WORKING PAPERS

N° 1245

September 2021

"Optimal Trade Mechanism with Adverse Selection and Inferential Mistakes"

Takeshi Murooka and Takuro Yamashita

Toulouse School of Economics

Optimal Trade Mechanisms with Adverse Selection and Inferential Mistakes*

Takeshi Murooka[†] Takuro Yamashita[‡]

September 4, 2021

Abstract

We study an adverse selection environment, where a rational seller can trade a good of which she privately knows its value to a buyer, and there are gains from trade. The buyer's types differ in their degree of inferential abilities: A rational type correctly infers the value of the good from the seller's offer, whereas a naive type under-appreciates the correlation between the seller's private information and offer. We characterize the optimal menu mechanism that maximizes the social surplus. Notably, no matter how severe the adverse selection is (in particular, even when no trade is the unique possible outcome if all agents are rational), *all* types of buyers trade in the optimal mechanism. The rational buyer's trade occurs at the expense of the naive buyer's losses. We also investigate a consumer-protection policy of limiting the losses and discuss its implications.

---

*We are grateful to Paul Heidhues, Botond Kőszegi, Klaus Schmidt, and seminar and conference audiences for helpful comments. Takeshi acknowledges financial support from JSPS KAKENHI (JP18H03640, JP19K01568, JP20K13451). Takuro acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (No. 714693), and ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). All remaining errors are our own.

†Osaka School of International Public Policy, Osaka University. murooka@osipp.osaka-u.ac.jp

‡Toulouse School of Economics, University of Toulouse, Capitole. takuro.yamashita@tse-fr.eu

1 Introduction

From the seminal work of Akerlof (1970), adverse selection and resulting market collapses have been extensively investigated. Samuelson (1984) and Attar, Mariotti, and Salanié (2011) show that, under a certain condition, a no-trade outcome is the only incentive-feasible allocation. These results build on a classical assumption that all parties are fully rational, in particular, an uninformed buyer can correctly infer the seller's type from observing her offer.[1] However, recent experimental and empirical evidence suggests that some buyers are systematically biased in that they fail to recognize the informational content of others' actions.[2] Eyster and Rabin (2005) propose a solution concept which incorporates such inferential mistakes (called a *cursed equilibrium*) in which a naive agent fails to infer other agents' private information from their actions.[3] Under severe adverse selection (i.e., an environment where only no-trade outcome is possible by rational agents) and the homogeneous type of buyer, Eyster and Rabin (2005) show that trade can occur between such naive buyers and (relatively) low-value sellers.[4]

Building on, but beyond these studies, we analyze one-sided adverse selection problems with *heterogeneous* inferential abilities. In the model, a rational seller has a single unit of a good $v \in [0, 1]$, which is her private information. Gains from trade is denoted by $\alpha > 1$ in that a buyer's value of the good is αv. Generalizing Eyster and Rabin (2005), we consider an envi-

[1]We use female pronouns to refer to a seller and male pronouns to refer to a buyer.

[2]See, for instance, Brown, Camerer, and Lovallo (2012), Enke and Zimmermann (2019), Enke (2020), and Jin, Luca, and Martin (2021). There is also accumulating evidence that, in some industries, a fraction of buyers make suboptimal inferences and purchase decisions; see Heidhues and Kőszegi (2018) on consumer behavior in markets and Beshears, Choi, Laibson, and Madrian (2018) on household decision makings.

[3]For theoretical works applying the concept of cursed equilibrium, see Eyster, Rabin, and Vayanos (2019), Kondor and Kőszegi (2017), and Ispano and Schwardmann (2017).

[4]As related theoretical studies, Jehiel (2005) and Jehiel and Koessler (2008) propose an *analogy-based expectation equilibrium* in which an agent bundles unobserved states into analogy classes. Spiegler (2016) develops a model of non-rational expectations based on a Bayesian-network approach. For reviews, see Eyster (2019) and Jehiel (2020).

ronment where the buyer is either a rational type (i.e., a standard Bayesian type) or a naive type (i.e., he under-appreciates the correlation between the seller's private information and action).

It turns out that environments with the co-existence of rational and naive types of buyers are fundamentally different from those with either a rational or naive type only. Notably, even with severe adverse selection where no trade would be the unique possible outcome if only rational agents existed, we show that, with the co-existence of rational and naive types, there exists a mechanism in which *both* types of buyers trade with positive probability. Intuitively, the trade between the naive buyer and the low-value seller generates *positive externalities* to other types: given that the (relatively) low-value seller trades with the naive buyer, there exists a mutually agreeable offer between the (relatively) high-value seller and the rational buyer, which generates a positive trade surplus.

We characterize optimal menu mechanisms when the buyer's type of inferential ability is heterogeneous. We show that the optimal mechanism always involves the rational buyer's trade with positive probability. The mechanism has the *double-separation property*: The seller types are separated so that the (relatively) low-value seller offers a low-price option that is traded with the naive buyer, and the higher-value seller takes a higher-price option that is traded with the rational buyer. When the buyer's value follows either the binary distribution or the power distribution (including the uniform distribution), the optimal menu mechanism has only two trading options. With a modest condition, the mechanism achieves a strictly higher expected total surplus than the one-option mechanism, and with cross-subsidization between different types of the buyer, ex ante Pareto improvement may also be possible.

Nevertheless, we show that it is impossible for both types of buyers to make non-negative expected payoffs. It implies that the naive buyer's loss is necessary for achieving a positive trade surplus.[5] Hence, as a potential

[5] In a companion paper (Murooka and Yamashita, 2020), we show that this impossibility

consumer-protection policy which can go together with the positive trade surplus, we investigate partial protection of the buyer in that the naive buyer's loss in terms of his (actual) expected payoff is limited to a certain amount. We characterize the optimal menu mechanism under the consumer protection. The optimal mechanism still exhibits the double-separation property, and it highlights a new trade-off between the naive buyer's protection and the social surplus.

This paper belongs to the growing literature on behavioral mechanism design and behavioral contract theory.[6] As most closely related to ours, Eliaz and Spiegler (2006, 2008) and Heidhues and Kőszegi (2010) investigate optimal contract designs with different degrees of agent naivety. In contrast to these studies, we focus on the market for lemons and shed light on the new positive externality to rational agents, as well as discussing its economic implications.

Our results are also related to the literature on behavioral industrial organization, in particular, interplays between rational and naive consumers. Most studies have focused on negative externalities imposed by rational consumers, e.g., cross-subsidization from naive to rational consumers (Gabaix and Laibson, 2006; Heidhues, Kőszegi, and Murooka, 2017; Jehiel, 2018).[7] For positive externalities from naive to rational consumers, de Clippel, Eliaz, and Rozen (2014) and Johnen (2020) show that equilibrium prices can be lower if there are more naive consumers in the market. Our result highlights a positive externality from naive to rational consumers too, but its logic is quite different: We show that the trade between a rational buyer and a (relatively) high-type seller occurs in the market for lemons, because a (relatively) low-type seller trades with a naive buyer. Based on this novel separation property, we investigate the welfare implications of the optimal

result holds for any kind of behavioral errors.

[6]See Kőszegi (2014) for a review.

[7]Herweg and Müller (2016) analyze the market for lemons with an overconfident buyer. Different from our results, they show that a high-quality good is traded with the overconfident buyer in equilibrium, whereas a rational buyer is crowded out of the market.

menu mechanism, the impossibility of full consumer protection, and the effect of partial consumer protection.

This paper is organized as follows. In Section 2, we set up a model of one-sided adverse selection and introduce each type of agents. In Section 3, we illustrate the properties of the optimal trade mechanism when the buyer's value follows the binary distribution. In Section 4, we analyze the optimal trade mechanism under general distributions and its properties. In Section 5, we investigate partial protection of the (naive) buyer. Section 6 discusses other extensions and concludes.

2 Model

A seller has a single unit of a good, whose value to her is $v \in [0,1]$ as her private information. A buyer's value is αv. The buyer does not know v, but only knows that it follows a distribution with cdf F. Let $\mu = E[v]$ denote its mean, and in case F admits a density, let f denote its density. They are risk neutral in monetary transfers. If the good is traded with probability $q \in [0,1]$ with monetary transfer $p \in \mathbb{R}$, then the seller's payoff is $p - vq$, and the buyer's payoff is $\alpha vq - p$. Hence, the social surplus from trade is given by $S = (\alpha - 1)vq$. We assume $\alpha > 1$ so that it is socially optimal to always make them trade. We also assume that α is not too large:

$$\alpha E[v|v \leq p] - p < 0, \tag{1}$$

for all $p > 0$.[8] As is well-known in the literature, Condition (1) implies *severe adverse selection* in the sense that only no-trade outcome can be incentive-feasible (if the agents are rational). To provide some intuition, note that, if the *reverse* of this inequality holds for some $p > 0$, then there is a simple posted-price scheme that induces some trading: Any seller type $v \leq p$ would wish to trade with price p, and the buyer would accept it as his expected

[8] For example, if v follows a uniform distribution, then the condition is equivalent to $\alpha < 2$.

payoff is $E[\alpha v - p | v \leq p] \geq 0$. Condition (1) says that, except for $p = 0$, such a mutually-agreeable price does not exist.

In the main part of the paper, we consider the following class of "menu mechanisms" as possible trading protocols. A *menu mechanism* is denoted by $(q_m, p_m)_{m=0}^M$, where $M \in \mathbb{N}$, $(q_0, p_0) = (0, 0)$, and for each $m = 1, \ldots, M$, $q_m \in [0, 1]$ and $p_m \in \mathbb{R}$. In the mechanism, the seller (knowing his v) chooses one option in the menu, say m. Then, the buyer, without knowing v but observing the seller's choice m, either accepts it or rejects it. If accepted, then the trade happens with probability q_m with monetary transfer p_m; otherwise, no trade occurs.[9] Our goal is to obtain insights about the mechanisms that are optimal in terms of the expected social surplus.

The key element of the analysis is the buyer's belief updating given the seller's choice in the menu. We assume that the buyer has one of the two types: a *rational* type, or a *naive* type. The rational type is a standard "Bayesian rational" type: Once he sees the seller's choice of (q_m, p_m), he correctly updates his belief about the seller's value v, and then makes an acceptance decision. The naive type, on the other hand, (incorrectly) believes that the seller's choice of (q_m, p_m) does not convey any information about v to him, and hence, he makes an acceptance decision without updating his belief. Let $\psi \in [0, 1]$ denote the fraction of the rational type among the buyers, and $1 - \psi$ denote the fraction of the naive type. Our formulation of the naive type coincides with the fully cursed agent in Eyster and Rabin (2005), as well as the coarsest analogy partition and the private information analogy partition in Jehiel and Koessler (2008). We discuss several kinds of extensions incorporating a partially-naive type in Section 6.1.

Given each v, let $\sigma(v) \in \{0, \ldots, M\}$ denote the seller's choice. Given the seller's choice $m \in \{0, \ldots, M\}$, let $d_r(m), d_n(m) \in \{0, 1\}$ denote the

[9] The class of mechanisms may appear restrictive; for example, the buyer cannot input any message other than his acceptance decision. Nevertheless, as long as a mechanism follows the *posterior implementability* developed by Green and Laffont (1987), the qualitative feature of the paper would not change with a more general class of mechanisms. We discuss this point in Section 6.2.

rational and naive buyer's acceptance decisions, respectively. A strategy profile $(\sigma(\cdot), d_r(\cdot), d_n(\cdot))$, or simply (σ, d), is a (Bayesian) equilibrium if it satisfies:

- (IC) Incentive compatibility for the seller: For each v and $m \in \{0, \ldots, M\}$,

$$\begin{aligned} & p_{\sigma(v)} - vq_{\sigma(v)}[\psi d_r(\sigma(v)) + (1-\psi)d_n(\sigma(v))] \\ \geq \; & p_m - vq_m[\psi d_r(m) + (1-\psi)d_n(m)]; \end{aligned}$$

- (IR-r) Individual rationality for the rational type of buyer: For each $m \in \{0, \ldots, M\}$,

$$\begin{aligned} & d_r(m) \int_0^1 (\alpha v q_m - p_m) 1_{\{\sigma(v)=m\}} dF \\ \geq \; & (1 - d_r(m)) \int_0^1 (\alpha v q_m - p_m) 1_{\{\sigma(v)=m\}} dF; \end{aligned}$$

- (IR-n) Individual rationality for the naive type of buyer: For each $m \in \{0, \ldots, M\}$,

$$d_n(m)(\alpha\mu q_m - p_m) \geq (1 - d_n(m))(\alpha\mu q_m - p_m).$$

Given an equilibrium, the expected social surplus is:

$$\int_0^1 (\alpha - 1) v q_{\sigma(v)} [\psi d_r(\sigma(v)) + (1-\psi) d_n(\sigma(v))] dF,$$

and therefore, an optimal mechanism is given by solving the following problem:

$$\begin{aligned} \max \quad & \int_0^1 (\alpha - 1) v q_{\sigma(v)} [\psi d_r(\sigma(v)) + (1-\psi) d_n(\sigma(v))] dF \\ \text{sub. to} \quad & \text{(IC), (IR-r), (IR-n).} \end{aligned}$$

Two benchmark results are well-known in the literature. First, if only the rational type exists ($\psi = 1$), then as in Samuelson (1984), a no-trade mechanism is optimal under Condition (1): $q(v) = p(v) = 0$ for all $v \in [0, 1]$.

Second, if only the naive type exists ($\psi = 0$), then as in Eyster and Rabin (2005), an optimal mechanism exhibits $M = 1$ (i.e., one-price mechanism) with $(q_1, p_1) = (1, \alpha\mu)$: In the mechanism, the seller with $v > \alpha\mu$ chooses $(q_0, p_0) = (0, 0)$, and hence no trade happens regardless of the buyer's choice; the seller with $v < \alpha\mu$ chooses $(q_1, p_1) = (1, \alpha\mu)$, and the (naive) buyer accepts it. The expected total surplus is:

$$\int_0^{\alpha\mu} (\alpha - 1) v dF \geq 0.$$

Despite its trade possibility, note that the buyer's expected payoff is necessarily negative due to Condition (1):

$$\begin{aligned} & \int_0^{\alpha\mu} (\alpha v - \alpha\mu) dF \\ = \; & \alpha E[v | v \leq \alpha\mu] - \alpha\mu < 0. \end{aligned}$$

3 Illustration with Binary-Value Distribution

Some of the main results of the paper (though not all) can be simply illustrated with a binary-value distribution: $v = 1$ with probability $\mu \in (0, 1)$ and $v = 0$ with probability $1 - \mu$ (and hence, $E[v] = \mu$). In this case, Condition (1) is equivalent to $\alpha\mu < 1$.

In the binary case, both benchmarks yield zero expected surplus: Even if only the naive type exists ($\psi = 0$), only the seller with $v = 0$ wishes to trade. So the resulting surplus is zero and the buyer earns a negative payoff $-\alpha\mu$.

Now we illustrate that, if $\psi \in (0, 1)$, then there exists a mechanism that achieves strictly positive expected surplus. In this sense, environments with co-existence of rational and naive types are fundamentally different from those with either a rational or naive type only.

We first analyze the case with $\psi < \frac{\alpha\mu}{1+\alpha\mu}$. Consider the menu mechanism where the seller chooses $(q, p) \in \{(0, 0), (1, 1), (1, \alpha\mu)\}$ (and the buyer accepts or rejects what the seller chooses). It is easy to verify that the following constitutes an equilibrium: The seller chooses $(q, p) = (1, 1)$ if $v = 1$, while

$(q, p) = (1, \alpha\mu)$ if $v = 0$; and the rational buyer accepts only $(q, p) = (1, 1)$, while the naive buyer accepts only $(q, p) = (1, \alpha\mu)$.[10]

The trading possibilities are summarized in the following table:

	naive	rational
$v = 0$	✓	
$v = 1$		✓

The expected surplus is now strictly positive: $\psi(\alpha - 1) > 0$. The intuition is simple. In this mechanism, the seller with $v = 0$ is matched with the naive buyer with price $\alpha\mu$. Given that they are out, now only the seller with $v = 1$ remains. Therefore, we can find a mutually agreeable price (e.g., $p = 1$) for the seller with $v = 1$ and the rational buyer, and their trade creates positive trade surplus. In this sense, the existence of the naive type generates a "positive externality" to the high-value seller and the rational buyer. On the other hand, note that this positive externality is at the cost of the naive buyer's loss.

The expected surplus is increasing in ψ as long as $\psi < \frac{\alpha\mu}{1+\alpha\mu}$, but this mechanism no longer works once ψ becomes higher than $\frac{\alpha\mu}{1+\alpha\mu}$. This is because, with $\psi > \frac{\alpha\mu}{1+\alpha\mu}$, even the low-value seller prefers $(q, p) = (1, 1)$ to $(1, \alpha\mu)$. In this case, we can modify the mechanism by partially suppressing the trade with price $p = 1$ in order to restore the incentive feasibility. More specifically, consider the modified menu mechanism where the seller chooses $(q, p) \in \{(0, 0), (q^*, 1), (1, \alpha\mu)\}$, where $q^* = \frac{(1-\psi)\alpha\mu}{\psi} < 1$. In this mechanism, the same "separating" strategy profile as above continues to be an equilibrium: The seller with $v = 0$ chooses $(q, p) = (1, \alpha\mu)$ which is only accepted by the naive buyer; while the seller with $v = 1$ chooses $(q, p) = (q^*, 1)$ which

[10] It is immediate that the naive buyer accepts only $(q, p) = (1, \alpha\mu)$. The rational buyer accepts only $(q, p) = (1, 1)$ because $(q, p) = (1, 1)$ means that the seller's type is $v = 1$ (and hence trading yields the rational buyer's expected payoff of $\alpha - 1 > 0$). Given the buyer's behavior, the seller with $v = 1$ earns zero payoff by choosing $(1, 1)$, while choosing $(1, \alpha\mu)$ yields a negative expected payoff; the seller with $v = 0$ prefers $(1, \alpha\mu)$ to $(1, 1)$ despite its lower price (recall $\alpha\mu < 1$), because $\psi < \frac{\alpha\mu}{1+\alpha\mu}$ is equivalent to $(1 - \psi)\alpha\mu > \psi$ and the seller with $v = 0$ prefers to sell the good to the naive buyer who only accepts $(1, \alpha\mu)$.

is only accepted by the rational buyer. In particular, note that the seller $v = 0$ (weakly) prefers $(q, p) = (1, \alpha\mu)$ to $(q^*, 1)$ because $(1 - \psi)\alpha\mu \geq \psi q^*$. The expected surplus is again strictly positive:

$$\psi q^*(\alpha - 1) = (1 - \psi)\alpha\mu(\alpha - 1) > 0.$$

Note that this is decreasing in ψ, and in particular, converging to zero as $\psi \to 1$, consistent with Samuelson (1984).

In both of the cases, the mechanisms are *doubly-separating*: The seller types are separated so that the low-value seller takes a low-price option that the high-value seller does not take, and the high-value seller takes a high-price option that the low-value seller does not take. The buyer types are separated too so that the naive type only accepts the low-price option, while the rational type only accepts the high-price option. The low-value seller is matched with the naive type, and the high-value seller is matched with the rational type. In the next section, we show that this double-separation property is a necessary feature of the optimal mechanism with a general class of distributions.

In terms of the induced trade surplus, there are several observations that do (or do not) continue to hold with more general distributions. First, because of the positive externality made by the naive buyer, the rational buyer earns a strictly higher expected payoff than under the "one-option" menu mechanism, $\{(q_m, p_m)\}_{m=0}^{M} = \{(0, 0), (1, \alpha\mu)\}$. On the other hand, the naive buyer and the seller continue to earn the same expected payoff. In this sense, relative to the optimal mechanisms in the two benchmarks, our mechanism attains Pareto improvement. With a general class of distributions, it is not necessarily the case, because the existence of the rational type may cause a *negative* externality to the naive type, which does not occur in the binary case.

Second, the expected surplus is non-monotone in ψ, and more importantly, the peak is with an interior ψ. We think it has an important implication in consumer education: The peak with $\psi > 0$ means that some consumer education could be socially beneficial, depending on the "current"

ψ. With a general class of distributions, the peak may or may not be with an interior ψ.

Finally, in the optimal mechanism above, the naive buyer earns a strictly negative payoff by matching with the low-value seller, while the rational buyer earns a strictly positive payoff. Thus, there might seem to be a room for cross-subsidization to make both buyer types better off (in the actual expected payoff).[11] However, we show that such cross-subsidization is impossible, regardless of the (binary or non-binary) distribution. That is, if a mechanism achieves strictly positive expected surplus, it necessarily implies that the naive buyer earns a strictly negative expected payoff.

4 Doubly-Separating Menu and Its Welfare Properties

In this section, we first show that the double-separation property is a necessary feature of the optimal mechanism with a general class of distributions. We then investigate welfare properties of the optimal mechanism.

4.1 Main theorem

We first state our main theorem that characterizes the optimal menu mechanism $\{(q_m, p_m)\}_{m=0}^{M}$, where $(q_0, p_0) = (0, 0)$ is (if any) chosen by the seller with the highest values and $(q_M, p_M) = (1, \alpha\mu)$ is chosen by the seller with the lowest values.

Theorem 1. An optimal menu mechanism $\{(q_m, p_m)\}_{m=0}^{M}$ has (i) $M \geq 1$, (ii) $(q_M, p_M) = (1, \alpha\mu)$, (iii) $\frac{p_m}{q_m} > \frac{p_{m+1}}{q_{m+1}}$ for $m = 1, \ldots, M-1$, and is (iv) associated with the following equilibrium: There exists $v^* \leq \alpha\mu$ such that the seller with $v < v^*$ chooses $m = M$, and the seller with $v > v^*$ chooses $m \neq M$;

[11] Recall that each type of the buyer's *perceived* expected payoff is always non-negative, although it is not correct for the naive type ex post.

the naive buyer accepts only (q_M, p_M), and the rational buyer accepts any (q_m, p_m) except (q_M, p_M).

As in the binary case, the mechanism separates the seller types based on some threshold type v^*. The optimal v^* depends on specific distributions. It also separates the naive buyer and the rational buyer. The fundamental intuition for this mechanism to achieve a positive surplus is similar to the binary case: The seller with $v < v^*$ is matched with the naive type by choosing a lower-price option, $(q_M, p_M) = (1, \alpha\mu)$. Given this, the seller with $v > v^*$ can trade with the rational buyer with higher-price options (q_m, p_m) where $m = 1, \ldots, M-1$.

In general, different from the binary distribution case, there may exist multiple options that further separate the high-value seller types. For a certain class of environments, however, at most two options (in addition to the no-trade option) would be enough.

Example 1. Let F be a power distribution, $F(v) = v^\gamma$ with $\gamma > 0$.[12] Then, $\mu = \frac{\gamma}{1+\gamma}$ and Condition (1) becomes $\alpha\mu < 1$.

In the power-distribution case, the optimal mechanism is a doubly-separating menu mechanism with only two trading options, as summarized in Proposition 1.

Proposition 1. Let F be a power distribution, $F(v) = v^\gamma$, with $\gamma > 0$ and $\alpha\mu < 1$. Then, the optimal menu mechanism is $\{(q_m, p_m)\}_{m=0}^M$ with $M = 2$, $(q_0, p_0) = (0, 0)$, $p_1 > \alpha\mu$, and $(q_2, p_2) = (1, \alpha\mu)$.

There are three properties in the optimal mechanism with the binary distribution: (i) The optimal mechanism (a two-option mechanism in the binary case) Pareto dominates the one-option mechanism; (ii) the expected total surplus is maximized at an interior ψ; and (iii) the naive-type buyer necessarily makes a loss. In the following subsections, we reexamine those properties in more general non-binary environments.

[12] A uniform distribution is a special case with $\gamma = 1$.

4.2 Suboptimality of one-price mechanism

The above theorem does not exclude a possibility that the one-option mechanism is optimal. The next proposition provides a sufficient condition with which a strictly better mechanism than the one-price mechanism exists.

Proposition 2. If $vf(v)$ is (locally) strictly increasing at $v = \alpha\mu$, then the optimal mechanism achieves strictly higher expected social surplus than the one-option mechanism.

Note that for any differentiable $f(\cdot)$, $(vf(v))' = f(v) + vf'(v)$, and the second term approaches 0 as v approaches 0. Hence, many popular distributions satisfy the condition if $\alpha\mu$ is sufficiently small. Proposition 2 implies that inducing the trade between a (relatively) high-value seller and a rational buyer is surplus-enhancing. To provide some intuition, consider the following two-option mechanism: $\{(0,0),(q,p),(1,\alpha\mu)\}$. The optimal one-option mechanism does not have the option (q,p), and thus, under the one-option mechanism, any seller with type $v < \alpha\mu$ trades with the naive buyer, while any seller type with $v > \alpha\mu$ does not trade. With the additional option (q,p) with $p > \alpha\mu q$, any seller type with $v < v^*$ trades with the naive buyer, and any seller type with $v \in (v^*, p)$ trades with the rational buyer. That $v^* < \alpha\mu$ may be interpreted as a "skimming" phenomenon: The seller with $v \in (v^*, \alpha\mu)$ corresponds to the high-end of those types who would trade with the naive buyer without the new option (q,p). With the new option, these sellers now trade with the rational buyer, and both the seller and the rational buyer are better off with this new option.

4.3 Pareto improvement relative to the one-option mechanism

Recall that, with a binary distribution, the two-option mechanism achieves "ex post" Pareto improvement relative to the one-option mechanism: The rational buyer is better off, and the naive buyer and the seller are indifferent.

With continuous distributions, it is impossible for the optimal mechanism to attain ex post Pareto improvement relative to the one-option mechanism, because the above skimming logic clearly implies that the naive buyer is worse off without cross-subsidization. However, it may be possible to achieve "ex ante" Pareto improvement: The buyer's average expected payoff of both types (with these types' probability weights) may be improved. And of course, if the buyer is ex ante Pareto improved, then appropriate cross-subsidization can make all the buyer types *ex post* Pareto improved.

Example 2.

1. By a continuity argument, even if F is continuous, if F is not too far from a binary distribution (e.g., a continuous bi-modal distribution with peaks close to $v = 0$ and $v = 1$), the optimal mechanism can achieve ex ante Pareto improvement relative to the one-option mechanism, and hence ex post Pareto improvement after appropriate cross-subsidization.

2. Let F be a power distribution, where $F(v) = v^\gamma$ with $\gamma > 0$. Then, under the optimal mechanism with any incentive-feasible cross-subsidization, the payoff of the naive type is strictly lower than the one-price mechanism.[13] Intuitively, this is because the seller receives most of the gains of the social surplus.

4.4 "Best mixture" of behavioral types and consumer education

As in the two benchmark cases, the expected surplus is zero if $\psi = 1$, and it is $\int_0^\alpha \mu(\alpha - 1)vdF$ if $\psi = 0$. Recall that, with a binary distribution, the social surplus is maximized with an interior ψ. This has an important implication for consumer education: If the "current" ψ is less than that achieves the

[13] Note that any cross-subsidy from a *seller* to a buyer is not possible in the optimal mechanism, as the seller's payoff for each v is uniquely pinned down up to a constant, and the seller with $v = 1$ earns zero payoff under the above optimal mechanism.

peak, then consumer education can increase the social surplus.[14]

Not all distributions induce an interior peak of ψ. For example, if $v \sim U(0,1)$ and $\alpha \in (\frac{4}{3}, 2)$, then the social surplus in the optimal mechanism is decreasing in ψ. Therefore, any consumer education reduces the social surplus in this case.

A sufficient condition for an interior peak is that the expected social surplus is increasing as a function of ψ, locally around $\psi = 0$:

Proposition 3. If there exists $p > \alpha\mu$ such that:

$$\int_{\alpha\mu}^{p} (vf(v) - \alpha\mu f(\alpha\mu))dv > \int_{0}^{\alpha\mu} vf(v)dv,$$

then the expected social surplus is maximized at some interior ψ.

To prove the result with some intuition, consider the following menu mechanism: $\{(q_m, p_m)\}_{m=0}^{M} = \{(0,0), (1,p), (1,\alpha\mu)\}$ with $p > \alpha\mu$. In the associated equilibrium, the seller with $v < v(\psi)$ chooses $(1, \alpha\mu)$, which is accepted only by the naive buyer; while the seller with $v \in (v(\psi), p)$ chooses $(1, p)$, which is accepted only by the rational buyer, where $v(\psi)$ represents the indifferent type:

$$\psi(p - v(\psi)) = (1-\psi)(\alpha\mu - v(\psi)) \;\Leftrightarrow\; v(\psi) = \alpha\mu - \frac{p - \alpha\mu}{1 - 2\psi}\psi,$$

which is positive if ψ is sufficiently close to 0.

The expected social surplus is increasing locally around $\psi = 0$ if:

$$\begin{aligned} &\frac{\partial}{\partial \psi}\left[(1-\psi)\int_{0}^{v(\psi)} (\alpha - 1)vdF + \psi \int_{v(\psi)}^{1} (\alpha-1)vdF\right]|_{\psi=0} \\ =&(\alpha - 1)\left[-\int_{0}^{\alpha\mu} vdF - \alpha\mu f(\alpha\mu)(p - \alpha\mu) + \int_{\alpha\mu}^{p} vdF\right] > 0. \end{aligned} \tag{2}$$

[14] However, a careful subsidy scheme may be necessary in its execution. Obviously, those who get educated are better off, while the other kinds of buyers, that is, those who are not educated (either naive before and after the policy, or rational before and after the policy), can be worse off.

The first term in the square brackets in 2 is a direct effect of decreasing a fraction of the naive buyer: The surplus created by the naive type trading with the seller with $v < \alpha\mu$ is lost, if a fraction of the naive buyer in the society is replaced by the rational buyer. Note that this negative effect does not exist in the binary case, because no surplus is directly created by the naive buyer (because he trades with the seller with $v = 0$ only). The third term in the square brackets in 2 is a direct effect of increasing the rational buyer, because this type of the buyer can enjoy a positive externality made by the naive type. The second term in the square brackets in 2 is negative externality made by the rational type on the naive type, based on the above "skimming" logic: With a higher fraction of the rational type in the society, the threshold type of the seller is lower, which implies a larger negative externality.

Example 3. Let F be a power distribution: $F(v) = v^\gamma$ with $\gamma > 0$. If $\gamma < \frac{1}{5}$ and $\alpha < \frac{5}{4}$, then the expected social surplus is maximized at $\psi > 0$.

4.5 Exploitation of naivete

As in the binary- or power-distribution cases, in the optimal menu mechanism, the naive buyer earns a negative expected payoff, while the rational buyer can earn a *strictly positive* expected payoff. This suggests a possibility of cross-subsidization from the rational buyer to the naive buyer, so that both types can enjoy non-negative expected payoffs.

The next result shows that this cannot be true in a strong sense. In *any* mechanism (not only the one that maximizes expected social surplus, but any other one, such as the one that maximizes only the buyer surplus) with non-trivial trading, the naive buyer's expected payoff must be strictly negative.

Proposition 4. If there exists a menu mechanism and its associated equilibrium where the buyer's ex-ante expected payoff is non-negative, then the expected social surplus must be 0, and in this sense, it is equivalent to the no-trade outcome.

Remark 1. The proof of the above result does not formally consider cross-subsidization, but it is immediate that the result implies impossibility of making both types earn non-negative expected payoffs with cross-subsidization. Indeed, if both types earn non-negative expected payoffs with cross-subsidization, then it must be the case that the buyer's ex ante expected payoff is non-negative, contradicting the proposition. In a companion paper (Murooka and Yamashita, 2020), we extend the above result: For any type of behavioral errors naive buyers make and for any incentive-feasible mechanism, either the expected payoff of some naive buyer must be strictly negative or the no-trade outcome occurs.

Proposition 4 implies that, to achieve a positive social surplus, we cannot "fully" protect the naive buyer by giving him a non-negative payoff. Still, it may be possible to "partially" protect the naive buyer by preventing him from a large negative payoff. We investigate this issue in Section 5.

5 (Partial) Protection of Naive Buyer

Proposition 4 in the previous section shows that a positive social surplus and a non-negative payoff of the naive buyer cannot go together, so we cannot "fully" protect the naive buyer under trade. However, it is possible to "partially" protect the naive type in the sense of avoiding large losses, by giving up (but not entirely) some trade surplus. In practice, a policymaker's goal may be to find a right balance between the social surplus and consumer protection.

In this section, we study the optimal mechanism under partial protection of the naive buyer. Specifically, consider the same surplus maximization problem as before, but with the additional constraint that the naive buyer's (actual) expected payoff is bounded from below by some fixed $-L \leq 0$:

$$\int_v (\alpha v q_{\sigma(v)} - p_{\sigma(v)}) d_c(\sigma(v)) \geq -L.$$

Call it Problem $(1; L)$. The previous sections consider the case with sufficiently large L, and Proposition 4 implies that $L = 0$ is possible only under the no-trade mechanism. We now investigate the case with intermediate $L > 0$.

An alternative, equivalent problem is that the principal maximizes a weighted sum of the social surplus and the naive buyer's (actual) expected payoff, where the weight is one for the social surplus and $\phi \geq 0$ for the naive buyer's expected payoff. Call it Problem $(2; \phi)$. A standard Lagrangian argument shows that each Problem $(1; L)$ is equivalent to Problem $(2; \phi)$ for some ϕ, and vice versa. This Problem $(2; \phi)$ can be interpreted as a situation where the principal has a higher "Pareto weight" for the naive buyer.

We show that the optimal menu mechanism is again doubly-separating, where the naive buyer trades with relatively low-value sellers (only) and the rational buyer trades with relatively high-value sellers (only). In this sense, the qualitative nature of the optimal mechanism is the same as in Theorem 1. However, in order to protect the naive buyer, the terms of trade may be modified.

Proposition 5. An optimal menu mechanism $\{(q_m, p_m)\}_{m=0}^{M}$ has $M \geq 1$, and there exists $m^* \in \{1, \ldots, M\}$ such that $(q_m, p_m) \leq (1, \alpha\mu)$ for $m \geq m^*$ and $p_m > \alpha\mu$ for $m < m^*$. It is associated with the following equilibrium: There exists v^* such that the seller with $v < v^*$ chooses some $m \geq m^*$, and the seller with $v > v^*$ chooses some $m < m^*$.

Example 4. Recall the binary example with $v \in \{0, 1\}$ with $\Pr(v = 1) = \mu$ and $\alpha\mu < 1$. The optimal menu mechanism with any L induces the same trading pattern: The seller with $v = 0$ trades with the naive buyer, and the seller with $v = 1$ trades with the rational buyer. However, the allocation that the low-value seller offers is not necessarily $(q, p) = (1, \alpha\mu)$, in order to protect the naive buyer.

Let $(q,p) \leq (1, \alpha\mu)$ be the allocation that the low-value seller offers and the naive buyer accepts; and let $(q', 1)$ be the allocation that the high-value seller offers and the rational buyer accepts.[15] Then, the surplus-maximization problem becomes as follows:

$$\begin{aligned} \max_{q,q',p} \quad & (\alpha - 1)\psi q' \\ \text{sub. to} \quad & (1-\psi) qp \geq q'\psi, \\ & p \leq \alpha\mu, \\ & -qp \geq -L, \end{aligned}$$

where the first constraint is the low-value seller's incentive compatibility, the second constraint is the naive buyer's individual rationality, and the third constraint corresponds to the protection of the naive buyer. The other constraints (i.e., the high-value seller's incentive compatibility and the rational buyer's individual rationality constraints) are omitted because they do not bind.

When L is sufficiently large (i.e., $L \geq \alpha\mu$), the constraint of protecting the naive buyer does not bind. Indeed, with $L \geq \alpha\mu$, the same optimal mechanism obtains: $q = 1$, $p = \alpha\mu$, and $q' = \min\{1, \frac{1-\psi}{\psi}\alpha\mu\}$.

When $L < \alpha\mu$, the constraint of protecting the naive buyer is binding. An optimal mechanism is $q = 1$, $p = L$, and $q' = \min\{1, \frac{1-\psi}{\psi} L\}$. First, this new mechanism protects the naive buyer by limiting the price to be L instead of $\alpha\mu$. Second, as a consequence, the trade probability between the high-value seller and the rational buyer can also be limited. Because of this second effect, the expected social surplus may become lower, but instead, the naive buyer is better off.

In Example 4, imposing L close to $\alpha\mu$ does not reduce the social sur-

[15] It is without loss to assume that the high-value seller's price in the optimal menu mechanism is 1. First, it cannot be strictly less than 1 because of the seller's individual rationality. Second, for any allocation for the high-value seller (and the rational buyer) with price strictly greater than 1, we can achieve a weakly higher expected surplus by lowering the price (and keeping the other parts of the allocations the same).

plus by much, while it increases the naive buyer's payoff. Indeed, when $\frac{1-\psi}{\psi}\alpha\mu < 1$, such a regulation affects neither the rational buyer's payoff nor the social surplus. In this sense, this consumer protection with a moderate level of $L > 0$ is in line with *asymmetric paternalism*: it benefits consumers who may make errors, while it imposes no (or relatively little) harm on rational consumers (Camerer, Issacharoff, Loewenstein, O'Donoghue, and Rabin, 2003).

The above consumer protection provides a rationale for caps on contract attributes and other regulations which moderately limit the ex-post losses of buyers. For example, the European Union's principle on unfair terms in consumer contracts prohibits a contractual term that "causes a significant imbalance [...] to the detriment of the consumer," so a transaction that is too disadvantageous to the consumer is disallowed.[16] Also, the Consumer Affairs Agency in Japan stipulates that consumer contract clauses with "a fixed penalty for contract cancellation in a total amount that exceeds the average amount of damages that the trader would incur from the cancellation" are void, which effectively limits the losses of buyers.[17]

6 Discussion and Concluding Remarks

In this section, we discuss (i) a model that incorporates a partially-naive buyer and (ii) analysis under more general mechanisms. While a thorough investigation of these exercises is beyond the scope of this paper, we illustrate a framework for the analysis and give a couple of remarks.

[16] Article 3 of directive 93/13/EEC of the Council of the European Commissions, available at `https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31993L0013:en:HTML#d1e245-29-1` (accessed September 1, 2021).

[17] Article 9 of the Consumer Contract Act (Japan), available at `http://www.japaneselawtranslation.go.jp/law/detail/?id=3578&vm=&re=` (accessed September 1, 2021).

6.1 Partial naivety

So far, we have analyzed the model in which a naive buyer is fully naive in the sense of Eyster and Rabin (2005). We now discuss how our analysis can be extended to cases in which a fraction of buyer is partially naive.

Motivated by Eyster and Rabin (2005), here we assume that a fraction $1-\psi$ of the buyers is *partially naive* in the following sense (and the other fraction ψ is rational, as before): Observing that the seller has chosen an option $m \in \{0,\dots,M\}$, the partially-naive buyer's anticipated value of the good is a weighted average of μ with weight $\chi \in [0,1]$ and $\tilde{\mu}_m$ with weight $1-\chi$, where $\tilde{\mu}$ is specified below. Note that our previous analysis covers the case with $\chi = 1$.

Then, (IR-n) becomes: For each $m \in \{0,\dots,M\}$,

$$\begin{aligned} & d_n(m)\{\alpha[\chi\mu + (1-\chi)\tilde{\mu}_m]q_m - p_m\} \\ \geq \ & (1-d_n(m))\{\alpha[\chi\mu + (1-\chi)\tilde{\mu}_m]q_m - p_m\}. \end{aligned}$$

The following specifications of $\tilde{\mu}_m$ may be possible depending on the contexts of interest. The first possibility is that $\tilde{\mu}_m$ is exogenously fixed and constant in m. In particular, when $\tilde{\mu}_m = E[v|v < \tilde{v}]$, it is as if the partially-naive buyer believes that the seller with $v > \tilde{v}$ does not trade at any price. The second, and perhaps more plausible, specification is that $\tilde{\mu}_m = E[v|v < p_m]$. In this case, with weight $1-\chi$, the naive buyer believes that only a seller whose value is lower than its associated price wishes to trade. In that sense, (with weight $1-\chi$) the naive buyer takes into account the fact that rational sellers never wish to trade at a price higher than own value. The third case is, in line with the definition of the partial cursedness in Eyster and Rabin (2005), that $\tilde{\mu}_m$ is based on the expected value. That is, $\tilde{\mu}_m = E[v|v \in V_m]$ where V_m is the set of the seller's types who choose option m in equilibrium, and in this sense, the partially-naive buyer's expectation is a weighted average of the fully-naive case (with weight χ) and rational case (with weight $1-\chi$).

In all of these three applications, our main messages and insights remain to hold: Naive buyers underestimate the selection and wish to trade at a relatively low price (which is lower than $\alpha\mu$ in these cases), while it creates a possibility of trades between (relatively) high-value sellers and rational buyers.

6.2 General mechanisms

In the main part of the paper, we only consider the class of menus. One may wonder if more general mechanisms, for example, by asking the buyer to reveal his behavioral type. In this section, we argue that the main qualitative results of the paper remain similar even if we consider more general mechanisms as long as they satisfy certain features that we think are relevant in the current trading context.

Formally, define a *general trading mechanism* by $(M_1, M_2; A_1, A_2; q, p)$ with finite M_1, M_2. Here, M_1 is interpreted as the seller's message set, M_2 is interpreted as the buyer's message set, $q : M_1 \times M_2 \to [0,1]$ is the trading probability, and $p : M_1 \times M_2 \to [0,1]$ is the transfer. For each $i \in \{1,2\}$, the binary set $A_i = \{0,1\}$ represents i's *acceptance* decision, and the trade with (q,p) happens if and only if both i choose $a_i = 1$. The timing of the mechanism is as follows: First, each agent i simultaneously chooses $m_i \in M_i$, and after observing $m = (m_1, m_2)$, each agent i simultaneously chooses $a_i \in A_i = \{0,1\}$. Consider a (weak) perfect-Bayesian equilibrium as the solution concept.

Although this formulation might look non-standard, this is to capture the key feature of the lemon's problem that the buyer makes an inference about the seller's type observing the seller's choice, and a trade is only upon mutual agreement.[18] See the seminal work by Green and Laffont (1987) which intro-

[18]Indeed, if a mechanism can ask both the seller and the buyer report both m_i and a_i *simultaneously* (which essentially means that the distinction between m_i and a_i is non-essential), then higher expected social surplus might be possible. As extensively discussed in Green and Laffont (1987), however, we think that the improvement is not well-justified within the lemon's trading context.

duces and studies some general properties of such *posterior implementation.*

This class of general trading mechanisms is richer than the class of menu mechanisms considered in the paper. Indeed, a menu mechanism is a special case of a general trading mechanism with the additional requirement of trivial M_2. In this sense, it is potentially possible that some non-menu mechanism strictly improves the optimal menu mechanism. Although further investigations of general mechanisms are left for future research, such mechanisms should possess similar qualitative features as in the menu mechanism. In particular, there must exist $v^* > 0$ such that the seller with $v < v^*$ trades only with the naive buyer. On the other hand, because the naive buyer does not accept any trade with price greater than $\alpha\mu$, the seller with $v > \alpha\mu$ can trade only with the rational buyer. In this sense, the outcome must exhibit some form of "double-separation" as in Theorem 1.

Proposition 6. In a (non-trivial) general trading mechanism $(M_1, M_2; A_1, A_2; q, p)$, there exists $v^* > 0$ such that the seller with $v < v^*$ trades only with the naive buyer, and the seller with $v > \alpha\mu$ trades only with the rational buyer.

For example, when a distribution is binary (i.e., $v \in \{0, 1\}$ with $\alpha\mu < 1$), Proposition 6 immediately implies that the menu mechanism considered in the previous sections is optimal among all general trading mechanisms.

References

Akerlof, G. A. (1970): "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84(3), 488–500.

Attar, A., T. Mariotti, and F. Salanié (2011): "Nonexclusive Competition in the Market for Lemons," *Econometrica*, 79(6), 1869–1918.

Beshears, J., J. J. Choi, D. Laibson, and B. C. Madrian (2018): "Behavioral Household Finance," in *Handbook of Behavioral Economics: Foundations and Applications 1*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, vol. 1, chap. 3, pp. 177–276. North Holland.

Brown, A. L., C. F. Camerer, and D. Lovallo (2012): "To Review or Not to Review? Limited Strategic Thinking at the Movie Box Office," *American Economic Journal: Microeconomics*, 4(2), 1–26.

Camerer, C. F., S. Issacharoff, G. Loewenstein, T. O'Donoghue, and M. Rabin (2003): "Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism"," *University of Pennsylvania Law Review*, pp. 1211–1254.

de Clippel, G., K. Eliaz, and K. Rozen (2014): "Competing for Consumer Inattention," *Journal of Political Economy*, 122(6), 1203–1234.

Eliaz, K., and R. Spiegler (2006): "Contracting with Diversely Naive Agents," *Review of Economic Studies*, 73(3), 689–714.

——— (2008): "Consumer Optimism and Price Discrimination," *Theoretical Economics*, 3, 459–497.

Enke, B. (2020): "What You See Is All There Is," *Quarterly Journal of Economics*, 135(3), 1363–1398.

Enke, B., and F. Zimmermann (2019): "Correlation Neglect in Belief Formation," *Review of Economic Studies*, 86(1), 313–332.

Eyster, E. (2019): "Errors in Strategic Reasoning," in *Handbook of Behavioral Economics: Foundations and Applications 2*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, vol. 2, chap. 3, pp. 187–259. North Holland.

Eyster, E., and M. Rabin (2005): "Cursed Equilibrium," *Econometrica*, 73(5), 1623–1672.

Eyster, E., M. Rabin, and D. Vayanos (2019): "Financial Markets Where Traders Neglect the Informational Content of Prices," *Journal of Finance*, 74(1), 371–399.

Gabaix, X., and D. Laibson (2006): "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 121(2), 505–540.

Green, J. R., and J.-J. Laffont (1987): "Posterior Implementability in a Two-Person Decision Problem," *Econometrica*, 55(1), 69–94.

Heidhues, P., and B. Kőszegi (2010): "Exploiting Naivete about Self-Control in the Credit Market," *American Economic Review*, 100(5), 2279–2303.

——— (2018): "Behavioral Industrial Organization," in *Handbook of Behavioral Economics: Foundations and Applications 1*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, vol. 1, chap. 6, pp. 517–612. North Holland.

Heidhues, P., B. Kőszegi, and T. Murooka (2017): "Inferior Products and Profitable Deception," *Review of Economic Studies*, 84(1), 323–356.

Herweg, F., and D. Müller (2016): "Overconfidence in the Markets for Lemons," *The Scandinavian Journal of Economics*, 118(2), 354–371.

Ispano, A., and P. Schwardmann (2017): "Competion over Cursed Consumers," Working Paper.

Jehiel, P. (2005): "Analogy-Based Expectation Equilibrium," *Journal of Economic Theory*, 123(2), 81–104.

——— (2018): "Investment Strategy and Selection Bias: An Equilibrium Perspective on Overoptimism," *American Economic Review*, 108(6), 1582–1597.

——— (2020): "Analogy-Based Expectation Equilibrium and Related Concepts: Theory, Applications, and Beyond," prepared for the World Congress of the Econometric Society, Milan 2020.

Jehiel, P., and F. Koessler (2008): "Revisiting Games of Incomplete Information with Analogy-Based Expectations," *Games and Economic Behavior*, 62(2), 533–557.

Jin, G. Z., M. Luca, and D. Martin (2021): "Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure," *American Economic Journal: Microeconomics*, 13(2), 141–73.

Johnen, J. (2020): "Dynamic Competition in Deceptive Markets," *RAND Journal of Economics*, 51(2), 375–401.

Kőszegi, B. (2014): "Behavioral Contract Theory," *Journal of Economic Literature*, 52(4), 1075–1118.

Kondor, P., and B. Kőszegi (2017): "Financial Choice and Financial Information," Working Paper.

Murooka, T., and T. Yamashita (2020): "A Note on Adverse Selection and Bounded Rationality," Working Paper.

Samuelson, W. (1984): "Bargaining under Asymmetric Information," *Econometrica*, 52(4), 995–1005.

Spiegler, R. (2016): "Bayesian Networks and Boundedly Rational Expectations," *Quarterly Journal of Economics*, 131(3), 1243–1290.

A Proofs

Proof of Theorem 1.

The following lemma collects some key properties of the equilibria. Fix any menu $(q_m, p_m)_{m=0}^M$ and any associated equilibrium (σ, d).

Let $V_m \subseteq [0,1]$ denote the set of the seller types who choose m, i.e., $V_m = \{v \in [0,1] | \sigma(v) = m\}$. By the single-crossing property, each V_m must be an interval.

Without loss of generality, assume that each V_m for $m = 1, \dots, M$ is non-trivial in that $\inf V_m < \sup V_m$; and that $m > m'$ implies $v \leq v'$ for all $v \in V_m$, $v' \in V_{m'}$.

Let $v_0 = 1$, $v_m = \sup V_m$ for each $m \neq 0$, and $v_{M+1} = 0$. Because each V_m is non-trivial for $m \neq 0$, we have $1 = v_0 \geq v_1 > \cdots > v_{M+1} = 0$.

For each $m \geq 1$, define:

$$x_m = (\psi d_r(m) + (1-\psi) d_n(m)) q_m.$$
$$y_m = (\psi d_r(m) + (1-\psi) d_n(m)) p_m.$$

Without loss of generality, assume $x_m > 0$ for all $m = 1, \dots, M$. Let $m^* = \max\{m | d_n(m) = 0\}$, which is well-defined by assuming $d_n(q_0, p_0) = 0$ without loss of generality.

Lemma 1.

1. $y_m - x_m v_m = y_{m-1} - x_{m-1} v_m$ for $m = 2, \dots, M$; and $y_1 - x_1 v_1 \geq 0$;
2. $x_{m+1} > x_m$ and $\frac{y_{m+1}}{x_{m+1}} < \frac{y_m}{x_m}$ for $m = 1, \dots, M-1$;
3. $d_r(M) = 0$, and hence, $x_M \leq 1 - \psi$;
4. $y_m \leq \alpha \mu x_m$ if $m > m^*$;
5. $y_m \geq \alpha \mu x_m$ if $1 \leq m \leq m^*$;
6. $x_{m^*} \leq \psi$;

7. $\alpha E[v|v \in V_m] \geq \frac{y_m}{x_m}$ if $m \leq m^*$.

Proof. The first property is implied by the seller's incentive compatibility (for $m \geq 2$) and individual rationality (for $m = 1$).

The second property is implied by the seller's incentive compatibility and the assumption that each V_m is non-trivial for $m \neq 0$.

For the third property, if $d_r(M) = 1$, then the rational buyer's individual rationality constraint implies:

$$0 \leq \alpha E[v|v \in V_M] - p_M \leq \alpha E[v|v \in (0, v_M)] - v_M,$$

contradicting our assumption of severe adverse selection given by Inequality (1).

For the fourth property, because the naive buyer cannot accept any price above $\alpha\mu$, we have $y_{m^*+1} \leq \alpha\mu x_{m^*+1}$. Then, the second property implies $y_m \leq \alpha\mu x_m$ for all $m > m^*$.

Conversely, for the fifth property, $d_n(m^*) = 0$ implies $y_{m^*} \geq \alpha\mu x_{m^*}$, and by the first property, $y_m > \alpha\mu x_m$ if $1 \leq m \leq m^*$. This also implies $d_n(m) = 0$ if $1 \leq m \leq m^*$.

The sixth property is immediate from $d_n(m^*) = 0$.

For the last property, because $m < m^*$, we have $d_n(m) = 0$. However, because $x_m > 0$, it must be that $d_r(m) > 0$. Thus, the rational buyer's individual rationality constraint implies $\alpha E[v|v \in V_m] q_m \geq p_m$. □

Now consider the problem of maximizing the expected social surplus among all possible $((v_m, x_m, y_m)_{m=1}^M, m^*)$ for some $M \leq \overline{M}$ that satisfy the seven properties above:

$$\begin{aligned} \max_{M \leq \overline{M};\ ((v_m, x_m, y_m)_{m=1}^M, m^*)} \quad & \sum_m \int_{v_{m+1}}^{v_m} x_m(\alpha - 1) v dF \\ \text{sub. to} \quad & m^* \in \{0, \ldots, M\},\ x_m \in [0,1],\ 0 = v_{M+1} < \ldots < v_1 \leq v_0 = 1 \\ & \text{(Properties 1-7 above).} \end{aligned}$$

Observe that the above problem admits a solution, even though some constraints are based on strict inequalities. First, consider a set of restricted problems where M is fixed (to any number between 1 and $\overline{M}$), and let $\text{Sol}(M)$ denote the solution of this restricted problem with fixed M, in case it exists. Let $\mathcal{M} \subseteq \{1, \ldots, \overline{M}\}$ denote the set of M where $\text{Sol}(M)$ exists. If this set is nonempty, then clearly the solution of the unrestricted problem (i.e., with flexible $M \leq \overline{M}$) is in $\{\text{Sol}(M) | M \in \mathcal{M}\}$; and indeed, it is nonempty, because $\text{Sol}(1)$ exists: By Eyster and Rabin (2005), a one-price menu mechanism $\{(0,0), (1, \alpha\mu)\}$ is known to be optimal.

It is straightforward to see that a solution to the above problem implies a menu mechanism that indeed achieves the maximized value.

Now we show that the solutions in the above (unrestricted) problem is a doubly-separating menu mechanism $(q_m, p_m)_{m=0}^{M}$ (for some M) with $(q_0, p_0) = (0, 0)$ and $(q_M, p_M) = (1, \alpha\mu)$.

Lemma 2. For the solution in the above problem, $m^* = M - 1$, $y_M = \alpha\mu x_M$, and $x_M = 1 - \psi$.

Proof. Let $((v_m, x_m, y_m)_{m=1}^{M}, m^*)$ denote a feasible policy in the above problem such that one of the stated conditions is violated. If $m^* < M - 1$, then $\alpha\mu \geq \frac{y_{M-1}}{x_{M-1}} > \frac{y_M}{x_M}$, and thus, in what follows, we focus on the two cases: either $y_M < \alpha\mu x_M$, or $x_M < 1 - \psi$.

Case 1: $y_M < \alpha\mu x_M$

In this case, consider the alternative policy $((\hat{v}_m, \hat{x}_m, \hat{y}_m)_{m=1}^{M}, \hat{m}^*)$ where, for sufficiently small $\varepsilon > 0$: $\frac{\hat{y}_M}{x_M} = \frac{y_M}{x_M} + \varepsilon < \min\{\frac{y_{M-1}}{x_{M-1}}, \alpha\mu\}$; $\hat{v}_M (\in (v_M, 1))$ is such that $\hat{y}_M - x_M \hat{v}_M = y_{M-1} - x_{M-1}\hat{v}_M$; and all the other components are the same as $((v_m, x_m, y_m)_{m=1}^{M}, m^*)$. Clearly, this new policy is feasible in the above problem, but achieves a strictly higher expected surplus, as the probability of trade increases from x_{M-1} to x_M for $v \in (v_M, \hat{v}_M)$.

Case 2: $x_M < 1 - \psi$.

In this case, consider the alternative policy $((\hat{v}_m, \hat{x}_m, \hat{y}_m)_{m=1}^{M}, \hat{m}^*)$ where, for sufficiently small $\varepsilon > 0$: $\hat{x}_M = x_M + \varepsilon (< 1 - \psi)$; $\hat{v}_M (\in (v_M, 1))$ is such that $\frac{y_M}{x_M}\hat{x}_M - \hat{x}_M \hat{v}_M = y_{M-1} - p_{M-1}\hat{v}_M$; and all the other components are the same as $((v_m, x_m, y_m)_{m=1}^M, m^*)$. Clearly, this new policy is feasible in the above problem, but achieves strictly higher expected surplus, as the probability of trade increases from x_M to $\hat{x}_M$ for $v < \hat{v}_M$. $\square$

Proof of Proposition 1.

Suppose that there is an optimal menu mechanism with $M \geq 3$, which we denote by $\{(q_m, p_m)\}_{m=0}^M$. It suffices to construct another optimal menu mechanism with $M = 2$.

It is without loss to assume that $\frac{p_m}{q_m}$ is decreasing in m, $\frac{p_M}{q_M} = \alpha\mu < \ldots < \frac{p_1}{q_1} \leq 1$, $q_M = 1$, and that there exists a sequence of threshold seller types $0 = v_{M+1} < \ldots < v_1 \leq v_0 = 1$ such that any seller type $v \in (v_{m+1}, v_m)$ chooses option (q_m, p_m) in the associated equilibrium. More specifically, we have $v_1 \leq \frac{p_1}{q_1}$, v_2 is the seller type who is indifferent between (q_1, p_1) and (q_2, p_2), and so on. Without loss, we can assume that the rational buyer's individual rationality given (q_2, p_2) is binding: $\alpha E[v | v \in (v_3, v_2)] = \frac{p_2}{q_2}$.[19]

Consider an alternative mechanism, which is exactly the same as the original one, except that, instead of (q_1, p_1) and (q_2, p_2), it offers (q_2', p_2') (together with $\{(q_m, p_m)\}_{m=3}^M$) where the same v_3 is indifferent between (q_2', p_2') and (q_3, p_3), and the rational buyer's individual rationality holds with equality, that is, $\alpha E[v | v \in (v_3, k_2')] = k_2'$, where $k_2' = \frac{p_2'}{q_2'}$. Let $k_2 = \frac{p_2}{q_2}$. In what follows, we focus on the case where $k_2' \leq 1$; the other case is similar.

The expected social surplus in the new mechanism is higher than in the

[19]If $\alpha E[v | v \in (v_3, v_2)] > \frac{p_2}{q_2}$, then we can be strictly better off by modifying the mechanism as follows. Fix $\varepsilon > 0$ small, and consider an alternative mechanism $\{(\hat{q}_m, \hat{p}_m)\}_{m=0}^M$, with $\hat{p}_2 = p_2 - \varepsilon$, $\hat{q}_2 = q_2 - \frac{\varepsilon}{v_3}$, and the rest is the same as $\{(q_m, p_m)\}_{m=0}^M$. The same thresholds obtain for the seller types, except that threshold v_2 becomes $\hat{v}_2 = \frac{p_2 - p_1 - \varepsilon}{q_2 - q_1 - \frac{\varepsilon}{v_3}}$. Note that all the constraints on the buyer side continue to be satisfied. The welfare change with small ε is approximately $(v_2 - v_3)v_2 f(v_2) - \int_{v_3}^{v_2} v dF > 0$ as $vf(v)$ is increasing for a power distribution.

original mechanism by at least Δ, where:

$$\begin{aligned}
\Delta &= q_2' \int_{v_3}^{k_2'} xdF - q_2 \int_{v_3}^{v_2} xdF - q_1 \int_{v_2}^{v_1} xdF \\
&\propto \int_{v_3}^{k_2'} xdF \frac{k_2 - v_3}{k_2' - v_3} - \int_{v_3}^{v_2} xdF - \int_{v_2}^{v_1} xdF \frac{k_2 - v_2}{v_1 - v_2}.
\end{aligned}$$

Note that the new mechanism has exactly one less option than in the original mechanism. Therefore, if $\Delta \geq 0$, then we complete the proof by induction.

From here on, the notation is modified as follows: We use v instead of v_3, βv instead of v_2, and δv instead of k_2'. Note that $\beta \geq 1$ is a free parameter, while δ is determined as a function of F and α:

$$\delta v = \frac{\alpha \int_v^{\delta v} xdF}{F(\delta v) - F(v)}.$$

By construction, we have $\delta \geq \beta \geq 1$.

We now show that $\Delta \geq 0$. First, observe that $v_1 \leq \delta\beta v$. Thus:

$$\Delta \propto \int_v^{\delta v} xdF \frac{k_2 - v}{\delta v - v} - \int_v^{\beta v} xdF - \int_{\beta v}^{v_1} xdF \frac{k_2 - \beta v}{v_1 - \beta v}.$$

Because $xf(x)$ is increasing, this is at least:

$$\begin{aligned}
& \int_v^{\delta v} xdF \frac{k_2 - v}{\delta v - v} - \int_v^{\beta v} xdF - \int_{\beta v}^{\delta\beta v} xdF \frac{k_2 - \beta v}{\delta\beta v - \beta v} \\
\propto\ & \beta \int_v^{\delta v} xdF(k_2 - v) - \beta v \int_v^{\beta v} xdF(\delta - 1) - \int_{\beta v}^{\delta\beta v} xdF(k_2 - \beta v) \\
\equiv\ & \Delta'.
\end{aligned}$$

Note that

$$\begin{aligned}
k_2 &= \frac{\alpha \int_v^{\beta v} xdF}{F(\beta v) - F(v)} \\
&= \frac{(F(\delta v) - F(v)) \int_v^{\beta v} xdF}{(F(\beta v) - F(v)) \int_v^{\delta v} xdF} \delta v,
\end{aligned}$$

and thus,

$$\begin{aligned}
k_2 - v &= \frac{\delta(F(\delta v) - F(v)) \int_v^{\beta v} xdF - (F(\beta v) - F(v)) \int_v^{\delta v} xdF}{(F(\beta v) - F(v)) \int_v^{\delta v} xdF} v, \\
k_2 - \beta v &= \frac{\delta(F(\delta v) - F(v)) \int_v^{\beta v} xdF - \beta(F(\beta v) - F(v)) \int_v^{\delta v} xdF}{(F(\beta v) - F(v)) \int_v^{\delta v} xdF} v.
\end{aligned}$$

Hence:

$$\begin{aligned}
\Delta' &\propto \beta \int_v^{\delta v} xdF \left(\delta(F(\delta v) - F(v)) \int_v^{\beta v} xdF - (F(\beta v) - F(v)) \int_v^{\delta v} xdF \right) \\
&\quad -\beta \int_v^{\beta v} xdF (\delta - 1)(F(\beta v) - F(v)) \int_v^{\delta v} xdF \\
&\quad - \int_{\beta v}^{\delta \beta v} xdF \left(\delta(F(\delta v) - F(v)) \int_v^{\beta v} xdF - \beta(F(\beta v) - F(v)) \int_v^{\delta v} xdF \right) \\
&= \beta \int_v^{\delta v} xdF \left(\delta(F(\delta v) - F(\beta v)) \int_v^{\beta v} xdF - (F(\beta v) - F(v)) \int_{\beta v}^{\delta v} xdF \right) \\
&\quad - \int_{\beta v}^{\delta \beta v} xdF \left(\delta(F(\delta v) - F(v)) \int_v^{\beta v} xdF - \beta(F(\beta v) - F(v)) \int_v^{\delta v} xdF \right).
\end{aligned}$$

Because $F(\delta v)F(\beta v) = F(\delta\beta v)F(v)$ and $\int_a^b xdF = c(bF(b) - aF(a))$ for any $a < b$ for some constant $c > 0$, we finally have:

$$\begin{aligned}
\Delta' &\propto F(v) \left(\delta(F(\delta v) - F(\beta v)) \int_v^{\beta v} xdF - (F(\beta v) - F(v)) \int_{\beta v}^{\delta v} xdF \right) \\
&\quad -F(\beta v) \left(\delta(F(\delta v) - F(v)) \int_v^{\beta v} xdF - \beta(F(\beta v) - F(v)) \int_v^{\delta v} xdF \right) \\
&= \delta F(v)(F(\delta v) - F(\beta v))(\beta v F(\beta v) - vF(v)) - F(v)(F(\beta v) - F(v))(\delta v F(\delta v) - \beta v F(\beta v)) \\
&\quad -\delta F(\beta v)(F(\delta v) - F(v))(\beta v F(\beta v) - vF(v)) - \beta F(\beta v)(F(\beta v) - F(v))(\delta v F(\delta v) - vF(v)) \\
&= 0. \quad \square
\end{aligned}$$

Proof of Proposition 2.

We consider a two-option mechanism of the form: $\{(0,0), (q,p), (1, \alpha\mu)\}$, where $q \simeq 0$, $p = kq$ with $k > \alpha\mu$. We show that, if $vf(v)$ is (locally) strictly increasing at $v = \alpha\mu$, then this two-option mechanism achieves strictly higher expected welfare than the (optimal) one-option mechanism $\{(0,0),(1,\alpha\mu)\}$.

Given the two-option mechanism, let v^* denote the seller type who is indifferent between offering (q,p) and $(1,\alpha\mu)$:

$$(1-\psi)(\alpha\mu - v^*) = \psi(kq - qv^*);$$

equivalently, $v^* = \frac{(1-\psi)\alpha\mu - \psi kq}{1-\psi-\psi q}$.

Therefore, the social welfare increases by:

$$\begin{aligned} & \psi q \int_{v^*}^{k} v dF - (1-\psi) \int_{v^*}^{\alpha\mu} v dF \\ \simeq \ & q[\psi \int_{\alpha\mu}^{k} v dF - v^* f(v^*)(1-\psi)\frac{dv^*}{dq}]_{q=0} \\ = \ & q[\psi \int_{\alpha\mu}^{k} v dF - \alpha\mu f(\alpha\mu)\psi(\alpha\mu - k)] \\ \propto \ & \int_{\alpha\mu}^{k} (vf(v) - \alpha\mu f(\alpha\mu))dv], \end{aligned}$$

which is positive if $vf(v)$ is (locally) strictly increasing at $v = \alpha\mu$. □

Proof of Proposition 3.

Consider the following menu mechanism: $\{(q_m,p_m)\}_{m=0}^{M} = \{(0,0),(1,p),(1,\alpha\mu)\}$ with $p > \alpha\mu$. In the associated equilibrium, the seller with $v < v(\psi)$ chooses $(1,\alpha\mu)$, which is accepted only by the naive buyer; while the seller with $v \in (v(\psi),p)$ chooses $(1,p)$, which is accepted only by the rational buyer, where $v(\psi)$ represents the indifferent type:

$$\psi(p - v(\psi)) = (1-\psi)(\alpha\mu - v(\psi)) \ \Leftrightarrow \ v(\psi) = \alpha\mu - \frac{p-\alpha\mu}{1-2\psi}\psi,$$

which is positive if ψ is sufficiently close to 0.

The expected social surplus is increasing locally around $\psi = 0$ if:

$$
\begin{aligned}
&\frac{\partial}{\partial \psi}\left[(1-\psi)\int_0^{v(\psi)}(\alpha-1)vdF + \psi\int_{v(\psi)}^{1}(\alpha-1)vdF\right]|_{\psi=0} \\
=&(\alpha-1)\left[-\int_0^{\alpha\mu} vdF - \alpha\mu f(\alpha\mu)(p-\alpha\mu) + \int_{\alpha\mu}^{p} vdF\right] > 0. \qquad (3)
\end{aligned}
$$

□

Proof of Proposition 4.

Fix any menu mechanism and its associated equilibrium. Let $x(v)$ denote the probability of trading conditional on the seller type v, and let $y(v)$ denote the expected transfer conditional on that the seller type is v.

Suppose that the buyer's ex ante expected payoff (i.e., the buyer's expected payoff before realizing his own type) is non-negative in this equilibrium:

$$\int_v (\alpha v x(v) - y(v))dF \geq 0.$$

It suffices to show that we must have $x(v) = 0$ for all v.

Recall that our assumption of severe adverse selection, Inequality (1), means that only the no-trade outcome is incentive-feasible if there is no naive buyer (Samuelson, 1984). Precisely, if $\{(x(v), y(v))\}_{v\in[0,1]}$ satisfies:

$$
\begin{aligned}
\text{(IC-r)}: &\quad y(v) - vx(v) \geq y(v') - vx(v'),\ \forall v, v' \\
\text{(IR-r)}: &\quad y(v) - vx(v) \geq 0,\ \forall v \\
\text{(EAIR-B)}: &\quad \int_v (\alpha v x(v) - y(v))dF \geq 0,
\end{aligned}
$$

then we must have $x(v) = 0$ for all v. □

Proof of Proposition 5.

Consider Problem $(2;\phi)$ for any given $\phi \geq 0$. Lemma 1 holds in this problem, because it is purely the implication of the incentive constraints of the seller and buyer, and does not depend on the principal's objective. Existence of m^* and v^* that satisfy the above conditions is then immediate. $\square$

Proof of Proposition 6.

Fix any mechanism in the class of general trading mechanisms and its associated equilibrium. Given the buyer's behavior in the equilibrium, the seller is facing the problem of choosing $m_1 \in M_1$ optimally (assuming that m_2, a_1, a_2 are optimally chosen), where each choice is associated with the expected continuation trade outcome denoted by $(q(m_1), p(m_1))$. Without loss of generality, assume that different choices of the seller induce different expected trade outcomes.

Then, the seller's choice of $m_1 \in M_1$ must induce the interval-partition structure because of the single-crossing payoff function. That is, there exist $\{v_m\}_{m=0}^M$ for some $M \in \mathbb{N}$ such that $0 = v_{M+1} > v_M > \ldots > v_0 = 1$ and those in the same interval $[v_{m+1}, v_m]$ choose the same message in M_1 in the equilibrium.

In particular, take the interval $[0, v_M]$. Given the seller's choice of m_1 and the rational buyer's choice of m_2, the severe adverse selection assumption implies that either the seller chooses $a_1 = 0$ or the rational buyer chooses $a_2 = 0$. Thus, trade is not possible between them. However, the mechanism being non-trivial implies that the seller must trade with the naive buyer. Therefore, we complete the proof by setting $v^* = v_M$. $\square$

B Optimal Mechanism and Its Properties under Power Distribution

This section investigates properties of the optimal menu mechanism under Power distribution (with $M = 2$, thanks to Proposition 1).

To investigate the naive type's payoff under the optimal mechanism with incentive-feasible cross-subsidization in Example 2.2, we focus on the case in which rational type's IR is not binding; when rational type's IR is binding, there is no incentive-feasible cross subsidization from the rational type to the naive type. In this case, $(q_1, p_1) = (1, 1)$ and $(q_2, p_2) = (1, \alpha\mu)$. The threshold type v^* is obtained from the seller's indifference condition: $(1-\psi)(\alpha\mu - v^*) = \psi(1 - v^*) \Leftrightarrow v^* = \frac{(1-\psi)\alpha\mu-\psi}{1-2\psi}$. Note that $v^* < \alpha\mu$ if and only if $\alpha\mu < 1$ and $\psi < \frac{1}{2}$. Note also that $\frac{dv^*}{d\psi} = -\frac{1-\alpha\mu}{(1-2\psi)^2} < 0$.

Under this mechanism, naive type's actual expected payoff is $\int_0^{v^*}(\alpha v - \alpha\mu)dF = \int_0^{v^*}(\alpha\gamma v^\gamma - \alpha\mu\gamma v^{\gamma-1})dv = [\alpha\frac{\gamma}{1+\gamma}v^{1+\gamma} - \alpha\mu v^\gamma]_0^{v^*} = -\alpha\mu(1-v^*)(v^*)^\gamma < 0$. Also, rational type's expected payoff is $\int_{v^*}^1(\alpha v - 1)dF = \int_{v^*}^1(\alpha\gamma v^\gamma - \gamma v^{\gamma-1})dv = [\alpha\frac{\gamma}{1+\gamma}v^{1+\gamma} - v^\gamma]_{v^*}^1 = \alpha\mu - 1 + (v^*)^\gamma - \alpha\mu(v^*)^{1+\gamma}$.

The expected buyer surplus is:

$$\begin{aligned}
& -(1-\psi)\alpha\mu(1-v^*)(v^*)^\gamma + \psi\left[\alpha\mu - 1 + (v^*)^\gamma - \alpha\mu(v^*)^{1+\gamma}\right] \\
= & -(1-2\psi)\alpha\mu(1-v^*)(v^*)^\gamma - \psi\alpha\mu(v^*)^\gamma + \psi\left[\alpha\mu - 1 + (v^*)^\gamma\right] \\
= & -(1-2\psi)\alpha\mu(1-v^*)(v^*)^\gamma - \psi(1-\alpha\mu)[1-(v^*)^\gamma] < 0.
\end{aligned}$$

Although the ex ante expected buyer surplus is always non-positive and maximized at $\psi = 1$ as in Proposition 4, it can be non-monotonic. For example, at $\psi = 0$, the ex ante expected buyer surplus is decreasing in ψ if α and γ are sufficiently small.

As discussed in Example 2.2, naive type's actual expected payoff with the maximal cross-subsidization is:

$$-\alpha\mu(1-v^*)(v^*)^\gamma + \frac{\psi}{1-\psi}\left[\alpha\mu - 1 + (v^*)^\gamma - \alpha\mu(v^*)^{1+\gamma}\right].$$

Its derivative with respect to ψ is:

$$
\begin{aligned}
&\frac{1}{(1-\psi)^2}\left[\alpha\mu - 1 + (v^*)^\gamma - \alpha\mu(v^*)^{1+\gamma}\right] \\
&- \frac{1-\alpha\mu}{(1-2\psi)^2}\left[-\alpha\mu\gamma(v^*)^{\gamma-1} + \alpha\mu(1+\gamma)(v^*)^\gamma + \frac{\psi}{1-\psi}\gamma(v^*)^{\gamma-1} - \frac{\psi}{1-\psi}\alpha\mu(1+\gamma)(v^*)^\gamma\right] \\
=& \frac{1}{(1-\psi)^2}\left[\alpha\mu - 1 + (v^*)^\gamma - \alpha\mu(v^*)^{1+\gamma}\right] \\
&- \frac{1-\alpha\mu}{(1-2\psi)^2}(v^*)^{\gamma-1}\underbrace{\left[-\alpha\mu\gamma + \frac{\psi}{1-\psi}\gamma - \frac{1-2\psi}{1-\psi}\alpha\mu(1+\gamma)\frac{(1-\psi)\alpha\mu-\psi}{1-2\psi}\right]}_{=\frac{1}{1-\psi}(\alpha-1)\gamma[(1-\psi)\alpha\mu-\psi]} \\
=& \frac{1}{(1-\psi)^2}\left[\alpha\mu - 1 + (v^*)^\gamma - \alpha\mu(v^*)^{1+\gamma}\right] - \frac{1-\alpha\mu}{(1-2\psi)}(v^*)^{\gamma-1}\frac{1}{1-\psi}(\alpha-1)\gamma v^* \\
=& \frac{1}{(1-\psi)^2}\left[\alpha\mu - 1 + (v^*)^\gamma - \alpha\mu(v^*)^{1+\gamma} - \underbrace{\frac{1-\psi}{1-2\psi}(1-\alpha\mu)}_{=1-v^*}(\alpha-1)\gamma(v^*)^\gamma\right] \\
=& \frac{1}{(1-\psi)^2}\left\{\alpha\mu - 1 + [1-(\alpha-1)\gamma-\gamma v^*](v^*)^\gamma\right\}.
\end{aligned}
$$

Because $[1-(\alpha-1)\gamma-\gamma v](v)^\gamma$ subject to $v \in (0,\alpha\mu)$ is maximized at $v = 1-\alpha\mu$, the derivative is at most

$$
\begin{aligned}
&\frac{1}{(1-\psi)^2}\left\{\alpha\mu - 1 + \underbrace{[1-(\alpha-1)\gamma-\gamma(1-\alpha\mu)]}_{=1-\alpha\mu}(1-\alpha\mu)^\gamma\right\} \\
=& -\frac{1}{(1-\psi)^2}(1-\alpha\mu)\left[1-(1-\alpha\mu)^\gamma\right] < 0.
\end{aligned}
$$

Hence, the naive type's actual expected payoff with the maximal cross-subsidization is decreasing in ψ, implying that the payoff of the naive type is strictly lower than the one-price mechanism.

We next turn to Example 3. The expected social surplus is:

$$
\begin{aligned}
&(1-\psi)\int_0^{v^*}(\alpha-1)vdF + \psi\int_{v^*}^1(\alpha-1)vdF \\
&= (1-\psi)(\alpha-1)[\mu v^{1+\gamma}]_0^{v^*} + \psi(\alpha-1)[\mu v^{1+\gamma}]_{v^*}^1 \\
&= (1-\psi)(\alpha-1)\mu(v^*)^{1+\gamma} - \psi(\alpha-1)\mu(v^*)^{1+\gamma} + \psi(\alpha-1)\mu \\
&= (\alpha-1)\mu\left[(1-2\psi)(v^*)^{1+\gamma} + \psi\right].
\end{aligned}
$$

The derivative at $\psi = 0$ is positive, for example, if $\gamma < \frac{1}{5}$ and $\alpha < \frac{5}{4}$.