Informationally Robust Welfare Predictions under Second-Degree Price Discrimination

Kai Hao Yang*

11/14/2019

Abstract

In an environment that features second-degree price discrimination, this paper fully characterizes the set of surplus divisions that can arise from all possible information consumers have about their valuation. By extending the techniques developed in a companion paper (Yang, 2019a), I show that the set of feasible surplus divisions can be characterized by a family of information structures that induce Pareto-distributed interim expected values. Unlike the linear model as in Roesler and Szentes (2017) where posted price is always optimal, the efficient frontier is generically not attainable under any information structures and there are environments in which a (nontrival) subset of the feasible surplus divisions collapses to a one-dimensional set. Nevertheless, the sets of feasible surplus divisions are stable around the linear environments.

Keywords: Surplus division, information structure, mechanism design, price discrimination, nonlinear pricing, revenue maximization, virtual valuation.

Jel classification: D42, D61, D82, D83

*Department of Economics, University of Chicago, khyang@uchicago.edu.

1 Introduction

In many industries, nonlinear pricing has been one of the most common pricing schemes used by producers. With nonlinear pricing, a monopolist can screen the consumers and engage in *second-degree price discrimination*. In the seminal paper, Mussa and Rosen (1978) consider an environment in which a monopolist can produce products of different qualities. When the marginal cost of production is increasing in qualities, they show that the monopolist would optimally engage in second-degree price discrimination. Inevitably, such optimal price discrimination scheme depends on the informational assumptions. That is, different information consumers possess greatly affects the exact form of the optimal scheme and therefore the welfare predictions. Motivated by this dependence, in a companion paper (Yang, 2019a), I study the buyer-optimal information structure in an environment that features nonlinear production technology and second-degree price discrimination, and explore how production technology affects the buyer's optimal surplus. Relatedly, another natural question motivated by the same dependence pertains to how consumers' information affects the division of surplus between the monopolist and consumers.

This paper studies possible surplus divisions that can be induced by consumers' information. Specifically, I consider an environment in which a monopolist has an increasing marginal cost and consumers can be informed about their valuation via arbitrary information structures. I show that the efficient surplus frontier is generically not attainable. In particular, only the full-extraction outcome is attainable by some information structure. In other words, under any information structure, whenever consumers retain positive surplus, there must be positive deadweight loss and the sum of consumer surplus and the monopolist's profit must be strictly below the efficient surplus. As such, the presence of deadweight loss is robust to consumers' information as long as the monopolist has an increasing marginal production cost. In addition, I further characterize the set of surplus divisions that can arise from all possible information structures. I show that when the prior is binary, the set of feasible surplus divisions can be represented by a family of information structures that induce Pareto-distributed interim expected values and can be described by finitely many parameters. With this characterization, the set of feasible surplus divisions can easily be computed. This characterization further implies that there exists an environment in which a subset of the feasible surplus divisions reduces to a one-dimensional set. Nevertheless, the characterization can also be applied to show that small perturbations of a constant-marginal cost model would never drastically change the set of feasible surplus divisions.

The main technique employed by this paper is developed in Yang (2019a), in which the buyer-optimal information structure under the environment where the monopolist has an increasing marginal cost is characterized. With the presence of second-degree price discrimination, the impact of information structures on consumer surplus is convoluted. Not only does an information structure affect the precision of consumers' estimation of the value of a product, it also determines how much information rent consumers should retain through their *virtual valuation*. To address this complexity, Yang (2019a) uses a local perturbation argument by showing that any information structure that is not in a particular family can be perturbed so that consumer surplus can be strictly improved. Together with a continuity result, Yang (2019a) is able to show that the buyer-optimal information structure must be in a particular family that can be described by finitely many parameters. This paper characterizes the set of feasible surplus divisions by maximizing weighed sums of consumer surplus and producer profit. Using similar arguments as in Yang (2019a), a particular family of information structures is first identified and any information structure that is not in this family is shown to permit a local perturbation that strictly improves the weighed sum. Along with a continuity result, the set of feasible surplus divisions is then identified by this particular family. As this family is simple and can be described by finitely many parameters, computing the feasible surplus divisions becomes tractable.

The rest of this paper is organized as follows: The next section summarizes the related literature; Section 3 provides the model; Section 4 states the main result together with some discussions; Section 5 concludes.

2 Related Literature

This paper is is closely related to the literature of mechanism design and the recent literature of information structures in mechanism design environments. The basic model that is fixed throughout this paper is similar to the framework of Mussa and Rosen (1978). Under any given information structure, the method developed by Myerson (1981) and generalized by Monteiro and Svaiter (2010) can be used to identify the monopolist's optimal price discrimination scheme. To explore the welfare implications of every possible information structure, this paper employs the method in Yang (2019a), in which a local perturbation argument is used to identify buyer-optimal information structures under a few related complex environments, including the environment studied in this paper.

Prior to this paper, several related works have explored the role of consumers' information about their own valuation in a monopolistic pricing setting. Lewis and Sappington (1991) solve for the seller-optimal information structure within a restricted class information structures the consumers can have about their valuations. Johnson and Myatt (2006) extend this result by expanding the restricted set. They show that within the restricted class of information structures, either full information or no information is optimal. Bergemann and Pesendorfer (2007) characterize the optimal selling mechanism and the seller-optimal information structure when there are multiple buyers with independent private values competing for a single object. They show that the optimal information structure is typically a partitional signal. Roesler and Szentes (2017) characterize the buyer-optimal information structure when the seller has a constant marginal cost and show that an information structure that induces Pareto-distributed interim expected values is optimal. Libgober and Mu (2019) characterize the buyer-optimal information structure in a dynamic setting under various assumptions about the timing. Yang (2019a) characterizes the buyer-optimal information structures in an environment where the monopolist has unknown production cost. Yang (2019b) also studies a scenario in which the seller's production cost is private information. In Yang (2019b), there is an informational intermediary whose objective is to maximize profit. Such intermediary extracts surplus from the seller by exploiting the technology that can provide the buyer different information about the product.

In terms of informationally robust welfare predictions, Bergemann et al. (2015) study a monopolistic pricing setting in which the monopolist has a constant marginal cost and characterize the set of surplus divisions that arise from all possible information the monopolist can have about the consumers' value. They show that essentially all surplus divisions—provided that the consumer surplus is nonnegative, the monopolist profit is above the uniform-pricing profit, and the sum of them is below the efficient surplus—is attainable by some information structure. Relatedly, Haghpanah and Siegel (2019) study the welfare implications of a monopolist's information about the consumers in nonlinear environments. Bergemann et al. (2017) provide a characterization of the set of surplus divisions that can arise from all possible information structures the bidders have in a first-price auction.

Among the aforementioned papers, this paper is the closest to Roesler and Szentes (2017), who examine a monopolistic pricing environment where the monopolist has a constant marginal cost and characterize the set of surplus divisions that can arise from all possible information the consumers have about their own valuation. The set-valued predictions also implies that essentially all surplus divisions can be attained by some information structure. A distinct feature of this paper is that the monopolist has an increasing marginal cost and therefore posted prices are no longer optimal.

3 Model

There is a seller (she) and a buyer (he). The seller produces and sells a product with quality $q \in [0,1]$ to the buyer. The buyer has outside option zero and a quasi-linear preference with valuation $v \in \{0,1\}$. If the quality of the product is q and the amount of payment is p, the buyer's ex-post payoff would be $qv - p$. Assume that the buyer's value follows a common prior so that $\mathbb{P}(v=1)=\mu\in(0,1)$.[1] On the other hand, the seller has an increasing marginal cost of producing higher-quality products. That is, the seller has a cost function $c:[0,1]\to\mathbb{R}_+$ so that $c(q)$ is the cost of producing a product with quality q. Assume that c is strictly increasing, strictly convex, twice differentiable and $c(0)=c'(0)=0$.[2] The buyer does not know about his own valuation *a priori*, instead, he learns about it through an *information structure*.[3] In sum, given an information structure, the timing of the events are as follows:

1. Nature draws $v\in\{0,1\}$ according to a common prior such that $\mathbb{P}(v=1)=\mu$.
2. The buyer is informed through an information structure.
3. The seller, without seeing the buyer's signal realizations, designs an optimal selling mechanism.
4. The buyer makes a participation decisions and takes actions in the mechanism if participating.

Since the buyer has a quasi-linear preference, the only payoff-relevant statistic of an information structure is the induced interim expected value. Using Blackwell's characterization (Blackwell, 1953), an information structure can be represented by the marginal distribution of interim expected value with mean μ.[4] Therefore, the set of all possible information structures can be represented by

$$\mathcal{I}_\mu := \left\{G:[0,1]\to[0,1]\,\middle|\,\int_0^1(1-G(x))dx=\mu,\;\; G \text{ is a CDF}\right\}.$$

[1]Similar to Yang (2019a), the binary support assumption is essential, as it allows a significant dimension reduction. Nevertheless, from Blackwell's perspective (Blackwell, 1953), any prior distribution with support on a subset of $[0,1]$ can be regarded as a garbling of a prior with support on $\{0,1\}$. As result, the characterization in this paper can be seen as an upper bound of what can be achieved under any other specifications of the prior.

[2]The assumption that $c(0)=c'(0)=0$ is not substantive but greatly simplifies the notations. For the case when $c'(0)>0$ or $c(0)>0$, the results are qualitatively similar but the characterization of possible surplus divisions becomes more complicated.

[3]An information structure is a Blackwell experiment that specifies the conditional distribution of signals given the realization of the true value.

[4]The same observation has been made in the recent literature, including Gentzkow and Kamenica (2016), Kolotilin et al. (2017), Roesler and Szentes (2017), Brooks and Du (2019), Dworczak and Martini (2019) and Yang (2019b). In these frameworks, the prior has support on the unit interval $[0,1]$, and therefore the second-order stochastic dominance constraint is needed. In this paper, however, as in Skreta and Perez-Richet (2018) and Yang (2019a), only the mean constraint needs to be considered because the prior has binary support.

Given an information structure $G \in \mathcal{I}_\mu$, when the seller designs the optimal selling mechanism, the revelation principle applies and thus it is without loss to restrict attention to incentive compatible direct mechanisms that ask the buyer to report his interim expected value and map the report into the quality and the amount of transfers. Therefore, the Myersonian approach can be employed and transfers can be pinned down by qualities up to a constant. Together with the individual rationality constraint, the seller's profit under an information structure G and an incentive compatible direct mechanism (q, p) is given by

$$\int_0^1 \Big(q(x)\psi(x|G) - c(q(x)) \Big) G(dx),$$

where $\psi(\cdot|G)$ is the virtual valuation induced by the information structure G.[5] As such, the seller's optimal mechanism is given by pointwise maximization. Notice that by strict convexity of c, for any $z \in (-\infty, 1]$, the maximization problem

$$\max_{q\in[0,1]} qz - c(q)$$

has a unique solution. Let $\gamma(z)$ denote this solution and let $V(z)$ denote the value. Notice that the assumptions on the cost function c imply that γ is differentiable on $(0, 1)$ and strictly increasing on $[0, 1]$, with $\gamma \equiv 0$ on $(-\infty, 0]$. With these notations, the seller's profit under information structure G is

$$\pi(G) := \int_0^1 V(\psi(x|G))G(dx),$$

whereas the buyer's surplus is

$$\sigma(G) := \int_0^1 \gamma(\psi(x|G))(1 - G(x))dx.$$

The main characterization of this paper is centered around the feasible payoff pairs

$$\Sigma := \{(\sigma(G), \pi(G))|G \in \mathcal{I}_\mu\}.$$

Another related object is the locus on which the buyer's surplus and the seller's profit sums to a constant. For any $s \in [0, 1]$, define

$$F^*(s) := \{(\sigma, \pi) \in \mathbb{R}^2_+|\sigma + \pi = s\}.$$

Notice that in this setting, the efficient surplus is given by $\mu V(1)$. Therefore, the efficient frontier is simply $F^*(\mu V(1))$.

[5]It is well-known that given a distribution G that is regular and has density $g > 0$, the induced virtual valuation is given by

$$x - \frac{1 - G(x)}{g(x)}.$$

Monteiro and Svaiter (2010) generalizes this definition to an arbitrary distribution. Together with a result in Toikka (2011), the Myersonian approach can be generalized to an arbitrary distribution G and the seller's profit can be written analogously, as presented above.

4 Informationally Robust Welfare Predictions

The main result of this paper characterizes the set Σ and explores some relevant economic properties. The first result considers whether and when the efficient frontier is attainable. In a linear environment as in Roesler and Szentes (2017), a rather surprising implication is that the efficient frontier is always attainable by some information structures. Furthermore, there are information structures that are Blackwell more informative than the completely uninformative ones but still attain the efficient frontier. This is contrary to the well-known intuition that asymmetric information creates distortion. In contrast, as shown in Proposition 1 below, such striking feature no longer exists in a nonlinear environment. According to Proposition 1, under the binary prior assumption, *any* information structure creates distortion except for the full information that enables the seller to extract all the surplus. Therefore, the efficient frontier is never attainable as long as the buyer retains positive rent.

Proposition 1. *Among all the surplus divisions on the efficient frontier, only the one in which the seller extracts all the surplus is attainable by some information structure. That is,*

$$\Sigma \cap F^*(\mu V(1)) = \{(0, \mu V(1))\}.$$

Proposition 1 implies that only the surplus division under which the seller extracts all the surplus is attainable. Such full extraction is only possible under the binary prior assumption since the only possible realizations of interim expected values are 0 and 1 under full information. When the interim expected value is zero, the product is always not sold to the buyer by individual rationality and by the assumption that $c(0) = 0$. When the interim expected value is 1, the buyer retains no information rent, which results from the standard "no distortion at the top" feature in quasi-linear environments. Since the efficient allocation also excludes the buyer with zero value, the optimal menu under full information is efficient and the seller extracts all the surplus. In fact, the arguments used in the proof of Proposition 1 can be easily generalized to arbitrary priors. For an arbitrary prior, even the full-surplus-extraction outcome is not attainable, which means that nonlinearity and the use of second-degree price discrimination distort the market even more severely so that there is always deadweight loss under any possible information structure.

The intuition is simple: In the linear environment, posted prices are always optimal and therefore it is possible to tailor an information structure that leaves the seller indifferent setting any prices in the support of the distribution of interim expected value. By selecting the smallest one, the efficient surplus can easily be attained even if the buyer has private information. In the nonlinear environment, however, the increasing marginal cost yields a unique optimal quality level given any interim expected value. As such, whenever the buyer retains positive information rents, the discrepancy between the virtual valuation and the actual valuation must create an inefficient allocation, regardless of which information structure is the buyer has.

In addition to exploring the efficient frontier, the next result provides a full characterization of the set of feasible surplus divisions Σ. In brief, Theorem 1 below states that although defined as an infinite-dimensional object, there exists a family of information structures, denoted as $\mathcal{I}_\mu^{**}$, that spans the whole feasible set. Moreover, the set $\mathcal{I}_\mu^{**}$ can be described by finitely many parameters, which further enables a tractable computation of the set Σ. To state Theorem 1 formally, the subset $\mathcal{I}_\mu^{*}$ must be defined. For any $k \in [0,1]$, any π, η, b such that $\pi/(1-\eta) + k \leq b \leq 1$, let

$$G_{\pi,k}^{\eta,b}(x) := \begin{cases} \eta, & \text{if } x \in \left[0, \frac{\pi}{1-\eta} + k\right) \\ 1 - \frac{\pi}{x-k}, & \text{if } x \in \left[\frac{\pi}{1-\eta} + k, b\right) \\ 1, & \text{if } x \in [b,1] \end{cases}.$$

Define

$$\mathcal{I}_\mu^{**} := \left\{ G_{\pi,k}^{\eta,b} \middle| \int_0^1 (1 - G_{\pi,k}^{\eta,b}(x))dx = \mu \right\}.$$

That is, $\mathcal{I}_\mu^{**}$ is a family of truncated Pareto distributions with jumps at 0 and some $b > 0$. More importantly, a special feature of an information structure in $\mathcal{I}_\mu^{**}$ is that the induced virtual valuation takes a simple form. Specifically, for any $G_{\pi,k}^{\eta,b} \in \mathcal{I}_\mu^{**}$,

$$\psi(x|G_{\pi,k}^{\eta,b}) = \begin{cases} \underline{\psi}, & \text{if } x \in \left[0, \frac{\pi}{1-\eta} + k\right) \\ k, & \text{if } x \in \left[\frac{\pi}{1-\eta} + k, b\right) \\ b, & \text{if } x \in [b,1] \end{cases},$$

for some $\underline{\psi} < 0$. With the definition of $\mathcal{I}_\mu^{**}$, the main result of this paper is stated as below.

Theorem 1. *The set of all possible surplus divisions can be generated by the information structures in $\mathcal{I}_\mu^{*}$. Moreover, the boundary of this set can only be attained by information structures in $\mathcal{I}_\mu^{**}$. That is,*

$$\partial\Sigma \subseteq \left\{ \left(\sigma\left(G_{\pi,k}^{\eta,b}\right), \pi\left(G_{\pi,k}^{\eta,b}\right)\right) \in \mathbb{R}_+^2 \middle| G_{\pi,k}^{\eta,b} \in \mathcal{I}_\mu^{**} \right\} = \Sigma.$$

Essentially, the proof of Theorem 1 extends the proof of the main result in Yang (2019a) by finding local perturbations that strictly improve the weighed sum of the buyer's surplus and the seller's profit for any $G \notin \mathcal{I}_\mu^{**}$; and then by showing that any of such weighed sum is upper-semicontinuous in information structures under the weak-* topology. Together, these imply that any weighed sum of the buyer's surplus and the seller's profit is only maximized by information structures in $\mathcal{I}_\mu^{**}$, which in turn implies that the boundary of the feasible payoff set, $\partial\Sigma$, can only be attained by information structures in $\mathcal{I}_\mu^{**}$. Finally, by continuity, it follows that any $(\sigma, \pi) \in \Sigma$ can be induced by an information structure in $\mathcal{I}_\mu^{**}$.

Using Theorem 1, the set of possible surplus divisions can be easily computed by the following procedure. First notice that under any $G^{\eta,b}_{\pi,k} \in \mathcal{I}^{**}_{\mu}$, the buyer's surplus can be written as

$$\gamma(k)(\mu - \pi - (1-\eta)k). \tag{1}$$

Let σ^* be the value of the maximization problem[6]

$$\begin{aligned}
&\max_{\pi,k,\eta,b} \gamma(k)(\mu - \pi - (1-\eta)k)\\
&\text{s.t. } \pi + (1-\eta)k + \pi \log\left(\frac{(1-\eta)(b-k)}{\pi}\right) = \mu.
\end{aligned}$$

and let

$$\Sigma_\sigma := \{\pi \in \mathbb{R}_+ | (\sigma, \pi) = (\sigma(G), \pi(G)), \text{ for some } G \in \mathcal{I}_\mu\}$$

for any $\sigma \in [0, \sigma^*]$. By Theorem 1, Σ_σ is an interval $[\underline{\pi}_\sigma, \bar{\pi}_\sigma]$, where $\bar{\pi}_\sigma$ and $\underline{\pi}_\sigma$ are solutions of the maximization problems

$$\begin{aligned}
&\max_{\pi,\eta,k,b} \left(1 - \frac{\pi}{b-k} - \eta\right) V(k) + \frac{\pi}{b-k} V(b)\\
&\text{s.t. } \pi + (1-\eta)k + \pi \log\left(\frac{(1-\eta)(b-k)}{\pi}\right) = \mu\\
&\gamma(k)(\mu - (\pi + (1-\eta)k)) = \sigma
\end{aligned}$$

and

$$\begin{aligned}
&\min_{\pi,\eta,k,b} \left(1 - \frac{\pi}{b-k} - \eta\right) V(k) + \frac{\pi}{b-k} V(b)\\
&\text{s.t. } \pi + (1-\eta)k + \pi \log\left(\frac{(1-\eta)(b-k)}{\pi}\right) = \mu\\
&\gamma(k)(\mu - (\pi + (1-\eta)k)) = \sigma,
\end{aligned}$$

respectively. Together, by computing $\underline{\pi}_\sigma$, $\bar{\pi}_\sigma$ pointwisely on $[0, \sigma^*]$, the set Σ is then obtained by "pasting up" the intervals $\{[\underline{\pi}_\sigma, \bar{\pi}_\sigma]\}_{\sigma \in [0,\sigma^*]}$.

Qualitatively, Theorem 1 and the procedure above imply that the feasible payoff set Σ is closed and connected (in $\mathbb{R}^2$). Moreover, the boundaries are smooth and typically nonlinear. This is in contrast to the triangular feature of the feasible surplus set in a linear environment as in Roesler and Szentes (2017). In particular, the information structure that gives the buyer-optimal surplus does not coincide with the worst information structure for the seller. Figure 1 above plots the set Σ under a parametrization $c(q) = q^\alpha/\alpha$, $\mu = \alpha/4(\alpha-1)$,[7] for $\alpha \in \{2, 3, 5\}$.

[6]Notice that σ^* is exactly the value given by the buyer-optimal information structure characterized in Yang (2019a).

[7]Such parameterization is chosen so that the efficient surplus, $\mu V(1)$, is fixed at $1/4$.

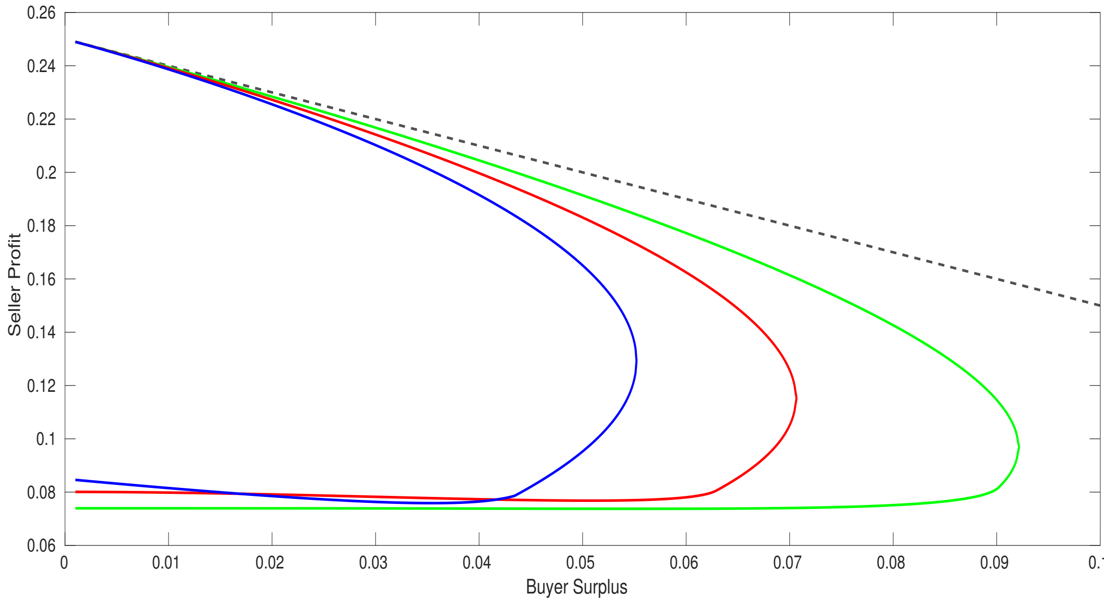

Figure 1: Possible Surplus Divisions

Note: The region enclosed by the blue curve is the possible surplus division when $\alpha = 2$, the region enclosed by the red curve is the possible surplus division when $\alpha = 3$, the region enclosed by the green curve is the possible surplus division when $\alpha = 5$. The dotted line represents the efficient frontier.

With Theorem 1, informationally robust welfare predictions can be made for any strictly increasing and strictly convex cost function. This additional flexibility allows further exploration of the set of feasible surplus divisions. Specifically, while essentially all surplus divisions are feasible in a linear environment, the results above show that many features no longer appear in a nonlinear environment. A further question is that whether there is an environment under which the set-valued prediction changes fundamentally. In particular, are there any environments under which the set of feasible surplus divisions "collapses" and becomes qualitatively different from the feasible payoff set in the linear environment? Proposition 2 below provides an answer to this question.

Proposition 2. *For any $\mu \in (0,1)$, there exists a sequence of strictly convex cost functions $\{c_n\}$ such that the induced set of possible surplus divisions, $\{\Sigma_n\}$, converges to a set Σ in which there exists a (non-degenerate) interval $[\underline{\sigma}, \bar{\sigma}]$ and a real-valued function f^* such that*

$$\{\sigma | (\sigma, \pi) \in \Sigma\} \cap [\underline{\sigma}, \bar{\sigma}] \neq \emptyset$$

and

$$\pi = f^*(\sigma),\ \forall (\sigma, \pi) \in \Sigma \quad s.t.\ \sigma \in [\underline{\sigma}, \bar{\sigma}].$$

For the purpose of providing robust welfare predictions, Proposition 2 can be regarded both as a positive and a negative result. On one hand, unlike the linear environment in which (surprisingly) every reasonable surplus division is attainable by some information structure,Proposition 2 shows that there exists nonlinear environments under which the set-valued predictions (asymptotically) becomes a lot more precise. In particular, even across all the possible information structures, the relation between the buyer's surplus and the seller's profit would only be affected minimally in such environments. On the other hand, this result illustrates the fragility of informationally robust predictions. Although the set-valued predictions are robust to mis-specification of information structures, the result in Proposition 2 implies that such set-valued predictions are sensitive to the assumptions about the environment. In the linear environment, the set of feasible surplus divisions is large and essentially all payoff pairs can be attained. However, in some other environments, as shown by Proposition 2, the set of feasible surplus divisions changes qualitatively and even sometimes collapses to a one-dimensional set. As such, the assumptions on production technology greatly affect the informationally robust predictions.

The fragility suggested by Proposition 2 naturally leads to another question: When perturbing the linear environment, are the predictions about surplus division robust? That is, is the set of surplus division under a linear environment also the limit of the set of surplus divisions induced by a sequence of environments that approximates this linear environment? Or is it possible that the set-valued predictions may collapse even with slight perturbation of the linear environment. Proposition 3 below provides a positive answer. That is, despite the fragility illustrated by Proposition 2, the set-valued predictions are stable around the linear environments. To state the result, let $\Sigma^L(c)$ be the set of surplus divisions in an environment where the seller has a constant marginal cost $c \in [0, 1]$.[8]

Proposition 3. *For any $\mu \in (0, 1)$, for any $c \in [0, 1]$ and for any sequence of increasing and strictly convex cost functions $\{c_n\}$ such that $\{c_n(q)\} \to cq$ for all $q \in [0, 1]$, there exists a subsequence such that the induced set of possible surplus divisions, $\{\Sigma_n\}$, converges to $\Sigma^L(c)$ under the Hausdorff distance as $n \to \infty$.*

By the result in Roesler and Szentes (2017), any $(\sigma, \pi) \in \Sigma^L(c)$ falls on a horizontal line where $\{(\sigma, \pi) | \sigma \in [0, \sigma(\pi)]\}$ for some $\sigma(\pi) \geq 0$. Although this construction heavily relies the indifference of the seller under a Pareto-shaped information structure and hence cannot be applied whenever there is a strictly convex cost function, in the limit, it is always possible to find sequences $\{k_n\}$ to approach any level of buyer's surplus $\sigma \in [0, \sigma(\pi)]$. This is possible because the optimal quantity γ when there is a constant marginal cost is discontinuous while the optimal quantity γ_n for a strictly convex cost function must be continuous, and hence the convergence $\{\gamma_n\} \to \gamma$ is never uniform. As a result, for an arbitrary number $\tilde{\gamma}$, there is always enough "space" for finding a sequence $\{k_n\}$ so that $\{k_n\} \to c$ and $\{\gamma_n(k_n)\} \to \tilde{\gamma}$.

[8]For each $c \in [0, 1]$, the set $\Sigma^L(c)$ is characterized in the online appendix of Roesler and Szentes (2017).

Together with (1), this then introduces an additional degree of freedom to "scale down" the buyer's surplus while keeping the seller's profit constant.

5 Conclusion

This paper considers a monopolistic pricing environment that features second-degree price discrimination. It explores the set of surplus divisions that can arise from all possible information structures the consumers have and shows that the efficient frontier is generically not attainable. Furthermore, it fully characterizes the set of feasible surplus divisions. Such feasible set is spanned by a family of information structures that can be described by finitely many parameters. Using this characterization, it also shows that a small perturbation of the environment with constant marginal cost can drastically affect the set of feasible surplus divisions.

Several extensions and related economic questions can be considered. First, the binary prior assumption can be relaxed. Under an arbitrary prior, a conjecture is that a critical family of information structures that can be described by partitions on the interval $[0, 1]$ spans the set of feasible surplus divisions. The exact form of this set and how this set contributes to computation of the set of feasible surplus divisions remain as future research topics.

Another related question considers a monopolist's optimal selling mechanism that is robust to mis-specification of a buyer's information structure when the seller has an increasing marginal production cost. That is, the seller designs an optimal selling mechanism to solve the max-min problem against Nature, who chooses the worst-case information structure for any mechanism selected. Theorem 1 provides a simple way to solve the associated min-max problem. Using Theorem 1, Nature's solution must lie in $\mathcal{I}_\mu^{**}$, which is a finite-dimensional constraint minimization problem. With this result, the remaining questions are whether this zero-sum game has a saddle point or whether the duality gap is zero.

Finally, while this paper provides informationally robust welfare predictions across all possible information the consumers have, a related question is to consider such set-valued predictions across all possible information the monopolist has about the consumers' valuation under a nonlinear environment. Bergemann et al. (2015) briefly introduce this question in their extension by considering a parameterized nonlinear preference. A natural question is to extend their characterization to a more general setting in which either the preference or the technology (or both) exhibits nonlineraity and characterize the set of possible surplus divisions as this paper does analogously. This can also be a topic for future studies.

References

Bergemann, D., B. Brooks, and S. Morris (2015): "The Limits of Price Discrimination," *American Economic Review*, 105, 921–957.

——— (2017): "First-Price Auctions with General Information Structures: Implications for Bidding and Revenue," *Econometrica*, 85, 107–143.

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

Blackwell, D. (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Brooks, B. and S. Du (2019): "Optimal Auction Design with Common Values: An Informationally-Robust Approach," Working Paper.

Dworczak, P. and G. Martini (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy,* forthcoming.

Gentzkow, M. and E. Kamenica (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review, Papers & Proceeding*, 106, 597–601.

Haghpanah, N. and R. Siegel (2019): "Pareto Improving Segmentation of Multi-product Markets," Working Paper.

Johnson, J. P. and D. P. Myatt (2006): "On the Simple Economics of Advertising, Marketing, and Product Design," *American Economic Review*, 96, 756–784.

Lewis, T. R. and D. E. M. Sappington (1991): "Supplying Information to Facilitate Price Discrimination," *International Economic Review*, 35, 309–327.

Libgober, J. and X. Mu (2019): "Informational Robustness in Intertemporal Pricing," Working Paper.

Milgrom, P. and I. Segal (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583–601.

Monteiro, P. K. and B. F. Svaiter (2010): "Optimal Auction with a General Distribution: Virtual Valuation without Densities," *Journal of Mathematical Economics*, 46, 21–31.

Mussa, M. and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

Myerson, R. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

Roesler, A.-K. and B. Szentes (2017): "Buyer-Optimal Learning and Monopoly Pricing," *American Economic Review*, 107, 2072–2080.

Skreta, V. and E. Perez-Richet (2018): "Test Design under Falsification," Working Paper.

Toikka, J. (2011): "Ironing without Control," *Journal of Economic Theory*, 146, 2510–2526.

Yang, K. H. (2019a): "Buyer-Optimal Information with Nonlinear Technology," Working Paper.

——— (2019b): "Equivalence in Business Models for Informational Intermediaries," Working Paper.

Appendix

A.1 Preliminaries

Before the proofs of the main results, several preliminaries and notations should be first introduced. Specifically, the virtual valuation $\psi(\cdot|G)$ should be formally definied. The following definition is from Monteiro and Svaiter (2010), who generalized the "virtual valuation" used in Myerson (1981) and provides similar properties.

Consider any CDF G with $\text{supp}(G) \subseteq [0, 1]$. Let $a := \inf\{x \in [0, 1] | G(x) > 0\}$ be the lower bound of $\text{supp}(G)$. Define

$$H(x|G) := \begin{cases} 0, & \text{if } x \in [0, a) \\ a - x(1 - G(x)), & \text{if } x \in [a, 1] \end{cases}.$$

Since G is increasing, $H(\cdot|G)$ is of bounded variation. Notice that for any measurable function $p : [0, 1] \to [0, 1]$,

$$\int_0^1 p(x)H(dx|G) = \int_0^1 p(x)xG(dx) - \int_0^1 p(x)(1 - G(x))dx,$$

where the integral on the left hand side is defined with respect to the signed measure induced by $H(\cdot|G)$.

Let $\Theta := \{(\alpha, \beta) \in \mathbb{R}^2 | \alpha + \beta G(x) \leq H(x|G), \forall x \in [0, 1]\}$ and let

$$\phi(x|G) := \sup\{\alpha + \beta G(x) | (\alpha, \beta) \in \Theta\},$$

we say that $w(x)$ is a *sub-gradient* of $\phi(\cdot|G)$ at $x \in [0, 1]$ if

$$\phi(z|G) - \phi(x|G) \geq w(x)(G(z) - G(x)), \forall z \in [0, 1].$$

For each $x \in [0, 1]$, let $\partial\phi(x|G)$ denote the set of sub-gradients of $\phi(\cdot|G)$ at x. Finally, let

$$\psi(x|G) := \inf \partial\phi(x|G).$$

$\psi(\cdot|G) : [0, 1] \to (-\infty, 1]$ is then defined as the *virtual valuation* induced by G.

With this definition, the following two lemmas are useful properties of the virtual valuation $\psi(\cdot|G)$. Lemma 1 can be found in Monteiro and Svaiter (2010). It gives an essential property of the virtual valuation $\psi(\cdot|G)$ for any CDF G. The statement is cited here for completeness. Lemma 2 can be found in Yang (2019a).

Lemma 1 (Monterio & Svaiter, 2010)**.** *For any CDF G, let $\sigma(\psi(\cdot|G))$ be the σ-algebra generated by $\psi(\cdot|G)$. That is, $\sigma(\psi(\cdot|G)) := \sigma(\{\psi^{-1}(B|G) | B \in \mathcal{B}([0, 1])\})$. Then for any function $m : [0, 1] \to [0, 1]$ that is measurable with respect to $\sigma(\psi(\cdot|G))$,*

$$\int_0^1 m(x)H(dx|G) = \int_0^1 m(x)\psi(x|G)G(dx).$$

In particular, for any measurable nondecreasing function $\gamma : (-\infty, 1] \to \mathbb{R}_+$,

$$\int_0^1 \gamma(\psi(x|G))xH(dx|G) = \int_0^1 \gamma(\psi(x|G))\psi(x|G)G(dx)$$

Lemma 2 (Yang, 2018a)**.** *Fix any CDF R with* $\text{supp}(R) \subseteq [0,1]$*. Suppose that $x_0 \in (0,1)$ is such that $\psi(x_0^+|R) > \psi(x|R)$ for all $x \in [0, x_0)$. Then for any $k \in [\psi(x_0^-|R), \psi(x_0^+|R)]$,*

$$(x-k)(1-R(x)) \leq (x-k)(1-R(x^-)) \leq (x_0-k)(1-R(x_0^-)),\ \forall x \in [0,1]$$

and the inequality is strict if $\phi(x|R) < H(x|R)$ or $\psi(x|R) \neq k$. In particular, let $r := \sup\{x \in [0,1] | \psi(x|R) < 0\}$. Then

$$x(1-R(x)) \leq x(1-R(x^-)) \leq r(1-R(r^-)),\ \forall x \in [0,1].$$

A.2 Auxiliary Results

The proof of Theorem 1 is conceptually similar to the proof of Theorem 1 in Yang (2019a). Specifically, notice that the boundary $\partial\Sigma$ can be characterized by solving planner problems that maximize weighed sums of buyer's surplus and the seller's profit. As such, we first argue that for any weights and for any distribution $R \notin \mathcal{I}_\mu^{**}$, there is a local perturbation $\widehat{R} \in \mathcal{I}_\mu$ such that the weighed sum can be strictly improved. Together with additional results that enable an alternative characterization of Σ and provide certain topological structures, this implies that $\partial\Sigma$ can only be attained by distributions in $\mathcal{I}_\mu^{**}$, which then proves Theorem 1. To begin with, for any $\alpha, \beta \in \mathbb{R}$, first define

$$W_{\alpha,\beta}(R) := \alpha\sigma(R) + \beta\pi(R),\ \forall H \in \mathcal{I}_\mu$$

as the weighed sum of the buyer's surplus and the seller's profit with weights α and β. Notice that since the seller's profit is maximized under the full information distribution $\bar{H}$, the boundary $\partial\Sigma$ can be characterized by the family of the planner's problem

$$\sup_{H \in \mathcal{I}_\mu} W_{\alpha,\beta}(R),$$

for pairs $\alpha > 0, \beta \geq 0$, $\alpha < 0, \beta < 0$, and $\alpha > 0, \beta \leq 0$. To this end, Lemma 3 and Lemma 4 first provide an alternative representation of the buyer's surplus and the seller's profit, which leads to an alternative representation of the set Σ. Then, Lemma 5 and Lemma 6 below show that for any $R \notin \mathcal{I}_\mu^*$, there must exist some $\widehat{R} \in \mathcal{I}_\mu$ such that $W_{\alpha,\beta}(R) < W_{\alpha,\beta}(\widehat{R})$. Finally, Lemma 7 proves the closedness of Σ and therefore the existence of solutions of the planner's problem. Together, these results imply that $\partial\Sigma$ is attainable and can only be attained by information structures in $\mathcal{I}_\mu^{**}$, which then proves Theorem 1.

Throughout this proof, for any $R \in \mathcal{I}_\mu$, take $\psi(\cdot|R)$ to be right-continuous.[9] Denote $\psi_R^{-1} : [0,1] \to [0,1]$ by $\psi_R^{-1}(z) := \sup\{x \in [0,1] | \psi(x|R) < z\}$ for all $z \geq 0$. Notice that for any $R \in \mathcal{I}_\mu$, by Lemma 2, $\psi_R^{-1}(c)$ is the largest element of $\text{argmax}_{x \in [0,1]}(x-c)(1-R(x^-))$ for all $c \in [0,1]$ and that $R \circ \psi_R^{-1}$ is also a valid CDF.

Below, we begin with two auxiliary results that provide an alternative characterization of the buyer's surplus the and the seller's profit. Lemma 3 below is a special case of Lemma 6 and Lemma

[9]From the payoff perspective, this is without loss, as there are at most countably many jumps of $\psi(\cdot|R)$.

7 in Yang (2019b) and therefore the proof is omitted. Using Lemma 3, Lemma 4 gives an alternative way to represent the feasible payoff set Σ, which also allows a way to re-write the seller's profit as a function of information structure that will be used later.

Lemma 3. *For any $R \in \mathcal{I}_\mu$, there exists a nondecreasing, left-continuous function $q : [0,1] \to [0,1]$ such that*

1. $q(c) = R(\psi_R^{-1}(c)^-)$ *for all* $c \in [0,1]$.
2. $\int_c^1 (1-q(z))dz/(1-q(c)) + c = \psi_R^{-1}(c)$ *for all* $c \in [0,1]$ *such that* $q(c) < 1$.
3. $\int_0^1 \max_{c\in[0,1]}[1 - \int_c^1(1-q(z))dz/(x-c)^+]^+ dx \leq 1-\mu$.

Conversely, for any nondecreasing, left-continuous function $q : [0,1] \to [0,1]$ such that $\int_0^1 \max_{c\in[0,1]}[1 - \int_c^1(1-q(z))dz/(x-c)^+]^+ dx \leq 1-\mu$, there exists $R \in \mathcal{I}_\mu$ such that:

1. $q(c) = R(\psi_R^{-1}(c)^-)$ *for all* $c \in [0,1]$.
2. $\int_c^1 (1-q(z))dz/(1-q(c)) + c = \psi_R^{-1}(c)$ *for all* $c \in [0,1]$ *such that* $q(c) < 1$.

Lemma 4. *Let $\mathcal{Q}$ be the collection of left-continuous and nondecreasing functions $q : [0,1] \to [0,1]$ such that $\int_0^1 \max_{c\in[0,1]}[1 - \int_c^1(1-q(z))dz/(x-c)^+]^+ dx \leq 1-\mu$. Define*

$$\widetilde{\sigma}(q) := \int_0^1 \gamma(z)\frac{\int_z^1(1-q(x))dx}{(1-q(z))}q^+(dz)$$
$$\widetilde{\pi}(q) := \int_0^1 V(z)q^+(dz),$$

for all $q \in \mathcal{Q}$, where $q^+(z) := \lim_{\delta\downarrow 0} q(z+\delta)$ for all $z \in [0,1]$. Then,

$$\Sigma = \{(\widetilde{\sigma}(q), \widetilde{\pi}(q)) | q \in \mathcal{Q}\}.$$

Proof. Take any $(\sigma, \pi) \in \Sigma$ and $R \in \mathcal{I}_\mu$ such that $\sigma(R) = \sigma$ and $\pi(R) = \pi$. By Lemma 3, there exists $q \in \mathcal{Q}$ such that $q(c) = R(\psi_R^{-1}(c)^-)$ for all $c \in [0,1]$ and $\psi_R^{-1}(c) = \int_c^1(1-q(z))dz/(1-q(c)) + c$ for all $c \in [0,1]$ with $q(c) < 1$. As such,

$$\pi(R) = \int_{\psi_R^{-1}(0)}^1 V(\psi(x|R))R(dx) = \int_0^1 V(z)R\circ\psi_R^{-1}(dz) = \int_0^1 V(z)q^+(dz) = \widetilde{\pi}(q).$$

On the other hand, since $\gamma \circ \psi(\cdot|R)$ is $\sigma(\psi(\cdot|R))$-measurable, by Lemma 1,

$$\begin{aligned}
\sigma(R) &= \int_{\psi_R^{-1}(0)}^1 \gamma(\psi(x|R))(1-R(x))dx \\
&= \int_{\psi_R^{-1}(0)}^1 \gamma(\psi(x|R))xR(dx) - \int_{\psi_R^{-1}(0)}^1 \gamma(\psi(x|R))\psi(x|R)R(dx) \\
&= \int_0^1 \gamma(z)(\psi_R^{-1}(z) - z)R\circ\psi_R^{-1}(dz) \\
&= \int_0^1 \gamma(z)\frac{\int_z^1(1-q(x))dx}{(1-q(z))}q^+(dz) = \widetilde{\sigma}(q)
\end{aligned}$$

Conversely, for any $q \in \mathcal{Q}$, by Lemma 3, there exists $R \in \mathcal{I}_\mu$ such that $q(c) = R(\psi_R^{-1}(c)^-)$ for all $c \in [0,1]$ and $\psi_R^{-1}(c) = \int_c^1 (1-q(z))dz/(1-q(c)) + c$ for all $c \in [0,1]$ with $q(c) < 1$. Thus,

$$\widetilde{\pi}(q) = \int_0^1 V(z) q^+(dz) = \int_0^1 V(z) R \circ \psi_R^{-1}(dz) = \int_{\psi_R^{-1}(0)}^1 V(\psi(x|R)) R(dx) = \pi(R)$$

and

$$\begin{aligned}
\widetilde{\sigma}(q) &= \int_0^1 \gamma(z) \frac{\int_z^1 (1-q(x))dx}{(1-q(z))} q^+(dz) \\
&= \int_0^1 \gamma(z)(\psi_R^{-1}(z) - z) R \circ \psi_R^{-1}(dz) \\
&= \int_{\psi_R^{-1}(0)}^1 \gamma(\psi(x|R)) x R(dx) - \int_{\psi_R^{-1}(0)}^1 \gamma(\psi(x|R)) \psi(x|R) R(dx) \\
&= \int_{\psi_R^{-1}(0)}^1 \gamma(\psi(x|R))(1 - R(x))dx = \sigma(R)
\end{aligned}$$

and therefore $(\widetilde{\sigma}(q), \widetilde{\pi}(q)) \in \Sigma$. This completes the proof. ■

An immediate corollary from Lemma 3 and Lemma 4 is a way to re-write the seller's profit in a more natural way.

Corollary 1. *For any $R \in \mathcal{I}_\mu$, for any $z \in [0,1]$ let*

$$\pi(z|R) := \max_{x \in [0,1]} (x - z)(1 - R(x^-))$$

be the optimal profit of the seller with a constant cost $z \in [0,1]$ under information structure R. Then, for any $R \in \mathcal{I}_\mu$

$$\pi(R) = \int_0^1 \pi(z|R) \gamma(dz).$$

Proof. For any $R \in \mathcal{I}_\mu$, $z \in [0,1]$, Lemma 2 and Lemma 3 imply that

$$\pi(z|R) = (\psi_R^{-1}(z) - z)(1 - R(\psi_R^{-1}(z))) = \int_z^1 (1 - q(x))dx.$$

Using integration by parts twice and the fact that $V'(z) = \gamma(z)$ for all $z \in [0,1]$, Lemma 4 then implies that

$$\pi(R) = \int_0^1 \pi(z|R) \gamma(dz).$$

■

We now begin the local perturbation arguments. Specifically, if $R \notin \mathcal{I}_\mu^{**}$, Lemma 5 below ensures that we can find a local perturbation that strictly improves the weighed sum $W_{\alpha,\beta}$ when $\alpha > 0$. Lemma 6 below then shows that a local perturbation can also be found when $\alpha < 0, \beta \leq 0$. The proofs are similar to the proof of Lemma 1 in Yang (2019a) except that several details are further tailored to incorporate both the seller's profit and the buyer's surplus.

Lemma 5. *Fix any $\alpha > 0$, $\beta \in \mathbb{R}$ such that $(\alpha, \beta) \neq (0,0)$. For any $R \in \mathcal{I}_\mu \backslash \mathcal{I}_\mu^{**}$, there exists $\widehat{R} \in \mathcal{I}_\mu$ such that*

$$W_{\alpha,\beta}(R) < W_{\alpha,\beta}(\widehat{R}).$$

Proof. Consider any $R \in \mathcal{I}_\mu \backslash \mathcal{I}_\mu^{**}$. Let $a := \inf\{x \in [0,1] | R(x) > 0\}$, $b := \sup\{x \in [0,1] | R(x) < 1\}$ be the lower bound and upper bound of $\text{supp}(R)$, respectively and $r := \inf\{x \in [0,1] | \psi(x|R) > 0\}$. If $\psi(x|R) = \bar{\psi}$ for all $x \in (r, b)$, let $\zeta := (r - \bar{\psi})(1 - R(r^-))$, then by Lemma 2

$$R(x) \geq 1 - \frac{\zeta}{x - \bar{\psi}}$$

for all $x \in [0,1]$. Furthermore, since $R \notin \mathcal{I}_\mu^{**}$, $\phi(x|R) < H(x|R)$ for some $x \in (r,b)$ and thus, by Lemma 2 again, the inequality is strict for some $x \in (r,b)$ such that $1 - \zeta/(x - \bar{\psi}) > 0$. In addition, since $\psi(x|R) = \bar{\psi}$ for all $x \in (r,b)$, it must be that $R(r) = R(r^-) = 1 - \zeta/(r - \bar{\psi})$ and $R(b^-) = 1 - \zeta/(b - \bar{\psi})$. Together, the intermediate value theorem implies that there exists $\hat{r} \in (r,b)$ and $\widetilde{R}$ with $1 - \zeta/(x - \bar{\psi}) \leq \widetilde{R}(x) \leq R(b^-)$ for all $x \in [\hat{r}, b)$ such that $\widehat{R} \in \mathcal{I}_\mu$, where

$$\widehat{R}(x) := \begin{cases} 1 - \frac{\xi}{\hat{r}}, & \text{if } x \in [0, \hat{r}) \\ \widetilde{R}(x), & \text{if } x \in [\hat{r}, b) \\ 1, & \text{if } x \in [b, 1] \end{cases},$$

and $\xi := r(1 - R(r))$. Notice that

$$\psi(x|\widehat{R}) = \begin{cases} \underline{\psi}, & \text{if } x \in [0, \hat{r}) \\ \widehat{\psi}, & \text{if } x \in [\hat{r}, b) \\ b, & \text{if } x \in [b, 1] \end{cases},$$

for some $\underline{\psi} < 0$ and $\widehat{\psi} \in [\bar{\psi}, b]$. Furthermore, $\widehat{R} \in \mathcal{I}_\mu$ implies that

$$\int_{\hat{r}}^1 (1 - \widehat{R}(x))dx + \hat{r}(1 - \widehat{R}(\hat{r})) = \mu = \int_r^1 (1 - R(x))dx + \int_0^r (1 - R(x))dx.$$

Together with $\hat{r}(1 - \widehat{R}(\hat{r})) = r(1 - R(r))$,

$$\int_{\hat{r}}^1 (1 - \widehat{R}(x))dx \geq \int_r^1 (1 - R(x))dx.$$

Therefore, as $\widehat{\psi} \geq \bar{\psi}$,

$$\int_0^1 \gamma(\psi(x|R))(1-R(x))dx = \gamma(\bar{\psi}) \int_r^1 (1-R(x))dx \leq \gamma(\widehat{\psi}) \int_{\hat{r}}^1 (1-\widehat{R}(x))dx = \int_0^1 \gamma(\psi(x|\widehat{R}))(1-\widehat{R}(x))dx.$$

On the other hand, by Corollary 1, the seller's profit under R and $\widehat{R}$ are $\int_0^1 \pi(z|R)\gamma(dz)$ and $\int_0^1 \pi(z|\widehat{R})\gamma(dz)$, respectively, where

$$\pi(z|R) = \begin{cases} \zeta + \bar{\psi} - z, & \text{if } z \in [0, \bar{\psi}] \\ \zeta \frac{b-z}{b-\bar{\psi}}, & \text{if } z \in (\bar{\psi}, b] \\ 0, & \text{if } z \in (b, 1]. \end{cases},$$

and

$$\pi(z|\widehat{R}) = \begin{cases} (\hat{r}-z)(1-\widehat{R}(\hat{r})) > \zeta + \bar{\psi} - z, & \text{if } z \in [0,\hat{\psi}] \\ (b-z)(1-\widehat{R}(b^-)) = \zeta\frac{b-z}{b-\hat{\psi}}, & \text{if } z \in (\hat{\psi}, b] \\ 0, & \text{if } z \in (b,1]. \end{cases}.$$

As such,

$$\pi(R) = \int_0^1 \pi(z|R)\gamma(dz) < \int_0^1 \pi(z|\widehat{R})\gamma(dz) = \pi(\widehat{R}).$$

Together, since $\alpha > 0$, under this $\widehat{R} \in \mathcal{I}_\mu$, it must be that $W_{\alpha,\beta}(\widehat{R}) > W_{\alpha,\beta}(R)$.

Now consider the case when there exists $x_1, x_2 \in (r,b)$ such that $\psi(x_2|R) > \psi(x_1|R) \geq 0$. Since $R \in \mathcal{I}_\mu$,

$$1-\mu-\int_0^1 R(x)dx = 1-\mu-\int_r^1 R(x)dx - \int_a^r R(x)dx = 0. \tag{2}$$

Since R is nondecreasing, it must be that

$$1-\mu-\int_r^1 R(x)dx - rR(r) \leq 1-\mu-\int_r^1 R(x)dx - (r-a)R(r) \leq 0.$$

Consider three cases separately:

Case 1: $R(r^-) > 0$ and

$$1-\mu-\int_r^1 R(x)dx - rR(r) < 0. \tag{3}$$

Fix $k > 0$, let $\rho := \inf\{x \in [0,1] | \psi(x|R) > k\}$. Clearly, $\rho \geq r$. From Lemma 2, $(x-k)(1-R(x)) \leq (\rho-k)(1-R(\rho^-))$ for all $x \in [0,1]$. Therefore,

$$R(x) \geq \max\left\{0, 1-\frac{\zeta}{x-k}\right\}. \tag{4}$$

for all $x \in [0,r]$, where $\zeta := (\rho-k)(1-R(\rho^-))$. Moreover, since $\psi(x|R) < 0$ for all $x \in [0,r)$ and $R(r^-) > 0$, for k close enough to 0, there exists $z \in [0,r)$ such that $R(x) > 1-\zeta/(x-k) > 0$ for all $x \in (z,r)$. Therefore, for k close enough to 0,

$$1-\mu-\int_\rho^1 R(x)dx - \int_0^\rho \left(1-\frac{\zeta}{(x-k)^+}\right)^+ dx > 1-\mu-\int_0^1 R(x)dx = 0.$$

Together with (3), by the intermediate value theorem, for such k, ρ, for any $\underline{x} < \rho$, there exists $z(\underline{x}), \in (0,\underline{x})$ such that

$$1-\mu-\int_\rho^1 R(x)dx - \int_{\underline{x}}^\rho \left(1-\frac{\zeta}{x-k}\right)dx - (\underline{x}-z(\underline{x}))\left(1-\frac{\zeta}{\underline{x}-k}\right) = 0 \tag{5}$$

with $1-\zeta/(\underline{x}-k) \geq 0$. Now define

$$\widehat{R}^{\underline{x}}(x) := \begin{cases} R(x), & \text{if } x \in [\rho,1] \\ 1-\frac{\zeta}{x-k}, & \text{if } x \in [\underline{x},\rho) \\ 1-\frac{\zeta}{\underline{x}-k}, & \text{if } x \in [z(\underline{x}),\underline{x}) \\ 0, & \text{if } x \in [0,z(\underline{x})) \end{cases}.$$

By (5), $\widehat{R}^{\underline{x}} \in \mathcal{I}_\mu$ for any $\underline{x} \in [0, \rho]$. Furthermore, by construction,

$$\psi(x|\widehat{R}^{\underline{x}}) = \begin{cases} \psi(x|R), & \text{if } x \in [\rho, 1] \\ k, & \text{if } x \in [\underline{x}, \rho) \\ \underline{\psi}, & \text{if } x \in [0, \underline{x}). \end{cases},$$

for some $\underline{\psi} < 0$. Therefore, define

$$\Psi(\underline{x}) := \int_0^1 \gamma(\psi(x|\widehat{R}^{\underline{x}}))(1 - \widehat{R}^{\underline{x}}(x))dx - \int_0^1 \gamma(\psi(x|R))(1 - R(x))dx.$$

Then $\Psi(\rho) = 0$. Moreover, for any $\underline{x} \in (0, \rho)$,

$$\begin{aligned}\Psi'(\underline{x}) &= \frac{d}{d\underline{x}} \int_0^1 \gamma(\psi(x|\widehat{R}^{\underline{x}}))(1 - \widehat{R}^{\underline{x}}(x))dx \\ &= \frac{d}{d\underline{x}} \left[\int_{\underline{x}}^{\rho} \gamma(k) \left(1 - \frac{\zeta}{x - k}\right) dx + \int_\rho^1 \gamma(\psi(x|R))(1 - R(x))dx \right] \\ &= -\gamma(k) \left(1 - \frac{\zeta}{\underline{x} - k}\right) < 0.\end{aligned}$$

As such, since Ψ is continuous, for $\underline{x} < \rho$ close enough to ρ, $\Psi(\underline{x}) > 0$ for $\underline{x}$ sufficiently close to ρ.

On the other hand, for such ρ, k, for each $\underline{x}$, let $\widehat{R}^{\underline{x}}$ be as constructed as above. By construction,

$$\left| \frac{\partial}{\partial \underline{x}} \pi(z|\widehat{R}^{\underline{x}}) \right| \leq 1$$

and

$$\lim_{\underline{x} \uparrow \rho} \frac{\partial}{\partial \underline{x}} \pi(z|\widehat{R}^{\underline{x}}) = 0, \ \forall z \in [0, 1].$$

Therefore, let

$$\Xi(\underline{x}) := \int_0^1 \pi(z|\widehat{R}^{\underline{x}})\gamma(dz) - \int_0^1 \pi(z|R)\gamma(dz),$$

we then have $\Xi(\rho) = 0$ and

$$\Xi'(\rho) = \lim_{\underline{x} \uparrow \rho} \Xi'(\underline{x}) = \lim_{\underline{x} \uparrow \rho} \int_0^1 \frac{\partial}{\partial \underline{x}} \pi(z|\widehat{R}^{\underline{x}})\gamma(dz) = 0,$$

by the Lebesgue dominant convergence theorem. Together, for $\underline{x}$ close enough to ρ, $W_{\alpha,\beta}(R) < W_{\alpha,\beta}(\widehat{R}^{\underline{x}})$.

Case 2:

$$1 - \mu - \int_r^1 R(x)dx - rR(r) = 0. \tag{6}$$

Then by (2) and (6),

$$0 = 1 - \mu - \int_r^1 R(x)dx - \int_a^r R(x) \geq 1 - \mu - \int_r^1 R(x)dx - (r - a)R(r) = aR(r) \geq 0$$

and hence $R(r) = R(r^-)$, $\int_a^r (R(r) - R(x))dx = 0$, $aR(r) = 0$ and therefore $R(r) = R(a)$.

Now consider the case when there exists some $x_0 \in (r, 1)$ and a sequence $\{x_n\}$ such that $x_n < x_{n+1} < x_0$ and $\psi(x_n|R) < \psi(x_{n+1}|R) < \psi(x_0^-|R)$ for all $n \in \mathbb{N}$ and that $\{x_n\} \uparrow x_0$. Since $\psi(x_{n+1}|R) > \psi(x_n|R)$ for all $n \in \mathbb{N}$ and $\psi(\cdot|R)$ is nondecreasing, we may take such $\{x_n\}$ so that $\psi(x_n|R) > \psi(y|R)$ for all $y \in [r, x_n)$ for all $n \in \mathbb{N}$. Let $k := \psi(x_0^-|R)$ and $\zeta_0 := (x_0 - k)(1 - R(x_0^-))$. Also, for each $n \in \mathbb{N}$, let $\zeta_n := (x_n - \psi(x_n|R))(1 - R(x_n^-))$. Then by Lemma 2, since $\psi(x_n|R) > \psi(y|R)$ for all $y \in [r, x_n)$ and since $\psi(x_{n+1}|R) > \psi(x_n|R)$, we must have

$$R(x) \geq \max\left\{1 - \frac{\zeta_0}{x-k},\, 1 - \frac{\zeta_n}{x - \psi(x_n|R)},\, 0\right\}, \tag{7}$$

for all $x \in [x_n, x_0]$, for all $n \in \mathbb{N}$, with strict inequality holding at some $x \in (x_n, x_0)$. Therefore, there exists $\bar{n} \in \mathbb{N}$ such that whenever $n > \bar{n}$, there exists $\underline{x}_n \in (r, x_n)$ and $\hat{x}_n \in (x_n, x_0)$ such that

$$\int_0^{x_0} R(x)dx = \underline{x}_n\left(1 - \frac{\pi_R}{\underline{x}_n - \psi(r|R)}\right) + \int_{\underline{x}_n}^{x_n} R(x)dx + \int_{x_n}^{x_0} \max\left\{1 - \frac{\zeta_0}{x-k},\, 1 - \frac{\zeta_n}{x - \psi(x_n|R)}\right\}dx \tag{8}$$

$$1 - \frac{\zeta_0}{\hat{x}_n - k} = 1 - \frac{\zeta_n}{\hat{x}_n - \psi(x_n|R)}, \tag{9}$$

where $\pi_R := (r - \psi(r|R))(1 - R(r^-))$.

As such, for any $n > \bar{n}$, define

$$\widehat{R}^{x_n}(x) := \begin{cases} R(x), & \text{if } x \in [x_0, 1] \\ 1 - \frac{\zeta_0}{x-k}, & \text{if } x \in [\hat{x}_n, x_0) \\ 1 - \frac{\zeta_n}{x - \psi(x_n|R)}, & \text{if } x \in [x_n, \hat{x}_n) \\ R(x), & \text{if } x \in [\underline{x}_n, x_n) \\ 1 - \frac{\pi_R}{\underline{x}_n - \psi(r|R)}, & \text{if } x \in [0, \underline{x}_n) \end{cases},$$

where $\hat{x}_n$ and $\underline{x}_n$ are uniquely defined by (8) and (9). Notice that by (7), $\hat{x}_n < \hat{x}_{n+1}$ for all $n \in \mathbb{N}$, $\{\hat{x}_n\} \uparrow x_0$ and $\underline{x}_n > \underline{x}_{n+1}$ for all $n \in \mathbb{N}$, $\{\underline{x}_n\} \downarrow r$.

By construction, for any $n > \bar{n}$, there exists $\underline{\psi}_n < 0$, $\tilde{x}_n \in [\underline{x}_n, x_n]$ such that for all $x \in [0, 1]$,

$$\psi(x|\widehat{R}^{x_n}) = \begin{cases} \psi(x|R), & \text{if } x \in [x_0, 1] \\ k, & \text{if } x \in [\hat{x}_n, x_0) \\ \psi(x_n|R), & \text{if } x \in [x_n, \hat{x}_n) \\ \psi(x|R), & \text{if } x \in [\tilde{x}_n, x_n) \\ \psi(\tilde{x}_n|R), & \text{if } x \in [\underline{x}_n, \tilde{x}_n) \\ \underline{\psi}_n, & \text{if } x \in [0, \underline{x}_n) \end{cases}.$$

Moreover, the sequence $\{\tilde{x}_n\}$ converges to r as $n \to \infty$. As such, the difference in the buyer's surplus between $\widehat{R}^{x_n}$ and R is

$$\begin{aligned}&\int_0^1 \gamma(\psi(x|\widehat{R}^{x_n}))(1-\widehat{R}^{x_n}(x))dx - \int_0^1 \gamma(\psi(x|R))(1-R(x))dx\\ \geq&\gamma(k)(r(1-R(r))-\underline{x}_n(1-R(\underline{x}_n))) + \int_r^{\underline{x}_n}(\gamma(k)-\gamma(\psi(x|R)))(1-R(x))dx\\ &-\int_{x_n}^{\hat{x}_n}\big((\gamma(k)-\gamma(\psi(x_n|R)))(1-\widehat{R}^{x_n}(x))\big)dx\\ &+\int_{x_n}^{x_0}\big(\gamma(k)-\gamma(\psi(x|R)))(1-R(x))\big)dx.\end{aligned}$$

Notice that by Lemma 2 and $R(r)=R(r^-)$, $r(1-R(r)) = r(1-R(r^-)) \geq \underline{x}_n(1-R(\underline{x}_n))$ for all $n>\bar{n}$.

On the other hand, for any $\underline{x} \in (r, x_0)$, let

$$\begin{aligned}\Psi(\underline{x}) :=& \int_r^{\underline{x}}\big(\gamma(k)-\gamma(\psi(x|R))\big)(1-R(x))dx\\ &-\int_{\bar{x}(\underline{x})}^{\hat{x}(\underline{x})}\big(\gamma(k)-\gamma(\psi(\bar{x}(\underline{x})|R)\big)\big(1-\widehat{R}^{\underline{x}}(x)\big)dx\\ &+\int_{\bar{x}(\underline{x})}^{x_0}\big(\gamma(k)-\gamma(\psi(x|R))\big)(1-R(x))dx,\end{aligned} \tag{10}$$

where $\hat{x}(\underline{x})$ and $\bar{x}(\underline{x})$ are uniquely defined by

$$\int_0^{x_0} R(x)dx = \underline{x}\left(1-\frac{\pi_R}{\underline{x}-\psi(r|R)}\right)+\int_{\underline{x}}^{\bar{x}(\underline{x})}R(x)dx+\int_{\bar{x}(\underline{x})}^{x_0}\max\left\{1-\frac{\zeta_0}{x-k}, 1-\frac{\zeta_{\bar{x}(\underline{x})}}{x-\psi(\bar{x}(\underline{x})|R)}\right\}dx \tag{11}$$

$$1-\frac{\zeta_0}{\hat{x}(\underline{x})-k} = 1-\frac{\zeta_{\bar{x}(\underline{x})}}{\hat{x}(\underline{x})-\psi(\bar{x}(\underline{x})|R)}, \tag{12}$$

and $\widehat{R}^{\underline{x}}$ is defined by

$$\widehat{R}^{\underline{x}}(x) := \begin{cases} R(x), & \text{if } x\in[x_0,1]\\ 1-\frac{\zeta_0}{x-k}, & \text{if } x\in[\hat{x}(\underline{x}),x_0)\\ 1-\frac{\zeta_{\bar{x}(\underline{x})}}{x-\psi(\bar{x}(\underline{x})|R)}, & \text{if } x\in[\bar{x}(\underline{x}),\hat{x}(\underline{x}))\\ R(x), & \text{if } x\in[\underline{x},\bar{x}(\underline{x}))\\ 1-\frac{\pi_R}{\underline{x}-\psi(r|R)}, & \text{if } x\in[0,\underline{x})\end{cases},$$

where $\zeta_x := (x-\psi(x|R))(1-R(x^-))$ for any $x\in[r,1]$. Let

Notice that by (8), (9), (11) and (12), for any $n>\bar{n}$, $x_n=\bar{x}(\underline{x}_n)$ and $\hat{x}_n=\hat{x}(\underline{x}_n)$. Also, $\hat{x}$ and $\bar{x}$ are decreasing in $\underline{x}$ and $\lim_{\underline{x}\to r}\bar{x}(\underline{x}) = \lim_{\underline{x}\to r}\hat{x}(\underline{x}) = x_0$, $\lim_{\underline{x}\to r}\psi(\bar{x}(\underline{x})|R) = \lim_{n\to\infty}\psi(x_n|R) = \psi(x_0^-|R)$.

Since $\psi(\cdot|R)$ and γ are nondecreasing, by (11) and (12), $\bar{x}$ and $\hat{x}$ are differentiable Lebesgue-almost everywhere and therefore Ψ is differentiable Lebesgue-almost everywhere. Thus, for Lebesgue almost all $\underline{x}$,

$$\Psi'(\underline{x}) = (\gamma(k) - \gamma(\psi(\underline{x}|R)))(1 - R(\underline{x})) - \hat{x}'(\underline{x})(\gamma(k) - \gamma(\psi(\bar{x}(\underline{x})|R)))(1 - \widehat{R}^{\underline{x}}(\hat{x}(\underline{x})))$$
$$- \int_{\bar{x}(\underline{x})}^{\hat{x}(\underline{x})} \frac{\partial}{\partial \underline{x}}(\gamma(k) - \gamma(\psi(\bar{x}(\underline{x})|R)))(1 - \widehat{R}^{\underline{x}}(x))dx.$$

Since $\lim_{\underline{x}\to r} \hat{x}(\underline{x}) = \lim_{\underline{x}\to r} \bar{x}(\underline{x}) = x_0$, $\lim_{\underline{x}\to r} \psi(\bar{x}(\underline{x})|R) = k$ and since $\hat{x}$ is decreasing, there exists $\delta > 0$ such that for $\underline{x}$ sufficiently close to r, whenever Ψ is differentiable at $\underline{x}$,

$$\Psi'(\underline{x}) > (\gamma(k) - \gamma(\psi(\underline{x}|R)))(1 - R(\underline{x})) - \delta > 0, \tag{13}$$

where the second inequality follows from that γ is strictly increasing. Therefore, since Ψ is continuous in $\underline{x}$ and $\Psi'(\underline{x}) > 0$ for $\underline{x}$ sufficiently close to r, there exists $\hat{r}$ such that $\Psi(\underline{x}) > 0$ for all $\underline{x} \in (r, \hat{r})$.

As such, for n sufficiently large so that $\underline{x}_n \in (r, \hat{r})$,

$$\begin{aligned} 0 < \Psi(\underline{x}_n) = &\int_r^{\underline{x}_n} (\gamma(k) - \gamma(\psi(x|R)))(1 - R(x))dx \\ &- \int_{x_n}^{\hat{x}_n} \big((\gamma(k) - \gamma(\psi(x_n|R)))(1 - \widehat{R}^{x_n}(x))\big)dx \\ &+ \int_{x_n}^{x_0} \big(\gamma(k) - \gamma(\psi(x|R)))(1 - R(x))\big)dx. \end{aligned}$$

Together, for n sufficiently large,

$$\int_0^1 \gamma(\psi(x|\widehat{R}^{x_n}))(1 - \widehat{R}^{x_n}(x))dx > \int_0^1 \gamma(\psi(x|R))(1 - R(x))dx,$$

which implies that the buyer's surplus can be strictly improved under $\widehat{R}^{x_n}$ when n is large. Now we show that the first-order difference of the seller's profit is zero. Indeed, by Corollary 1, the difference in seller's profit between R and $\widehat{R}^{x_n}$ is

$$\int_0^1 \pi(z|R)dz - \int_0^1 \pi(z|\widehat{R}^{x_n})\gamma(dz).$$

Notice that for each $n \in \mathbb{N}$, any $z \in [0,1]\backslash(\psi(x_n|R), k) \cup (\psi(r|R), \psi(\tilde{x}_n|R))$, $\pi(z|R) = \pi(z|\widehat{R}^{x_n})$. Therefore,

$$\begin{aligned} &\int_0^1 \pi(z|R)\gamma(dz) - \int_0^1 \pi(z|\widehat{R}^{x_n})\gamma(dz) \\ =&\int_{\psi(r|R)}^{\psi(\tilde{x}_n|R)} \gamma'(z)(\pi(z|R) - \pi(z|\widehat{R}^{x_n}))dz + \int_{\psi(x_n|R)}^{k} \gamma'(z)(\pi(z|R) - \pi(z|\widehat{R}^{x_n}))dz. \end{aligned}$$

Let

$$\Xi(\underline{x}) := \int_{\psi(r|R)}^{\psi(\tilde{x}(\underline{x})|R)} \gamma'(z)(\pi(z|R) - \pi(z|\widehat{R}^{\underline{x}}))dz + \int_{\psi(\bar{x}(\underline{x})|R)}^{k} \gamma'(z)(\pi(z|R) - \pi(z|\widehat{R}^{\underline{x}}))dz. \tag{14}$$

Notice that $\Xi(r)=0$. Furthermore, since $\psi(\tilde{x}(\underline{x})|R)\to\psi(r|R)$ and $\psi(\bar{x}(\underline{x})|R)\to k$ as $\underline{x}\to r$, and since $\pi(\psi(r|R)|R)=\pi(\psi(r|R)|\widehat{R}^{\underline{x}})$ and $\pi(k|R)=\pi(k|\widehat{R}^{\underline{x}})$ for all $\underline{x}$, we have

$$\Xi'(r)=\gamma'(\psi(r|R))(\pi(\psi(r|R))-\pi(\psi(r|R)))+\gamma'(k)(\pi(k|R)-\pi(k|R))=0. \tag{15}$$

Together, for $n\in\mathbb{N}$ large enough,

$$W_{\alpha,\beta}(R)<W_{\alpha,\beta}(\widehat{R}^{x_n}),$$

as desired.

If, on the other hand, for any $x\in(r,1)$ and for any sequence $\{x_n\}$ such that $\{x_n\}\uparrow x$, there exists $\bar{n}\in\mathbb{N}$ such that $\psi(x_n|R)=\psi(x^-|R)$ for all $n>\bar{n}$, then for any $x\in(r,1)$, there exists $\delta>0$ such that $\psi(y|R)=\psi(x^-|R)$ for all $y\in(x-\delta,x)$. Let $\delta_x:=\sup\{\delta>0|\psi(y|R)=\psi(x^-|R),\forall y\in(x-\delta,x)\}$. Then $\delta_x>0$ for all $x\in(r,1)$. Moreover, for any $x,y\in(r,1)$, if $\psi(x^-|R)\neq\psi(y^-|R)$, then it must be that $(x-\delta_x,x)\cap(y-\delta_y,y)=\emptyset$. Therefore, $\{\psi(x^-|R)\}_{x\in(r,1)}$ is at most countable. Since $\psi(\cdot|R)^+$ is nondecreasing, it must be a step function. Consider first the case when for any $\delta>0$, there exists $x,y\in(b-\delta,b)$ with $x<y$ such that $\psi(x|R)<\psi(y|R)<\psi(b^-|R)$. Since $\psi(\cdot|R)^+$ is a step function, $\psi(\cdot|R)^+$ has countably infinitely many jumps and therefore we may represent $\psi(\cdot|R)^+$ as

$$\psi(x|R)=\sum_{n=1}^{\infty}\psi_n\mathbf{1}\{x\in(\alpha_n,\beta_n)\},\ \forall x\in[r,1]\backslash[\{\alpha_n\}_{n=1}^{\infty}\cup\{\beta_n\}_{n=1}^{\infty}].$$

for some $\{\alpha_n\}$, $\{\beta_n\}$ such that $\alpha_n<\beta_n$ for all $n\in\mathbb{N}$ and some $\{\psi_n\}_{n=1}^{\infty}$ such that $\psi_n>\psi(r|R)$ for all $n\in\mathbb{N}$, $n\geq 2$. Since for any $\delta>0$, there exists $x,y\in(b-\delta,b)$, $x<y$, such that $\psi(x|R)<\psi(y|R)<\psi(b^-|R)$, there exists a sequence $\{x_j\}$ such that $x_j\in\{\alpha_n\}_{n=1}^{\infty}\cup\{\beta_n\}_{n=1}^{\infty}$, $x_j<x_{j+1}$, $\psi_j:=\psi(x_j|R)=\psi(x_{j+1}^-|R)<\psi(x_{j+1}|R)=:\psi_{j+1}$ for all $j\in\mathbb{N}$, and that $\{x_j\}\uparrow b$ and $\{\psi_j\}\uparrow\psi(b^-|R)$.

Since $\psi(\cdot|R)^+$ is a step function and $\psi(x|R)=b$ for all $x\in[b,1]$, we must have $\psi(b^-|R)<\psi(b|R)=b$ and $R(b^-)<1$. As such, let $\zeta_b:=(b-\psi(b^-|R))(1-R(b^-))$. Then $\zeta_b>0$. Also, for each $j\in\mathbb{N}$, let $\zeta_j:=(x_j-\psi_j)(1-R(x_j^-))$, by Lemma 2, since $\psi_{j-1}<\psi_j<\psi_{j+1}$, we have

$$R(x)>\max\left\{1-\frac{\zeta_b}{x-\psi(b^-|R)},1-\frac{\zeta_j}{x-\psi_j}\right\},$$

for all $x\in(x_j,b)$.

As such, there exists a sequence $\{r_j\}$ such that $\{r_j\}\downarrow r$ and $\widehat{R}^j\in\mathcal{I}_\mu$ for j large enough, where

$$\widehat{R}^j(x):=\begin{cases}R(x), & \text{if } x\in[b,1]\\ 1-\frac{\zeta_b}{x-\psi(b^-|R)}, & \text{if } x\in[\hat{x}_j,b)\\ 1-\frac{\zeta_j}{x-\psi_j}, & \text{if } x\in[\bar{x}_j,\hat{x}_j)\\ R(x), & \text{if } x\in[r_j,\bar{x}_j)\\ 1-\frac{\pi_R}{r_j-\psi(r|R)}, & \text{if } x\in[0,r_j)\end{cases}.$$

and $r_j < \hat{x}_j < \bar{x}_j < b$ are uniquely defined by

$$\int_0^b \widehat{R}^j(x)dx = \int_0^b R(x)dx$$
$$1 - \frac{\zeta_b}{\hat{x}_j - \psi(b^-|R)} = 1 - \frac{\zeta_j}{\hat{x}_j - \psi_j},$$

Similar to the previous case, for each $j \in \mathbb{N}$ such that $\widehat{R}^j \in \mathcal{I}_\mu$, the difference in the buyer's surplus between R and $\widehat{R}^j$ is at least

$$\gamma(\psi(b^-|R))(r(1-R(r)) - r_j(1-R(r_j))) + \int_r^{r_j} (\gamma(\psi(b^-|R)) - \gamma(\psi(x|R)))(1-R(x))dx$$
$$- \int_{\bar{x}_j}^{\hat{x}_j} \big(\gamma(\psi(b^-|R)) - \gamma(\psi_j)\big)(1-\widehat{R}^j(x))dx$$
$$+ \int_{\bar{x}_j}^{x_j} \big(\gamma(\psi(b^-|R)) - \gamma(\psi(x|R))\big)(1-R(x))dx.$$

As noted above, from Lemma 2, $r(1-R(r)) = r(1-R(r^-)) \geq r_j(1-R(r_j))$ for all $j \in \mathbb{N}$. Also, as shown in the previous case, from (10) and (13), with $x_0 = b$ and $k = \psi(b^-|R)$, there eixsts $\delta > 0$ such that for $\underline{x}$ sufficiently close to r, $\Psi'(\underline{x}) > \gamma(\psi(b^-|R)) - \gamma(\psi(r|R)) - \delta > 0$ since $\psi_j > \psi(r|R)$ for all $j \in \mathbb{N}$ and γ is strictly increasing. Thus, since by (11) and (12), $\hat{x}_j = \hat{x}(r_j)$ and $\bar{x}_j = \bar{x}(r_j)$, and since $\{r_j\} \downarrow r$, for j sufficiently large,

$$0 < \Psi(r_j) = \int_r^{r_j} (\gamma(\psi(b^-|R)) - \gamma(\psi(x|R)))(1-R(x))dx$$
$$- \int_{\bar{x}_j}^{\hat{x}_j} \big((\gamma(\psi(b^-|R)) - \gamma(\psi_j)\big)(1-\widehat{R}^j(x))dx$$
$$+ \int_{\bar{x}_j}^{x_j} \big(\gamma(\psi(b^-|R)) - \gamma(\psi(x|R)))\big)(1-R(x))dx.$$

Together, for j large enough, $\widehat{R}^j \in \mathcal{I}_\mu$ and

$$\int_0^1 \gamma(\psi(x|\widehat{R}^j))(1-\widehat{R}^j(x))dx > \int_0^1 \gamma(\psi(x|R))(1-R(x))dx,$$

which means that the buyer's surplus can be strictly improved under $\widehat{R}^j$, for j large enough. On the other hand, the difference in seller's profit between R and $\widehat{R}^j$ is

$$\int_0^1 \pi(z|R)\gamma(dz) - \int_0^1 \pi(z|\widehat{R}^j)\gamma(dz).$$

Similar to the previous case, for all $j \in \mathbb{N}$, $\pi(z|R) = \pi(z|\widehat{R}^j)$ for all $z \in [0,1]\backslash(\psi_j, \psi(b^-|R)) \cup (\psi(r|R), \psi(\tilde{r}_j|R))$, for some $\tilde{r}_j \in [r_j, \bar{x}_j]$. Moreover, $\{r_j\} \to r$ and $\{\psi(r_j|R)\} \to \psi(r|R)$ as $j \to \infty$. Thus, by (14) and (15), $\Xi'(r) = 0$ implies that the first order difference in seller's profit is zero. Together, there exists $\widehat{R}^j \in \mathcal{I}_\mu$ such that

$$W_{\alpha,\beta}(R) < W_{\alpha,\beta}(\widehat{R}^j),$$

as desired.

Finally, if there exists $\delta > 0$ such that for any $x, y \in (b-\delta, b)$, $\psi(x|R) = \psi(y|R) = \psi(b^-|R)$. Let $\underline{b} := \inf\{b_0 \in [r,1) | \psi(x|R) = \psi(y|R), \forall x, y \in (b_0, b)\}$ and let $\underline{\psi} := \psi(b^-|R)$. Then $\underline{b} < b$ and $\underline{\psi} > \psi(r|R)$ since $\psi(\cdot|R)$ is not a constant on $[r, b)$. We claim that it is without loss to suppose that $\phi(x|R) = H(x|R)$ for all $x \in (\underline{b}, b)$. Indeed, if there exists $x \in (\underline{b}, b)$ such that $\phi(x|R) < H(x|R)$, let $\zeta := (\underline{b} - \underline{\psi})(1 - R(\underline{b}^-))$. Lemma 2 then ensures that

$$R(x) \geq 1 - \frac{\zeta}{x - \underline{\psi}},$$

for all $x \in (\underline{b}, b)$ with strict inequality for some $x \in (\underline{b}, b)$ and therefore, there exists $\hat{b} \in (\underline{b}, b)$ such that

$$\int_{\underline{b}}^{b} (1 - R(x))dx = \int_{\underline{b}}^{\hat{b}} \frac{\zeta}{x - \underline{\psi}} dx.$$

Therefore, for

$$\widehat{R}(x) := \begin{cases} R(x), & \text{if } x \in [0, \underline{b}) \\ 1 - \frac{\zeta}{x - \underline{\psi}}, & \text{if } x \in [\underline{b}, \hat{b}) \\ 1, & \text{if } x \in [\hat{b}, 1] \end{cases},$$

$\widehat{R} \in \mathcal{I}_\mu$,

$$\int_0^1 \gamma(\psi(x|R))(1 - R(x))dx = \int_0^1 \gamma(\psi(x|\widehat{R}))(1 - \widehat{R}(x))dx$$
$$\int_0^1 \pi(z|R)\gamma(dz) = \int_0^1 \pi(z|\widehat{R})\gamma(dz).$$

and $H(x|\widehat{R}) = \phi(x|\widehat{R})$ for all $x \in (\underline{b}, b)$.

Therefore, as $H(x|R) = \phi(x|R)$ for all $x \in (\underline{b}, b)$ and $\psi(x|R) = \underline{\psi}$ for all $x \in (\underline{b}, b)$, we must have

$$R(x) = 1 - \frac{\zeta}{x - \underline{\psi}}, \; \forall x \in [\underline{b}, b),$$

for some $\zeta > 0$. Now take and fix any $\bar{x} \in (\underline{b}, b)$ notice that for any $h \in (\psi(\underline{b}^-|R), \underline{\psi})$ and any $k \in (\underline{\psi}, b)$, let $\zeta(k) := (\bar{x} - k)(1 - R(\bar{x}^-))$ and $\zeta(h) := (\underline{b} - h)(1 - R(\underline{b}^-))$, we must have

$$R(x) > \max\left\{1 - \frac{\zeta(k)}{x - k}, 1 - \frac{\zeta(h)}{x - h}\right\},$$

for all $x \in (\underline{b}, \bar{x})$. Thus, for any $\underline{x} > r$ that is close enough to r, there exists $h(\underline{x}) \in (\psi(\underline{b}^-|R), \underline{\psi})$,$k(\underline{x}) \in (\underline{\psi}, b)$ and $\hat{x}(\underline{x}) \in (\underline{b}, b)$ such that $\lim_{\underline{x} \downarrow r} h(\underline{x}) = \lim_{\underline{x} \downarrow r} k(\underline{x}) = \underline{\psi}$,

$$\frac{\zeta(k(\underline{x}))}{\hat{x}(\underline{x}) - k(\underline{x})} = \frac{\zeta(h(\underline{x}))}{\hat{x}(\underline{x}) - h(\underline{x})}$$

and

$$\int_0^{\bar{x}} R(x)dx = \int_0^{\bar{x}} \widehat{R}^{\underline{x}}(x)dx,$$

where

$$\widehat{R}^{\underline{x}}(x) := \begin{cases} R(x), & \text{if } x \in [\bar{x}, 1] \\ 1 - \frac{\zeta(k(\underline{x}))}{x - k(\underline{x})}, & \text{if } x \in [\hat{x}(\underline{x}), \bar{x}) \\ 1 - \frac{\zeta(h(\underline{x}))}{x - h(\underline{x})}, & \text{if } x \in [\underline{b}, \hat{x}(\underline{x})) \\ R(x), & \text{if } x \in [\underline{x}, \underline{b}) \\ 1 - \frac{\pi_R}{\underline{x} - \psi(r|R)}, & \text{if } x \in [0, \underline{x}) \end{cases}$$

and thus $\widehat{R}^{\underline{x}} \in \mathcal{I}_\mu$. Moreover, such $h(\cdot)$ and $k(\cdot)$ can be selected so that $\hat{x}(\underline{x})$ is decreasing in $\underline{x}$ and $\lim_{\underline{x} \downarrow r} \hat{x}(\underline{x}) = \bar{x}$.

Notice that for any such $\widehat{R}^{\underline{x}}$,

$$\psi(x|\widehat{R}^{\underline{x}}) = \begin{cases} \underline{\psi}, & \text{if } x \in [0, \underline{x}) \\ \psi(\tilde{x}(\underline{x})|R), & \text{if } x \in [\underline{x}, \tilde{x}(\underline{x})) \\ \psi(x|R), & \text{if } x \in [\underline{x}, \underline{b}) \\ h(\underline{x}), & \text{if } x \in [\underline{b}, \hat{x}(\underline{x})) \\ \tilde{k}(\underline{x}), & \text{if } x \in [\bar{x}(\underline{x}), b) \\ b, & \text{if } x \in [b, 1]. \end{cases},$$

for some $\underline{\psi} < 0$, some $\tilde{x}(\underline{x}) \in [\underline{x}, \underline{b}]$ with $\lim_{\underline{x} \to r} \tilde{x}(\underline{x}) = r$, and some $\tilde{k}(\underline{x}) \in (\underline{\psi}, k(\underline{x}))$. As such, for any $\underline{x}$ such that $\widehat{R}^{\underline{x}} \in \mathcal{I}_\mu$, the difference in the buyer's surplus between R and $\widehat{R}^{\underline{x}}$ is

$$\begin{aligned} &\int_0^1 \gamma(\psi(x|\widehat{R}^{\underline{x}}))(1 - \widehat{R}^{\underline{x}}(x))dx - \int_0^1 \gamma(\psi(x|R))(1 - R(x))dx \\ >&\gamma(\tilde{k}(\underline{x}))(r(1 - R(r)) - \underline{x}(1 - R(\underline{x}))) + \int_r^{\underline{x}} (\gamma(\tilde{k}(\underline{x})) - \gamma(\psi(x|R)))(1 - R(x))dx \\ -&\int_{\underline{b}}^{\hat{x}(\underline{x})} (\gamma(\tilde{k}(\underline{x})) - \gamma(h(\underline{x})))(1 - \widehat{R}^{\underline{x}}(x))dx, \end{aligned}$$

where the stritc inequality follows from $\tilde{k}(\underline{x}) > \underline{\psi}$ and that γ is strictly increasing.

Again, by Lemma 2, $r(1 - R(r)) \geq \underline{x}(1 - R(\underline{x}))$ for all $\underline{x} \geq r$. Also, let

$$\Psi(\underline{x}) := \int_r^{\underline{x}} (\gamma(\tilde{k}(\underline{x})) - \gamma(\psi(x|R)))(1 - R(x))dx - \int_{\underline{b}}^{\hat{x}(\underline{x})} (\gamma(\tilde{k}(\underline{x})) - \gamma(h(\underline{x})))(1 - \widehat{R}^{\underline{x}}(x))dx.$$

Then Ψ is differentiable Lebesgue-almost everywhere and the derivative converges to

$$(\gamma(\psi(\underline{b}^-|R)) - \gamma(\psi(r|R)))(1 - R(r)) - \lim_{k,h \to \underline{\psi},\, k > h} (\gamma(k) - \gamma(h)) \cdot \lim_{\underline{x} \downarrow r} \hat{x}'(\underline{x}) \cdot (1 - R(\bar{x})) > 0$$

as $\underline{x}$ approaches r, since $\hat{x}(\underline{x})$ is decreasing in $\underline{x}$, $\psi(b^-|R) > \psi(r|R)$ and γ is strictly increasing. Thus, $\Psi(\underline{x}) > 0$ for $\underline{x}$ sufficiently close to r.

Together, for $\underline{x}$ close enough to r, there exists $\widehat{R}^{\underline{x}} \in \mathcal{I}_\mu$ such that

$$\int_0^1 \gamma(\psi(x|\widehat{R}^{\underline{x}}))(1 - \widehat{R}^{\underline{x}}(x))dx > \int_0^1 \gamma(\psi(x|R))(1 - R(x))dx.$$

On the other hand, let

$$\Xi(\underline{x}) := \int_0^1 \pi(z|R)\gamma(dz) - \int_0^1 \pi(z|\widehat{R}^{\underline{x}})\gamma(dz).$$

Again, for any $\underline{x} > r$, for any $z \in [0,1]\backslash(\psi(r|R), \psi(\tilde{x}(\underline{x})|R)) \cup (h(\underline{x}), \tilde{k}(\underline{x}))$, $\pi(z|R) = \pi(z|\widehat{R}^{\underline{x}})$, which in turn, as in the previous case, implies that

$$\Xi'(r) = 0.$$

Together, for $\underline{x}$ close enough to r, there exists $\widehat{R}^{\underline{x}} \in \mathcal{I}_\mu$ such that

$$W_{\alpha,\beta}(R) < W_{\alpha,\beta}(\widehat{R}).$$

Case 3: $R(r^-) = 0$

Notice that the arguments in **case 2** rely only on the observations that $R(r) = R(r^-) = R(a)$ and $aR(r) = 0$. As such, if $R(r) = R(r^-) = 0$, this is exactly **case 2**. On the other hand, if $R(r) > R(r^-) = 0$, we may define

$$G(x) := \begin{cases} R(r), & \text{if } x \in [0,r) \\ R(x), & \text{if } x \in [r,1]. \end{cases}.$$

Then since the the arguments in **case 2** do not depend on particular value of $\mu = \int_0^1 (1 - R(x))dx$, by the same arguments, there exists $\widehat{G} \in \mathcal{I}_{\bar{\mu}}$ such that

$$W_{\alpha,\beta}(G) < W_{\alpha,\beta}(\widehat{G}),$$

and that $G(\underline{x}) = \widehat{G}(\underline{x}) = \widehat{G}(0)$ for some $\underline{x} \in (r,1)$, where $\bar{\mu} := \int_0^1 (1 - G(x))dx$. Since $\underline{x} > r$ and $G(\underline{x}) \geq G(r) = R(r)$, there exists $\epsilon > 0$ such that $\underline{x}(\widehat{G}(\underline{x}) - \epsilon) = rR(r)$ and therefore $\widehat{R} \in \mathcal{I}_\mu$, where

$$\widehat{R}(x) := \begin{cases} \widehat{G}(\underline{x}) - \epsilon, & \text{if } x \in [0,\underline{x}) \\ \widehat{G}(x), & \text{if } x \in [\underline{x},1] \end{cases}.$$

Furthermore, by construction, $\psi(x|\widehat{G}) = \psi(x|\widehat{R})$ for all $x \in (\underline{x},1]$, $\psi(x|G) = \psi(x|R)$ for all $x \in (r,1]$, $\psi(x|\widehat{G}) < 0$ if and only if $\psi(x|\widehat{R}) < 0$ and $\psi(x|G) < 0$ if and only if $\psi(x|R) < 0$. Together, we have

$$W_{\alpha,\beta}(R) = W_{\alpha,\beta}(G) < W_{\alpha,\beta}(\widehat{G}) = W_{\alpha,\beta}(\widehat{R})$$

for some $\widehat{R} \in \mathcal{I}_\mu$, as desired ■

Lemma 6. *Fix any $\alpha < 0$, $\beta < 0$. For any $R \in \mathcal{I}_\mu \backslash \mathcal{I}_\mu^{**}$, there exists $\widehat{R} \in \mathcal{I}_\mu$ such that*

$$W_{\alpha,\beta}(R) < W_{\alpha,\beta}(\widehat{R}).$$

Proof. Consider any $R \in \mathcal{I}_\mu \backslash \mathcal{I}_\mu^{**}$. Let $a := \inf\{x \in [0,1] | R(x) > 0\}$, $b := \sup\{x \in [0,1] | R(x) < 1\}$ be the lower bound and upper bound of supp(R), respectively and $r := \inf\{x \in [0,1] | \psi(x|R) > 0\}$. Since $R \in \mathcal{I}_\mu$,

$$1 - \mu - \int_0^1 R(x)dx = 1 - \mu - \int_r^1 R(x)dx - \int_a^r R(x)dx = 0. \tag{16}$$

Since R is nondecreasing, it must be that

$$1 - \mu - \int_r^1 R(x)dx - rR(r) \leq 1 - \mu - \int_r^1 R(x)dx - (r-a)R(r) \leq 0.$$

Consider three cases separately:

Case 1: $R(r^-) > 0$ and

$$1 - \mu - \int_r^1 R(x)dx - rR(r) < 0. \tag{17}$$

By Lemma 2 , $x(1 - R(x)) \leq r(1 - R(r^-)))$ for all $x \in [0,1]$. Thus,

$$R(x) \geq \max\left\{0, 1 - \frac{\zeta}{x}\right\}$$

for all $x \in [0,r]$, where $\zeta := r(1 - R(r^-))$. Moreover, since $\psi(x|R) < 0$ for all $x \in [0,r)$ and $R(r^-) > 0$, there exists $\underline{x} \in [0,r)$ such that $R(x) > 1 - \zeta/x > 0$ for all $x \in (\underline{x}, r)$. Together with (16), by the intermediate value theorem, there exists $\bar{b} \in (\underline{x}, b)$ such that

$$\int_{\underline{x}}^{\bar{b}} R(x)dx + \underline{x}R(\underline{x}) + (1 - \bar{b}) = 1 - \mu. \tag{18}$$

Now define

$$\widehat{R}(x) = \begin{cases} 0, & \text{if } x \in [0, \underline{x}) \\ 1 - \frac{\zeta}{x}, & \text{if } x \in [\underline{x}, r) \\ R(x), & \text{if } x \in [r, \bar{b}) \\ 1, & \text{if } x \in [\bar{b}, 1] \end{cases}.$$

By (18), $\widehat{R} \in \mathcal{I}_\mu$. Furthermore, by construction,

$$\psi(x|\widehat{R}) = \begin{cases} \underline{\psi}, & \text{if } x \in [0, \underline{x}) \\ 0, & \text{if } x \in [\underline{x}, r) \\ \psi(x|R), & \text{if } x \in [r, \bar{b}) \\ \bar{b}, & \text{if } x \in [\bar{b}, 1]. \end{cases},$$

for some $\underline{\psi} < 0$. As such,

$$\begin{aligned} \sigma(R) &= \int_r^1 \gamma(\psi(x|R))(1 - R(x))dx \\ &> \int_r^{\bar{b}} \gamma(\psi(x|R))(1 - R(x))dx \\ &= \int_r^1 \gamma(\psi(x|\widehat{R}))(1 - \widehat{R}(x))dx = \sigma(\widehat{R}). \end{aligned}$$

On the other hand,

$$\pi(z|\widehat{R}) = \begin{cases} \pi(z|R), & \text{if } z \in [0, \bar{b}] \\ 0, & \text{if } z \in (\bar{b}, 1] \end{cases},$$

which implies that

$$\pi(R) = \int_0^1 \pi(z|R)\gamma(dz) > \int_0^1 \pi(z|\widehat{R})\gamma(dz) = \pi(\widehat{R}).$$

Together, since $\alpha < 0, \beta < 0$,

$$W_{\alpha,\beta}(R) < W_{\alpha,\beta}(\widehat{R}),$$

as desired.

Case 2:

$$1 - \mu - \int_r^1 R(x)dx - rR(r) = 0. \tag{19}$$

Then by (16) and (19),

$$0 = 1 - \mu - \int_r^1 R(x)dx - \int_a^r R(x) \geq 1 - \mu - \int_r^1 R(x)dx - (r-a)R(r) = aR(r) \geq 0$$

and hence $R(r) = R(r^-)$, $\int_a^r (R(r) - R(x))dx = 0$, $aR(r) = 0$ and therefore $R(r) = R(a)$.

Suppose first that $\psi(x|R) = \bar{\psi}$ for all $x \in (r, b)$, let $\zeta := (r - \bar{\psi})(1 - R(r^-))$, then by Lemma 2

$$R(x) \geq 1 - \frac{\zeta}{x - \bar{\psi}}$$

for all $x \in [0, 1]$. Furthermore, since $R \notin \mathcal{I}_\mu^{**}$, $\phi(x|R) < H(x|R)$ for some $x \in (r, b)$ and thus, by Lemma 2 again, the inequality is strict for some $x \in (r, b)$ such that $1 - \zeta/(x - \bar{\psi}) > 0$. Therefore, there exists $\bar{b} \in [\mu, b)$, $\widehat{R} \in \mathcal{I}_\mu$ such that

$$G_{\zeta,\bar{\psi}}^{R(r^-),\bar{b}} \in \mathcal{I}_\mu^* \subseteq \mathcal{I}_\mu$$

and that

$$\begin{aligned}
\sigma(R) =& \int_0^1 \gamma(\psi(x|R))(1 - R(x))dx \\
=& \gamma(\bar{\psi}) \left(\mu - \int_0^r (1 - R(x))dx \right) \\
=& \gamma(\bar{\psi})(\mu - \zeta - \bar{\psi}(1 - R(r^-))) \\
=& \int_0^1 \gamma\left(\psi\left(x|G_{\zeta,\bar{\psi}}^{R(r^-),\bar{b}}\right)\right)\left(1 - G_{\zeta,\bar{\psi}}^{R(r^-),\bar{b}}(x)\right)dx = \sigma(G_{\zeta,\bar{\psi}}^{R(r^-),\bar{b}}).
\end{aligned}$$

On the other hand, since $\psi(x|R) = \bar{\psi}$ for all $x \in (r, b)$, by Lemma 2 ,

$$\pi(z|R) = \begin{cases} (r - z)(1 - R(r^-)) = \zeta + \bar{\psi} - z, & \text{if } z \in [0, \bar{\psi}] \\ (b - z)(1 - R(r^-)) = \zeta\frac{b-z}{b-\bar{\psi}}, & \text{if } z \in (\bar{\psi}, b] \\ 0, & \text{if } z \in (b, 1] \end{cases}.$$

By construction of $G^{R(r^-),\bar{b}}_{\zeta,\bar{\psi}}$, $\pi(z|R) = \pi(z|G^{R(r^-),\bar{b}}_{\zeta,\bar{\psi}})$ for all $z \in [0, \bar{\psi}]$ and $\pi(z|G^{R(r^-),\bar{b}}_{\zeta,\bar{\psi}}) < \pi(z|R)$ for all $z \in (\bar{\psi}, b]$. Therefore, by Corollary 1,

$$\pi(R) = \int_0^1 \pi(z|R)\gamma(dz) > \int_0^1 \pi(z|G^{R(r^-),\bar{b}}_{\zeta,\bar{\psi}})\gamma(dz) = \pi(G^{R(r^-),\bar{b}}_{\zeta,\bar{\psi}}).$$

Together, as $\beta < 0$, there exists some $\widehat{R} \in \mathcal{I}_\mu$ such that $W_{\alpha,\beta}(\widehat{R}) > W_{\alpha,\beta}(R)$.

Now consider the case when there exists $x_1, x_2 \in (r, b)$ such that $\psi(x_2|R) > \psi(x_1|R) \geq 0$.

Consider first the case when there exists some $x_0 \in (r, 1)$ and a sequence $\{x_n\}$ such that $x_n < x_{n+1} < x_0$ and $\psi(x_n|R) < \psi(x_{n+1}|R) < \psi(x_0^-|R)$ for all $n \in \mathbb{N}$ and that $\{x_n\} \uparrow x_0$. Since $\psi(x_{n+1}|R) > \psi(x_n|R)$ for all $n \in \mathbb{N}$ and $\psi(\cdot|R)$ is nondecreasing, we may take such $\{x_n\}$ so that $\psi(x_n|R) > \psi(y|R)$ for all $y \in [r, x_n)$ for all $n \in \mathbb{N}$. Let $k := \psi(x_0^-|R)$ and $\zeta_0 := (x_0 - k)(1 - R(x_0^-))$. Also, for each $n \in \mathbb{N}$, let $\zeta_n := (x_n - \psi(x_n|R))(1 - R(x_n^-))$. Then by Lemma 2 , since $\psi(x_n|R) > \psi(y|R)$ for all $y \in [r, x_n)$ and since $\psi(x_{n+1}|R) > \psi(x_n|R)$, we must have

$$R(x) \geq \max\left\{1 - \frac{\zeta_0}{x - k}, 1 - \frac{\zeta_n}{x - \psi(x_n|R)}, 0\right\}, \tag{20}$$

for all $x \in [x_n, x_0]$, for all $n \in \mathbb{N}$, with strict inequality holding at some $x \in (x_n, x_0)$. Therefore, there exists $\bar{n} \in \mathbb{N}$ such that whenever $n > \bar{n}$, there exists $b_n \in (x_0, b)$ and $\hat{x}_n \in (x_n, x_0)$ such that:

$$1 - \mu = \int_0^{x_n} R(x)dx + \int_{x_n}^{x_0} \max\left\{1 - \frac{\zeta_0}{x - k}, 1 - \frac{\zeta_n}{x - \psi(x_n|R)}\right\} dx + \int_{x_0}^{b_n} R(x)dx + 1 - b_n \tag{21}$$

$$1 - \frac{\zeta_0}{\hat{x}_n - k} = 1 - \frac{\zeta_n}{\hat{x}_n - \psi(x_n|R)}. \tag{22}$$

As such, for any $n > \bar{n}$, define

$$\widehat{R}^{x_n}(x) := \begin{cases} 1, & \text{if } x \in [b_n, 1] \\ R(x), & \text{if } x \in [x_0, b_n) \\ 1 - \frac{\zeta_0}{x-k}, & \text{if } x \in [\hat{x}_n, x_0) \\ 1 - \frac{\zeta_n}{x-\psi(x_n|R)}, & \text{if } x \in [x_n, \hat{x}_n) \\ R(x), & \text{if } x \in [0, x_n) \end{cases},$$

where $\hat{x}_n$ and b_n are uniquely defined by (21) and (22). Notice that by (20), $\hat{x}_n < \hat{x}_{n+1}$ for all $n \in \mathbb{N}$, $\{\hat{x}_n\} \uparrow x_0$ and $b_n < b_{n+1}$ for all $n \in \mathbb{N}$, $\{b_n\} \uparrow b$.

By construction, for any $n > \bar{n}$,

$$\psi(x|\widehat{R}^{x_n}) = \begin{cases} b_n, & \text{if } x \in [b_n, 1] \\ \psi(x|R), & \text{if } x \in [x_0, b_n) \\ k, & \text{if } x \in [\hat{x}_n, x_0) \\ \psi(x_n|R), & \text{if } x \in [x_n, \hat{x}_n) \\ \psi(x|R), & \text{if } x \in [0, x_n) \end{cases}.$$

As such, the difference in the buyer's surplus between R and $\widehat{R}^{x_n}$ is

$$\int_0^1 \gamma(\psi(x|R))(1-R(x))dx - \int_0^1 \gamma(\psi(x|\widehat{R}^{x_n}))(1-\widehat{R}^{x_n}(x))dx$$
$$= \int_{x_n}^{x_0} \gamma(\psi(x|R))(1-R(x))dx - \int_{x_n}^{x_0} \gamma(\psi(x|\widehat{R}^{x_n}))(1-\widehat{R}^{x_n}(x))dx + \int_{b_n}^{b} \gamma(\psi(x|R))(1-R(x))dx.$$

For any $\bar{x} \in (r, x_0)$, let

$$\Psi(\bar{x}) := \int_{\bar{x}}^{x_0} \gamma(\psi(x|R))(1-R(x))dx - \int_{\bar{x}}^{x_0} \gamma(\psi(x|\widehat{R}^{\bar{x}}))(1-\widehat{R}^{\bar{x}}(x))dx, \tag{23}$$

where $\widehat{R}^{\bar{x}}$ is defined by

$$\widehat{R}^{\bar{x}}(x) := \begin{cases} 1, & \text{if } x \in [b(\bar{x}), 1)] \\ R(x), & \text{if } x \in [x_0, b(\bar{x})) \\ 1 - \frac{\zeta_0}{x-k}, & \text{if } x \in [\hat{x}(\bar{x}), x_0) \\ 1 - \frac{\zeta_{\bar{x}(\underline{x})}}{x - \psi(\bar{x}(\underline{x})|R)}, & \text{if } x \in [\bar{x}, \hat{x}(\bar{x})) \\ R(x), & \text{if } x \in [0, \bar{x}) \end{cases},$$

with $\hat{x}(\bar{x})$ and $b(\bar{x})$ being uniquely defined by

$$1-\mu = \int_0^{\bar{x}} R(x)dx + \int_{\bar{x}}^{x_0} \max\left\{1 - \frac{\zeta_0}{x-k}, 1 - \frac{\zeta_x}{x - \psi(\bar{x}|R)}\right\} dx + \int_{x_0}^{b(\bar{x})} R(x)dx + 1 - b(\bar{x}) \tag{24}$$

$$1 - \frac{\zeta_0}{\hat{x}(\bar{x}) - k} = 1 - \frac{\zeta_{\bar{x}}}{\hat{x}(\bar{x}) - \psi(\bar{x}|R)}, \tag{25}$$

and $\zeta_x := (x - \psi(x|R))(1 - R(x^-))$ for any $x \in [r, 1]$.

Notice that by (21), (22), (24) and (25), for any $n > \bar{n}$, $\hat{x}_n = \hat{x}(\bar{x}_n)$. Also, $\hat{x}$ is increasing in $\bar{x}$ and $\lim_{\bar{x}\to x_0} \hat{x}(\bar{x}) = x_0$, $\lim_{\bar{x}\to x_0} \psi(\hat{x}(\bar{x})|R) = \lim_{n\to\infty} \psi(x_n|R) = \psi(x_0^-|R)$.

Since $\psi(\cdot|R)$ and γ are nondecreasing, by (25), $\hat{x}$ is differentiable Lebesgue-almost everywhere and therefore Ψ is differentiable Lebesgue-almost everywhere. Thus, for Lebesgue almost all $\bar{x}$,

$$\Psi'(\bar{x}) = \gamma(\psi(\bar{x}|\widehat{R}^{\bar{x}}(\bar{x})))(1-\widehat{R}^{\bar{x}}(\bar{x})) - \gamma(\psi(\bar{x}|R))(1-R(\bar{x})) + \int_{\bar{x}}^{x_0} \frac{\partial}{\partial \bar{x}} \gamma(\psi(x|\widehat{R}^{\bar{x}}))(1-\widehat{R}^{\bar{x}}(x))dx.$$

By construction, $\psi(\bar{x}|\widehat{R}^{\bar{x}}) = \psi(\bar{x}|R)$ and $R(\bar{x}) = \widehat{R}^{\bar{x}}(\bar{x})$ for all $\bar{x}$. Moreover,

$$\frac{\partial}{\partial \bar{x}} \gamma(\psi(x|\widehat{R}^{\bar{x}}))(1-\widehat{R}^{\bar{x}}(x))$$

is uniformly bounded. As such,

$$\lim_{\bar{x}\uparrow x_0} \Psi'(\bar{x}) = 0. \tag{26}$$

Together, for n sufficiently large,

$$\int_{x_n}^{x_0} \gamma(\psi(x|R))(1-R(x))dx - \int_{x_n}^{x_0} \gamma(\psi(x|\widehat{R}^{x_n}))(1-\widehat{R}^{x_n}(x))dx + \int_{b_n}^{b} \gamma(\psi(x|R))(1-R(x))dx > 0$$

Thus, for n sufficiently large,

$$\int_0^1 \gamma(\psi(x|\widehat{R}^{x_n}))(1-\widehat{R}^{x_n}(x))dx < \int_0^1 \gamma(\psi(x|R))(1-R(x))dx,$$

On the other hand, for $n \in \mathbb{N}$ large enough, the difference in seller's profit between R and $\widehat{R}^{x_n}$ is

$$\int_0^1 \pi(z|R)dz - \int_0^1 \pi(z|\widehat{R}^{x_n})\gamma(dz).$$

Notice that for each $n \in \mathbb{N}$, any $z \in [0,1]\backslash(\psi(x_n|R), k) \cup (b_n, b)$, $\pi(z|R) = \pi(z|\widehat{R}^{x_n})$. Therefore,

$$\begin{aligned}&\int_0^1 \pi(z|R)\gamma(dz) - \int_0^1 \pi(z|\widehat{R}^{x_n})\gamma(dz)\\ =&\int_{\psi(x_n|R)}^k \gamma'(z)(\pi(z|R)-\pi(z|\widehat{R}^{x_n}))dz + \int_{b_n}^b \gamma'(z)(\pi(z|R)-\pi(z|\widehat{R}^{x_n}))dz.\end{aligned}$$

Moreover, $\pi(z|\widehat{R}^{x_n}) < \pi(z|R)$ for all $z \in (b_n, b)$. Now let

$$\Xi(\bar{x}) := \int_{\psi(\bar{x}|R)}^k \gamma'(z)(\pi(z|R)-\pi(z|\widehat{R}^{\bar{x}}))dz. \tag{27}$$

Then

$$\Xi'(\bar{x}) = \int_{\psi(\bar{x}|R)}^k \frac{\partial}{\partial \bar{x}}\gamma'(z)(\pi(z|R)-\pi(z|\widehat{R}^{\bar{x}}))dz,$$

as $\pi(\psi(\bar{x}|R)|R) = \pi(\psi(\bar{x}|R)|\widehat{R}^{\bar{x}})$. Thus,

$$\lim_{\bar{x}\uparrow x_0} \Xi'(\bar{x}) = 0.$$

Together, for $n \in \mathbb{N}$ large enough,

$$\int_0^1 \pi(z|R)\gamma(dz) - \int_0^1 \pi(z|\widehat{R}^{x_n})\gamma(dz) > 0.$$

Combining, since $\alpha < 0$ and $\beta < 0$, for $n \in \mathbb{N}$ sufficiently large, $\widehat{R}^{x_n} \in \mathcal{I}_\mu$ and

$$W_{\alpha,\beta}(R) < W_{\alpha,\beta}(\widehat{R}^{x_n}),$$

as desired.

If, on the other hand, for any $x \in (r,1)$ and for any sequence $\{x_n\}$ such that $x_n < x_{n+1} < x$, $\{x_n\} \uparrow x$, there exists $\bar{n} \in \mathbb{N}$ such that $\psi(x_n|R) = \psi(x^-|R)$ for all $n > \bar{n}$, then for any $x \in (r,1)$, there exists $\delta > 0$ such that $\psi(y|R) = \psi(x^-|R)$ for all $y \in (x-\delta, x)$. Let $\delta_x := \sup\{\delta > 0 | \psi(y|R) = \psi(x^-|R), \forall y \in (x-\delta, x)\}$. Then $\delta_x > 0$ for all $x \in (r,1)$. Moreover, for any $x, y \in (r,1)$, if $\psi(x^-|R) \neq \psi(y^-|R)$, then it must be that $(x-\delta_x, x) \cap (y-\delta_y, y) = \emptyset$. Therefore, $\{\psi(x^-|R)\}_{x\in(r,1)}$ is at most countable. Since $\psi(\cdot|R)^+$ is nondecreasing, it must be a step function. Consider first the case when for any $\delta > 0$, there exists $x, y \in (b-\delta, b)$ with $x < y$ such that $\psi(x|R) < \psi(y|R) <$

$\psi(b^-|R)$. Since $\psi(\cdot|R)^+$ is a step function, $\psi(\cdot|R)^+$ has countably infinitely many jumps and therefore we may represent $\psi(\cdot|R)^+$ as

$$\psi(x|R) = \sum_{n=1}^{\infty} \psi_n \mathbf{1}\{x \in (\alpha_n, \beta_n)\}, \; \forall x \in [r,1] \backslash [\{\alpha_n\}_{n=1}^{\infty} \cup \{\beta_n\}_{n=1}^{\infty}].$$

for some $\{\alpha_n\}$, $\{\beta_n\}$ such that $\alpha_n < \beta_n$ for all $n \in \mathbb{N}$ and some $\{\psi_n\}_{n=1}^{\infty}$ such that $\psi_n > \psi(r|R)$ for all $n \in \mathbb{N}$, $n \geq 2$. Since for any $\delta > 0$, there exists $x, y \in (b-\delta, b)$, $x < y$, such that $\psi(x|R) < \psi(y|R) < \psi(b^-|R)$, there exists a sequence $\{x_j\}$ such that $x_j \in \{\alpha_n\}_{n=1}^{\infty} \cup \{\beta_n\}_{n=1}^{\infty}$, $x_j < x_{j+1}$, $\psi_j := \psi(x_j|R) = \psi(x_{j+1}^-|R) < \psi(x_{j+1}|R) =: \psi_{j+1}$ for all $j \in \mathbb{N}$, and that $\{x_j\} \uparrow b$ and $\{\psi_j\} \uparrow \psi(b^-|R)$.

Since $\psi(\cdot|R)^+$ is a step function and $\psi(x|R) = b$ for all $x \in [b,1]$, we must have $\psi(b^-|R) < \psi(b|R) = b$ and $R(b^-) < 1$. As such, let $\zeta_b := (b - \psi(b^-|R))(1 - R(b^-))$. Then $\zeta_b > 0$. Also, for each $j \in \mathbb{N}$, let $\zeta_j := (x_j - \psi_j)(1 - R(x_j^-))$, by Lemma 2 , since $\psi_{j-1} < \psi_j < \psi_{j+1}$, we have

$$R(x) > \max\left\{1 - \frac{\zeta_b}{x - \psi(b^-|R)}, 1 - \frac{\zeta_j}{x - \psi_j}\right\},$$

for all $x \in (x_j, b)$.

As such, there exists a sequence $\{b_j\}$ such that $\{b_j\} \uparrow b$ and $\widehat{R}^j \in \mathcal{I}_\mu$ for j large enough, where

$$\widehat{R}^j(x) := \begin{cases} 1, & \text{if } x \in [b_j, 1] \\ 1 - \frac{\zeta_b}{x - \psi(b^-|R)}, & \text{if } x \in [\hat{x}_j, b_j) \\ 1 - \frac{\zeta_j}{x - \psi_j}, & \text{if } x \in [\bar{x}_j, \hat{x}_j) \\ R(x), & \text{if } x \in [0, \bar{x}_j) \end{cases}.$$

and $r < \hat{x}_j < \bar{x}_j < b_j$ are uniquely defined by

$$\int_0^{b_j} \widehat{R}^j(x)dx = \int_0^b R(x)dx$$
$$1 - \frac{\zeta_b}{\hat{x}_j - \psi(b^-|R)} = 1 - \frac{\zeta_j}{\hat{x}_j - \psi_j},$$

Similar to the previous case, for each $j \in \mathbb{N}$ such that $\widehat{R}^j \in \mathcal{I}_\mu$, the difference in the buyer's surplus between R and $\widehat{R}^j$ is

$$\int_{\bar{x}_j}^{b_j} \gamma(\psi(x|R))(1 - R(x))dx - \int_{\bar{x}_j}^{b_j} \gamma(\psi(x|\widehat{R}^j))(1 - \widehat{R}^j(x))dx + \int_{b_j}^{b} \psi(x|R)(1 - R(x))dx.$$

As shown in the previous case, from (23), (24), (25) and (26), with $x_0 = b$ and $k = \psi(b^-|R)$, $\lim_{b_j \uparrow b} \Psi'(b_j) = 0$. Thus, since by (24) and (25), $\hat{x}_j = \hat{x}(\bar{x}_j)$ and since $\{\bar{x}_j\} \uparrow b$, for j sufficiently large, $\widehat{R}^j \in \mathcal{I}_\mu$ and

$$\int_0^1 \gamma(\psi(x|\widehat{R}^j))(1 - \widehat{R}^j(x))dx < \int_0^1 \gamma(\psi(x|R))(1 - R(x))dx,$$

On the other hand, the difference in seller's profit between R and $\widehat{R}^j$ is

$$\int_0^1 \pi(z|R)\gamma(dz) - \int_0^1 \pi(z|\widehat{R}^j)\gamma(dz).$$

Similar to the previous case, for all $j \in \mathbb{N}$, $\pi(z|R) = \pi(z|\widehat{R}^j)$ for all $z \in [0,1]\backslash(\psi_j, b_j)$. Moreover, $\{b_j\} \to b$ as $j \to \infty$. Thus, by (27), $\lim_{\bar{x}\to x_0} \Xi'(\bar{x}) = 0$ implies

$$\int_0^1 \pi(z|R)\gamma(dz) - \int_0^1 \pi(z|\widehat{R}^j)\gamma(dz) > 0.$$

Together, there exists $\widehat{R}^j \in \mathcal{I}_\mu$ such that

$$W_{\alpha,\beta}(R) < W_{\alpha,\beta}(\widehat{R}^j),$$

as desired.

Finally, if there exists $\delta > 0$ such that for any $x, y \in (b-\delta, b)$, $\psi(x|R) = \psi(y|R) = \psi(b^-|R)$. Let $\underline{b} := \inf\{b_0 \in [r,1) | \psi(x|R) = \psi(y|R), \forall x, y \in (b_0, b)\}$ and let $\underline{\psi} := \psi(b^-|R)$. Then $\underline{b} < b$ and $\underline{\psi} > \psi(r|R)$ since $\psi(\cdot|R)$ is not a constant on $[r, b)$. We claim that it is without loss to suppose that $\phi(x|R) = H(x|R)$ for all $x \in (\underline{b}, b)$. Indeed, if there exists $x \in (\underline{b}, b)$ such that $\phi(x|R) < H(x|R)$, let $\zeta := (\underline{b} - \underline{\psi})(1 - R(\underline{b}^-))$. Lemma 2 then ensures that

$$R(x) \geq 1 - \frac{\zeta}{x - \underline{\psi}},$$

for all $x \in (\underline{b}, b)$ with strict inequality for some $x \in (\underline{b}, b)$ and therefore, there exists $\hat{b} \in (\underline{b}, b)$ such that

$$\int_{\underline{b}}^{b} (1 - R(x))dx = \int_{\underline{b}}^{\hat{b}} \frac{\zeta}{x - \underline{\psi}} dx.$$

Therefore, for

$$\widehat{R}(x) := \begin{cases} R(x), & \text{if } x \in [0, \underline{b}) \\ 1 - \frac{\zeta}{x - \underline{\psi}}, & \text{if } x \in [\underline{b}, \hat{b}) \\ 1, & \text{if } x \in [\hat{b}, 1] \end{cases},$$

$\widehat{R} \in \mathcal{I}_\mu$,

$$\int_0^1 \gamma(\psi(x|R))(1 - R(x))dx = \int_0^1 \gamma(\psi(x|\widehat{R}))(1 - \widehat{R}(x))dx$$
$$\int_0^1 \pi(z|R)\gamma(dz) = \int_0^1 \pi(z|\widehat{R})\gamma(dz).$$

and $H(x|\widehat{R}) = \phi(x|\widehat{R})$ for all $x \in (\underline{b}, b)$.

Therefore, as $H(x|R) = \phi(x|R)$ for all $x \in (\underline{b}, b)$ and $\psi(x|R) = \underline{\psi}$ for all $x \in (\underline{b}, b)$, we must have

$$R(x) = 1 - \frac{\zeta}{x - \underline{\psi}}, \forall x \in [\underline{b}, b),$$

for some $\zeta > 0$. Now take and fix any $\bar{x} \in (\underline{b}, b)$ and notice that for any $h \in (\psi(\underline{b}^-|R), \underline{\psi})$ and any $k \in (\underline{\psi}, b)$, let $\zeta(k) := (\bar{x} - k)(1 - R(\bar{x}^-))$ and $\zeta(h) := (\underline{b} - h)(1 - R(\underline{b}^-))$, we must have

$$R(x) > \max\left\{1 - \frac{\zeta(k)}{x-k}, 1 - \frac{\zeta(h)}{x-h}\right\},$$

for all $x \in (\underline{b}, \bar{x})$. Thus, for any $\bar{b} \in (\underline{b}, b)$ that is close enough to b, there exists $h(\bar{b}) \in (\psi(\underline{b}^-|R), \underline{\psi})$,$k(\bar{b}) \in (\underline{\psi}, b)$ and $\hat{x}(\bar{b}) \in (\underline{b}, \bar{b})$ such that $\lim_{\bar{b}\uparrow b} h(\bar{b}) = \lim_{\bar{b}\uparrow b} k(\bar{b}) = \underline{\psi}$,

$$\frac{\zeta(k(\bar{b}))}{\hat{x}(\bar{b}) - k(\bar{b})} = \frac{\zeta(h(\bar{b}))}{\hat{x}(\bar{b}) - h(\bar{b})}$$

and

$$\int_0^b R(x)dx = \int_0^b \widehat{R}^{\underline{x}}(x)dx,$$

where

$$\widehat{R}^{\bar{b}}(x) := \begin{cases} 1, & \text{if } x \in [\bar{b}, 1] \\ R(x), & \text{if } x \in [\bar{x}, \bar{b}) \\ 1 - \frac{\zeta(k(\bar{b}))}{x - k(\bar{b})}, & \text{if } x \in [\hat{x}(\bar{b}), \bar{x}) \\ 1 - \frac{\zeta(h(\bar{b}))}{x - h(\bar{b})}, & \text{if } x \in [\underline{b}, \hat{x}(\bar{b})) \\ R(x), & \text{if } x \in [0, \underline{b}) \end{cases}$$

and thus $\widehat{R}^{\bar{b}} \in \mathcal{I}_\mu$. Moreover, such $h(\cdot)$ and $k(\cdot)$ can be selected so that $\hat{x}(\bar{b})$ is decreasing in $\bar{b}$ and $\lim_{\bar{b}\uparrow b} \hat{x}(\bar{b}) = \bar{x}$.

Notice that for any such $\widehat{R}^{\bar{b}}$,

$$\psi(x|\widehat{R}^{\bar{b}}) = \begin{cases} \psi(x|R), & \text{if } x \in [0, \underline{b}) \\ h(\bar{b}), & \text{if } x \in [\underline{b}, \hat{x}(\bar{x})) \\ \tilde{k}(\bar{b}), & \text{if } x \in [\hat{x}(\bar{b}), b) \\ \bar{b}, & \text{if } x \in [\bar{b}, 1]. \end{cases},$$

for some $\underline{\psi} < 0$, some $\tilde{k}(\bar{b}) \in (\underline{\psi}, k(\bar{b}))$. As such, for any $\bar{b}$ such that $\widehat{R}^{\bar{b}} \in \mathcal{I}_\mu$, the difference in the buyer's surplus between R and $\widehat{R}^{\bar{b}}$ is

$$\begin{aligned} &\int_0^1 \gamma(\psi(x|R))(1 - R(x))dx - \int_0^1 \gamma(\psi(x|\widehat{R}^{\bar{b}}))(1 - \widehat{R}^{\bar{b}}(x))dx \\ =&\int_{\underline{b}}^{\bar{b}} \gamma(\psi(x|R))(1 - R(x))dx - \int_{\underline{b}}^{\bar{b}} \gamma(\psi(x|\widehat{R}^{\bar{b}}))(1 - \widehat{R}^{\bar{b}}(x))dx + \int_{\bar{b}}^{b} \gamma(\psi(x|R))(1 - R(x))dx \end{aligned}$$

Let

$$\Psi(\bar{b}) := \int_{\underline{b}}^{\bar{b}} \gamma(\psi(x|R))(1 - R(x))dx - \int_{\underline{b}}^{\bar{b}} \gamma(\psi(x|\widehat{R}^{\bar{b}}))(1 - \widehat{R}^{\bar{b}}(x))dx$$

Then Ψ is differentiable Lebesgue-almost everywhere and

$$\lim_{\bar{b}\uparrow b} \Psi'(\bar{b}) = 0.$$

Together, for $\bar{b}$ close enough to b, there exists $\widehat{R}^{\bar{b}} \in \mathcal{I}_\mu$ such that

$$\int_0^1 \gamma(\psi(x|\widehat{R}^{\bar{b}}))(1-\widehat{R}^{\bar{b}}(x))dx < \int_0^1 \gamma(\psi(x|R))(1-R(x))dx.$$

On the other hand, let

$$\Xi(\bar{b}) := \int_0^1 \pi(z|R)\gamma(dz) - \int_0^1 \pi(z|\widehat{R}^{\bar{b}})\gamma(dz).$$

Again, for any $\bar{b} < b$, for any $z \in [0,1]\backslash(\psi(\underline{b}^-|R), b)$, $\pi(z|R) = \pi(z|\widehat{R}^{\bar{x}})$. Moreover, $\pi(\psi(\underline{b}^-|R)|R) = \pi(\psi(\underline{b}^-|R|\widehat{R}^{\bar{b}}))$ and $\pi(z|R) > \pi(z|\widehat{R}^{\bar{b}})$ for all $z \in (\bar{b}, b)$. Together,

$$\lim_{\bar{b}\uparrow b} \Xi'(\bar{b}) = 0$$

Together, for $\bar{b}$ close enough to b, there exists $\widehat{R}^{\bar{b}} \in \mathcal{I}_\mu$ such that

$$W_{\alpha,\beta}(R) < W_{\alpha,\beta}(\widehat{R}).$$

as desired.
Case 3: $R(r^-) = 0$

Notice that the arguments in **case 2** rely only on the observations that $R(r) = R(r^-) = R(a)$ and $aR(r) = 0$. As such, if $R(r) = R(r^-) = 0$, this is exactly **case 2**. On the other hand, if $R(r) > R(r^-) = 0$, we may define

$$G(x) := \begin{cases} R(r), & \text{if } x \in [0,r) \\ R(x), & \text{if } x \in [r,1]. \end{cases}.$$

Then since the the arguments in **case 2** do not depend on particular value of $\mu = \int_0^1 (1-R(x))dx$, by the same arguments, there exists $\widehat{G} \in \mathcal{I}_{\bar{\mu}}$ such that

$$W_{\alpha,\beta}(G) < W_{\alpha,\beta}(\widehat{G}),$$

and that $G(\underline{x}) = \widehat{G}(\underline{x}) = \widehat{G}(0)$ for some $\underline{x} \in (r,1)$, where $\bar{\mu} := \int_0^1 (1-G(x))dx$. Since $\underline{x} > r$ and $G(\underline{x}) \geq G(r) = R(r)$, there exists $\epsilon > 0$ such that $\underline{x}(\widehat{G}(\underline{x}) - \epsilon) = rR(r)$ and therefore $\widehat{R} \in \mathcal{I}_\mu$, where

$$\widehat{R}(x) := \begin{cases} \widehat{G}(\underline{x}) - \epsilon, & \text{if } x \in [0,\underline{x}) \\ \widehat{G}(x), & \text{if } x \in [\underline{x},1] \end{cases}.$$

Furthermore, by construction, $\psi(x|\widehat{G}) = \psi(x|\widehat{R})$ for all $x \in (\underline{x},1]$, $\psi(x|G) = \psi(x|R)$ for all $x \in (r,1]$, $\psi(x|\widehat{G}) < 0$ if and only if $\psi(x|\widehat{R}) < 0$ and $\psi(x|G) < 0$ if and only if $\psi(x|R) < 0$. Together, we have

$$W_{\alpha,\beta}(R) = W_{\alpha,\beta}(G) < W_{\alpha,\beta}(\widehat{G}) = W_{\alpha,\beta}(\widehat{R})$$

for some $\widehat{R} \in \mathcal{I}_\mu$, as desired ■

With the above two lemmas, we know that if the solution of the planner's problem

$$\sup_{R\in\mathcal{I}_\mu} W_{\alpha,\beta}(R)$$

exists, then it must be in the family $\mathcal{I}_\mu^{**}$. Lemma 7 below establishes closedness of the set Σ, which means that the planner's problem always exists.

Lemma 7. *Let $\mathcal{Q}$ be endowed with the weak-* topology on the set of probability measures induced by $\{q^+|q\in\mathcal{Q}\}$. Then the function $\widetilde{\pi}$ is continuous and the function $\widetilde{\sigma}$ is lower-semicontinuous. In particular, the set Σ is closed.*

Proof. Continuity of $\widetilde{\pi}$ immediately follows from the Portmanteau lemma as V is bounded and continuous. Now consider a sequence $\{q_n\}\subset\mathcal{Q}$ such that $\{q_n^+\}\to q^+$ for some $q\in\mathcal{Q}$ under the weak-* topology. For each $n\in\mathbb{N}$, let $Q_n(z):=\gamma(z)\int_z^1(1-q(x))dx/(1-q(z))$ and let $Q(z):=\gamma(z)\int_z^1(1-q(x))dx/(1-q(z))$ for all $z\in[0,1]$ with $q_n(z)<1$ and $q(z)<1$ for all $n\in\mathbb{N}$. Notice that Q is lower-semicontinuous since q is left-continuous. By the Portmanteau lemma,

$$\liminf_{n\to\infty}\int_0^1 Q(z)q_n^+(dz)\geq\int_0^1 Q(z)q^+(dz). \tag{28}$$

On the other hand, since $\{q_n^+\}\to q^+$ under the weak-* topology and since γ is continuous, $\{Q_n\}\to Q$ q^+-almost everywhere. Thus, by Egoroff's theorem, for any $\varepsilon>0$, there exists a closed set $A\subset[0,1]$ such that $\nu_q(A)<\varepsilon$ and that $\{Q_n\}\to Q$ uniformly on A, where ν_q is the probability measure associated with q^+. As such,

$$\begin{aligned}
&\liminf_{n\to\infty}\int_0^1 Q_n(z)q_n^+(dz)-\int_0^1 Q(z)q^+(dz)\\
=&\liminf_{n\to\infty}\int_0^1(Q_n(z)-Q(z))q_n^+(dz)+\int_0^1 Q(z)q_n^+(dz)-\int_0^1 Q(z)q^+(dz)\\
=&\liminf_{n\to\infty}\left[\int_{[0,1]\setminus A}(Q_n(z)-Q(z))q_n^+(dz)+\int_A(Q_n(z)-Q(z))q_n^+(dz)\right]+\liminf_{n\to\infty}\left[\int_0^1 Q(z)q_n^+(dz)-\int_0^1 Q(z)q^+(dz)\right]\\
\geq&\liminf_{n\to\infty}\inf_{x\in[0,1]\setminus A}(Q_n(z)-Q(z))-\limsup_{n\to\infty}\nu_{q_n}(A)+\liminf_{n\to\infty}\left[\int_0^1 Q(z)q_n^+(dz)-\int_0^1 Q(z)q^+(dz)\right]\\
\geq&-\nu_q(A)\\
\geq&-\varepsilon,
\end{aligned}$$

where the last inequality follows from the uniform convergence of $\{Q_n\}$ on $[0,1]\setminus A$, the Portmanteau lemma, which establishes that $\limsup_{n\to\infty}\nu_{q_n}(A)\leq\nu_q(A)$ as A is closed, and from (28). Since ε is arbitrary, we have

$$\liminf_{n\to\infty}\int_0^1 Q_n(z)q_n^+(dz)\geq\int_0^1 Q(z)q^+(dz),$$

as desired.

Finally, closedness of Σ then follows from Lemma 2 in Yang (2019a), Lemma 4 and the continuity results obtained above. ■

A.3 Proofs of Main Results

A.3.1 Proof of Proposition 1

Proof of Proposition 1. Notice that by Lemma 1, for any $G \in \mathcal{I}_\mu$,

$$\begin{aligned}\sigma(G) + \pi(G) &= \int_0^1 \gamma(\psi(x|G))(1 - G(x))dx + \int_0^1 V(\psi(x|G))G(dx) \\ &= \int_0^1 \gamma(\psi(x|G))(x - \psi(x|G))G(dx) + \int_0^1 \big[\gamma(\psi(x|G))\psi(x|G) - c(\gamma(\psi(x|G)))\big]G(dx) \\ &= \int_0^1 \big[\gamma(\psi(x|G))x - c(\gamma(\psi(x|G)))\big]G(dx).\end{aligned}$$

For any G such that $\sigma(G) > 0$, it must be that $\mu_G((0,1)) > 0$ and that $\mu_G(\{x \in [0,1]|\psi(x|G) \geq 0\}) > 0$, where $\mu_G \in \Delta([0,1])$ is the probability measure induced by G. Moreover, by definition of γ and by strict convexity of c, it must be that

$$\gamma(\psi(x|G))x - c(\psi(x|G)) < V(x),$$

for all $x \in [0,1]$ such that $\psi(x|G) \neq x$. Together, since $\psi(x|G) \neq x$ for all $0 < x < \sup \mathrm{supp}(G)$ and $\mu_G(\{x \in [0,1]|\psi(x|G) \geq 0\}) > 0$,

$$\int_0^1 \big[\gamma(\psi(x|G))x - c(\gamma(\psi(x|G)))\big]G(dx) < \int_0^1 V(x)G(dx).$$

Finally, since $V'(z) = \gamma(z)$ for almost all $z \in [0,1]$ by the envelope theorem (Milgrom and Segal, 2002) and since γ is increasing by convexity of c, V is convex. Together, since any $G \in \mathcal{I}_\mu$ is a mean-preserving contraction of the distribution $\bar{G} := \mu\delta_{\{1\}} + (1-\mu)\delta_{\{0\}}$ and since $V(0) = 0$ as $c(0) = 0$, we have

$$\begin{aligned}\sigma(G) + \pi(G) &= \int_0^1 \big[\gamma(\psi(x|G))x - c(\gamma(\psi(x|G)))\big]G(dx) \\ &< \int_0^1 V(x)G(dx) \\ &\leq \mu V(1).\end{aligned}$$

Moreover, notice that under $\bar{G}$, $\sigma(\bar{G}) = 0$ and $\pi(\bar{G}) = \mu V(1)$. Therefore, $\Sigma \cap S^* = (0, \mu V(1))$, as desired. ■

A.3.2 Proof of Theorem 1

Proof of Theorem 1. Lemma 5, Lemma 6 and Lemma 7 imply that

$$\partial\Sigma \subseteq \{(\sigma(R), \pi(R))|R \in \mathcal{I}_\mu^{**}\} \subseteq \Sigma. \tag{29}$$

Define

$$\begin{aligned}\sigma^* := &\max_{\pi,\eta,k,b} \gamma(k)(\mu - (\pi + (1-\eta)k)) \\ &\text{s.t. } \pi + (1-\eta)k + \pi \log\left(\frac{(1-\eta)(b-k)}{\pi}\right) = \mu.\end{aligned}$$

For any $\sigma \in [0, \sigma^*]$, define

$$\Sigma_\sigma := \{\pi \in \mathbb{R}_+ | \pi = \pi(R), \sigma = \sigma(R), \text{ for some } R \in \mathcal{I}_\mu\}$$

and let $\bar{\pi}_\sigma$ and $\underline{\pi}_\sigma$ be the value of the problems

$$\begin{aligned}\max_{\pi,\eta,k,b} & \left(1 - \frac{\pi}{b-k} - \eta\right) V(k) + \frac{\pi}{b-k} V(b) \\ \text{s.t. } & \pi + (1-\eta)k + \pi \log\left(\frac{(1-\eta)(b-k)}{\pi}\right) = \mu \\ & \gamma(k)(\mu - (\pi + (1-\eta)k)) = \sigma\end{aligned}$$

and

$$\begin{aligned}\min_{\pi,\eta,k,b} & \left(1 - \frac{\pi}{b-k} - \eta\right) V(k) + \frac{\pi}{b-k} V(b) \\ \text{s.t. } & \pi + (1-\eta)k + \pi \log\left(\frac{(1-\eta)(b-k)}{\pi}\right) = \mu \\ & \gamma(k)(\mu - (\pi + (1-\eta)k)) = \sigma\end{aligned}$$

respectively. (29) then implies that

$$\Sigma_\sigma = [\underline{\pi}_\sigma, \bar{\pi}_\sigma], \forall \sigma \in [0, \sigma^*],$$

which in turn, by continuity of V, implies that

$$\Sigma = \{(\sigma(R), \pi(R) | R \in \mathcal{I}_\mu^{**}\},$$

as desired. ■

A.3.3 Proof of Proposition 2

Proof of Proposition 2. Take any fix any $\varepsilon \in (0, 1)$. Let Γ_ε be the collection of continuous and nondecreasing functions on $[\varepsilon, 1]$ with range being a subset of $[\varepsilon, 1]$. Endow Γ_ε with the topology induced by the sup-norm on $C([\varepsilon, 1])$. Also, let

$$\Theta_\varepsilon := \left\{(\pi, k, \eta, b) \in [0, 1]^4 \middle| k \in [\varepsilon, 1], \, \pi + (1-\eta)k + \pi \log\left(\frac{(b-k)(1-\eta)}{\pi}\right) = \mu\right\}.$$

Define the operators $\Phi_\sigma^\varepsilon : \Gamma_\varepsilon \to \mathbb{R}^{\Theta_\varepsilon}$ and $\Phi_p^\varepsilon : \Gamma_\varepsilon \to \mathbb{R}^{\Theta_\varepsilon}$ as the following

$$[\Phi_\sigma^\varepsilon(\gamma)](\pi, k, \eta, b) := \gamma(k)(\mu - (\pi + (1-\eta)k))$$

and

$$[\Phi_p(\gamma)](\pi, k, \eta, b) := (1-\eta)\left(\frac{\varepsilon^2 \gamma(\varepsilon)}{2} + \int_\varepsilon^k \gamma(z) dz\right) + \frac{\pi}{b-k} \int_k^b \gamma(z) dz,$$

for all $\gamma \in \Gamma_\varepsilon$ and for all $(\pi, k, \eta, b) \in \Theta_\varepsilon$. Let $Y_\sigma^\varepsilon := \Phi_\sigma^\varepsilon(\Gamma_\varepsilon)$ and $Y_p^\varepsilon := \Phi_p^\varepsilon(\Gamma_\varepsilon)$. Endow Y_p^ε and Y_σ^ε the topology induced by the sup norm on $C(\Theta_\varepsilon)$. Also, let $\mathcal{F}_\varepsilon$ be the collection of continuous linear maps $F : Y_p^\varepsilon \to Y_\sigma^\varepsilon$ and endow $\mathcal{F}_\varepsilon$ with the pointwise topology. Finally, for any $y_\sigma \in Y_\sigma^\varepsilon$ and for any $y_p \in Y_p^\varepsilon$, say that y_σ is ε-measurable with respect to y_p if there exists $y_p' \in N_\varepsilon(y_p)$ such that y_σ is measurable with respect to y_p'.

Now let

$$Z_\varepsilon := \{(\gamma, F) \in \Gamma_\varepsilon \times \mathcal{F}_\varepsilon | \Phi_\sigma^\varepsilon(\gamma) \text{ is } \varepsilon\text{-measurable with respect to } \Phi_p^\varepsilon(\gamma)\}.$$

Notice that Z_ε is nonempty,[10] convex and compact.[11]

Since $\gamma(k) \in [\varepsilon, 1]$ for all $k \in [\varepsilon, 1]$, for any $F \in \mathcal{F}_\varepsilon$, the correspondence $(\Phi_p^\varepsilon)^{-1}(F(\Phi_\sigma^\varepsilon(\cdot))) : \Gamma_\varepsilon \rightrightarrows \Gamma_\varepsilon$ is nonempty-valued and lower-hemicontinuous and hence admits a continuous selection $\phi_\varepsilon(\cdot, F) : \Gamma_\varepsilon \to \Gamma_\varepsilon$. Moreover, since Φ_σ^ε and Φ_p^ε are continuous on Γ_ε, ϕ_ε is continuous in (γ, F).

Now define an operator $\mathbb{W}_\varepsilon$ as

$$\mathbb{W}_\varepsilon(\gamma, F) := \begin{pmatrix} \phi_\varepsilon(\gamma, F) \\ F \end{pmatrix}.$$

Then, $\mathbb{W}_\varepsilon : Z_\varepsilon \to Z_\varepsilon$ is a continuous self-map. Moreover, since Z_ε is a nonempty, compact and convex subset of a locally convex topological vector space, by Schauder's fixed point theorem, $\mathbb{W}_\varepsilon$ has a fixed point $(\gamma_\varepsilon, F_\varepsilon)$. Thus

$$\Phi_p^\varepsilon(\gamma_\varepsilon) = F_\varepsilon(\Phi_\sigma^\varepsilon(\gamma_\varepsilon)).$$

and $\Phi_\sigma^\varepsilon(\gamma_\varepsilon)$ is ε-measurable with respect to $F_\varepsilon(\Phi_p^\varepsilon(\gamma_\varepsilon))$. For this fixed point $(\gamma_\varepsilon, F_\varepsilon)$, continuously extend γ_ε so that it is defined on $[0, 1]$ by letting

$$\gamma_\varepsilon(k) := \frac{\gamma(\varepsilon)}{\varepsilon} k, \ \forall k \in [0, \varepsilon].$$

By Helley's selection theorem, after possibly taking a subsequence, as $\varepsilon \to 0$, $\{\Phi_\sigma^\varepsilon\} \to \Phi_\sigma$, $\{\Phi_p^\varepsilon\} \to \Phi_p$, $\{\gamma_\varepsilon\} \to \gamma^*$ and $\{F_\varepsilon\} \to F^*$, where

$$[\Phi_\sigma(\gamma)](\pi, k, \eta, b) = \gamma(k)(\mu - (\pi + (1 - \eta)k)),$$

$$[\Phi_p(\gamma)](\pi, k, \eta, b) = (1 - \eta)\left(\int_0^k \gamma(z)dz\right) + \frac{\pi}{b - k}\int_k^b \gamma(z)dz,$$

for all $\gamma \in \cap_{\varepsilon>0}\Gamma_\varepsilon$ and for all $(\pi, k, \eta, b) \in \cap_{\varepsilon>0}\Theta_\varepsilon$ and $\gamma^* : [0, 1] \to [0, 1]$ is nondecreasing and continuous. Furthermore, since $(\gamma_\varepsilon, F_\varepsilon)$ is a fixed point of ϕ^ε for all $\varepsilon > 0$, and since $\Phi_\sigma(\gamma_\varepsilon)$ is ε-measurable with respect to $F_\varepsilon(\Phi_p^\varepsilon(\gamma_\varepsilon))$, for all $\varepsilon > 0$,

$$(1-\eta)\left(\int_0^k \gamma^*(z)dz\right) + \frac{\pi}{b - k}\int_k^b \gamma^*(z)dz = f^*\left(\gamma_\varepsilon(k)(\mu - (\pi + (1 - \eta)k))\right), \forall (\pi, k, \eta, b) \in \Theta^* := \bigcap_{\varepsilon>0} \Theta_\varepsilon, \tag{30}$$

[10]For instance, for $\bar{\gamma} \in \Gamma_\varepsilon$ such that $\bar{\gamma}(k) = 1$ for all $k \in [\varepsilon, 1]$ and $\bar{F} \in \mathcal{F}_\varepsilon$ such that $[\bar{F}(y)](\pi, k, \eta, b) = (\mu - (\pi + (1 - \eta))k)$ for all $y \in Y_p^\varepsilon$ and for all $(\pi, k, \eta, b) \in \Theta_\varepsilon$, $(\bar{\gamma}, \bar{F}) \in Z_\varepsilon$.

[11]Convexity follows from linearity of any $F \in \mathcal{F}$ and the operators Φ_σ^ε, Φ_p^ε and compactness follows from continuity of Φ_p, Φ_σ and F.

for some real-valued function f^*.[12]

Since γ^* is continuous and nondecreasing on $[0,1]$, by the Stone-Weierstrass theorem, there exists a sequence of differentiable and strictly increasing functions $\{\gamma_n\}$ such that $\{\gamma_n\} \to \gamma^*$ uniformly. For each $n \in \mathbb{N}$, let

$$c_n(z) := \int_0^z \gamma_n^{-1}(x)dx.$$

Then c_n is strictly convex and c_n'' exists. Also, let

$$\Sigma_n := \{(\sigma_n(G), \pi_n(G)) | G \in \mathcal{I}_\mu\},$$

where for any $G \in \mathcal{I}_\mu$ and for any $n \in \mathbb{N}$,

$$\sigma_n(G) := \int_0^1 \gamma_n(\psi(x|G))(1 - G(x))dx,$$

$$\pi_n(G) := \int_0^1 V_n(\psi(x|G))G(dx),$$

and

$$V_n(z) := \max_{q \in [0,1]} qz - c_n(q).$$

By Theorem 1, for any $(\sigma_n, p_n) \in \Sigma_n$,

$$\sigma_n = \gamma_n(k_n)(\mu - (\pi_n + (1-\eta_n)k_n)$$

and

$$p_n = (1-\eta_n)V_n(k_n) + \frac{\pi_n}{b_n - k_n}(V_n(b_n) - V_n(k_n)).$$

Moreover, by the envelope theorem,

$$V_n(z) = \int_0^z \gamma_n(x)dx, \; \forall z \in [0,1].$$

Since Σ_n is closed for all $n \in \mathbb{N}$, which in turn is implied by Lemma 7, and can be nested in a compact set in $\mathbb{R}_+^2$, by possibly taking a subsequence, $\{\Sigma_n\} \to \Sigma \subseteq \mathbb{R}_+^2$ under the Hausdorff distance. Furthermore, since $\{\gamma_n\} \to \gamma^*$ uniformly, $\{c_n\} \to c^*$ and $\{V_n\} \to V^*$ uniformly, for some nondecreasing and convex $c^* : [0,1] \to \mathbb{R}_+$ and some V^* such that

$$V^*(z) = \int_0^z \gamma^*(x)dx, \; \forall z \in [0,1].$$

As a result, for any $\{(\pi_n, k_n, \eta_n, b_n)\} \subset \Theta^*$, there exists $(\sigma, p) \in \Sigma$ and that

$$\sigma = \lim_{n \to \infty} \gamma_n(k_n)(\mu - (1-\eta_n)k_n) = \gamma^*(k)(\mu - (1-\eta)k)$$

[12]More specifically, for each $y \in [\Phi_\sigma(\gamma^*)](\Theta^*)$, take and fix any $\theta(y) \in [\Phi_\sigma(\gamma^*)]^{-1}(y)$ and let $f^*(y) := [F^*(\Phi_\sigma(\gamma^*))](\theta(y))$. Then $[F^*(\Phi_\sigma(\gamma^*))](\theta) = f^*([\Phi_\sigma(\gamma^*)](\theta))$ for all $\theta \in \Theta$ since $\Phi_\sigma(\gamma^*)$ is measurable with respect to $F^*(\Phi_\sigma * (\gamma^*))$, which follows from the fact that $\Phi_\sigma(\gamma_\varepsilon)$ is ε-measurable with respect to $F_\varepsilon(\Phi_p^\varepsilon(\gamma_\varepsilon))$.

and

$$p = \lim_{n\to\infty}(1-\eta_n)V_n(k) + \frac{\pi_n}{b_n - k_n}(V_n(b_n) - V_n(k_n)) = (1-\eta)\left(\int_\varepsilon^k \gamma(z)dz\right) + \frac{\pi}{b-k}\int_k^b \gamma^*(z)dz.$$

for some $(\pi, k, \eta, b) \in \Theta^*$. Furthermore, since $\gamma^*(\varepsilon) > 0$, for all $\varepsilon > 0$, there is a positive Lebesgue measure of such σ. Together, by (30), there exists an interval $[\underline{\sigma}, \bar{\sigma}]$ with $\bar{\sigma} > \underline{\sigma}$ such that

$$\{\sigma|(\sigma, p) \in \Sigma\} \cap [\underline{c}, \bar{c}] \neq \emptyset$$

and for any $(\sigma, p) \in \Sigma$ with $\sigma \in [\underline{c}, \bar{c}]$,

$$\sigma = f^*(p)$$

for some real-valued function f^*, as desired. ■

A.3.4 Proof of Proposition 3

Proof of Proposition 3. Take and fix and $c \in [0, 1]$. Consider any sequence $\{c_n\}$ of strictly increasing and strictly convex functions with $c_n(0) = 0$ such that $\{c_n(q)\} \to cq$ for all $q \in [0, 1]$. Since c_n is strictly increasing and strictly convex,

$$\gamma_n(z) := \operatorname*{argmax}_{q\in[0,1]} qz - c_n(q)$$

is a well-defined, strictly increasing function. Moreover, $\{\gamma_n\} \to \gamma$ pointwisely, where

$$\gamma(z) = \begin{cases} 0, & \text{if } z < c \\ 1, & \text{if } z > c \end{cases}$$

and $\gamma(c) \in [0, 1]$. For each $n \in \mathbb{N}$, let

$$\sigma_n(G) := \int_0^1 \gamma_n(\psi(x|G))(1 - G(x))dx,$$

$$\pi_n(G) := \int_0^1 V_n(\psi(x|G))G(dx)$$

for all $G \in \mathcal{I}_\mu$, where

$$V_n(z) := \max_{q\in[0,1]} qz - c_n(q), \forall z \in [0, 1].$$

Now let Σ_n be

$$\Sigma_n := \{(\sigma_n(G), \pi_n(G))|G \in \mathcal{I}_\mu\}.$$

Since Σ_n can be nested in a compact set in $\mathbb{R}^2_+$ and is closed (see Lemma 7), by possibly taking a subsequence, $\{\Sigma_n\} \to \Sigma \subseteq \Sigma(c)^L$ under the Hausdorff distance. It then suffices to show that $\Sigma = \Sigma^L(c)$. Indeed, consider any $(\sigma, \pi) \in \Sigma^L(c)$. By the result of Roesler and Szentes (2017), $\sigma \in [0, \mu - (1 - \eta)c]$ for some $\eta \in [0, 1 - \mu]$ such that $G^{\eta,1}_{\pi,c} \in \mathcal{I}^*_\mu$. Let

$$\tilde{\gamma} := \frac{\sigma}{\mu - \pi - (1 - \eta^*)c}.$$

Then $\tilde{\gamma} \in [0,1]$. For each $n \in \mathbb{N}$, let $k_n := \gamma_n^{-1}(\tilde{\gamma})$. Then it must be that $\{k_n\} \to c$. Now and take any sequence $\{\pi_n, \eta_n, b_n\}$ such that $\{\pi_n\} \to \pi$, $\{\eta_n\} \to \eta$ and $G^{\eta_n,b_n}_{\pi_n,k_n}$ (such sequences always exist as $G^{\eta,1}_{\pi,k} \in \mathcal{I}^*_\mu$)). Then, $(\sigma_n, p_n) \in \Sigma_n$, where

$$\sigma_n := \sigma_n(G^{\eta_n,b_n}_{\pi_n,k_n}) = \gamma_n(k_n)(\mu - (\pi_n + (1-\eta_n)k_n)$$

and

$$p_n := \pi_n(G^{\eta_n,b_n}_{\pi_n,k_n}) = (1-\eta_n)V_n(k_n) + \frac{\pi_n}{b_n - k_n}(V_n(b_n) - V_n(k_n)),$$

for all $n \in \mathbb{N}$. As a result, since by definition, $\gamma_n(k_n) = \tilde{\gamma}$ for all $n \in \mathbb{N}$, we have

$$\lim_{n\to\infty} \sigma_n = \tilde{\gamma}(\mu - (\pi + (1-\eta)c)$$

and

$$\lim_{n\to\infty} p_n = \pi.$$

Therefore, since $(\sigma_n, \pi_n) \in \Sigma_n$ for all $n \in \mathbb{N}$, it must be that $\Sigma^L(c) \subseteq \Sigma$. This completes the proof. ■