde{A},N}) + \delta$. Setting $N = N^*_A$, one then obtains $\mathcal{U}_P(\mathcal{M}^*_A) \leq \mathcal{U}_P(\mathcal{M}^*_{\tilde{A},N^*_A}) + \delta < \mathcal{U}_P(\mathcal{M}^*_{\tilde{A}})$. A contradiction.

References

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. 2016. "The Economics of Privacy." *Journal of Economic Literature,* 54 (2):442–492.

Aguirre, Inaki, Simon Cowan, and John Vickers. 2010. "Monopoly Price Discrimination and Demand Curvature." *American Economic Review,* 100 (4):1601–1615.

Alonso, Ricardo and Niko Matouschek. 2008. "Optimal Delegation." *Review of Economic Studies,* 75 (1):259–293.

Amador, Manuel and Kyle Bagwell. 2013. "The Theory of Optimal Delegation with An Application to Tariff Caps." *Econometrica,* 81 (4):1541–1599.

———. 2020. "Money Burning in the Theory of Delegation." *Games and Economic Behavior,* 121:382–412.

Amador, Manuel, Iván Werning, and George-Marios Angeletos. 2006. "Commitment v.s. Flexibility." *Econometrica,* 74 (2):365–396.

Argenziano, Rossella and Alessandro Bonatti. 2020. "Information Revelation and Privacy Protection." Mimeo, University of Essex and MIT Sloan School of Management.

Baron, David P. and Roger B. Myerson. 1982. "Regulating A Monopolist with Unknown Costs." *Econometrica,* 50 (4):911–930.

Bénabou, Roland and Jean Tirole. 2002. "Self-Confidence and Personal Motivation." *The Quarterly Journal of Economics,* 117 (3):871–915.

Bergemann, Dirk and Stephen Morris. 2016. "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games." *Theoretical Economics,* 11 (2):487–522.

Bergemann, Dirk and Martin Pesendorfer. 2007. "Information Structures in Optimal Auctions." *Journal of Economic Theory,* 137 (1):580–609.

Carrillo, Juan D. and Thomas Mariotti. 2000. “Strategic Ignorance as a Self-Disciplining Device.” *Review of Economic Studies,* 67 (3):529–544.

Cowan, Simon. 2016. “Welfare-Increasing Third-Degree Price Discrimination.” *RAND Journal of Economics,* 47 (2):326–340.

Crawford, Vincent P. and Joel Sobel. 1982. “Strategic information transmission.” *Econometrica,* 50 (6):1431–1451.

Dworczak, Piotr and Giorgio Martini. 2019. “The Simple Economics of Optimal Persuasion.” *Journal of Political Economy,* 127 (5):1993–2048.

Frankel, Alexander. 2014. “Aligned Delegation.” *American Economic Review,* 104 (1):66–83.

Holmström, Bengt. 1977. “On Incentives and Controls in Organizations.” Phd thesis, Stanford University.

———. 1984. “On the Theory of Delegation.” *in M. Boyer, and R. Kihlstrom (eds.) Bayesian Models in Economic Theory (New York: North-Holland)* :115–141.

Ivanov, Maxim. 2021. “Optimal Monotone Signals in Bayesian Persuasion Mechanisms.” *Economic Theory,* 72:955–1000.

Kamenica, Emir and Matthew Gentzkow. 2011. “Bayesian Persuasion.” *American Economic Review,* 101 (6):2590–2615.

Kolotilin, Anton. 2018. “Optimal Information Disclosure: A Linear Programming Approach.” *Theoretical Economics,* 13 (2):607–635.

Martimort, David and Aggey Semenov. 2006. “Continuity in Mechanism Design without Transfers.” *Economic Letters,* 93 (2):182–189.

Melumad, Nahum D. and Toshiyuki Shibano. 1991. “Communication in Settings with No Transfers.” *RAND Journal of Economics,* 22 (2):173–198.

Mensch, Jeffrey. 2021. “Monotone Persuasion.” *Games and Economic Behavior* 130:521–542.

Pigou, Arthur C. 1920. *The Economics of Welfare.* London: Macmillan.

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

Robinson, Joan. 1933. *The Economics of Imperfect Competition.* London: Macmillan.

Roesler, Anne-Katrin and Balázs Szentes. 2017. "Buyer-Optimal Learning and Monopoly Pricing." *American Economic Review,* 107 (7):2072–2080.

Schmalensee, Richard. 1981. "Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination." *American Economic Review,* 71 (1):242–247.

Strotz, Robert H. 1995. "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies,* 23 (3):165–180.

Varian, Hal R. 1985. "Price Discrimination and Social Welfare." *American Economic Review,* 75 (4):870–875.