Mechanism Design with Aftermarkets: Optimal Mechanisms under Binary Actions

Piotr Dworczak[*]

February 18, 2019

Abstract

I study a mechanism design problem of allocating a single good to an agent when the mechanism is followed by a post-mechanism game (aftermarket) played between the agent and a third-party. The aftermarket is beyond the direct control of the designer. However, she can influence the information structure of the post-mechanism game by disclosing information about the agent's type elicited by the mechanism. Under the simplifying assumption that the third party takes a binary action, I characterize the optimal Bayesian mechanism for a large set of objective functions of the mechanism designer. I relate optimal transparency of the mechanism to the properties of the aftermarket game.

Keywords: Mechanism Design, Information Design, Privacy, Transparency

JEL codes: D44, D47, D82, D83.

[*]Northwestern University, Department of Economics piotr.dworczak@northwestern.edu. I would like to thank Darrell Duffie, Paul Milgrom, Michael Ostrovsky, Alessandro Pavan, and Andy Skrzypacz for helpful discussions.

1 Introduction

Mechanism design theory often abstracts away from the broader economic environment in which the mechanism design problem is embedded. While this is a useful abstraction, real-life mechanisms do not take place in a vacuum. Auctions are often run in markets where bidders can resell the items they bought. Procurement mechanisms are followed by bargaining between winners of contracts and subcontractors. Mergers and acquisitions influence subsequent strategic interactions between players in the underlying market. While such post-mechanism games may be beyond the direct control of the designer, their presence has important consequences for the design problem. One such consequence is that the transparency of the mechanism matters: By carefully disclosing the information elicited by the mechanism, the designer can indirectly influence the outcomes of the post-mechanism game. For example, by revealing the price paid by the winning bidder in a first price auction, an auctioneer changes the information structure of a post-auction resale game. This can lead to a different split of surplus in the aftermarket, and hence different behavior of agents in the mechanism. Thus, in mechanism design with aftermarkets, the designer chooses not only the allocation rule and transfers but also the information disclosure rule.

In this paper, I propose a tractable framework for mechanism design with aftermarkets that allows me to highlight how the form of the aftermarket influences the structure of the optimal Bayesian mechanism, and in particular its transparency. For example, what is the optimal level of privacy in the mechanism depending on whether agents resell their objects in the aftermarket or buy complementary goods from third-party sellers? How to structure an M&A transaction to maximize revenue for the seller or total surplus in the market, depending on the structure of competition and the potential for entry to the industry? The answer identified by this paper is that the form of the optimal mechanism depends on the properties of the agent's payoff from participating in the aftermarket. That payoff function depends on the type of the agent, and the aftermarket belief about her type induced by the signal disclosed in the mechanism. I show that what matters is how the willingness to pay for beliefs changes with the type of the agent. Other factors, such as whether the agent benefits from signaling a high or a low type, are less relevant.

The baseline model that I study features three players: the mechanism designer, the agent, and a third party (the analysis extends to the case of multiple agents in the mechanism under the assumption that only the winner interacts in the aftermarket).

The mechanism designer controls a single good, and designs a mechanism consisting of an allocation rule, a disclosure rule, and a transfer rule, under full commitment. The agent has a one-dimensional type, with higher types corresponding to higher willingness to pay for the good, and a utility that is linear in the type. The agent first participates in the mechanism, and then, conditional on acquiring the good, interacts in the aftermarket with the third party. The third party observes the signal disclosed by the mechanism, and takes a binary action which affects all players' payoffs. Crucially, the designer cannot design the aftermarket, and cannot interact with the third party directly (except by influencing the third party's beliefs via signals revealed in the mechanism).

The linearity of the agent's utility function and the binary action of the third party are the main simplifications that make the model tractable. Despite the strong restrictions imposed by these stylized assumptions, the model is flexible enough to qualitatively capture the main features of economically relevant aftermarkets, including resale, acquisition of a complementary good, investment decisions, or entry games. Thus, the model sacrifices mathematical generality in order to tractably analyze a range of aftermarkets differing in their key economic properties, in line with the goal of the paper.

The key notion, introduced first in a more general setting by a companion paper Dworczak (2019), is that of *submodular* and *supermodular* aftermarkets. Suppose that agent's types are ordered by willingness to pay for the good, and beliefs are ordered by, for example, first-order stochastic dominance (an order in which a belief is higher if it puts more mass on higher types). Then, an aftermarket is submodular if the agent's aftermarket payoff, viewed as a function of her type and the posterior belief about her type, is submodular. That is, lower types have a higher willingness to pay for higher beliefs. For example, resale markets are submodular because low types benefit relatively more from high resale prices. In the opposite case, when higher types value higher beliefs more than low types, the aftermarket is supermodular. The companion paper Dworczak (2019) points out that submodular aftermarket make information disclosure in the mechanism "difficult," while supermodular aftermarkets make information disclosure "easy."

The first result of this paper (Theorem 1) shows that this claim holds in a very strong sense in the current simple setting. I prove that if the allocation rule is constant, no incentive-compatible mechanism can disclose any non-trivial information (except for the fact that trade took place) when the aftermarket is submodular. In contrast, information disclosure is possible if the aftermarket is supermodular, and indeed even the exact type of the agent can be revealed in an incentive-compatible way.

The main results of the paper (Theorems 2 and 3) demonstrate that the distinction between submodular and supermodular aftermarkets remains relevant for the identification of the optimal Bayesian mechanism. Under certain additional conditions, *cutoff mechanisms* are optimal (or nearly optimal) in submodular aftermarket, while *partitional mechanisms* are optimal in supermodular aftermarkets. Cutoff mechanisms (introduced and analyzed extensively by Dworczak, 2019) are defined as mechanisms that only disclose information about the random threshold that the agent must outbid in order to acquire the object. For example, a designer might use a random reserve price, and disclose some information about its realization which indirectly reveals information about the agent's type conditional on the agent acquiring the good. Partitional mechanisms reveal information about the type of the agent directly by partitioning the type space, and disclosing the element of the partition that contains that type.

Exact optimality of cutoff mechanisms is obtained under a relatively strong form of submodularity of the aftermarket. I show that when this notion is relaxed, cutoff mechanisms may not be optimal but they remain "nearly optimal" in the sense that by using a cutoff mechanism, the designer can guarantee a large fraction of the payoff that she would obtain by employing the fully optimal mechanism.

These results, although established in a restrictive setting, have implications for optimal transparency of real-life mechanisms. Dworczak (2019) shows that when the designer optimizes within the class of cutoff mechanisms, the optimal mechanism does not disclose any information when there is only a single agent in the mechanism. Combined with the results of this paper, this means that privacy may be desirable in bilateral transactions followed by resale aftermarkets, or other submodular aftermarkets, even if the designer has intrinsic preferences for more information being disclosed. In contrast, transparency may be a better design when the mechanism is followed by a supermodular aftermarket, assuming that the designer benefits from releasing more information.[1] When there are more agents in the mechanism, and the designer runs an auction (this case is handled in Appendix A), the conclusion is that in submodular aftermarket only information about the second highest bid should be disclosed, while in supermodular aftermarket, the auctioneer can disclose information about the winning bid directly, without upsetting existence of a separating equilibrium.

[1] This finding complements the main result of Calzolari and Pavan (2006b) who also investigate whether privacy is optimal in bilateral contracting using a different setting and a set of conditions focusing on other aspects of the problem.

Other papers studied the consequences of post-mechanism interactions with third parties for the design of transparency in auctions and other allocation mechanisms. Progress has been made under the assumption that the aftermarket has a special structure under which full disclosure is feasible. Most notably, Zhong (2002), Goeree (2003), Das Varma (2003), Katzman and Rhodes-Kropf (2008), and Hu and Zhang (2017) study the effects of different bid disclosure policies on revenue in auctions followed by Bertrand, Cournot, or other forms of competition.[2,3] The special structure guarantees existence of separating equilibria even in transparent auctions, and these papers point out that no such separating equilibrium exists if the assumption is relaxed. I generalize the relevant condition and identify it as supermodularity of the aftermarket.[4] Submodular aftermarkets are much more difficult to study, precisely because they do not allow to support full disclosure in the mechanism. Some progress has been made on resale aftermarkets but under restrictive assumptions on the type space and other features of the environment: For example, relying on a binary type space, Calzolari and Pavan (2006a) derive the optimal mechanism for the case when the agent gets an offer from a third party in the aftermarket, and Engelbrecht-Wiggans and Kahn (1991) construct non-monotone equilibria of an auction followed by resale.

In relation to this literature, this paper is the first to solve for the optimal Bayesian mechanism allowing for both supermodular and submodular aftermarkets. By doing so, I uncover the fundamental distinction between these two cases, and explain the previous results in a single framework. Section 5.3 elaborates on this point further, and Section 6.2 discusses in detail the relationship to Calzolari and Pavan (2006a). The approach of this paper is complementary to the approach of the companion paper Dworczak (2019) which does not impose as many simpliying assumptions but restricts attention to cutoff mechanisms. Here, I restrict the setting (by assuming a linear utility for the agent and a binary action for a single third party) but derive the optimal Bayesian mechanism. This allows me to identify cases when restricting attention to cutoff mechanisms is without loss of optimality. The relevant condition is the submodularity of the aftermarket.[5]

[2] See also Giovannoni and Makris (2014) who assume that bidders have reduced-form reputational concerns, and Back, Liu and Teguia (2018) who examine the impact of transparency on efficiency and dealers' revenues in financial over-the-counter markets.

[3] With the exception of Hu and Zhang, 2017 who also employ a mechanism design approach, this literature compares a small number of auction formats (e.g. first-price, second-price) and disclosure rules (e.g. full revelation of bids, revelation of the winning bid).

[4] For example, Assumption 2 in Goeree (2003) is essentially supermodularity of the aftermarket but for the case of degenerate beliefs (fully revealing signals).

[5] In Dworczak (2019), I show that submodularity of the aftermarket implies that cutoff mechanisms

An important related problem is when bidders interact *with each other* after the auction. If the aftermarket is modeled as a general game, the problem is difficult to study because agents in the first-stage mechanism consider not only the signaling effect of their behavior (which is present already when they interact with a third party), but also how much they learn about others. An easier special case is when the aftermarket stems from the inability of the auctioneer to prohibit inter-bidder resale (as in Gupta and Lebrun, 1999, Zheng, 2002, Haile, 2003, Hafalir and Krishna, 2008, Hafalir and Krishna, 2009, Zhang and Wang, 2013, Carroll and Segal, 2018).[6] In this case, the focus is on whether the designer can implement the optimal allocation rule, and the disclosure rule typically plays a secondary role in the analysis (e.g., it is beyond the control of the designer). Balzer and Schneider (2017) analyze a different case in which players engage in a costly conflict if agreement is not reached in the mechanism.

The presence of an aftermarket motivates studying information disclosure *after* the mechanism. There is a large literature on information disclosure *before* the auction (when the information is contorolled by the designer and released to bidders to refine the estimates of their values) or *during* the auction (with correlated values, observing how much other bidders bid is a signal about the common value compenent of the good). Notable examples include Milgrom and Weber (1982), Eső and Szentes (2007), Lauermann and Virág (2012), Bergemann and Wambach (2015), Li and Shi (2017), and more recently Smolin (2019).

Finally, aftermarkets create allocative and informational externalities in mechanisms which have been studied by Jehiel, Moldovanu and Stacchetti (1996) and Jehiel and Moldovanu (2001, 2006).

The rest of the paper is organized as follows. In Section 2, I introduce the model. Although I study a version with one agent, all results extend to the multiple-agent case which is formally considered in Appendix A. In Section 3, I present the auxiliary result assuming that the allocation rule is constant. Section 4 contains the main results on optimality of cutoff and partitional mechanisms, and Section 5 analyzes a few applications. In Section 6, I discuss cases not covered by the assumptions of the main results. In particular, I show that the optimal allocation rule can sometimes be non-monotone, and derive robust payoff bounds from using cutoff mechanisms.

are the only feasible class of mechanisms under additional regularity conditions and a strong notion of implementation.

[6] Calzolari and Pavan (2006a) consider an extension of their model to inter-bidder resale but they do not derive the optimal mechanism in that framework.

2 The baseline model

A *mechanism designer* is a seller who chooses a mechanism to sell an indivisible object to an *agent*. The agent has a private type $\theta \in \Theta$. I normalize $\Theta = [0, 1]$, and assume that θ is distributed according to a continuous full-support distribution F with density f. If the agent acquires the good, she participates in the aftermarket which is an interaction with a *third party*. The mechanism designer cannot contract with the third party, neither can she design the aftermarket game.

The market game consists of two stages: *(1)* implementation of the mechanism, and *(2)* post-mechanism interaction between the agent and the third party (the aftermarket). In the first stage, the seller chooses and publicly announces a direct mechanism (x, π, t), where $x : \Theta \to [0, 1]$ is an allocation function, $\pi : \Theta \to \Delta(\mathcal{S})$ is a signal function with some finite signal set $\mathcal{S}$, and $t : \Theta \to \mathbb{R}$ is a transfer function.[7] If the agent reports $\hat{\theta}$, she receives the good with probability $x(\hat{\theta})$ and pays $t(\hat{\theta})$. Conditional on selling the good, the designer draws and publicly announces a signal $s \in \mathcal{S}$ according to distribution $\pi(\cdot|\,\hat{\theta})$. For technical reasons, I assume that $x(\theta)$ must be right-continuous, and $\pi(s|\,\theta)$ is measurable in θ. I call (x, π) a *mechanism frame*.

In the second stage, which only takes place if the good was allocated, the third party observes the signal realization s, and Bayes-updates her beliefs. I let F^s denote the cdf of the updated belief over the agent's type (conditional on the event that the agent acquired the good and signal s was observed). The third party then takes a binary decision $a \in \{l, h\}$ to maximize the expectation of an upper semi-continuous function $v^a(\theta) : \Theta \to \mathbb{R}$,

$$a^\star(F^s) = \underset{a\in\{l,h\}}{\operatorname{argmax}} \int_0^1 v^a(\theta) dF^s(\theta).$$

When the third party is indifferent, it is assumed that the selection from the argmax correspondence can be made by the designer.

The agent's payoff, net of transfers, conditional on acquiring the good in the mechanism is given by

$$U^a(\theta) = u^a\,\theta + c^a, \tag{2.1}$$

where $u^a \geq 0$ and c^a are constants. Linearity of $U^a(\theta)$ in θ is assumed for tractability. The important assumption is that the action of the third party influences the slope of the utility function of the agent. The agent's payoff is normalized to zero if the agent

[7] Using a direct mechanism is without loss of generality, by the Revelation Principle, see for example Myerson (1982).

does not acquire the good. The agent's final utility is quasi-linear in transfers, and the agent is an expected-utility maximizer.

Although very restrictive, this form of the utility function allows me to capture a variety of different economic application. We can interpret θ as the probability with which the agent has an ex-post high type $\theta_H > 0$. With complementary probability $1-\theta$, the agent has a low type $\theta_L \in (0, \theta_H)$. The agent learns the ex-post type only after acquiring the object (but before the aftermarket game). If the payoffs of all players in the aftermarket only depend on the winner's ex-post type, the ex-ante expected utility of the agent takes the form (2.1). I will refer to this interpretation as the ex-post binary-type model.

The mechanism designer's ex-post utility is given by the function $V^a(\theta)$ if the good is allocated and the third party takes action a. If the good is not allocated, the payoff is normalized to zero. For clarity of exposition, I assume that the designer always weakly prefers the third party to take the high action, i.e., $V^h(\theta) \geq V^l(\theta)$, for all $\theta \in \Theta$. To avoid uninteresting cases, I also assume that when the good is always allocated and no signal is sent, the third party will not choose the high action:

$$\int_0^1 (v^h(\theta) - v^l(\theta))f(\theta)d\theta < 0. \tag{2.2}$$

To guarantee existence of solutions to all problems considered in this paper, I assume that $V^h(\theta)$, $V^l(\theta)$, and $V^h(\theta)-V^l(\theta)$ are all upper semi-continuous in θ. The mechanism designer maximizes expected utility.[8]

To avoid trivial cases, unless explicitly stated, I assume that $u^h \neq u^l$, and that $v^h(\theta) - v^l(\theta)$ takes both strictly positive and strictly negative values, on sets of points of non-zero measure.

2.1 Aftermarkets

In this subsection, I introduce a classification of aftermarkets, and give examples.

Definition 1. A function $\Phi : \{l, h\} \times \Theta \to \mathbb{R}$ is *submodular* if $\Phi^h(\theta) - \Phi^l(\theta)$ is non-increasing in θ, and is *supermodular* if $\Phi^h(\theta) - \Phi^l(\theta)$ is non-decreasing in θ.

[8] The designer's utility does not depend explicitly on transfers but this is without loss of generality: By the payoff equivalence theorem, expected transfers in any incentive-compatible mechanism are uniquely pinned down (up to a constant) by the mechanism frame, and therefore the setting allows for objectives such as expected revenue maximization.

Definition 2. The preferences of two players are *co-modular* if their respective utility functions are either both submodular or both supermodular. If one of the utility functions is supermodular and the other one is submodular, the preferences of the two players are called *counter-modular*.

We can now define the key condition on the aftermarkets.

Definition 3. I will say that the aftermarket is *submodular* if the agent's and the third party's utility functions are counter-modular; The aftermarket is *supermodular* if the agent's and the third party's utility functions are co-modular.

The distinction between submodular and supermodular aftermarkets has been introduced in a companion paper Dworczak (2019). To understand that terminology, let $\hat{U}(\theta,\, F) = U^{a^\star(F)}(\theta)$ be the expected aftermarket utility of an agent with type θ when the belief over her type is given by cdf F. If the aftermarket is submodular, then $\hat{U}(\theta,\, F)$ is submodular in $(\theta,\, F)$ with the usual order on Θ and the first-order stochastic dominance order on $\Delta(\Theta)$. If the aftermarket is supermodular, then $\hat{U}(\theta,\, F)$ is supermodular in $(\theta,\, F)$.[9]

Under a submodular aftermarket, lower types of the agent have a higher willingness to pay for high beliefs (in the first-order stochastic dominance order). Supermodularity implies the opposite relationship: High types value higher beliefs more than low types (in relative terms). This property of the aftermarket turns out to be key for understanding which mechanisms are feasible: Since the mechanism sends informative signals about the agent's type, it is as if it allocated posterior beliefs as "goods". Thus, sub- or supermodularity of the aftermarket decides about the direction of the single-crossing condition.

For many results, I will need a stronger definition that additionally requires the agent's utility to be strongly sub- or supermodular.

Definition 4. The function $U^a(\theta) = u^a\,\theta + c^a$ is *strongly submodular* if $u^h = 0$ and $u^l > 0$, and is *strongly supermodular* if $u^h > 0$ and $u_l = 0$.

Strong submodularity implies submodularity which only requires that $u^l > u^h$.

Definition 5. The aftermarket is strongly submodular (supermodular) if it is submodular (supermodular) and the agent's utility is either strongly submodular or strongly supermodular.

[9] In Dworczak (2019), I use the stronger monotone likelihood ratio order on $\Delta(\Theta)$.

I conclude with some examples of aftermarkets.

Example 1. In cases (a) and (b) below, I explore the ex-post binary-type model: θ is the ex-ante probability of having a high ex-post type θ_H; with probability $1-\theta$, the ex-post type is low, θ_L. The agent learns her ex-post type after acquiring the object in the mechanism.

(a) [Resale][10] The ex-post type is the value for the good, with $\theta_H > \theta_L > 0$. The third party has full bargaining power and values the good at $v > \theta_H$. Without loss, the price offered by the third party is either high (θ_H) or low (θ_L). The aftermarket takes place with probability $\lambda > 0$. This corresponds to $u^h = (1-\lambda)(\theta_H - \theta_L)$, $c^h = \lambda\theta_H + (1-\lambda)\theta_L$, $u^l = \theta_H - \theta_L$, $c^l = \theta_L$ for the agent, and $v^h(\theta) = v - \theta_H$, $v^l(\theta) = (1-\theta)(v-\theta_L)$ for the third party.[11]

(b) [Acquiring a complementary good] After the agent acquires the object in the mechanism, she can buy a complementary good from a monopolist (third party) who produces at zero marginal cost. The combined value for holding both objects is high (θ_H) or low (θ_L). If the agent fails to acquire the complementary good in the aftermarket, she gets a liquidation value $r \geq 0$. It is assumed that $r < \theta_L < \theta_H$. The monopolist quotes an optimal monopoly price which is either $\theta_L - r$ or $\theta_H - r$. Because I adopted the convention that the designer always prefers the high action of the third party, I label $\theta_L - r$ the high action (leading to high volume of trade in the aftermarket) and $\theta_H - r$ the low action. Then, we have $u^h = \theta_H - \theta_L$, $c^h = r$, $u^l = 0$, $c^l = r$ for the agent and $v^h(\theta) = \theta_L - r$, $v^l = \theta(\theta_H - r)$ for the third party.

(c) [Investment game] The first stage is an acquisition of a company by a private equity firm (agent). In the second stage, the agent raises additional capital from a potential new investor (the third party). If capital is raised, the expected profit of the company is given by $\bar{\sigma}_0 + \bar{\sigma}_1\theta$; otherwise it is $\underline{\sigma}_0 + \underline{\sigma}_1\theta$, with $\bar{\sigma}_1 \geq \underline{\sigma}_1 \geq 0$, $\bar{\sigma}_0 \geq 0$, and $\bar{\sigma}_0 \geq \underline{\sigma}_0$. If the third party invests, she gets a share $\alpha \in (0, 1 - \underline{\sigma}_1/\bar{\sigma}_1)$ of the company's profits; if she doesn't, she gets an outside option worth $R > 0$. We have $u^h = (1-\alpha)\bar{\sigma}_1$, $c^h = (1-\alpha)\bar{\sigma}_0$, $u^l = \underline{\sigma}_1$, $c^l = \underline{\sigma}_0$ for the agent, and $v^h(\theta) = \alpha(\bar{\sigma}_0 + \bar{\sigma}_1\theta)$, $v^l(\theta) = R$ for the third party.

(d) [Entry game] A regulator decides whether to allocate a license to a start-up offering a new service. Granting the license has an opportunity cost c to the regulator

[10] This example is also analyzed in Dworczak (2019), and it is similar to the baseline model of Calzolari and Pavan (2006a).

[11] The third party's expected utility is defined conditionally on the aftermarket taking place.

but the new service (if offered) generates a social surplus of $\pi > c$. The start-up will serve the market with probability θ if the permission is granted. An incumbent in this market (the third party) decides whether to offer a similar service at a cost $k < \pi$ conditional on the license being granted. If only one player offers the new service, she extracts all the surplus but if both players offer the service, the entire surplus is captured by consumers.[12] In this setting, $u^h = c^h = 0$, $u^l = \pi$, $c^l = 0$, and $v^h(\theta) = (1-\theta)\pi - k$, $v^l(\theta) = 0$.

In Example 1(a), the agent's utility is submodular (low types benefit more from high resale prices) while the third party's utility is supermodular (the third party quotes a high price when she believes that the agent's type is high), making the aftermarket submodular. The aftermarket is strongly submodular if it takes place with probability one ($\lambda = 1$). The aftermarket in Example 1(b) is also strongly submodular, although for a different reason: Here, the agent's utility is supermodular (high types benefit more from the high action – low price – in the aftermarket) but the third party's utility is submodular (the third party chooses the high action – low price – when she believes the type of the agent to be low). The investment game from Example 1(c) is supermodular – both the agent's and the third party's utility functions are supermodular. It is strictly supermodular when $\underline{\sigma}_1 = 0$, that is, when the profit of the company does not depend on the agent's type when additional capital is not raised. Finally, the entry game from Example 1(d) induces a strongly supermodular aftermarket: Both the agent's and the third party's utility functions are submodular. Note that sub- and supermodularity of the aftermarket has nothing to do with whether the agent prefers to signal a high or a low type. In cases (a) and (c), the agent benefits from inducing a high belief about her type, while in cases (b) and (d) – from inducing a low belief. What matters is how the willingness to pay for high beliefs changes with the type.

2.2 Implementability

Because the action of the third party is binary, it is without loss of generality to assume that the signal sent by the optimal mechanism is also binary (see for example Myerson, 1982). Moreover, the signal can be labeled by the action that it induces, that is, $\mathcal{S} = \{l, h\}$. I denote $\pi_h(\theta) = \pi(h|\,\theta)$. From now on, a mechanism frame is represented by the pair (x, π_h), where $x(\theta)$ is the allocation rule, and $\pi_h(\theta)$ is the conditional

[12] While this is obviously stylized, one can think of Bertrand-style competition in a market with perfect price discrimination.

probability of recommending the high action conditional on type θ acquiring the good in the mechanism.

Definition 6. A mechanism frame (x, π_h) is *implementable* if there exist transfers t such that the agent participates and reports truthfully in the first-stage mechanism, taking into account the continuation payoff from the aftermarket,

$$U^h(\theta)\pi_h(\theta)x(\theta) + U^l(\theta)(1 - \pi_h(\theta))x(\theta) - t(\theta) \geq 0, \qquad (IR)$$

$$\theta \in \underset{\hat{\theta}}{\operatorname{argmax}}\{U^h(\theta)\pi_h(\hat{\theta})x(\hat{\theta}) + U^l(\theta)(1 - \pi_h(\hat{\theta}))x(\hat{\theta}) - t(\hat{\theta})\}, \qquad (IC)$$

for all $\theta \in \Theta$, and the third party obeys the recommendations,

$$\int_0^1 [v^h(\theta) - v^l(\theta)]\pi_h(\theta)x(\theta)f(\theta)d\theta \geq 0, \qquad (OB_h)$$

$$\int_0^1 [v^h(\theta) - v^l(\theta)](1 - \pi_h(\theta))x(\theta)f(\theta)d\theta \leq 0. \qquad (OB_l)$$

The tractability of the binary model stems from the fact that the (IR) and (IC) constraints admit a simple representation (which is not typically the case in the presence of an aftermarket). By a standard argument (see Myerson, 1981), it can be shown that there exists a transfer function t such that (IR) and (IC) are satisfied if and only if

$$\pi_h(\theta)x(\theta)u^h + (1 - \pi_h(\theta))x(\theta)u^l \text{ is non-decreasing in } \theta. \qquad (M)$$

Fact 1. *A mechanism frame (x, π_h) is implementable if and only if conditions (M), (OB_h), and (OB_l) hold.*

An interesting consequence of Fact 1 is that a decreasing allocation rule x may be implemented (in contrast to the classical setting of Myerson, 1981) if information is disclosed in an appropriate way. This is possible whenever $u^h \neq u^l$, that is, whenever the agent's types have differential preferences over the actions of the third party in the aftermarket. For example when $u^h = 0$ and $u^l > 0$, implementability is equivalent to monotonicity of $(1 - \pi_h(\theta))x(\theta)$, and therefore $x(\theta)$ can be decreasing as long as this is offset by a sufficient increase in $1 - \pi_h(\theta)$, the probability of sending the low signal. This is further explored in Section 6.1 which shows that a non-monotone allocation rule may be chosen in the optimal mechanism.

I say that an implementable (x, π_h) *reveals no information* if π_h is constant. That is, there is no correlation between the agent's type and the action taken by the third party. Two comments are in place. First, different no-information-revealing mechanisms may induce different actions taken by the third party in the aftermarket because the third party conditions on the event that the agent acquired the good. If x is non-constant, the posterior belief will be different from the prior belief f even if the mechanism does not send any explicit signals. Second, I focus on payoff relevant consequences of information disclosure: A mechanism could send signals that are informative about the agent's type but are payoff irrelevant in that they don't change the action taken by the third party – I classify such mechanisms as not revealing any information.

3 Pure information intermediation

In a general mechanism, the designer screens the agent's type by exploiting the differences in values that different types have for *(i)* the acquisition of the good, and *(ii)* different posteriors beliefs in the aftermarket. In this section, I study an auxiliary problem by shutting down the first of these channels: I assume that the allocation rule x implemented by the mechanism is constant. This means that the designer engages in *pure information intermediation*: she elicits reports from the agent and sends messages to the third party under commitment, without the possibility to influence the allocation. Thus, incentives in the mechanism must come from differences in valuations for posterior beliefs across the agent's types. This exercise allows me to better understand the role of sub- and supermodularity of the aftermarket.

Theorem 1. *Suppose that $x(\theta) = 1$ for all θ. If the aftermarket is submodular, then all implementable mechanism frames (x, π_h) reveal no information.*

If the aftermarket is supermodular, there exist mechanism frames (x, π_h) that reveal information. In particular, full disclosure of the agent's type is implementable.

The theorem implies that if the designer attempts to send non-redundant signals in a mechanism followed by a submodular aftermarket, the agent will misreport making the signals uninformative. The only mechanisms consistent with truth-telling are ones that reveal no information about the agent's type that could influence the action taken by the third party. In contrast, supermodular aftermarket impose relatively few constraints on the informativeness of signals in the mechanism. In particular, full disclosure of the agent's type is consistent with incentive-compatibility (and influences the action taken by the third party).

To gain intuition for Theorem 1 in the submodular case, consider the scenario when the agent's utility is submodular (then, the third party's utility is supermodular), such as in a resale aftermarket. Low types of the agent have a higher willingness to pay for signals that lead to a high action of the third party (low types benefit relatively more from resale at a high resale price). However, the third party takes a high action only when she believes that the agent's type is high, that is, after seeing a signal that is chosen more often by high types. We get a contradiction: If the designer sets a relatively high price for a signal that leads to a high action, then only low types want to choose that signal but then that signal cannot lead to a high price. And if the price for the high signal is relatively low, then all types choose it, and hence the signal is uninformative. In the opposite case when the agent's utility is supermodular, low types have a higher willingness to pay for signals that lead to a low action, but a low action is only taken when the third party believes the type of the agent to be high (the third party's utility is submodular).

The contradiction can be seen at a more abstract level by considering posterior beliefs as "goods" allocated by the mechanism. I will informally refer to a belief being "high" if it puts relatively more mass on high types.[13] There is a conflict between Bayes rule and incentive-compatibility: Bayes rule implies that higher types must be (at least on average) associated with higher beliefs; Incentive-compatibility in a submodular aftermarket implies that higher beliefs must be "allocated" to lower types.

The above conflict disappears in a supermodular aftermarket which reverses the direction of single crossing between types and beliefs: In incentive-compatible mechanisms, higher beliefs must be allocated to higher types. Because this is consistent with Bayes rule, the mechanism can disclose very precise information about the agent's type, up to full disclosure.[14]

Theorem 1 extends to non-constant allocation rules that take the form of a threshold rule, $x(\theta) = \mathbf{1}_{\{\theta \geq \bar{\theta}\}}$. That is, when the aftermarket is submodular, a threshold allocation rule cannot be accompanied by an informative disclosure rule. However, the result does not extend any further: as will be demonstrated soon, information disclosure is possible

[13] This could be formalized using the first-order stochastic dominance order.

[14] Crawford and Sobel (1982) study a related problem but with cheap talk. That is, in their model, communication cannot be supported with transfers, and no player has commitment power. Their results on (im)possibility of communication have similar intuition.

for any non-threshold allocation rule even if the aftermarket is submodular. Intuitively, this is because the designer can use differences in probabilities of obtaining the good to screen types.

4 Optimal mechanisms

4.1 Cutoff mechanisms and partitional mechanisms

Before stating the main results on optimality, I define two classes of mechanisms that will play a key role in subsequent analysis.

Cutoff mechanisms were defined and discussed in the companion paper Dworczak (2019). I restate their definition in the context of the binary model.

Definition 7. A mechanism frame $(x,\, \pi_h)$ is a *cutoff rule* if x is non-decreasing, and the signal function π_h can be represented as

$$\pi_h(\theta)x(\theta) = \int_0^\theta \gamma(c)dx(c), \tag{4.1}$$

$$(1-\pi_h(\theta))x(\theta) = \int_0^\theta (1-\gamma(c))dx(c), \tag{4.2}$$

for almost all $\theta \in \Theta$, for some measurable signal function $\gamma:\, \Theta \to [0,\, 1]$, where $dx(\cdot)$ denotes the measure induced by the non-decreasing right-continuous function $x(\theta)$.

Definition 7 has the following interpretation. Any non-decreasing allocation rule x can be extended to a cumulative distribution function on the type space Θ.[15] (I will abuse notation slightly by using the same symbol x to denote that cdf.) Let c_x be a random variable, called the *cutoff*, with realizations in Θ and a distribution given by the cdf x. Then, the allocation rule x can be implemented by drawing a realization c of the random cutoff c_x, and giving the good to the agent if and only if the reported type exceeds c. Indeed, we have

$$\mathbb{P}(\text{type } \theta \text{ gets the good}) = \mathbb{P}(\theta \geq c_x) = x(\theta).$$

In a cutoff mechanism, the signal distribution is determined (through the function γ) by the realization of the random cutoff c_x representing the allocation rule x. That

[15] This is because x was assumed to be right-continuous.

is, the agent reports a type $\hat{\theta}$, the designer draws a cutoff c, the agent gets the good if and only if $\hat{\theta} \geq c$, and then the designer draws and announce a signal s whose distribution depends only on c. Because the third party takes a binary action, it is without loss of generality to focus on binary signals. Therefore, the signal function γ is one-dimensional: $\gamma(c)$ is the probability of recommending the high action conditional on cutoff realization c.

Conditional on the cutoff c_x, the signal from a cutoff rule is independent of the agent's type. However, because the third party conditions on the event $\theta \geq c_x$, the signal realization is informative about the agent's type if it refines the third party's belief about c_x. Thus, informative signals can be sent by a cutoff rule when the distribution of c_x is non-degenerate (not deterministic). A non-degenerate distribution of c_x is equivalent to the allocation rule x being a non-threshold rule.

An equivalent definition of a cutoff rule can be stated as follows.

Fact 2. *A mechanism frame* (x, π_h) *is a cutoff rule if and only if* $x(\theta)\pi_h(\theta)$ *and* $x(\theta)(1-\pi_h(\theta))$ *are both non-decreasing in* θ.

One direction of the equivalence is immediate: If a cutoff rule satisfies conditions (4.1)-(4.2), then it clearly satisfies the monotonicity conditions from Fact 2. The other direction can be shown using the Radon-Nikodym theorem.[16]

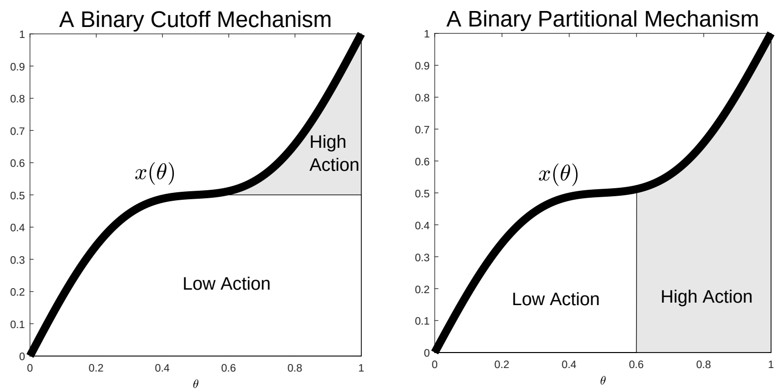
Fact 2 has an immediate corollary: Cutoff rules always satisfy incentive-compatibility constraints (with appropriately chosen transfers).

Fact 3. *Any cutoff rule satisfies the monotonicity condition* (M).

By Fact 1, a cutoff rule is implementable if and only if it satisfies the obedience constraints (OB_h) and (OB_l). In other words, any information disclosure about the cutoff is consistent with incentive-compatibility constraints. In particular, as long as the allocation rule is not a threshold rule (so that the distribution of the cutoff is non-degenerate), information about the agent's type can be disclosed regardless of the aftermarket. Intuitively, a cutoff rule discloses information that separates the agent's types $\theta > \hat{\theta}$ in proportion to the difference in their allocation probabilities $x(\theta)$ and $x(\hat{\theta})$. Note that when $x(\theta) = x(\hat{\theta})$, the types θ and $\hat{\theta}$ must receive exactly the same signal distributions in a cutoff rule, and hence the same distribution of actions of the third party (this can be seen from either (4.1) - (4.2) or Fact 2). When $\Delta = x(\theta) - x(\hat{\theta}) > 0$, there is a probability Δ that a cutoff realization will lie between θ and $\hat{\theta}$, and in this

[16] See Dworczak (2019) for a formal proof.

Fig. 4.1: A Cutoff Mechanisms versus a Partitional Mechanism: The shaded area under the allocation rule represents the probability of sending a recommendation conditional on a certain type (third party's utility is supermodular)



case the mechanism can send a signal that will separate types θ and $\hat{\theta}$ (e.g., by fully disclosing the cutoff realization, the mechanism induces a posterior belief that truncates the prior above the lower type $\hat{\theta}$ but below the higher type θ).

In contrast to cutoff mechanisms, *partitional mechanisms* reveal information about the agent's type directly.

Definition 8. A mechanism frame (x, π_h) is a *partitional rule* if the signal is a deterministic function of the type: $\pi_h(\theta) \in \{0, 1\}$, for all $\theta \in \Theta$.

A partitional mechanism defines a binary partition of the type space, with the two sets in the partition corresponding to the two possible actions of the third party. In an implementable partitional mechanism, the partition will consist of two intervals. A partitional mechanism is typically not a cutoff mechanism, and vice versa (the only exceptions are boundary cases, for example, a mechanism that makes no announcements is both a partitional and a cutoff mechanism). Figure 4.1 contrasts a typical shape of a cutoff and a partitional mechanism.

The following fact (whose proof is straightforward and thus skipped) confirms the intuition developed in Section 3.

Fact 4. *The only partitional mechanisms that are implementable for a strongly submodular aftermarket are ones that reveal no information.*

When the aftermarket is supermodular, there are many partitional rules that are implementable: For example, full disclosure (which induces a binary monotone partition of the type space in terms of π_h) is a partitional mechanism, and is implementable when the allocation rule x is non-decreasing.

4.2 Main results

Under the assumptions of Section 2, the mechanism designer solves

$$\max_{x,\,\pi_h} \int_0^1 \left(V^h(\theta)\pi_h(\theta)x(\theta) + V^l(\theta)(1-\pi_h(\theta))x(\theta)\right) f(\theta)d\theta \tag{4.3}$$

subject to (M), (OB_h), and (OB_l).

I introduce optimal mechanisms in two steps. In the first step, I consider a problem in which the allocation rule x is fixed, and the designer optimizes over disclosure rules (the crucial difference to Section 3 is that x is not necessarily constant in the type). In the second step, I consider joint optimization.

4.2.1 Optimal information disclosure

Fix a non-decreasing allocation rule x and consider optimizing (4.3) only over disclosure rules.[17] The problem can be simplified to

$$\max_{\pi_h} \int_0^1 \left(V^h(\theta) - V^l(\theta)\right) \pi_h(\theta)x(\theta)f(\theta)d\theta. \tag{4.4}$$

Theorem 2. *Fix a non-decreasing allocation rule x, and suppose that the designer's and the third party's preferences are co-modular. The following mechanism achieves the maximal value of (4.4) subject to (M), (OB_h), and (OB_l):*

- *A cutoff rule, when the aftermarket is strongly submodular;*
- *A partitional rule, when the aftermarket is supermodular.*

[17] The results of this section easily generalize to the case of allocation rules that are not non-decreasing. The optimal solution in this case can be obtained by replacing $x(\theta)$ in the monotonicity constraint (derived from condition M) by its lower monotone envelope, denoted $\underline{x}(\theta)$, and defined as

$$\underline{x}(\theta) = \sup\{\chi(\theta) : \chi(\theta) \leq x(\theta), \forall\theta, \chi \text{ is non-decreasing}\}.$$

Theorem 2 makes a strong prediction about the form of the optimal disclosure rule depending on the form of the aftermarket: In (strongly) submodular aftermarkets, the optimal mechanism discloses information about the cutoff, while in supermodular aftermarkets, the optimal mechanism discloses information about the type of the agent directly. The intuition follows from the results of Section 3: Submodular aftermarkets make information disclosure difficult; however, information about the cutoff can always be revealed (Fact 3); thus, disclosing the cutoff is enough to achieve the optimum. In contrast, supermodular aftermarkets make information disclosure easy: thus, the designer can send very informative signals to induce the optimal distribution of actions in the aftermarket.

Theorem 2 makes two assumptions that are worth discussing. First, it is assumed that the third party's and the designer's utility functions are co-modular (either both supermodular, or both submodular). This condition is not essential but significantly simplifies the problem: It is possible to improve any suboptimal mechanism by shifting probability mass under π_h in the direction that increases the objective function (4.4) while preserving the obedience constraints. Formally, I prove the theorem by defining an order (similar to first-order stochastic dominance) on the set of feasible disclosure rules, and by arguing that the optimal mechanism must be a maximal point of that order. When the preferences of the designer and the third party are not co-modular, this simple approach does not work.[18] In Section 5, I analyze several optimization problems in which this assumption is violated but a cutoff/ partitional mechanism is nevertheless optimal. Second, I assume that the aftermarket is *strongly* submodular. The theorem would not be true if I only assumed that the aftermarket is submodular but I argue in Section 6.2 that cutoff mechanisms are likely to remain *approximately* optimal when the strong version of submodularity is replaced with the weaker version. The reason for the asymmetry in the statement of the theorem (I did not assume *strong* supermodularity in the second part) is that in supermodular aftermarkets the monotonicity constraint (M) turns out to be slack. In contrast, constraint (M) binds in submodular aftermarkets, and thus more structure is needed to obtain a simple optimal mechanism.

[18] Instead, I could employ optimal control techniques but because the monotonicity constraint (M) will typically bind at the optimal solution, the problem is not always tractable.

4.3 Joint optimization

I turn attention to joint optimization over allocation and disclosure rules. First, I provide conditions under which the optimal allocation rule is non-decreasing.

Lemma 1. *If either (i) all player's preferences are supermodular and $V^l(\theta)$ is non-decreasing; or (ii) the agent's utility function is supermodular, the third party's and the designer's preferences are submodular, and $V^l(\theta)$ is non-decreasing; or (iii) the agent's utility function is submodular, the third party's and the designer's preferences are supermodular, and $V^h(\theta)$ is non-decreasing; then the optimal allocation rule is non-decreasing.*

Lemma 1 assumes that the designer's and the third party's utility functions are co-modular, and considers three (out of four possible) cases. The first bullet point in Lemma 1 corresponds to a supermodular aftermarket, while the second and the third to submodular aftermarkets. The conclusion of Lemma 1 does not hold in the fourth case of a supermodular aftermarket but with all player's utility functions being submodular (this structure obtains in the aftermarket from Example 1(d)). In Section 6.1, I construct an example of an optimal partitional mechanism that has a non-monotone allocation rule in this case. Intuitively, this is a mechanism in which the designer excludes high types from trading to induce a high action in the aftermarket (since the third party chooses the high action when she believes the type of the agent to be low). Such a mechanism is incentive-compatible because lower types have a higher willingness to pay for signals that lead to the high action.

Lemma 1 gives sufficient conditions for optimality of a non-decreasing allocation rule, while Theorem 2 derives conditions under which optimal disclosure takes a particular form for any non-decreasing allocation rule. Thus, I can provide the following sufficient conditions for optimality of cutoff and partitional mechanisms.

Theorem 3. *Suppose that the designer's and the third party's preferences are co-modular, and that $V^h(\theta)$ and $V^l(\theta)$ are non-decreasing in θ.*[19] *The following mechanism achieves the maximal value of (4.3) subject to (M), (OB_h), and (OB_l):*

- *A cutoff rule, when the aftermarket is strongly submodular;*
- *A partitional rule, when all players' utility functions are supermodular.*

[19] It is enough that $V^{\bar{a}}(\theta)$ is non-decreasing, where $\bar{a} = l$ when the designer's utility is supermodular, and $\bar{a} = h$ when the designer's utility is submodular.

Proof. The theorem follows directly from Theorem 2 and Lemma 1. □

Examples satisfying the assumptions of Theorem 3 are provided in Section 5 which considers applications of the model.

An optimal partitional rule typically discloses information in the form of a binary signal. However, an optimal cutoff mechanism never discloses information:

Fact 5 (Dworczak (2019))**.** *If the designer is constrained to use a cutoff mechanisms, and optimizes jointly over allocation and disclosure rules, then the optimal mechanism does not reveal any information (in the single-agent model).*

Therefore, Theorem 3 makes a strong prediction about the optimal transparency of the mechanism depending on the structure of the underlying aftermarket: When the aftermarket is strongly submodular, privacy is always optimal. When the aftermarket is supermodular, information disclosure can (and often will be) optimal. This conclusion relies on the single-agent case: In Appendix A, I extend the model to the case of multiple agents and show that Theorem 3 is still true. However, Fact 5 is not. Therefore, the prediction must be appropriately modified.

5 Applications

In this section, I analyze the simple applications introduced by Example 1.

For the designer's objective, I focus on two cases: total surplus maximization, which requires $V^a(\theta) = U^a(\theta) + v^a(\theta)$, and revenue maximization which can be shown to correspond to

$$V^a(\theta) = c^a + u^a J(\theta), \text{ where } J(\theta) = \theta - \frac{1 - F(\theta)}{f(\theta)}, \tag{5.1}$$

is the virtual surplus function which I assume is non-decreasing throughout.

I derive the optimal mechanism explicitly in two cases: for the model from Example 1(a) and the model from Example 1(c). Case (b) is similar to (a) and discussed only informally. Case (d) is different and will be discussed in Section 6.1.

5.1 Resale

In this subsection, I consider the model from Example 1(a): After buying the good in the mechanism, the agent learns her ex-post value which is high (θ_H) with probability

θ and low (θ_L) with probability $1-\theta$, and then receives a take-it-or-leave-it from a third party with value v. I consider the case $\lambda = 1$, that is, when the offer is made with probability one. This means that the aftermarket is strongly submodular (the general case is considered in Section 6.2).

Under the assumption that $V^h(\theta)$ is non-decreasing in θ, let $\underline{\theta}_h$ be the smallest number in [0, 1] such that $V^h(\theta) \geq 0$ for all $\theta \geq \underline{\theta}_h$. To avoid trivial cases, I assume that,

$$\int_{\underline{\theta}_h}^{1} (v^h(\theta) - v^l(\theta)) f(\theta) d\theta < 0. \tag{5.2}$$

If condition (5.2) fails, the optimal mechanism is trivial: The allocation rule $x(\theta) = \mathbf{1}_{\{\theta \geq \underline{\theta}_h\}}$ followed by no disclosure implements the high action of the third party with probability one, and because the designer always prefers the high action, achieves the upper bound on the designer's payoff. Assumption (5.2) implies that this upper bound is not achievable.

5.1.1 Total surplus maximization

Suppose that the mechanism designer maximizes efficiency in the market. When the resale price is high, the total surplus (conditional on allocating the good) is v, and otherwise it is $(1-\theta)v + \theta\theta_H$ because only low-value agents resell for a low price. Thus, $V^h(\theta) = v$, and $V^l(\theta) = (1-\theta)v + \theta\theta_H$, i.e., the designer's utility is supermodular. This means that the designer's and the third party's preferences are co-modular. By Theorem 2, a cutoff mechanism is optimal for any fixed non-decreasing allocation rule. By Theorem 3, because $V^h(\theta)$ is non-decreasing in θ, a cutoff mechanism is also optimal for the joint optimization problem. Moreover, using the proofs of Theorems 2 and 3, I can pin down the exact form of the optimal mechanism.[20]

Claim 1. *Fix a non-decreasing allocation rule x. Suppose that*

$$\int_0^1 (\theta(v - \theta_L) - (\theta_H - \theta_L)) x(\theta) f(\theta) d\theta \geq 0. \tag{5.3}$$

Then, the total-surplus-maximizing mechanism implementing x reveals no information (and induces the high price in the aftermarket).

[20] The proofs of Claims 1 - 2 are skipped – they follow directly from the proofs of Theorems 2 and 3.

If condition (5.3) fails, define $\bar{x}_{res}$ *as the smallest solution to the equation*

$$\int_0^1 (\theta(v-\theta_L) - (\theta_H - \theta_L)) \max\{x(\theta) - \bar{x}_{res}, 0\} f(\theta) d\theta = 0. \tag{5.4}$$

Then, the total-surplus-maximizing mechanism implementing x *has a disclosure rule given by* $\pi_h^\star(\theta) = \max\{1 - \bar{x}_{res}/x(\theta), 0\}$.

When condition (5.3) holds, it is optimal not to send any signals in the mechanism because in the absence of additional information, the third party takes the high action. To understand the case when (5.3) fails, define $\bar{\theta}_{\text{res}}$ by

$$\bar{\theta}_{\text{res}} = \sup\{\theta \in \Theta : x(\theta) \leq \bar{x}_{\text{res}}\}. \tag{5.5}$$

When x is continuous, $\bar{x}_{\text{res}} = x(\bar{\theta}_{\text{res}})$. The optimal mechanism recommends the low action with conditional probability one when the good is allocated to a type θ below $\bar{\theta}_{\text{res}}$, and with conditional probability $\bar{x}_{\text{res}}/x(\theta)$ when the good is allocated to a type θ above $\bar{\theta}_{\text{res}}$. The intuition is that the mechanism has to exclude enough low types from the high signal to induce the high action of the third party. This is possible when $x(\theta)$ is non-constant and non-decreasing: The unconditional probability of sending a high signal for type $\theta \geq \bar{\theta}_{\text{res}}$ is equal to $x(\theta) - \bar{x}_{\text{res}}$, so that higher types have a higher probability of receiving a high price in the aftermarket. The unconditional probability of sending the low signal is constant and (in general) non-zero for all types above $\bar{\theta}_{\text{res}}$. This is necessary to keep the mechanism incentive-compatible. If the highest type received the high price with probability one, low types would deviate and report a high type. The non-zero probability of a low signal provides the necessary separation between low and high types (only when the low signal is sent, high types have a strictly higher value for winning the object).

I now turn to the problem of joint optimization over allocation and disclosure rules.

Claim 2. *One of the two following mechanisms maximizes total surplus:*

(a) $x^\star(\theta) = 1$ *and* $\pi_h^\star(\theta) = 0$, *for all* θ,

(b) $x^\star(\theta) = \mathbf{1}_{\{\theta \geq \theta^\star\}}$, *and* $\pi_h^\star(\theta) = 1$, *where* $\theta^\star$ *is defined by*

$$\mathbb{E}_f[\theta | \, \theta \geq \theta^\star] = \frac{\theta_H - \theta_L}{v - \theta_L}.\text{[21]} \tag{5.6}$$

[21] This is just a special case of the equation $\int_{\theta^\star}^1 (v^h(\theta) - v^l(\theta)) f(\theta) d\theta = 0$.

Mechanism (a) is optimal if and only if $\mathbb{E}_f[\theta] \leq \frac{v}{v-\theta_H} F(\theta^\star)$.

In both cases, the optimal mechanism sends no signals (see Fact 5), so it either always induces the low price (case a) or always induces the high price (case b). If the designer is not trying to affect the default price in the aftermarket (which is low under assumption 5.2), it is optimal to always allocate the object because the objective function is non-negative. To induce the high price, the designer excludes low types from trading: The threshold $\theta^\star$ is chosen so that the third party is exactly indifferent between the high and the low price (and quotes the high price).

5.1.2 Revenue maximization

I now derive the profit-maximizing mechanism. Equation (5.1) implies that revenue-maximization corresponds to $V^h(\theta) = \theta_H$ and $V^l(\theta) = \theta_L + (\theta_H - \theta_L)J(\theta)$. Because the distribution F was assumed to be regular, $V^a(\theta)$ is submodular. This means that the designer's and the third party's preferences are counter-modular, and hence I cannot apply Theorem 2 directly. Intuitively, the reason for the difficulty is that the third party must be persuaded to offer a high price. This requires that the mechanism sends the high recommendation only when the type of the agent is relatively high (on average). On the other hand, the designer benefits more from giving the high resale price to low types (the gap between $V^h(\theta) - V^l(\theta)$ is decreasing in the type). The resulting trade-off makes the problem of finding the optimal mechanism harder.

Under additional regularity assumptions, the problem can be solved by applying optimal control techniques. Let $\bar{J} = \max\{J'(\theta) : \theta \in \Theta\}$, and assume that $\bar{J}$ is well defined and finite.

Claim 3. *Fix a non-decreasing and absolutely continuous allocation rule x.*[22] *Suppose that*

$$\frac{\theta_H - \theta_L}{v - \theta_L} \leq \mathbb{E}[\theta | \, \theta \geq \bar{\theta}_{res}] + \frac{1 - \bar{\theta}_{res}}{\bar{J}}, \tag{5.7}$$

where $\bar{\theta}_{res}$ was defined in (5.5). Then, the highest expected revenue over all mechanisms implementing x can be obtained by a cutoff mechanism. Moreover, in this case, the profit-maximizing mechanism coincides with the welfare-maximizing mechanism from Claim 1. If F is the uniform distribution on $[0, 1]$, *condition* (5.7) *holds.*

[22] It is enough if x is absolutely continuous on $\{\theta \in [0, 1] : x(\theta) > 0\}$, i.e. x can be equal to zero in some initial interval.

Under the assumptions of Claim 3, the revenue-maximizing disclosure rule coincides with the surplus-maximizing disclosure rule. Intuitively, this is because efficient disclosure does not create additional information rents for the agent in the submodular aftermarket – it is the low types that have a higher willingness to pay for the signals that lead to high resale prices.

I now turn to the problem of optimizing jointly over allocation and disclosure rules. This is an optimal control problem with two control variables and a monotonicity constraint, so obtaining a general solution is challenging. I focus on the case in which I can solve the problem by relaxing the monotonicity constraint.[23] I do not have to assume regularity of the distribution F but instead (for technical reasons) I assume that the virtual surplus function $J(\theta)$ is convex.

Claim 4. *Suppose that J is convex. Let $\theta^\star$ be defined by (5.6). If $\theta_L + (\theta_H - \theta_L)J(\theta) \leq 0$ for all $\theta \leq \theta^\star$, then the profit-maximizing mechanism is given by $x^\star(\theta) = \mathbf{1}_{\{\theta \geq \theta^\star\}}$, and $\pi_h^\star(\theta) = 1$.*

The optimal mechanism under the assumption of Claim 4 reveals no information. The allocation rule excludes just enough low types from trading so that a high price is quoted in the aftermarket conditional on trade in the mechanism. This allows the designer to charge a price in the mechanism equal to the resale price, and extract all the surplus from the agent.

For an illustration of what might happen when the assumptions of Claim 4 fail, suppose that F is the uniform distribution on $[0, 1]$, and $\theta_L = 0$. Then, $\theta^\star$ defined by (5.6) is given by $\theta^\star = 2\theta_H/v - 1$ and lies in $(0, 1]$ under assumption (5.2). When $\theta_H/v \leq 3/4$, it is optimal to sell to all types above $\theta^\star$ and reveal no information. When $\theta_H/v > 3/4$, so that $\theta^\star > 1/2$, it becomes more difficult to induce a high price in the aftermarket, in the sense that the mechanism must exclude relatively high types from trading (types with a positive virtual surplus). When $\theta^\star$ is high enough, it becomes optimal not to induce a high price at all – in this case, the optimal mechanism is identical to the one that would arise in the absence of an aftermarket (i.e., the seller sells to all types with non-negative virtual surplus).

The aftermarket from Example 1(b) (acquiring a complementary good) is strongly submodular, and thus the analysis is very similar to the one above. In particular, a cutoff mechanisms maximizes total surplus, and maximizes revenue under additional

[23] In other cases, the monotonicity constraint will bind at the optimal solution, and standard optimal control techniques cannot be applied.

conditions. The nature of the aftermarket is quite different because the agent wants to pretend that her type is low – if she induces a low belief of the third party, she can acquire the complementary good cheaper. Nevertheless, what matters for the properties of optimal mechanisms is not the direction of preferences over beliefs but rather how the willingness to pay for favorable beliefs depends on the type of the agent.

5.2 The investment game

In this subsection, I solve the investment model from Example 1(c). The mechanism designer sells a company to an agent (e.g., a private equity firm) who then hopes to attract an investment from the third party in the aftermarket. The profitability of the company is increased from $\underline{\sigma}_0 + \underline{\sigma}_1\theta$ to $\bar{\sigma}_0 + \bar{\sigma}_1\theta$ when the investment is made, and the investor is rewarded with a fraction $\alpha \in (0, 1 - \underline{\sigma}_1/\bar{\sigma}_1)$ of the company profits (not investing yields an outside option of R).

5.2.1 Total surplus maximization

In the case of maximizing efficiency in the market, the designer's utility function is equal to the total profit of the company: $V^h(\theta) = \bar{\sigma}_0 + \bar{\sigma}_1\theta$ and $V^l(\theta) = \underline{\sigma}_0 + \underline{\sigma}_1\theta$. All players' preferences are supermodular. Moreover, $V^l(\theta)$ is non-decreasing. By Theorem 2, optimal disclosure for a fixed allocation rule takes a partitional form. By Theorem 3, a partitional mechanism with a non-decreasing allocation rule is also optimal overall.

Claim 5. *Fix a non-decreasing allocation rule x. The total-surplus-maximizing mechanism implementing x has a disclosure rule $\pi_h^\star(\theta) = \mathbf{1}_{\{\theta \geq \theta_x^\star\}}$, where $\theta_x^\star$ is defined as*

$$\theta_x^\star = \min\left\{\hat{\theta} : \int_{\hat{\theta}}^1 \left(\alpha(\bar{\sigma}_0 + \bar{\sigma}_1\theta) - R\right) x(\theta) f(\theta) d\theta \geq 0\right\}. \tag{5.8}$$

Claim 6. *Let $\theta_1^\star$ be defined by (5.8) when $x \equiv 1$. The total-surplus-maximizing mechanism is a partitional rule given by $x^\star(\theta) = \mathbf{1}_{\{\theta \geq \theta_{inv}^\star\}}$, where*

$$\theta_{inv}^\star = \min\left\{\theta_1^\star,\ -\frac{\underline{\sigma}_0}{\underline{\sigma}_1}\right\}$$

and has a disclosure rule $\pi_h^\star(\theta) = \mathbf{1}_{\{\theta \geq \theta_{x^\star}^\star\}}$, where $\theta_{x^\star}^\star$ is defined by (5.8).

I omit the proofs of Claims 5 and 6 because they follow immediately from the proofs of Theorems 2 and 3. The optimal mechanisms are very intuitive: Because in a supermodular aftermarket high types have a higher willingness to pay for high beliefs, the mechanism charges types above the threshold $\theta^\star_x$ a higher price for the good in exchange for sending a signal that induces investment in the aftermarket. Types below a threshold pay a discounted price but since this fact is revealed, they do not benefit from the investment in the aftermarket. The surplus-maximizing allocation rule is simply $x^\star \equiv 1$ when the profits of the company are non-negative even without investment and for the lowest type of the agent (that is, when $\underline{\sigma}_0 \geq 0$). In the opposite case, the designer may exclude from trading types $\theta \leq \theta_0 \equiv (-\underline{\sigma}_0/\underline{\sigma}_1)$, where θ_0 is the marginal type that breaks even in the absence of investment. If $\theta_0 \geq \theta^\star_1$, then all types that trade pay the high price and receive investment in the aftermarket.

5.2.2 Revenue maximization

By formula (5.1), maximizing revenue of the designer from selling the company to the agent corresponds to $V^h(\theta) = (1-\alpha)\bar{\sigma}_0 + (1-\alpha)\bar{\sigma}_1 J(\theta)$ and $V^l(\theta) = \underline{\sigma}_0 + \underline{\sigma}_1 J(\theta)$. Thus, because I assumed that $(1-\alpha)\bar{\sigma}_1 \geq \underline{\sigma}_1$, the designer's utility is supermodular – Theorems 2 and 3 imply optimality of partitional rules. Moreover, the proof can be used to deduce the exact form of the optimal mechanism.

Claim 7. *For a fixed non-decreasing allocation rule x, the revenue-maximizing mechanism implementing x coincides with the surplus-maximizing mechanism implementing x (from Claim 5).*

Intuitively, if the allocation is fixed, a revenue-maximizing designer prefers to induce investment as often as possible, because investment in the aftermarket raises the virtual surplus for all types of the agent. The difference in revenue and surplus maximization is reflected in the choice of the allocation rule:

Claim 8. *Let $\underline{\theta}_a$ be the smallest solution to $V^a(\theta) = 0$.*[24] *Let $\theta^\star_1$ be defined by (5.8) when $x \equiv 1$. The revenue-maximizing mechanism is a partitional rule given by $x^\star(\theta) = \mathbf{1}_{\{\theta \geq \theta^\star_{inv}\}}$, where*

$$\theta^\star_{inv} = \min\left\{\max\{\underline{\theta}_h,\, \theta^\star_1\},\, \underline{\theta}_l\right\}$$

and has a disclosure rule $\pi^\star_h(\theta) = \mathbf{1}_{\{\theta \geq \theta^\star_{x^\star}\}}$, where $\theta^\star_{x^\star}$ is defined by (5.8).

The allocation rule in the revenue-maximizing mechanism excludes more types than the surplus-maximizing mechanism. When $\underline{\theta}_l$ is high, so that the seller makes losses

[24] If $V^a(\theta)$ does not cross zero at all, then $\underline{\theta}_a = 0$ if $V^a \geq 0$, and $\underline{\theta}_a = 1$ if $V^a \leq 0$.

when selling to most types of the agent when there is no investment, it is optimal to set a single high price that always induces investment. The threshold type is the larger of $\theta_1^\star$ (which is just enough to induce investment) and $\underline{\theta}_h$ (which is the marginal type with exactly zero virtual surplus for the seller conditional on investment in the aftermarket). When $\underline{\theta}_l$ is low, the mechanism posts two prices, announces which price was paid, and induces investment precisely when the buyer chooses the higher price.

5.3 Comparing submodular and supermodular aftermarkets

Sections 5.1 and 5.2 illustrate the differences between submodular and supermodular aftermarkets. Information is difficult to disclose in a submodular aftermarket: It would not be incentive-compatible to disclose the type of the agent directly because low types could not be disincentiveized from pretending to be high types. As a result, the optimal mechanism only discloses information about the cutoff which provides coarse information about the agent's type. If the allocation and disclosure rules are chosen jointly, it is optimal not to disclose anything (except for the fact that trade took place). In contrast, in a supermodular aftermarket, it is feasible to disclose the type of the agent directly, and the optimal mechanism exploits that by partitioning the type space and charging higher types for the higher signal. Typically, there will be non-trivial information disclosure even when the allocation and disclosure rules are chosen jointly.

A more subtle difference that connects these results to the results from Section 3 is that supermodular aftermarkets allow the mechanism designer to essentially sell the good and the signals separately. In a partitional mechanism, the buyer could be seen as paying a constant price for the good and an extra fee for acquiring the high signal (a signal inducing the high action in the aftermarket). In contrast, a cutoff mechanism crucially relies on "bundling" signals with the allocation probabilities. If the allocation is constant, only a constant (uninformative) signal can be sent (see Theorem 1). When types are separated by different allocation probabilities, they can also receive different signals but only in proportion to these differences: If $\Delta x = x(\theta) - x(\hat{\theta})$, then types θ and $\hat{\theta}$ must receive the same signal with probability $1 - \Delta x$ in a cutoff mechanism. Effectively, the allocation rule can be used to introduce slack in the monotonicity constraint (M) which creates the possibility to send informative signals until the constraint starts binding again.

Due to their tractability, supermodular aftermarkets has been studied quite extensively in the literature (without being formally distinguished from submodular aftermarkets); see for example Zhong (2002), Goeree (2003), Das Varma (2003), Katzman and Rhodes-Kropf (2008), and Hu and Zhang (2017). Consider the case when the aftermarket is a Cournot competition between the agent and the third party. Supermodularity of that aftermarket obtains because lower-cost (high-type) agents benefit more than high-cost (low-type) agents from inducing a belief that they have a low cost. The way these papers use that assumption is by showing that the auction in the first stage has a separating equilibrium even when the winning bids (or all the bids) are disclosed. When the aftermarket is submodular (e.g., if it is a resale aftermarket), a separating equilibrium does not exist in a transparent auction, in line with the above discussion. However, separating equilibria exist if the auction only reveals the cutoff (which corresponds in this case to the second-highest bid). I formally extend my analysis to multi-agent mechanisms in Appendix A.

6 Limits of the main results

6.1 Beyond the case of a non-decreasing allocation rule

In Lemma 1, I did not include the case when the aftermarket is supermodular but all payers' preferences are submodular. The reason is that in this case the optimal allocation rule can be decreasing, even when $V^a(\theta)$ is non-decreasing.

For a simple example, consider case (d) from Example 1: A regulator (designer) decides whether to allocate a license to a start-up (agent) offering a new service with social value π. The type θ is the probability that the entrant will offer the service conditional on getting the license. An incumbent (third party) can also offer the service at cost k. If there is a monopolist provider of the service, that player gets the entire surplus, but if there is a duopoly, Bertrand competition shifts the entire surplus to consumers. Suppose that the regulator puts weight $\alpha \leq 1$ on the entrant's profits, and weight 1 on the joint surplus of the incumbent and the consumers. Then, because granting the license has cost c, we have $V^h(\theta) = \pi - c > 0$ and $V^l(\theta) = \alpha\theta\pi - c$ (note that both $V^a(\theta)$ are non-decreasing in θ). To limit the number of cases, assume that $1 - k/\pi < c/(\alpha\pi)$ (e.g., α is relatively low or k is relatively high). I provide the following claim without a proof because it follows relatively easily from previous proofs.

Claim 9. *The optimal mechanism is partitional:* $x^\star(\theta) = \mathbf{1}_{\{\theta \leq \underline{\theta}^\star\}} + \mathbf{1}_{\{\theta \geq \bar{\theta}^\star\}}$, $\pi_h^\star(\theta) = \mathbf{1}_{\{\theta \leq \underline{\theta}^\star\}}$, *where* $\underline{\theta}^\star = 1 - k/\pi$ *and* $\bar{\theta}^\star = c/(\alpha\pi)$.

The allocation rule is non-monotone because it excludes intermediate types ($\theta \in (\underline{\theta}^\star, \bar{\theta}^\star)$). The intuition behind the result is as follows. If the incumbent believes that the type of the agent is relatively low, she will enter (this gives the regulator her firs-best ex-post). Moreover, low types of the agent are hurt less by the entry decision of the incumbent because they are less likely to enter themselves – it is thus possible to support an outcome in which the low action is implemented when the type of the agent is low. The optimal mechanism allocates the license to types below $\underline{\theta}^\star$ (for free) and reveals that fact to induce entry by the incumbent. Conditional on types above $\underline{\theta}^\star$, the designer knows that she cannot induce entry by the incumbent. Thus, she only wants to allocate the license if the probability of the agent's entry conditional on allocation is high enough to cover the designer's cost c. As a result, she allocates the license to types above $\bar{\theta}^\star$ at a strictly positive fee equal to $\bar{\theta}^\star \pi = c/\alpha$. Intuitively, intermediate types of the agent are excluded because they are high enough to deter the incumbent from entering but too low to cover the cost for the regulator.

Note that when c/α is high (above 1), then because $\Theta \equiv [0, 1]$, the unique optimal allocation rule is decreasing: $x^\star(\theta) = \mathbf{1}_{\{\theta \leq \underline{\theta}^\star\}}$. Without information disclosure, there would always exist an optimal allocation rule that is non-decreasing because the agent's payoff in the aftermarket is non-decreasing in her type regardless of the action taken by the third party.

6.2 Beyond the case of a strongly submodular aftermarket

In this section, I investigate the robustness of the main results to relaxing the assumption that the aftermarket is *strongly* submodular (that is, I only assume that it is submodular). I show that the optimal mechanism may sometimes lie outside of the class of cutoff mechanisms. However, a cutoff mechanism is guaranteed to achieve at least half of the optimal value in the worst case regardless of the objective function of the designer (subject only to some general conditions), and achieves a much better optimality guarantee in more restricted environments.

Throughout, I focus (without loss of generality) on the case when the agent's utility function is submodular. I use the resale model from Example 1(a) as an illustration of the results – note that this aftermarket is always submodular but is strongly submodular only when the resale stage happens with probability $\lambda = 1$.

I assume that $V^l(\theta)$ and $V^h(\theta)$ are both non-decreasing, and that the preferences of the designer and the third party are co-modular. Recall that $\underline{\theta}_a \in [0, 1]$ was defined as the (smallest) point at which $V^a(\theta) = 0$, for $a \in \{l, h\}$.[25] I assume that a slightly stronger version of condition (2.2) holds:

$$\int_{\underline{\theta}_h}^{1} (v^h(\theta) - v^l(\theta)) f(\theta) d\theta < 0. \tag{6.1}$$

If assumption (6.1) fails, the optimal mechanism is trivial (and is a cutoff mechanism): Allocate to all types above $\underline{\theta}_h$ and reveal no information. Similarly as in (5.6), I define $\theta^\star$ as the (smallest) solution to

$$\int_{\theta^\star}^{1} (v^h(\theta) - v^l(\theta)) f(\theta) d\theta = 0. \tag{6.2}$$

That is, when the good is allocated only to types above $\theta^\star$, it is optimal for the third party to choose the high action (assumption 6.1 implies that $\underline{\theta}_h < \theta^\star$).

Proposition 1. *Under the above assumptions, the optimal mechanism takes one of the two possible forms:*

1. $x^\star(\theta) = \mathbf{1}_{\{\theta \geq \underline{\theta}_l\}}$, $\pi^\star_h(\theta) = 0$ *(no information is revealed and the low action is taken by the third party),*

2. *The allocation rule $x^\star(\theta)$ is a two-step function:*

$$x^\star(\theta) = \begin{cases} 0 & \theta < \underline{\theta}_l, \\ \frac{u^h}{u^l} & \underline{\theta}_l \leq \theta < \theta^\star, \\ 1 & \theta^\star \leq \theta, \end{cases}$$

   and $\pi^\star_h(\theta) = \mathbf{1}_{\{\theta \geq \theta^\star\}}$.

Mechanism 1 is optimal if and only if

$$\int_{\underline{\theta}_l}^{1} V^l(\theta) f(\theta) d\theta \geq \frac{u^h}{u^l} \int_{\underline{\theta}_l}^{\theta^\star} V^l(\theta) f(\theta) d\theta + \int_{\theta^\star}^{1} V^h(\theta) f(\theta) d\theta.$$

Mechanisms *1* is a cutoff mechanism but mechanism *2* is not (in fact, it is a partitional mechanism). In mechanism *2*, types below $\underline{\theta}_l$ do not get the object, types in

[25] If $V^a(\theta)$ does not cross zero at all, then $\underline{\theta}_a = 0$ if $V^a \geq 0$, and $\underline{\theta}_a = 1$ if $V^a \leq 0$.

$[\underline{\theta}_l, \theta^\star]$ receive the object with probability $\frac{u^h}{u^l}$ and always face a low action in the aftermarket, and types in $[\theta^\star, 1]$ receive the object with probability 1 and always face a high action in the aftermarket.

To better understand the intuition for optimality of mechanism *2*, it is useful to compare it to the optimal cutoff mechanism. Within the class of cutoff mechanisms, the derivation of the optimal mechanism is straightforward – the proof is omitted.[26]

Claim 10. *Under the above assumptions, the optimal cutoff mechanism is either mechanism 1 from Proposition 1, or mechanism 3:* $x^\star(\theta) = \mathbf{1}_{\{\theta \geq \theta^\star\}}$, $\pi_h^\star(\theta) = 1$ *(no information is revealed and a high action is taken in the aftermarket) with* $\theta^\star$ *defined by* (6.2).

Mechanism *2* from Proposition 1 converges to the mechanism *3* from Claim 10 as $u^h \to 0$. When $u^h = 0$, the aftermarket becomes strongly submodular, and the two mechanisms coincide (which is consistent with Theorem 3 which predicts that a cutoff mechanism is optimal). However, for any u^h, mechanism *2* dominates. Both mechanisms induce a high action for types above $\theta^\star$. It is not possible to induce a high action for types below $\theta^\star$ because, by definition, $\theta^\star$ is the smallest threshold that still induces the high action. However, mechanism *2* allocates to types below $\theta^\star$ and recommends a low action, while mechanism *3* does not allocate the good at all.

To understand the difference, I will use the resale model from Example 1(a). In this case, $u^h/u^l = 1 - \lambda$, where λ is the probability that the resale stage happens. Suppose first that $\lambda = 1$ (the aftermarket is strongly submodular). Both mechanisms allocate to types above $\theta^\star$ and reveal no information. When the aftermarket takes place with probability one, and the high price is always quoted by the third party, the agent's endogenous value for acquiring the object is equal to the resale price θ_H. In particular, the value no longer depends on the agent's type θ. As a consequence, the designer has to charge exactly θ_H for reporting a type above $\theta^\star$, leaving the agent with no information rents (if the agent received a positive information rent, types below $\theta^\star$ would have an incentive to misreport to also receive a positive rent). It follows that it is not incentive-compatible to offer the good and recommend the low price to agents below $\theta^\star$. Otherwise, higher types would have to be offered a strictly positive information rent, contradicting the above reasoning.

Suppose now that $\lambda < 1$. In this case, even when a high resale price is quoted in the aftermarket, higher types of the agent have a strictly higher value for the object (driven

[26] That proof is easy because, by Fact 5, the optimal cutoff mechanism reveals no information. Therefore, the problem reduces to finding the optimal allocation rule in two cases: when the high action is always taken, or when the low action is always taken.

by the event that the aftermarket does not take place and the agent is the final holder of the good). Thus, types in $[\theta^\star, 1]$ receive a positive information rent, proportional to the probability of the aftermarket not taking place, $1-\lambda$. As a consequence, the designer can sell the good to types below $\theta^\star$ but only with sufficiently small probability so that higher types do not want to deviate to reporting a type below $\theta^\star$. In the optimal mechanism *2*, the good is offered exactly with probability $1-\lambda$ (the probability of the aftermarket not taking place) to types below $\theta^\star$. A low price is recommended for these types, so it remains optimal for the third party to quote a high price for types above $\theta^\star$.[27]

Intuitively, the slope u^a of the agent's utility in her type determines the ability of the designer to screen types by offering the good with different probabilities. When the utility is flat (e.g., $u^h = 0$), it is impossible to screen. When the slope is strictly positive ($u^h > 0$), more information can be elicited and revealed by offering different allocation probabilities. With the aftermarket, the slope depends on the action of the third party. When $u^h > 0$, the slope is bounded away from zero regardless of the action taken by the third party ($u^l > u^h$). The optimal mechanism *2* fully exploits this feature to screen the agent's type by offering the good with different probabilities. Cutoff mechanisms, by design, reveal information about the cutoff, and are hence incentive-compatible regardless of the slope of the agent's utility in the aftermarket (Fact 3). However, this also means that cutoff mechanisms do not fully exploit the possibility to screen when the slope of the agent's utility is bounded away from zero.

Example 1(a) is similar to the baseline model considered by Calzolari and Pavan (2006a). The differences are that *(i)* Calzolari and Pavan assume that the agent's type is binary (I model it as a continuous variable), *(ii)* Calzolari and Pavan consider revenue maximization (I study general objective functions), and *(iii)* Calzolari and Pavan consider a stochastic binary value of the third party but assume that the aftermarket always happens (I assume that the third party has a fixed value which is higher than the value of the agent but the aftermarket happens with interior probability). The effect of a stochastic value of the third party is similar to the effect of an interior probability of the aftermarket: Instead of assuming that the aftermarket does not take place, I could assume that the value of the third party is below the value of the agent. As a result, the optimal mechanisms from Proposition 1 have a similar structure to the

[27] This would not be possible in a cutoff mechanism because the probability of recommending the low price would have to be non-decreasing in the type, by Fact 2. It is not non-decreasing: the low price is never recommended for types above $\theta^\star$.

optimal mechanisms identified by Calzolari and Pavan (2006a) (with some differences stemming from the discrete versus continuous types space).

One insight from Proposition 1 and Claim 10 complementary to Calzolari and Pavan (2006a) is that the optimal mechanism reveals information only when there is non-zero probability that there will be no gains from trade after the mechanism: either because the aftermarket does not happen (as in my model), or because the third party has a low value (as in the model of Calzolari and Pavan, 2006a).

6.2.1 Robust payoff bounds

Proposition 1 and Claim 10 imply that a cutoff mechanism may sometimes be sub-optimal. However, these results do not directly show how much value the designer may lose by using a mechanism in this class instead of the optimal mechanism. Under the assumptions of the previous subsection, I provide robust payoff bounds on the performance of cutoff mechanisms.

Proposition 3a. *Under the assumptions of Section 6.2, regardless of the functional form of $V^a(\theta)$, a cutoff mechanism achieves at least a fraction $u^l/(u^l + u^h)$ of the value of the optimal mechanism.*

Because $u^l > u^h$ in a submodular aftermarket, a cutoff mechanism achieves at least half of the optimal value in the worst case but that fraction can be much larger when the "deviation" from a strongly submodular aftermarket is small (u^h is small). For instance, in Example 1(a), $u^l/(u^l + u^h) = 1/(2 - \lambda)$, so if the aftermarket happens with probability at least 1/2, then a cutoff mechanism achieves at least 2/3 of the optimal value, and when it happens with probability 3/4, the guarantee rises to 4/5.

For the simple Example 1(a), it is possible to obtain even better guarantees for particular objective functions of the designer.

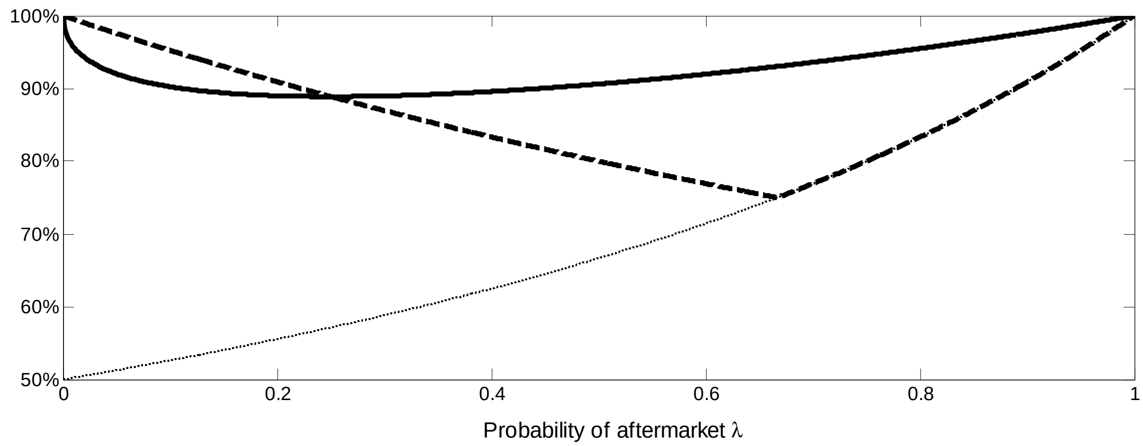
Proposition 3b. *If the designer maximizes total expected surplus in Example 1(a), then for any fixed $\lambda \in [0, 1]$ the optimal cutoff mechanism achieves at least a fraction*

$$\frac{1}{1 + \frac{1}{2}(\sqrt{\lambda} - \lambda)}$$

of the social surplus of the optimal mechanism.

In particular, regardless of λ, f, l, h, v, a cutoff mechanism is guaranteed to yield

Fig. 6.1: Robust payoff guarantees from using a cutoff mechanism in Example 1(a): social surplus (solid line), revenue for $\theta_H/\theta_L \leq 1.5$ (thick dotted line) and arbitrary objective (thin dotted line)



more than 88% of the optimal surplus. Moreover, a cutoff mechanism is optimal at $\lambda = 0$ and at $\lambda = 1$ (see Figure 6.1).

Proposition 3c. *If the designer maximizes expected revenue in Example 1(a), and $r \equiv \theta_H/\theta_L$, then for any fixed $\lambda \in [0, 1]$ the optimal cutoff mechanism achieves at least a fraction*

$$\max\left\{\frac{1}{2-\lambda}, \frac{1}{1-\lambda(1-r)}\right\}.$$

of the expected revenue of the optimal mechanism.

For example, if $r = 1.5$, then a cutoff mechanisms achieves at least 75% of the optimal expected revenue in all cases. For any finite r, the gap disappears close to $\lambda = 0$ and $\lambda = 1$ (see Figure 6.1).

6.2.2 Discussion

Summarizing, the assumption of strongly submodular aftermarket is needed for Theorems 2 and 3 but the designer might not be losing much by looking at cutoff mechanisms. Moreover, cutoff mechanisms remain relatively simple when the problem gets more complicated (see Dworczak, 2019). In contrast, I conjecture that the exactly optimal mechanism *2* identified by Proposition 1 is likely to increase in complexity when the relatively stylized assumptions are relaxed because its construction relies on satisfying the monotonicity constraint with equality whenever possible.

References

Back, K., Liu, R. and Teguia, A. (2018). Signaling in otc markets: Benefits and costs of transparency. *Journal of Financial and Quantitative Analysis*, **forthcoming**.

Balzer, B. and Schneider, J. (2017). Managing a conflict. Working Paper.

Bergemann, D. and Wambach, A. (2015). Sequential Information Disclosure in Auctions. *Journal of Economic Theory*, **159**, 1074–1095.

Border, K. C. (1991). Implementation of reduced form auctions: A geometric approach. *Econometrica*, **59** (4), 1175–1187.

Calzolari, G. and Pavan, A. (2006a). Monopoly with resale. *RAND Journal of Economics*, **37** (2), 362–375.

— and — (2006b). On the optimality of privacy in sequential contracting. *Journal of Economic Theory*, **130** (1), 168–204.

Carroll, G. and Segal, I. (2018). Robustly optimal auctions with unknown resale opportunities. *Review of Economic Studies*, **forthcoming**.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, **50** (6), 1431–1451.

Das Varma, G. (2003). Bidding for a process innovation under alternative modes of competition. *International Journal of Industrial Organization*, **21** (1), 15–37.

Dworczak, P. (2019). Mechanism Design with Aftermarkets: Cutoff Mechanisms. Working Paper.

Engelbrecht-Wiggans, R. and Kahn, C. M. (1991). Protecting the winner. second-price versus oral auctions. *Economics Letters*, **35**, 243 – 248.

Eső, P. and Szentes, B. (2007). Optimal Information Disclosure in Auctions and the Handicap Auction. *Review of Economic Studies*, **74** (3), 705–731.

Giovannoni, F. and Makris, M. (2014). Reputational bidding. *International Economic Review*, **55** (3), 693–710.

GOEREE, J. (2003). Bidding for the future: signaling in auctions with an aftermarket. *Journal of Economic Theory*, **108** (2), 345–364.

GUPTA, M. and LEBRUN, B. (1999). First Price Auctions with Resale. *Economics Letters*, **64**, 181–185.

HAFALIR, I. and KRISHNA, V. (2008). Asymmetric Auctions with Resale. *American Economic Review*, **98** (1), 87–112(26).

— and — (2009). Revenue and efficiency effects of resale in first-price auctions. *Journal of Mathematical Economics*, **45** (9-10), 589–602.

HAILE, P. A. (2003). Auctions with private uncertainty and resale opportunities. *Journal of Economic Theory*, **108** (1), 72 – 110.

HU, A. X. and ZHANG, J. (2017). Optimal mechanism design with aftermarket competitions. Working Paper.

JEHIEL, P. and MOLDOVANU, B. (2001). Efficient design with interdependent valuations. *Econometrica*, **69** (5), 1237–1259.

— and — (2006). Allocative and informational externalities in auctions and related mechanisms. In R. Blundell, W. Newey and T. Persson (eds.), *The Proceedings of the 9th World Congress of the Econometric Society*, Cambridge University Press.

—, — and STACCHETTI, E. (1996). How (not) to sell nuclear weapons. *The American Economic Review*, **86** (4), 814–829.

KATZMAN, B. E. and RHODES-KROPF, M. (2008). The consequences of information revealed in auctions. *Applied Economics Research Bulletin*, **2**, 53–87.

LAUERMANN, S. and VIRÁG, G. (2012). Auctions in Markets: Common Outside Options and the Continuation Value Effect. *American Economic Journal: Microeconomics*, **4** (4).

LI, H. and SHI, X. (2017). Discriminatory information disclosure. *American Economic Review*, **107** (11), 3363–85.

MATTHEWS, S. A. (1984). On the implementability of reduced form auctions. *Econometrica*, **52** (6), 1519–1522.

MILGROM, P. and WEBER, R. (1982). A Theory of Auctions and Competitive Bidding. *Econometrica*, **50**, 1089–1122.

MYERSON, R. (1981). Optimal auction design. *Mathematics of Operations Research*, **6** (1), 58–73.

MYERSON, R. B. (1982). Optimal coordination mechanisms in generalized principal-agent problems. *Journal of Mathematical Economics*, **10** (1), 67–81.

SEIERSTAD, A. and SYDSAETER, K. (1987). *Optimal Control Theory with Economic Applications*. Elsevier, North-Holland, Advanced Textbooks in Economics, volume 24.

SMOLIN, A. (2019). Disclosure and pricing of attributes. Working Paper.

ZHANG, J. and WANG, R. (2013). Optimal mechanism design with resale via bargaining. *Journal of Economic Theory*, **148** (5), 2096 – 2123.

ZHENG, C. Z. (2002). Optimal Auctions with Resale. *Econometrica*, **70** (6).

ZHONG, H. (2002). Information revelation, post-bid market interaction and auction choice. Working Paper.

A Multi-agent mechanisms

In this section, I extend the model to multi-agent symmetric settings.[28] I maintain the assumption that only the agent who acquires the object in the mechanism interacts in the aftermarket.

There are N ex-ante identical agents, indexed by $i = 1, 2, ..., N$. Each agent i has a privately observed type $\theta_i \in \Theta$. Types are distributed i.i.d. according to a full-support distribution f on $\Theta \equiv [0, 1]$. The payoff of the designer and the third party may depend on the type of the agent in the aftermarket but not on the agent's identity. Therefore, the utility functions $V^a(\theta)$ and $v^a(\theta)$ take the same form as in the one-agent model.

Under these assumptions, it is without loss of generality to look at symmetric mechanisms. I consider Bayesian implementation. Symmetric N-agent direct mechanisms can be represented by their reduced forms (x, π, t), where $x : \Theta \to [0, 1]$, $\pi : \Theta \to \Delta(\mathcal{S})$,

[28] Symmetry is assumed solely for expositional simplicity.

and $t : \Theta \to \mathbb{R}$ are all one-dimensional functions, subject to the constraint that x is feasible under f for some joint N-dimensional allocation rule $\boldsymbol{x}$ with $\sum_{i=1}^{N} \boldsymbol{x}_i(\boldsymbol{\theta}) = 1$ for all $\boldsymbol{\theta} \in \Theta^N$:

$$x(\theta) = \int\limits_{\Theta^{N-1}} \boldsymbol{x}_i(\theta, \boldsymbol{\theta}_{-i}) \prod_{j \neq i} f(\theta_j) d\theta_j. \tag{A.1}$$

For non-decreasing interim allocation rules x, condition (A.1) is equivalent to the so-called Matthews-Border condition:

$$\int_{\tau}^{1} x(\theta) f(\theta) d\theta \leq \frac{1 - F^N(\tau)}{N}, \forall \tau \in [0, 1], \tag{MB}$$

If the allocation rule x is not non-decreasing, condition (MB) is necessary but no longer sufficient (see Matthews, 1984, and Border, 1991). I will consider the relaxed problem with constraint (MB). Thus, the obtained solution is guaranteed to be feasible only if the optimal x turns out to be non-decreasing.

As before, it is without loss of generality to look at binary signal spaces, $\mathcal{S} = \{l, h\}$. Then, letting $\pi_h(\theta)$ denote the probability of sending the high signal conditional on type θ and allocating the object, any mechanism can be represented by its reduced form (x, π_h, t).

Under Bayesian implementation, conditions for implementability of (x, π_h) are formally identical to those from the one-agent model. That is, equations (M), (OB_l), (OB_h) together with (MB) fully characterize all implementable pairs (x, π_h).

Finally, the objective function of the mechanism designer is given by

$$N \int_{0}^{1} \left(V^h(\theta)\pi_h(\theta)x(\theta) + V^l(\theta)(1 - \pi_h(\theta))x(\theta) \right) f(\theta) d\theta \tag{A.2}$$

which only differs from the one-agent objective function in that it is multiplied by N.

A symmetric mechanism in the multi-agent setting is defined as a cutoff (partitional) mechanism if its reduced form is a one-agent cutoff (partitional) mechanism.

For a fixed interim allocation rule, we obtain a problem that is formally identical to the one-agent problem considered in the previous sections (the additional constraint MB only pertains to the allocation rule). Therefore, Theorem 2 applies without any modifications (with x interpreted as an interim allocation rule). Theorem 3 can be extended as well, although this requires a proof.

Theorem 4. *For any N, suppose that the designer's and the third party's preferences* *are co-modular, and and that $V^h(\theta)$ and $V^l(\theta)$ are non-decreasing in θ. The following mechanism achieves the maximal value of* (4.3) *subject to* (M), (OB_h), (OB_l), (MB)*:*

- *A cutoff rule, when the aftermarket is strongly submodular;*
- *A partitional rule, when all players' utility functions are supermodular.*

A.1 Discussion

The implications of the analysis with multiple agents are qualitatively similar to the ones from the single-agent model. I briefly comment on the differences below.

The problem of optimal information disclosure for a fixed allocation rule is identical mathematically. The only difference is in the interpretation: In the multi-agent model, the role of $x(\theta)$ from the single-agent model is played by the interim expected allocation rule. In a partitional mechanism, this is mostly irrelevant. In a cutoff mechanism, the implication is that the distribution of the cutoff can depend on the distribution of other agents' information. For example, in a monotone equilibrium of an auction, the cutoff corresponds to the second highest bid, and thus results about optimality of disclosure should be interpreted as pertaining to the disclosure of the second highest bid.

The problem of joint optimization over allocation and disclosure rules is altered by the presence of the Matthews-Border condition. Practically, this means that the optimal (interim expected) allocation rule will typically not be a simple threshold rule. This has consequences for the optimal design of cutoff mechanisms which might disclose non-trivial information about the cutoff when there are multiple agents (which never happens with one agent). This is explained in detail in the companion paper Dworczak (2019) that focuses entirely on cutoff mechanisms.

In the resale example of Section 5.1, the surplus-maximizing mechanism for the case $N \geq 2$ is a cutoff mechanism that can take two possible forms:[29] Either *(i)* it is an auction with a reserve price that is sufficiently high so that conditional on allocation the high resale price is always induced (no signals are sent), or *(ii)* it is a mechanism that sends a binary signal that depends on the second highest type, and allocates the good to the highest bidder when the second highest type is above a threshold and to a random agent (regardless of the type) when the second highest type is below that threshold. A revenue-maximizing mechanism could take an even more complicated form but at least in some cases mechanism *(i)* is also revenue-maximizing.

[29] This case is solved formally in the Online Appendix of Dworczak (2019).

In the investment game of Section 5.2, the surplus-maximizing mechanism for the case $N \geq 2$ is a partitional mechanism: The designer runs an efficient auction, and then discloses whether the bid of the winner was above or below a threshold. To maximize revenue, the designer would additionally introduce a reserve price (but the disclosure rule would remain the same).

B Proofs

B.1 An auxiliary optimization problem

This appendix formulates and proves some key lemmas that will be used in the subsequent proofs. I consider the maximization problem

$$\max_{y\colon \Theta \to [0,1]} \int_0^1 \phi(\theta) y(\theta) f(\theta) d\theta \tag{B.1}$$

subject to

$$y(\theta) \text{ is non-decreasing,} \tag{B.2}$$

$$y(\theta) \leq x(\theta),\ \forall \theta \in \Theta, \tag{B.3}$$

$$\int_0^1 \psi(\theta) y(\theta) f(\theta) d\theta \geq c, \tag{B.4}$$

for some upper semi-continuous functions $\phi : \Theta \to \mathbb{R}$, $\psi : \Theta \to \mathbb{R}$, a non-decreasing right-continuous allocation rule x, and a constant $c \in \mathbb{R}$. The function ϕ is either non-negative or non-positive.

To avoid trivial cases, I make the following assumptions. There exists at least one feasible y. In the case when ϕ is non-negative, (B.1) is upper-bounded by taking $y(\theta) = x(\theta)$, and this upper bound cannot be achieved:

$$\int_0^1 \psi(\theta) x(\theta) f(\theta) d\theta < c.$$

In the case when ϕ is non-positive, (B.1) is upper-bounded by taking $y(\theta) = 0$, and this upper bound cannot be achieved: $c > 0$.

Lemma 2. *Consider the maximization problem* (B.1) - (B.4).

If ϕ and ψ are both non-increasing, the optimal solution takes the form $y(\theta) =$

$\min\{x(\theta), \bar{x}\}$ *for* $\bar{x} \in [0, 1]$ *such that* (B.4) *holds with equality.*

If ϕ *and* ψ *are both non-decreasing, the optimal solution takes the form* $y(\theta) = x(\theta)\mathbf{1}_{\{\theta \geq \theta^\star\}}$ *for* $\theta^\star \in [0, 1]$ *such that* (B.4) *holds with equality.*

Proof of Lemma 2. Consider two candidate solutions y_1 and y_2. I say that y_1 *dominates* y_2 if *(i)* $\int_0^1 y_1(\theta)f(\theta)d\theta = \int_0^1 y_1(\theta)f(\theta)d\theta$, and *(ii)* $\int_0^\alpha y_1(\theta)f(\theta)d\theta \leq \int_0^\alpha y_1(\theta)f(\theta)d\theta$ for all $\alpha \in [0, 1]$.

Take an optimal solution $y^\star$ to problem (B.1) - (B.4).[30] Let $\alpha \equiv \int_0^1 y^\star(\theta)f(\theta)d\theta$. Then, $y^\star$ solves the problem (B.1) - (B.4) with the additional constraint

$$\alpha = \int_0^1 y(\theta)f(\theta)d\theta. \tag{B.5}$$

Suppose that ϕ and ψ are non-decreasing. Consider $y'(\theta)$ which satisfies (B.2), (B.3), (B.5), and dominates $y^\star$. Then, $Y'(\theta) = \int_0^\theta y'(\tau)f(\tau)d\tau$ first-order stochastically dominates $Y^\star(\theta) = \int_0^\theta y^\star(\tau)f(\tau)d\tau$. It follows that y' satisfies (B.4) and achieves a higher value of the objective (B.1) than $y^\star$, so y' is optimal.

Thus, an optimal solution can be found among functions y that satisfy (B.2), (B.3), and are not dominated by any other function satisfying these constraints with the same value of $\int_0^1 y(\theta)f(\theta)d\theta$. Because functions of the form $y(\theta) = x(\theta)\mathbf{1}_{\{\theta \geq \theta^\star\}}$ dominate all functions satisfying (B.2) and (B.3) with equal $\int_0^1 y(\theta)f(\theta)d\theta$, $y(\theta) = x(\theta)\mathbf{1}_{\{\theta \geq \theta^\star\}}$ is optimal for some $\theta^\star \in [0, 1]$.

I only have to prove that $\theta^\star$ is set in such a way that condition (B.4) holds with equality. By assumption, there exists a feasible solution y'. Some function of the form $y(\theta) = x(\theta)\mathbf{1}_{\{\theta \geq \theta^\star\}}$ dominates y', so there exists a feasible solution of the form $y(\theta) = x(\theta)\mathbf{1}_{\{\theta \geq \theta^\star\}}$. Again by assumption, whether ϕ is non-negative or non-positive, the upper-bound solutions $y \equiv x$ and $y \equiv 0$ are not feasible, respectively. Thus, (B.4) has to hold with equality.

In the case when ϕ and ψ are non-increasing, consider y' satisfying (B.2), (B.3), (B.5), which is dominated by $y^\star$. By a similar reasoning, y' is optimal. Thus, an optimal solution is a function y that is dominated by all functions satisfying (B.2) and (B.3). Such a function takes the form $y(\theta) = \min\{x(\theta), \bar{x}\}$ for some constant $\bar{x} \in [0, 1]$. By the same reasoning as in the previous case, $\bar{x}$ has to be such that constraint (B.4) holds with equality. □

[30] The solution exists because the domain is compact in the product topology, and the objective function is upper semi-continuous.

For the remaining results in this appendix, I will use tools from optimal control theory. The following lemma establishes sufficient conditions for a feasible candidate solution y to be optimal for problem (B.1)-(B.4).

Lemma 3. *Suppose that y is an absolutely continuous function which satisfies constraints (B.2) - (B.4). Suppose that there exist (i) a function $u(\theta) \geq 0$ such that $y(\theta) - y(0) = \int_0^\theta u(\tau)d\tau$, for all $\theta \in [0, 1]$, (ii) a piece-wise continuous $\lambda(\theta) \geq 0$, (iii) a piece-wise differentiable and piece-wise continuous $p(\theta) \leq 0$, and (iv) a constant η, such that:*

1. $u(\theta) > 0 \implies p(\theta) = 0$,
2. $x(\theta) > y(\theta) \implies \lambda(\theta) = 0$,
3. $p'(\theta) = (\lambda(\theta) - \phi(\theta) - \eta\psi(\theta)) f(\theta)$, *whenever it is differentiable,*
4. *p can jump up at finitely many points τ, $p(\tau^-) < p(\tau^+)$, provided that $x(\tau) = y(\tau)$.*
5. $p(0) = 0$ *and* $p(1) = 0$,
6. *if constraint (B.4) does not bind, then $\eta = 0$; if constraint (B.4) is assumed to bind, then η is unrestricted; in other cases, $\eta \geq 0$.*

Then, y solves the problem (B.1) - (B.4).

Proof. Follows by applying Theorem 1 on page 317 of Seierstad and Sydsaeter (1987) to the problem (B.1) - (B.4). The condition $p(\theta) \leq 0$ and condition (1) come from the requirement that the non-negative control function $u(\theta)$ maximizes the Hamiltonian (which is linear in the control function). Condition (2) is a complementary-slackness condition on the Lagrange multiplier $\lambda(\theta)$ that comes from the constraint $y(\theta) \leq x(\theta)$. Condition (3) is the standard law of motion for the multiplier p, and condition (4) follows from the fact that the endpoints of y are not restricted. Sometimes p might have a jump discontinuity at 1, which is captured by condition (5). The constraint (B.4) is incorporated by defining an auxiliary state variable $\Gamma(\theta)$ with $\Gamma'(\theta) = \psi(\theta)y(\theta)f(\theta)$, $\Gamma(0) = 0$, and $\Gamma(1) \geq c$. Because Γ does not appear in the Hamiltonian, the corresponding multiplier η is constant. Condition (6) summarizes the properties of η depending on the whether $\Gamma(1) \geq c$ is binding or not. The concavity assumptions are satisfied because the problem and the constraints are linear in y. □

I now apply Lemma 3 to obtain sufficient conditions for a cutoff mechanism to solve the problem (B.1) - (B.4).

Lemma 4. *Consider problem* (B.1)-(B.4) *and assume additionally that* x *is absolutely continuous. Suppose that* $\bar{x}$ *is well-defined by*

$$\int_0^1 \psi(\theta)\min\{x(\theta),\, \bar{x}\}f(\theta)d\theta = c.$$

Define $\theta^\star = \max\{\theta \in [0,\,1] :\, x(\theta) \le \bar{x}\}$*, and* η *by*

$$\int_{\theta^\star}^1 (\phi(\theta) + \eta\psi(\theta))f(\theta)d\theta = 0.$$

If $\phi(\theta) + \eta\psi(\theta)$ *crosses zero once and from above, then* $y(\theta) = \min\{x(\theta),\, \bar{x}\}$ *solves the problem* (B.1)-(B.4)*.*

Proof. By assumption, $y(\theta) = \min\{x(\theta),\, \bar{x}\}$ is a feasible candidate solution. To prove that y is optimal, let $\lambda(\theta) = \phi(\theta) + \eta\psi(\theta)$ for $\theta \le \theta^\star$, and $\lambda(\theta) = 0$ for $\theta > \theta^\star$. Next, define p as in Lemma 3 with $p(0) = 0$. By the choice of λ, $p(\theta) = 0$ for $\theta \le \theta^\star$. To guarantee that $p(1) = 0$, we need

$$\int_{\theta^\star}^1 (\phi(\theta) + \eta\psi(\theta))f(\theta)d\theta = 0,$$

and this pins down η. Suppose that $\phi(\theta) + \eta\psi(\theta)$ crosses zero once and from above. The function $\phi(\theta) + \eta\psi(\theta)$ has to cross zero to the right of $\theta^\star$ because otherwise the above equality could not hold. Thus, $\phi(\theta) + \eta\psi(\theta)$ is positive for $\theta \le \theta^\star$. This means that $\lambda(\theta) \ge 0$ for all θ. Finally, $p(\theta)$ is non-positive because its derivative on $[\theta^\star,\, 1]$, equal to $-(\phi(\theta) + \eta\psi(\theta))$, is first non-positive, and then non-negative. □

B.2 Proof of Theorem 1

I will not assume that (2.2) holds to show that the proof extends to the opposite case as well.

Since $x(\theta) = 1$ for all θ, constraint (M) becomes

$$\pi_h(\theta)(u^h - u^l) \text{ is non-decreasing in } \theta.$$

Take for concreteness the case when the agent's utility is submodular, $u^l > u^h$ (the other case is fully analogous, and $u^h = u^l$ is ruled out by assumption). Then, π_h has to be non-increasing. By condition (OB_h), the mean value theorem for integrals, and the assumption that v is supermodular (since the aftermarket is submodular),

$$0 \leq \int_0^1 (v^h(\theta) - v^l(\theta))\pi_h(\theta)f(\theta)d\theta = \pi_h(0^+)\int_0^\alpha (v^h(\theta) - v^l(\theta))f(\theta)d\theta,$$

for some $\alpha \in [0, 1]$. If $\pi_h(0^+) = 0$, then $\pi_h \equiv 0$ (which would end the proof) because π_h is non-increasing. Otherwise, we have

$$\int_0^\alpha (v^h(\theta) - v^l(\theta))f(\theta)d\theta \geq 0$$

which, by the fact that v is supermodular, implies that

$$\int_\beta^1 (v^h(\theta) - v^l(\theta))f(\theta)d\theta \geq 0,$$

for all $\beta \in [0, 1]$. Then, by condition (OB_l), and again by the mean value theorem,

$$0 \geq \int_0^1 (v^h(\theta) - v^l(\theta))(1 - \pi_h(\theta))f(\theta)d\theta = (1 - \pi_h(1^-))\int_\gamma^1 (v^h(\theta) - v^l(\theta))f(\theta)d\theta,$$

for some $\gamma \in [0, 1]$. Unless $\int_\gamma^1 (v^h(\theta) - v^l(\theta))f(\theta)d\theta = 0$, this implies that $\pi_h(1^-) = 1$, and because π_h is non-increasing, we must have $\pi_h \equiv 1$ (which would end the proof).

If $\int_\gamma^1 (v^h(\theta) - v^l(\theta))f(\theta)d\theta = 0$, then because $\int_0^\alpha (v^h(\theta) - v^l(\theta))f(\theta)d\theta \geq 0$, and v is supermodular, it must be that $\int_0^1 (v^h(\theta) - v^l(\theta))f(\theta)d\theta = 0$. Since v^h and v^l are not identical ($v^h \neq v^l$), π_h must be constant.

Now suppose that the aftermarket is supermodular. First, consider the case when both the agent's and the third party's utilities are supermodular. Then, full disclosure of the type, understood as $\mathcal{S} = [0, 1]$, $\pi(\cdot|\theta) = \delta_{\{s=\theta\}}$, corresponds to implementing $\pi_h(\theta) = \mathbf{1}_{\{\theta \geq \hat{\theta}\}}$, where $\hat{\theta}$ is defined as the smallest solution to $v^h(\theta) - v^l(\theta) = 0$ (due to supermodularity of the third party's objective function). (OB_h) holds by construction, and (M) holds because $\pi_h(\theta)$ is non-decreasing. When both player's utility functions are submodular, then full disclosure corresponds to $\pi_h(\theta) = \mathbf{1}_{\{\theta < \hat{\theta}\}}$, and ($M$) holds because $\pi_h(\theta)$ is non-increasing (and agent's utility is submodular, so that $u^l \geq u^h$).

B.3 Proof of Theorem 2

Part I: Submodular aftermarkets. I prove the theorem assuming that agent's utility is strongly supermodular – the opposite case when the agent's utility is strongly submodular is analogous.[31] Under this assumption and the assumptions of Theorem 2, $u^h > 0$, and $u^l = 0$, and the third party's and designer's utility functions are submodular. The optimal design problem takes the form

$$\max_{\pi_h} \int_0^1 \left[\pi_h(\theta)V^h(\theta) + (1-\pi_h(\theta))V^l(\theta)\right] x(\theta)f(\theta)d\theta \tag{B.6}$$

subject to

$$0 \le \pi_h(\theta) \le 1,\ \forall \theta \in \Theta, \tag{B.7}$$

$$\pi_h(\theta)x(\theta) \text{ is non-decreasing in } \theta, \tag{B.8}$$

$$\int_0^1 v^h(\theta)\pi_h(\theta)x(\theta)f(\theta)d\theta \ge \int_0^1 v^l(\theta)\pi_h(\theta)x(\theta)f(\theta)d\theta, \tag{B.9}$$

$$\int_0^1 v^l(\theta)(1-\pi_h(\theta))x(\theta)f(\theta)d\theta \ge \int_0^1 v^h(\theta)(1-\pi_h(\theta))x(\theta)f(\theta)d\theta. \tag{B.10}$$

I consider two cases. First, suppose that

$$\int_0^1 v^h(\theta)x(\theta)f(\theta)d\theta \ge \int_0^1 v^l(\theta)x(\theta)f(\theta)d\theta, \tag{B.11}$$

i.e., the third party takes the high action in the absence of any additional information. Then, revealing no information ($\pi_h \equiv 0$) is a feasible mechanism, and it is optimal because, by assumption, $V^h \ge V^l$. In this case, the optimal mechanism does not reveal any information, and is thus trivially a cutoff mechanism.

Consider the opposite case in which condition (B.11) fails. Let $y(\theta) \equiv \pi_h(\theta)x(\theta)$. I consider a relaxed problem without the constraint (B.10) (I will verify ex-post that this constraint is satisfied at the solution of the relaxed problem). By deleting terms

[31] In that case, in the derivation that follows, I would still consider (B.6) - (B.10) but with (B.8) replaced by

$$(1-\pi_h(\theta))x(\theta) \text{ is non-decreasing in } \theta,$$

and $y(\theta) \equiv (1-\pi_h(\theta))x(\theta)$.

that do not affect the value of the objective function we obtain

$$\max_{y} \int_0^1 \left[V^h(\theta) - V^l(\theta)\right] y(\theta) f(\theta) d\theta$$

subject to

$$0 \leq y(\theta) \leq x(\theta),\ \forall \theta \in \Theta,$$

$$y(\theta) \text{ is non-decreasing in } \theta,$$

$$\int_0^1 \left[v^h(\theta) - v^l(\theta)\right] y(\theta) f(\theta) d\theta \geq 0.$$

Under assumptions of Theorem 2, define $\phi(\theta) = V^h(\theta) - V^l(\theta)$, $\psi(\theta) = v^h(\theta) - v^l(\theta)$. Then, the problem becomes identical to the auxiliary optimization problem considered in Appendix B.1. The functions V and v are submodular, so ϕ and ψ are non-increasing. By Lemma 2, the optimal mechanism takes the form $y(\theta) = \min\{x(\theta), \bar{x}\}$ for some $\bar{x} \in [0, 1]$. This means that both $\pi_h(\theta)x(\theta)$ and $(1 - \pi_h(\theta))x(\theta) = x(\theta) - y(\theta) = \max\{x(\theta) - \bar{x}, 0\}$, are non-decreasing which corresponds to a cutoff mechanism (by Fact 2). The problem becomes

$$\max_{\bar{x}} \int_0^1 \left[V^h(\theta) - V^l(\theta)\right] \min\{x(\theta), \bar{x}\} f(\theta) d\theta$$

subject to

$$\int_0^1 \left[v^h(\theta) - v^l(\theta)\right] \min\{x(\theta), \bar{x}\} f(\theta) d\theta \geq 0.$$

Setting $\bar{x} = \max_\theta\{x(\theta)\}$ is not feasible when condition (B.11) fails, and the objective function is non-decreasing in $\bar{x}$. Thus, the constraint must bind at the optimal solution:

$$\int_0^1 \left[v^h(\theta) - v^l(\theta)\right] \min\{x(\theta), \bar{x}\} f(\theta) d\theta = 0.$$

Finally, I have to verify that the solution to the relaxed problem is feasible for the original problem, i.e., that

$$\int_0^1 v^l(\theta)(1 - \pi_h(\theta))x(\theta) f(\theta) d\theta \geq \int_0^1 v^h(\theta)(1 - \pi_h(\theta))x(\theta) f(\theta) d\theta.$$

We know that

$$\int_0^1 [v^l(\theta) - v^h(\theta)] x(\theta) f(\theta) d\theta > 0.$$

From the above,

$$
\begin{aligned}
0 < & \int_0^1 [v^l(\theta) - v^h(\theta)]x(\theta)f(\theta)d\theta \\
& = \underbrace{\int_0^1 [v^l(\theta) - v^h(\theta)]y(\theta)f(\theta)d\theta}_{0} + \int_0^1 [v^l(\theta) - v^h(\theta)](x(\theta) - y(\theta))f(\theta)d\theta \\
& \qquad = \int_0^1 [v^l(\theta) - v^h(\theta)](1 - \pi_h(\theta))x(\theta)f(\theta)d\theta, \quad \text{(B.12)}
\end{aligned}
$$

so constraint (B.10) holds. Thus, a cutoff mechanism is optimal.

Part II: Supermodular aftermarkets. The proof follows almost exactly the same steps as in the previous case, so I omit most details, and only highlight the differences.

Assume that agent's utility is supermodular (again, the opposite case if fully analogous and thus skipped). That is, $u^h > u^l$. Then, the assumptions of Theorem 2 imply that the third party's and designer's utility functions are supermodular. As in the case of a submodular aftermarket, I consider the optimization problem (B.6) – (B.10), however, with constraint (B.8) replaced by the original monotonicity constraint (M). Define $y(\theta) = \pi_h(\theta)x(\theta)$, and consider the relaxed problem

$$\max_y \int_0^1 \left[V^h(\theta) - V^l(\theta)\right] y(\theta)f(\theta)d\theta$$

subject to

$$0 \leq y(\theta) \leq x(\theta), \forall \theta \in \Theta, \tag{B.13}$$

$$\int_0^1 \left[v^h(\theta) - v^l(\theta)\right] y(\theta)f(\theta)d\theta \geq 0. \tag{B.14}$$

The relaxation is that I dropped both the monotonicity constraint (M) and the obedience constraint (B.10). We can define $\phi(\theta) = V^h(\theta) - V^l(\theta)$, and $\psi(\theta) = v^h(\theta) - v^l(\theta)$. This time, these functions are non-decreasing. Thus, Lemma 2 from Appendix B.1 yields the conclusion that the optimal $y(\theta) = \pi_h(\theta)x(\theta)$ takes the form $y(\theta) = x(\theta)\mathbf{1}_{\{\theta \geq \theta^\star\}}$ for some $\theta^\star \in [0, 1]$ (technically, the problem in Lemma 2 has a monotonicity constraint but dropping it does not change the solution). It is immediate that because x is non-decreasing, condition (M) is satisfied with this π_h, and so is (B.10) for the optimal choice of $\theta^\star$ (see the proof of Lemma 2). Therefore, the partitional mechanism that solves the relaxed problem is also feasible (and hence optimal) for the original problem.

B.4 Proof of Lemma 1

I prove the result for cases *(i)* and *(ii)*. Case *(iii)* is fully analogous to case *(ii)* and thus skipped.[32] In cases *(i)* and *(ii)*, the agent's utility is supermodular ($u^h > u^l$). Let $y(\theta) = x(\theta)\pi_h(\theta)$. The (relaxed[33]) problem of the designer can be written as

$$\max_{x,y} \int_0^1 (V^h(\theta) - V^l(\theta))y(\theta)f(\theta)d\theta + \int_0^1 V^l(\theta)x(\theta)f(\theta)d\theta \tag{B.15}$$

subject to

$$0 \leq y(\theta) \leq x(\theta), \forall\theta \in \Theta, \tag{B.16}$$

$$\int_0^1 \left[v^h(\theta) - v^l(\theta)\right] y(\theta)f(\theta)d\theta \geq 0, \tag{B.17}$$

$$(u^h - u^l)y(\theta) + u^l x(\theta) \text{ is non-decreasing in } \theta. \tag{B.18}$$

The rest of the proof proceeds differently depending on the cases. I consider case *(ii)* first.

Case (ii). Define $z(\theta) = \frac{u^h - u_l}{u^h} y(\theta) + \frac{u^l}{u^h} x(\theta)$. Constraint (B.18) implies that z is non-decreasing in a feasible mechanism. Thus, fixing any non-decreasing z, we can consider an auxiliary problem of choosing x optimally:

$$\max_{x} \int_0^1 \left[V^l(\theta) + (V^l(\theta) - V^h(\theta))\frac{u^l}{u^h - u^l}\right] x(\theta)f(\theta)d\theta \tag{B.19}$$

subject to

$$0 \leq z(\theta) \leq x(\theta), \forall\theta \in \Theta, \tag{B.20}$$

$$u^h \int_0^1 \left[v^h(\theta) - v^l(\theta)\right] z(\theta)f(\theta)d\theta \geq \int_0^1 u^l \left[v^h(\theta) - v^l(\theta)\right] x(\theta)f(\theta)d\theta. \tag{B.21}$$

[32] The difference is that I would define $y(\theta) = x(\theta)(1 - \pi_h(\theta))$ instead of $y(\theta) = x(\theta)\pi_h(\theta)$.

[33] As usually, we can drop the constraint (OB_l) because it is never binding in the optimal solution.

Because the aftermarket is submodular, both $V^h(\theta) - V^l(\theta)$ and $v^h(\theta) - v^l(\theta)$ are non-increasing. $V^l(\theta)$ is non-decreasing by assumption. This means that the objective function $V^l(\theta) + (V^l(\theta) - V^h(\theta))\frac{u^l}{u^h - u^l}$ is non-decreasing, and the constraint (B.21) is preserved when x is modified by "shifting mass to the right." By the reasoning in Lemma 2 in Appendix B.1, the optimal $x^\star(\theta)$ must be equal to $z(\theta)$ for $\theta < \theta^\star$ and to 1 for $\theta \geq \theta^\star$. In particular, $x^\star$ is non-decreasing.

Case (i). In this case, the proof is easier. Because the aftermarket is supermodular, $V^h(\theta) - V^l(\theta)$ and $v^h(\theta) - v^l(\theta)$ are both non-decreasing. Because $V^l(\theta)$ is non-decreasing by assumption, the objective function (B.15) is increased by shifting mass to the right under both x and y (in the sense explained in the proof of Lemma 2). Moreover, such a simultaneous shift preserves constraints (B.17) and (B.18). It follows that both y and x must be non-decreasing in the optimal mechanism. In fact, in this case, we can pin down the shapes of the optimal $x^\star$ and $y^\star$ more precisely – they must both have a threshold structure: $x^\star(\theta) = \mathbf{1}_{\{\theta \geq \theta^\star\}}$ and $y^\star(\theta) = \mathbf{1}_{\{\theta \geq \hat{\theta}\}}$ for some $\hat{\theta} \geq \theta^\star \geq 0$.

B.5 Proof of Claim 3

Deleting terms that do not depend on π_h, and letting $y(\theta) \equiv (1 - \pi_h(\theta))x(\theta)$, the design problem can be written as

$$\max_y \int_0^1 [J(\theta) - 1]\, y(\theta) f(\theta) d\theta$$

subject to

$$0 \leq y(\theta) \leq x(\theta),\ \forall \theta \in \Theta,$$

$$y(\theta) \text{ is non-decreasing in } \theta,$$

$$\int_0^1 [(\theta_H - \theta_L) - \theta(v - \theta_L)]\, y(\theta) f(\theta) d\theta \geq \int_0^1 [(\theta_H - \theta_L) - \theta(v - \theta_L)]\, x(\theta) f(\theta) d\theta.$$

Let $\phi(\theta) = J(\theta) - 1$, and $\psi(\theta) = (\theta_H - \theta_L) - \theta(v - \theta_L)$. Let $\bar{x}$ and $\theta^\star$ be defined as in Lemma 4. We can apply Lemma 4 which provides sufficient conditions for $y(\theta) = \min\{x(\theta), \bar{x}\}$ to be optimal. Define η as a solution to equation $\int_{\theta^\star}^1 (\phi(\theta) + \eta\psi(\theta)) f(\theta) d\theta = 0$. That is,

$$\eta = \frac{\int_{\theta^\star}^1 (J(\theta) - 1)\, f(\theta) d\theta}{\int_{\theta^\star}^1 (\theta(v - \theta_L) - (\theta_H - \theta_L))\, f(\theta) d\theta} = \frac{(1 - \theta^\star)}{(\theta_H - \theta_L) - (v - \theta_L)\mathbb{E}[\theta |\, \theta \geq \theta^\star]}$$

By Lemma 4, it is enough to prove that $\Lambda(\theta) \equiv \phi(\theta) + \eta\psi(\theta)$ crosses zero once and from above. A sufficient condition is that $\Lambda'(\theta) \leq 0$ for all θ. This gives us the condition

$$\eta \geq \frac{\bar{J}}{v - \theta_L},$$

where recall that $\bar{J} = \max_\theta J'(\theta)$. Let $\Delta \equiv (\theta_H - \theta_L)/(v - \theta_L)$. Plugging in the definition of η, and simplifying, a sufficient condition for optimality of $y(\theta) = \min\{x(\theta), \bar{x}\}$ is

$$\Delta - \frac{(1-\theta^\star)}{\bar{J}} \leq \mathbb{E}[\theta | \, \theta \geq \theta^\star] \leq \Delta \tag{B.22}$$

By definition of $\bar{x}$, the obedience constraint holds with equality at y, which means that

$$0 = \int_{\theta^\star}^1 [\Delta - \theta] \, (x(\theta) - \bar{x}) f(\theta) d\theta.$$

Because $x(\theta)$ is non-decreasing, the mean value theorem for integrals implies that

$$\int_{\theta^\star}^1 [\Delta - \theta] \, f(\theta) d\theta \geq 0.$$

And this condition implies that $\Delta \geq \mathbb{E}[\theta | \, \theta \geq \theta^\star]$. Thus, the sufficient condition (B.22) boils down to

$$\Delta \leq \mathbb{E}[\theta | \, \theta \geq \theta^\star] + \frac{1-\theta^\star}{\bar{J}}.$$

For uniform distribution, we have $\bar{J} = 2$, and $\mathbb{E}[\theta | \, \theta \geq \theta^\star] = (1+\theta^\star)/2$. Consequently,

$$\Delta \leq \frac{1+\theta^\star}{2} + \frac{1-\theta^\star}{2} = 1,$$

and condition (B.22) always holds.

B.6 Proof of Claim 4

I denote $y(\theta) \equiv (1 - \pi_h(\theta))x(\theta)$ and $z(\theta) \equiv \pi_h(\theta)x(\theta)$. Then the joint design problem can be written as

$$\max_{y \geq 0, \, z \geq 0} \theta_H \int_0^1 z(\theta) f(\theta) d\theta + \int_0^1 [\theta_H - (\theta_H - \theta_L)(1 - J(\theta))] \, y(\theta) f(\theta) d\theta \tag{B.23}$$

subject to

$$y(\theta) \text{ is non-decreasing in } \theta, \tag{B.24}$$

$$0 \leq y(\theta) + z(\theta) \leq 1, \, \forall \theta, \tag{B.25}$$

$$\int_0^1 z(\theta)\phi(\theta)f(\theta)dt \geq 0, \tag{B.26}$$

where $\phi(\theta) \equiv (\theta_H - \theta_L) - (v - \theta_L)(1 - \theta)$. I will consider a relaxed problem where I omit constraint (B.24), and verify at the end that this constraint holds.

After constraint (B.24) is dropped, we can apply standard optimal control techniques to solve the problem (B.23) subject to (B.25) - (B.26) treating $y(\theta)$ and $z(\theta)$ as control variables chosen at any θ from the set $U = \{(y, z) \in (0, 1)^2 : y + z \leq 1\}$. The Hamiltonian is

$$H = \left[\underbrace{(\theta_H + \lambda\phi(\theta))}_{\lambda_z(\theta)} z(\theta) + \underbrace{(\theta_L + (\theta_H - \theta_L)J(\theta))}_{\lambda_y(\theta)} y(\theta) \right] f(\theta),$$

Let $\underline{\theta}$ be the unique point such that $\phi(\underline{\theta}) = 0$, and let $\theta^\star$ be defined by (5.6). Then, we have $\theta^\star < \underline{\theta}$, and there exists a unique $\lambda > 0$ such that $\lambda_z(\theta)$ goes through the point $(\theta^\star, 0)$. By assumption, $\lambda_y(\theta) \leq 0$ for all $\theta \leq \theta^\star$. Moreover, for $\theta \geq \underline{\theta}$, we have $\lambda_z(\theta) \geq \lambda_y(\theta)$, by direct inspection. Because $\lambda_z(\theta)$ is affine in θ, and $\lambda_y(\theta)$ is convex in θ, this implies that $\lambda_z(\theta) \geq \lambda_y(\theta)$ for all $\theta \geq \theta^\star$. Therefore, $y(\theta) = 0$ and $z(\theta) = \mathbf{1}_{\{\theta \geq \theta^\star\}}$ maximizes the Hamiltonian point-wise subject to $(y(\theta), z(\theta)) \in U$. Moreover, $z(\theta)$ satisfies constraint (B.26) with equality, by definition of $\theta^\star$. Because $y(\theta) = 0$ is non-decreasing in θ, constraint (B.24) also holds, and thus we have obtained an optimal solution, by the Maximum Principle for optimal control.[34]

B.7 Proof of Proposition 1

The problem of the designer is given by

$$\max_{x, \pi_h} \int_0^1 \left[\pi_h(\theta)V^h(\theta) + (1 - \pi_h(\theta))V^l(\theta)\right] x(\theta)f(\theta)d\theta \tag{B.27}$$

[34] See for example Seierstad and Sydsaeter (1987).

subject to

$$0 \leq \pi_h(\theta) \leq 1, \; \forall \theta \in \Theta, \tag{B.28}$$

$$\pi_h(\theta)x(\theta)u^h + (1 - \pi_h(\theta))x(\theta)u^l \text{ is non-decreasing in } \theta, \tag{B.29}$$

$$\int_0^1 v^h(\theta)\pi_h(\theta)x(\theta)f(\theta)d\theta \geq \int_0^1 v^l(\theta)\pi_h(\theta)x(\theta)f(\theta)d\theta, \tag{B.30}$$

$$\int_0^1 v^l(\theta)(1 - \pi_h(\theta))x(\theta)f(\theta)d\theta \geq \int_0^1 v^h(\theta)(1 - \pi_h(\theta))x(\theta)f(\theta)d\theta. \tag{B.31}$$

The last constraint (B.31) can be ignored because it will always be slack at the optimal solution (due to the assumption that $V^h(\theta) \geq V^l(\theta)$ for all θ).

I will solve the problem in three steps. In the first two steps, I solve two auxiliary problems in which some choice variables are fixed. This allows me to derive restrictions on the structure of the optimal solution. In the last step, I optimize in the class of candidate solutions that satisfy these restrictions.

First, I consider an auxiliary problem for a fixed non-decreasing allocation rule x (I will show later that a non-decreasing x is optimal). If the high action is an optimal response for the third party given x when no further information is reveled, then it is clearly optimal not to reveal any information. Thus, I can focus on the case when the low action is chosen if no further information is revealed:

$$\int_0^1 (v^h(\theta) - v^l(\theta))x(\theta)f(\theta)d\theta < 0.$$

Defining $y(\theta) = (1 - \pi_h(\theta))x(\theta)$, and

$$\lambda \equiv (u^l - u^h)/u^l$$

(the symbol λ has already been used in Example 1(a) but the definition above coincides with the meaning of λ in that example), the auxiliary problem is

$$\min_y \int_0^1 (V^h(\theta) - V^l(\theta))y(\theta)f(\theta)d\theta \tag{B.32}$$

subject to

$$0 \leq y(\theta) \leq x(\theta),\ \forall\theta, \tag{B.33}$$

$$(1-\lambda)x(\theta) + \lambda y(\theta) \text{ is non-decreasing in } \theta, \tag{B.34}$$

$$\int_0^1 (v^h(\theta) - v^l(\theta))x(\theta)f(\theta)d\theta \geq \int_0^1 (v^h(\theta) - v^l(\theta))y(\theta)f(\theta)d\theta, \tag{B.35}$$

Because both $V^h(\theta) - V^l(\theta)$ and $v^h(\theta) - v^l(\theta)$ are non-decreasing (by submodularity of the aftermarket, and co-modularity of the designer's and the third party's preferences), the optimal y pushes mass as far as possible to the left (in the sense explained in Appendix B.1, in the proof of Lemma 2). Formally, consider the following candidate solution, for some $\alpha \in [0,\, 1]$: $y_\alpha(\theta) = \max\{0,\, \tilde{y}_\alpha(\theta)\}$, where

$$\tilde{y}_\alpha(\theta) = \begin{cases} x(\theta) & \text{if } \theta < \alpha \\ x(\alpha) - \frac{1-\lambda}{\lambda}(x(\theta) - x(\alpha)) & \text{if } \theta \geq \alpha. \end{cases}$$

That is, $y_\alpha(\theta)$ is first equal to $x(\theta)$, then it is such that constraint (B.34) holds with equality, and then $y_\alpha(\theta) = 0$. I claim that for any feasible y', there exists α such that y_α achieves a weakly lower value of the objective function (B.32). Indeed, given y', define a function of the form $y_\alpha(\theta)$ such that

$$\int_0^1 y_\alpha(\theta)f(\theta)d\theta = \int_0^1 y'(\theta)f(\theta)d\theta.$$

Given a feasible $y'(\theta)$, an α that gives rise to the above equality can always be found.[35] Then, y' first-order stochastically dominates y_α, in the sense defined in previous proofs. Therefore, y^α is feasible, and achieves a weakly lower value of the objective function (B.32). Therefore, an optimal y can always be found in the form of y_α, for some $\alpha \in [0,\, 1]$.

Note that when $\lambda = 1$, this is the same solution as appeared in the proof of Theorem 2. However, when $\lambda < 1$, the optimal function $y_\alpha(\theta)$ is first non-decreasing, and then it might be strictly decreasing (if $x(\theta)$ is strictly increasing). When $y_\alpha(\theta)$ is decreasing, it decreases at exactly the rate which makes the constraint (B.34) bind.

In the second step, I prove that the optimal $x(\theta)$ is non-decreasing, and derive

[35] In particular, if $y'(\theta)$ satisfies condition (B.35), then the condition is preserved as mass is shifted to the left, because the function $v^h(\theta) - v^l(\theta)$ is non-decreasing.

further necessary conditions on the structure of the solution. First, I change variables in the problem (B.27) - (B.30). Let $z(\theta) = (1-\lambda)x(\theta) + \lambda y(\theta)$. Then, the problem becomes

$$\max_{x,\,z} \int_0^1 \left[\lambda V^h(\theta) + (1-\lambda)(V^h(\theta) - V^l(\theta))\right] x(\theta) f(\theta) d\theta - \int_0^1 \left[V^h(\theta) - V^l(\theta)\right] z(\theta) f(\theta) d\theta \tag{B.36}$$

subject to

$$z(\theta) \leq x(\theta) \leq \max\left\{1,\, \frac{1}{1-\lambda} z(\theta)\right\},\ \forall \theta, \tag{B.37}$$

$$z(\theta) \text{ is non-decreasing in } \theta, \tag{B.38}$$

$$\int_0^1 (v^h(\theta) - v^l(\theta)) x(\theta) f(\theta) d\theta \geq \int_0^1 (v^h(\theta) - v^l(\theta)) z(\theta) f(\theta) d\theta, \tag{B.39}$$

I fix an arbitrary feasible z, and show that the optimal x must be non-decreasing. By the usual argument (using the assumption that both $V^h(\theta)$ and $V^h(\theta) - V^l(\theta)$ are non-decreasing), the optimal $x(\theta)$ pushes mass as far as possible to the right. That is, the optimal $x(\theta)$ must be first equal to the lower bound $z(\theta)$, and then to the upper bound $max\left\{1,\, \frac{1}{1-\lambda} z(\theta)\right\}$. Because $z(\theta)$ is non-decreasing as well, it follows that $x(\theta)$ is non-decreasing.

Moreover, using the definition of z, the optimal solution has the following structure: for some $0 \leq \alpha \leq \beta \leq 1$,

$$\begin{cases} x(\theta) = y(\theta) \text{ is non-decreasing} & \theta < \alpha, \\ y(\theta) = 0 \text{ and } x(\theta) \text{ is non-decreasing} & \alpha < \theta < \beta, \\ x(\theta) = 1 \text{ and } y(\theta) \text{ is non-decreasing} & \beta < \theta. \end{cases}$$

Moreover, at the points α and β, monotonicity of z must be preserved. In particular, because y drops down to zero at α, x has to jump up at α, except for cases when $\alpha \in \{0,\, 1\}$.

Since we know that the optimal $x(\theta)$ is non-decreasing, we can now combine the structural insights about the solution from both auxiliary problems considered above. We know that $y(\theta)$ is non-decreasing and equal to $x(\theta)$ on $[0,\, \alpha]$, and then non-increasing until it hits zero. Thus, we can refine the structure of the optimal solution:

$$\begin{cases} x(\theta) = y(\theta) \text{ is non-decreasing} & \theta < \alpha, \\ y(\theta) = 0 \text{ and } x(\theta) \text{ is non-decreasing} & \alpha < \theta < \beta, \\ y(\theta) = 0,\ x(\theta) = 1 & \beta < \theta. \end{cases}$$

Given the above structure of the optimal solution, the optimization problem can be, without loss of generality, formulated as

$$\max_{x,\,\alpha} \int_0^\alpha V^l(\theta)x(\theta)f(\theta)d\theta + \int_\alpha^1 V^h(\theta)x(\theta)f(\theta)d\theta \tag{B.40}$$

subject to

$$x(\theta) \text{ is non-decreasing on } [0,\, \alpha) \cup (\alpha,\, 1] \tag{B.41}$$

$$x(\alpha^+) \geq \frac{1}{1-\lambda}x(\alpha^-), \tag{B.42}$$

$$\int_\alpha^1 (v^h(\theta) - v^l(\theta))x(\theta)f(\theta)d\theta \geq 0. \tag{B.43}$$

In the above problem, $x(\alpha^+)$ and $x(\alpha^-)$ denote the right and the left limits of $x(\theta)$ at α, respectively, where (by convention) $x(0^-) = 0$ and $x(1^+) = 1$. If $\alpha = 0$ and $x(\theta) = 1$ satisfy constraint (B.43), then this is clearly the optimal solution (given that $V^h(\theta) \geq V^l(\theta)$). However, such a solution is precluded by assumption (6.1).

Recall that $\underline{\theta}_l$ and $\underline{\theta}_h$ denote the points where V^l and V^h cross zero, respectively. I consider three candidate solutions, depending on whether *(i)* $\alpha = 0$, *(ii)* $\alpha = 1$, or *(iii)* $\alpha \in (0,\, 1)$ in the optimal solution. In the last step of the proof, I derive conditions for each of these three candidate solutions to be optimal.

Cases *(i)* and *(ii)* are relatively straightforward. In case *(i)*, because V^h is non-decreasing, the optimal solution is $x(\theta) = \mathbf{1}_{\{\theta \geq \gamma\}}$ for some γ. Let $\theta^\star$ denote the solution to the equation

$$\int_{\theta^\star}^1 (v^h(\theta) - v^l(\theta))f(\theta)d\theta \geq 0.$$

Then, $\gamma = \max\{\theta^\star,\, \underline{\theta}_h\}$. Assumption (6.1) implies that $\theta^\star \geq \underline{\theta}_h$, so that $\gamma = \theta^\star$. The optimal y is equal to 0 everywhere in this case. I will refer to this mechanism as mechanism *0*.

In case *(ii)*, it is optimal to set $x(\theta) = y(\theta) = \mathbf{1}_{\{\theta \geq \underline{\theta}_l\}}$. This is mechanism *1* from

Proposition 1.

Finally, I consider case *(iii)*. By considering the auxiliary problem in which $x(\alpha)$ is fixed, the problem can be decomposed into two independent parts, on $[0, \alpha]$, and on $[\alpha, 1]$. On $[0, \alpha]$, the optimal x must take the form $x(\theta) = x(\alpha)\mathbf{1}_{\{\theta \geq \underline{\theta}_l\}}$. On $[\alpha, 1]$, the optimal x takes the form $x(\theta) = \mathbf{1}_{\{\theta \geq \max\{\underline{\theta}_h, \theta^\star\}\}}$. If $\underline{\theta}_l \geq \alpha$, then we conclude that the optimal solution must take the form from case *(i)*. Therefore, because $\underline{\theta}_h \leq \underline{\theta}_l$, we must have $\underline{\theta}_l \leq \alpha \leq \theta^\star$, and $x(\theta) = \mathbf{1}_{\{\theta \geq \theta^\star\}}$ on $[\alpha, 1]$. However, it is not possible in case *(iii)* that x drops at α, and thus we must have $\alpha \geq \theta^\star$. Finally, inequality (B.34) must hold with equality at the optimal solution in case *(iii)*. Summarizing, we obtained,

$$x(\theta) = \begin{cases} 0 & \theta < \underline{\theta}_l, \\ 1 - \lambda & \underline{\theta}_l \leq \theta < \theta^\star, \\ 1 & \theta^\star \leq \theta. \end{cases}$$

The corresponding optimal y is equal to x on $[0, \theta^\star]$, and equal to $1 - \lambda$ on $[\theta^\star, 1]$, so that $y(\theta) = (1 - \lambda)\mathbf{1}_{\{\underline{\theta}_l \leq \theta \leq \theta^\star\}}$. This is mechanism *2* from Proposition 1.

Summing up, the optimal solution is one of three candidate solutions derived above. One can directly compare the expected payoffs from these three mechanisms to find the optimal one. The expected payoffs in cases *(i)* - *(iii)* are, respectively,

$$\int_{\theta^\star}^{1} V^h(\theta) f(\theta) d\theta, \tag{B.44}$$

$$\int_{\underline{\theta}_l}^{1} V^l(\theta) f(\theta) d\theta, \tag{B.45}$$

$$(1 - \lambda) \int_{\underline{\theta}_l}^{\theta^\star} V^l(\theta) f(\theta) d\theta + \int_{\theta^\star}^{1} V^h(\theta) f(\theta) d\theta. \tag{B.46}$$

Mechanism *0* (corresponding to the expected payoff (B.44)) is never strictly optimal because mechanism *2* always yields a weakly higher expected payoff. Therefore, either mechanism *1* or mechanism *2* is optimal.

B.8 Proof of Propositions 3b, 3c and 3a

The goal is calculate (the inverse of)

$$r^\star(\lambda) \equiv \sup \frac{\text{exp. payoff of the optimal mechanism}}{\text{exp. payoff of the optimal cutoff mechanism}},$$

where the supremum is taken over all parameters of the model subject to constraints imposed in Section 6.2, for fixed $\lambda \equiv (u^l - u^h)/u^h$. I use the notation from the proof of Proposition 1. Using Proposition 1 and Claim 10, I can write

$$r^\star(\lambda) \le \sup \frac{(1-\lambda)\int_{\underline{\theta}_l}^{\theta^\star} V^l(\theta) f(\theta) d\theta + \int_{\theta^\star}^1 V^h(\theta) f(\theta) d\theta}{\max\{\int_{\underline{\theta}_l}^1 V^l(\theta) f(\theta) d\theta,\ \int_{\theta^\star}^1 V^h(\theta) f(\theta) d\theta\}}.$$

Proof of Proposition 3a: The proof of Proposition 3a is immediate:

$$r^\star(\lambda) \le \frac{(1-\lambda)\int_{\underline{\theta}_l}^{\theta^\star} V^l(\theta) f(\theta) d\theta + \int_{\theta^\star}^1 V^h(\theta) f(\theta) d\theta}{\max\{\int_{\underline{\theta}_l}^1 V^l(\theta) f(\theta) d\theta,\ \int_{\theta^\star}^1 V^h(\theta) f(\theta) d\theta\}} \le \sup_{a \ge 0,\, b \ge 0} \frac{(1-\lambda)a + b}{\max\{a,\, b\}} = 2 - \lambda,$$

using the fact that $V^a(\theta)$ is non-decreasing.

For the remaining two cases, I assume that the model is that of Example 1(a).

Proof of Proposition 3b: For the problem of maximizing total surplus in Example 1(a), we have $\underline{\theta}_l = 0$. It is easy to see that in the worst-case scenario, the two candidate optimal cutoff mechanisms must yield exactly the same surplus, that is,

$$\int_0^1 V^l(\theta) f(\theta) d\theta = \int_{\theta^\star}^1 V^h(\theta) f(\theta) d\theta. \tag{B.47}$$

This allows me to formulate the problem as

$$\sup_{\theta_L, \theta_H, v, f} \frac{(1-\lambda)\int_0^{\theta^\star} V^l(\theta) f(\theta) d\theta + \int_{\theta^\star}^1 V^h(\theta) f(\theta) d\theta}{\int_{\theta^\star}^1 V^h(\theta) f(\theta) d\theta}, \tag{B.48}$$

subject to (B.47).

I can normalize one of parameters θ_L, θ_H, v because the numerator and the denominator can be divided by a constant without changing the value of the ratio. I choose a normalization such that $v = 1$ (then, $0 \le \theta_L \le \theta_H \le 1$).

Using the form of the objective function that arises under total surplus maximization, and in particular linearity in θ, I can write

$$\int_0^1 V^l(\theta)f(\theta)d\theta = V^l(\mathbb{E}_f[\theta]), \tag{B.49}$$

$$\int_{\theta^\star}^1 V^h(\theta)f(\theta)d\theta = (1-F(\theta^\star))V^h(\mathbb{E}_f[\theta|\theta \geq \theta^\star]), \tag{B.50}$$

$$\int_0^{\theta^\star} V^l(\theta)f(\theta)d\theta = F(\theta^\star)V^l(\mathbb{E}_f[\theta|\theta \leq \theta^\star]). \tag{B.51}$$

Therefore, the dependence of the ratio on f is only through three parameters: $\beta \equiv F(\theta^\star)$, $\bar{\theta} \equiv \mathbb{E}_f[\theta|\theta \geq \theta^\star])$, and $\underline{\theta} \equiv \mathbb{E}_f[\theta|\theta \leq \theta^\star])$. In particular, $\mathbb{E}_f[\theta] = \beta\underline{\theta} + (1-\beta)\bar{\theta}$. Moreover, by the definition of $\theta^\star$ (see equation 6.2), we have

$$(v-\theta_H)(1-F(\theta^\star)) = \int_{\theta^\star}^1 (v-\theta_L)(1-\theta)f(\theta)d\theta,$$

so that

$$\bar{\theta} = \frac{\theta_H - \theta_L}{v-\theta_L}.$$

The distribution parameters are only constrained by $\underline{\theta} \leq \bar{\theta}$, and in particular the ratio no longer depends explicitly on $\theta^\star$.

The next step is to solve for $\underline{\theta}$ using equality (B.47). Because (B.47) is linear in $\underline{\theta}$, we get a unique solution that we can plug back into (B.48) to obtain

$$r^\star(\lambda) \leq \sup_{\theta_L,\theta_H} \frac{(3\theta_H\lambda - 3\lambda - 2\theta_H - \theta_H\lambda^2 + \lambda^2 + 1)\theta_L + \theta_H^2\lambda^2 - 2\theta_H^2\lambda + \theta_H^2 - \theta_H\lambda^2 + \theta_H\lambda + \lambda}{(2\theta_H\lambda - 2\lambda - 2\theta_H + 1)\theta_L + \lambda - \theta_H^2\lambda + \theta_H^2}.$$

The derivative of the above expression with respect to θ_L is

$$\frac{\lambda(\theta_H-1)^2(\lambda-1)(\theta_H+\lambda-\theta_H\lambda)}{(\theta_L+\lambda-2\theta_H\theta_L-2\theta_L\lambda-\theta_H^2\lambda+\theta_H^2+2\theta_H\theta_L\lambda)^2} \leq 0,$$

because $\lambda \leq 1$, $\theta_H + \lambda \geq \theta_H\lambda$, and the denominator is non-negative. Therefore, it is optimal to set $\theta_L = 0$. Plugging this above, we can conclude that

$$r^\star(\lambda) \leq \sup_{\theta_H} \frac{\theta_H^2\lambda^2 - 2\theta_H^2\lambda + \theta_H^2 - \theta_H\lambda^2 + \theta_H\lambda + \lambda}{\lambda - \theta_H^2\lambda + \theta_H^2}.$$

Since an optimal mechanism is trivially a cutoff mechanism when either $\theta_H = \theta_L$ or $\theta_H = v$, there must be an interior solution for θ_H in the above problem. From the first-order condition, we obtain the following condition for the optimal θ_H:

$$\frac{2\theta_H + \lambda - 4\theta_H\lambda + 2\theta_H\lambda^2 - \lambda^2}{\lambda - \theta_H^2\lambda + \theta_H^2} = \frac{(2\theta_H - 2\theta_H\lambda)(\theta_H^2\lambda^2 - 2\theta_H^2\lambda + \theta_H^2 - \theta_H\lambda^2 + \theta_H\lambda + \lambda)}{(\lambda - \theta_H^2\lambda + \theta_H^2)^2}$$

Simplifying,

$$\lambda - 2h\lambda + \theta_H^2\lambda - \theta_H^2 = 0.$$

The above quadratic equation has only one solution in the feasible region, and it is equal to

$$\theta_H = \frac{\sqrt{\lambda}}{\sqrt{\lambda} + 1}.$$

Plugging the solution back into the ratio, we obtain,

$$r^\star(\lambda) \leq 1 + \frac{1}{2}(\sqrt{\lambda} - \lambda).$$

The inverse of this expression is the ratio in the statement of Proposition 3b. Finally, notice that the gap is largest for λ that maximizes $\sqrt{\lambda} - \lambda$, that is, for $\lambda = 1/4$. We have $r^\star(1/4) = 9/8$.

Proof of Proposition 3c: For the case of revenue maximization, we have $V^l(\theta) = \theta_L + (\theta_H - \theta_L)J(\theta)$, where $J(\theta)$ is the virtual surplus function (non-decreasing by assumption), and $V^h(\theta) = (1 - \lambda)V^l(\theta) + \lambda\theta_H$. Thus,

$$\int_{\theta^\star}^1 V^l(\theta)f(\theta)d\theta = (1 - F(\theta^\star))(\theta_L + (\theta_H - \theta_L)\theta^\star),$$

and we can write

$$\int_{\theta^\star}^1 V^h(\theta)f(\theta)d\theta = \left[(1 - \lambda) + \frac{\lambda\theta_H}{\theta_L + (\theta_H - \theta_L)\theta^\star}\right] \int_{\theta^\star}^1 V^l(\theta)f(\theta)d\theta.$$

Therefore,

$$r^\star(\lambda) \le \frac{(1-\lambda)\int_{\underline{\theta}_l}^{\theta^\star} V^l(\theta)f(\theta)d\theta + \int_{\theta^\star}^1 V^h(\theta)f(\theta)d\theta}{\max\{\int_{\underline{\theta}_l}^1 V^l(\theta)f(\theta)d\theta,\ \int_{\theta^\star}^1 V^h(\theta)f(\theta)d\theta\}}$$

$$\le \frac{(1-\lambda)\int_{\underline{\theta}_l}^{\theta^\star} V^l(\theta)f(\theta)d\theta + \left[(1-\lambda) + \frac{\lambda\theta_H}{\theta_L+(\theta_H-\theta_L)\theta^\star}\right]\int_{\theta^\star}^1 V^l(\theta)f(\theta)d\theta}{\int_{\underline{\theta}_l}^1 V^l(\theta)f(\theta)d\theta}$$

$$\le 1-\lambda+\frac{\lambda\theta_H}{\theta_L} = 1-\lambda(1-r). \quad \text{(B.52)}$$

Combining the bound established above with the bound from the proof of Proposition **3a**, we obtain that for revenue maximization:

$$r^\star(\lambda) \le \min\{2-\lambda,\ 1-\lambda(1-r)\}.$$

Taking the inverse of the above expression yields the conclusion of Proposition **3c**.

B.9 Proof of Theorem 4

The proof is similar to the proof of Lemma 1 except for the presence of an additional constraint – the Matthews-Border condition (MB). (The intuition behind the proof is identical but the formal argument must be appropriately adjusted.) As before, I will show that the optimal allocation rule x has to be non-decreasing. Then, the conclusion of Theorem 4 follows from Theorem 2.

I will only consider the case when the agent's utility is strongly supermodular and the aftermarket is submodular – this is the most difficult case, and the remaining cases are either analogous or straightforward given the proof of Lemma 1.

Under strong submodularity, at the optimal solution, $y(\theta) \equiv \pi_h(\theta)x(\theta)$ must be non-decreasing. Consider the problem of optimizing over x for a fixed non-decreasing y (such that there exists at least one feasible x; otherwise such y cannot be part of the optimal solution):

$$\max_x \int_0^1 V^l(\theta)x(\theta)f(\theta)d\theta,$$

subject to

$$y(\theta) \le x(\theta) \le 1,\ \forall\theta, \quad \text{(B.53)}$$

$$\int_{\tau}^{1} x(\theta)f(\theta)d\theta \leq \frac{1-F^N(\tau)}{N}, \forall \tau \in [0,\, 1]. \tag{B.54}$$

Notice that y must satisfy (B.54). Otherwise, because x must be point-wise higher than y, no x would be feasible for the above problem. Conditions (B.53) and (B.54) jointly imply that for any $\beta \leq \tau$,

$$\int_{\beta}^{\tau} y(\theta)d\theta + \int_{\tau}^{1} x(\theta)f(\theta)d\theta \leq \frac{1-F^N(\beta)}{N}.$$

Therefore, condition (B.54) can be sharpened to

$$\int_{\tau}^{1} x(\theta)f(\theta)d\theta \leq \Gamma(\tau) \equiv \min_{\beta\leq\tau} \underbrace{\left[\frac{1-F^N(\beta)}{N} - \int_{\beta}^{1} y(\theta)f(\theta)d\theta\right]}_{g(\beta)} + \int_{\tau}^{1} y(\theta)f(\theta)d\theta.$$

I denote by $\bar{x}(\tau)$ the function that is obtained by solving

$$\int_{\tau}^{1} x(\theta)f(\theta)d\theta = \Gamma(\tau) \tag{B.55}$$

point-wise for all τ ($\bar{x}$ is determined uniquely up to a zero-measure set of points). By assumption, $V^l(\theta)$ is non-decreasing. Thus, the optimal $x^\star$ pushes all mass to the right (in the formal sense explained in the proof of Lemma 2). This implies that there exists some threshold $\theta^\star$ such that $x^\star(\theta) = y(\theta)$ for $\theta < \theta^\star$ and $x^\star(\theta) = \bar{x}(\theta)$ for $\theta \geq \theta^\star$. Thus, to prove that $x^\star(\theta)$ is non-decreasing, I have to show that $\bar{x}(\theta)$ is non-decreasing and that it lies above $y(\theta)$.

Denote by $\beta^\star(\tau) = \max\{\text{argmin}_{\beta\leq\tau}\, g(\beta)\}$ the (largest) solution to the inner optimization problem in $\Gamma(\tau)$. We have either $\beta^\star(\tau) \in \{0,\, \tau\}$ (boundary solution) or $F^{N-1}(\beta^\star(\tau)) = y(\beta^\star(\tau))$ (an interior solution satisfying the first-order condition). As τ grows, larger values of β are feasible but the objective function does not depend on τ, so $\beta^\star(\tau)$ is non-decreasing. When $\beta^\star(\tau) < \tau$ in some neighborhood, then $\beta^\star(\tau)$ does not depend on τ, and differentiating (B.55) yields $\bar{x}(\theta) = y(\theta)$. Similarly, when $\beta^\star(\tau) = \tau$ in some neighborhood, we conclude that $\bar{x}(\theta) = F^{N-1}(\theta)$. In the case of a boundary solution, the (left) derivative of $g(\beta)$ at $\beta = \tau$ must be non-positive (otherwise the optimal β would be smaller than τ), that is, $F^{N-1}(\tau) \geq y(\tau)$. This shows that $\bar{x}(\theta) \geq y(\theta)$ for all θ. Finally, to show that $\bar{x}(\theta)$ is non-decreasing, I have to show that it is non-decreasing whenever it switches between $y(\theta)$ and $F^{N-1}(\theta)$ (I have to rule out the possibility of a downward jump). When $\bar{x}(\theta)$ switches from $y(\theta)$ to $F^{N-1}(\theta)$, this follows from the argument shown above. Suppose that $\bar{x}(\theta)$ switches from $F^{N-1}(\theta)$ to $y(\theta)$ at θ. Then, we must have $\beta^{\star}(\tau) = \theta$ for all τ in some right neighborhood of θ (because $\beta^{\star}(\theta) = \theta$, $\beta^{\star}(\tau)$ is non-decreasing and does not depend on τ when the solution is interior). It follows that the first-order condition holds at $\beta^{\star}(\theta) = \theta$ so that $F^{N-1}(\theta) = y(\theta)$. Thus, $\bar{x}(\theta)$ is in fact continuous at θ.

This proves that the optimal $x^{\star}(\theta)$ is non-decreasing. The conclusion now follows by applying Theorem 2.