NON-DISCRIMINATORY PERSONALIZED PRICING

By

Philipp Strack and Kai Hao Yang

May 2026

COWLES FOUNDATION DISCUSSION PAPER NO. 2447R1

Non-Discriminatory Personalized Pricing*

Philipp Strack† Kai Hao Yang‡

May 28, 2026

Abstract

A monopolist offers personalized prices to consumers with unit demand. Consumers differ in their values, costs, and *protected characteristics*—such as race or gender. The seller is subject to a non-discrimination constraint: consumers with the same cost, but different protected characteristics must face identical price distributions. Such regulations are present in markets like credit or insurance. We characterize the optimal pricing rule. Under this rule, surplus accrues to both protected groups, but only to those with intermediate values. Strengthening the constraint to cover transaction prices redistributes surplus, harming the low-value group and benefiting the high-value group. Meanwhile, prohibiting the use of protected characteristics as pricing inputs instead of regulating outputs harms the low-value group.

Keywords: Price discrimination, personalized pricing, discrimination, market segmentation, protected characteristics, optimal transport

JEL classification: D42, D63, D82

---

*We thank the anonymous referees, Dirk Bergemann, Alessandro Bonatti, Ben Brooks, Roberto Corrao, Laura Doval, Yingni Guo, Nima Haghpanah, Elliot Lipnowski, Ellen Muir, Barry Nalebuff, Harry Pei, Alessandro Pavan, Jacopo Perego, Benjamin Polak, Phil Reny, Vasiliki Skreta, Alex Smolin, Rakesh Vohra, Jonas von Wangenheim, Frank Yang, Jidong Zhou, and various seminar audiences, for their valuable comments and suggestions. We also thank Tsong-Hong Tenn for his research assistance. All errors are our own.

†Department of Economics, Yale University, Email: philipp.strack@yale.edu

‡School of Management, Yale University, Email: kaihao.yang@yale.edu

1 Introduction

Motivation Advances in data collection have enabled firms to tailor prices to consumers based on a wide range of observable characteristics. In many markets, sellers now have access to rich datasets that allow for increasingly fine-grained segmentation, often approaching fully personalized pricing. At the same time, legal frameworks prohibit discrimination based on protected characteristics—such as gender, race, or age.[1] This raises questions about how anti-discrimination laws should apply in data-rich environments, where pricing algorithms operate on rich data.

In such settings, legal prohibitions on *disparate treatment*—that is, the explicit use of protected characteristics in decision-making—may have little effect: Because many non-protected variables can serve as proxies, firms may reproduce discriminatory outcomes even if protected characteristics are not used explicitly.[2] This motivates a different, distribution-based regulatory standard, often associated with disparate impact: rather than only restricting the inputs to pricing decisions, one may ask whether consumers with different protected characteristics face the same price distributions, perhaps conditional on characteristics that are "essential for business."[3]

In this paper, we study how a seller maximizes profit through personalized pricing while

[1]For instance, in the U.S., Title VII of the Civil Rights Act and the Equal Pay Act (EPA) protect workers from gender-based wage discrimination; the Fair Housing Act (FHA) prohibits housing discrimination based on protected characteristics such as race and gender; the Equal Credit Opportunity Act (ECOA) prevents lenders from offering different loan terms to borrowers based on protected characteristics; and recent legislation in California (AB1287) explicitly prohibits businesses from price discrimination based on gender.

[2]As noted in The White House (2015): "Big data naturally raises concerns among groups that have historically been victims of discrimination. Given hundreds of variables to choose from, it is easy to imagine that statistical models could be used to hide more explicit forms of discrimination by generating customer segments that are closely correlated with race, gender, ethnicity, or religion [...], even if the profit motive is different from, and in many cases fundamentally inconsistent with, the sort of prejudice that our anti-discrimination laws seek to prohibit."

[3]For example, according to The White House (2015): "It is often straightforward to conduct statistical tests for disparate impact by asking whether the prices generated by a particular algorithm are correlated with variables such as race, gender or ethnicity." Likewise, in credit markets, prior to the CFPB final rule in 2026, 12 CFR Regulation B stated: "The [ECO] Act and regulation may prohibit a creditor practice that is discriminatory in effect because it has a disproportionately negative impact on a prohibited basis, even though the creditor has no intent to discriminate and the practice appears neutral on its face." In the context of employment, Title VII of the Civil Rights Act holds employers accountable for "practice that causes a disparate impact on the basis of race, color, religion, sex, or national origin." In the context of housing, the FHA and its regulations (cf., 24 CFR) establish "liability [...] based on a practice's discriminatory effect, even if not motivated by discriminatory intent."

satisfying a distribution-based non-discrimination constraint motivated by disparate impact concerns. We show that the seller's problem reduces to a non-standard optimal transport, in which consumers across different protected groups are matched and face equal prices. We solve for the profit-maximizing non-discriminatory pricing rule, characterize how non-discrimination constraints shape the welfare consequences, and compare the welfare outcomes across different notions of non-discrimination regulations.

Specifically, we consider a model where a monopolist faces a unit mass of consumers with unit demand, different values, and different costs of being served. Consumers differ along a binary protected characteristic. Conditional on having the same cost, consumers in the "l" group have lower values and are more elastic, while consumers in the "h" group have higher values and are less elastic. The seller can charge consumers personalized prices, but the prices must satisfy a non-discrimination constraint: among consumers who are equally costly to serve, the price distributions faced by the two protected groups must be identical.

Results We first show that the seller's problem can be decomposed into two stages: a *matching stage*, where the seller matches consumers in different protected groups into pairs; and a *pricing stage*, where the seller chooses a common price for each matched pair. Intuitively, as the non-discrimination constraint prevents the seller from charging each consumer a different price, some consumers must be charged the same price. Because the protected characteristic is binary, it suffices to pair each consumer in one group with a consumer in the other and charge them the same price. Formally, this decomposition represents the non-discriminatory pricing problem as an optimal transport problem (Proposition 1).

Our main result, Theorem 1, solves this optimal transport problem and explicitly constructs a profit-maximizing non-discriminatory pricing rule. Under this pricing rule, consumers with intermediate values are matched assortatively and face a price equal to the lower value of the matched pair when trade occurs; while consumers with high values are matched with consumers from the other protected group who have low values, and face a price equal to the higher value of the matched pair when trade occurs. Technically, the transport problem is non-standard: its objective is neither supermodular nor submodular and combines opposing assortative and anti-assortative forces. The closed-form solution therefore provides a new tractable class of optimal transport problems.

We then use the characterization to study welfare. Optimal non-discriminatory pricing generally leaves positive surplus to consumers in both protected groups and creates deadweight loss by excluding low-value consumers in both groups, with larger losses for the lower-value group. Which group benefits more from the non-discrimination regulation depends both on value distributions and population shares of protected groups: as a group becomes larger, the seller can tailor prices more finely to that group, reducing its surplus. We also show that non-discriminatory personalized pricing can generate much stronger profit guarantees than uniform pricing, while its effect on consumer surplus is ambiguous: in some environments both groups prefer it to uniform pricing, and in others they do not.

Finally, we compare different non-discrimination regulations. Requiring non-discriminatory transaction outcomes, rather than non-discriminatory offered-price distributions, changes how the seller can exclude consumers and reduces surplus for the l-group while increases surplus for the h-group. Similarly, regulations that prohibit only disparate treatment—the use of protected characteristics as inputs—also reduce the surplus for the l-group. Such disparate treatment regulations are effective only when non-protected characteristics are sufficiently uncorrelated with protected characteristics. In this case, disparate treatment restrictions can meaningfully limit the seller's ability to tailor prices and lead to lower seller profit relative to disparate impact restrictions.

These results highlight fundamental distinctions between regulatory approaches and suggest that these differences do not merely stem from different philosophies, legal bases, or regulatory practices. Rather, different regulatory approaches would lead to fundamentally different welfare outcomes. It is thus imperative for regulators to understand the economic implications when designing non-discrimination policies.

Illustrative Example We illustrate the main economic forces with a simple example. There is a unit mass of consumers with unit demand for a single product. Consumers' values are 1, 2, or 3, each with equal probability in the population. There are two protected groups, men and women, each with a mass of $1/2$. Their value distributions are $(1/6, 1/3, 1/2)$ for men and $(1/2, 1/3, 1/6)$ for women. In particular, men are more likely to have high values than women. Production is costless, and the seller observes values and protected characteristics. Without regulation, the seller can charge each consumer their value, extract all surplus, and obtain profit $\frac{1+2+3}{3} = 2$. The outcome is efficient and consumer surplus is zero.

Now suppose prices must be non-discriminatory: the distribution of prices faced by men and women must be the same. Clearly, uniform pricing is feasible. The optimal uniform price is 2, yielding profit $4/3$. Consumers with value 1 are excluded, creating deadweight loss

W \ M	$(v=1)$ $\frac{1}{2}$	$(v=2)$ $\frac{1}{3}$	$(v=3)$ $\frac{1}{6}$
$(v=1)$ $\frac{1}{6}$	$\frac{1}{6}\,(p=1)$	0	0
$(v=2)$ $\frac{1}{3}$	$\frac{1}{3}\,(p=1)$	0	0
$(v=3)$ $\frac{1}{2}$	0	$\frac{1}{3}\,(p=2)$	$\frac{1}{6}\,(p=3)$

(a) Assortative

W \ M	$(v=1)$ $\frac{1}{2}$	$(v=2)$ $\frac{1}{3}$	$(v=3)$ $\frac{1}{6}$
$(v=1)$ $\frac{1}{6}$	0	0	$\frac{1}{6}\,(p=3)$
$(v=2)$ $\frac{1}{3}$	0	$\frac{1}{3}\,(p=2)$	0
$(v=3)$ $\frac{1}{2}$	$\frac{1}{2}\,(p=3)$	0	0

(b) Anti-Assortative

W \ M	$(v=1)$ $\frac{1}{2}$	$(v=2)$ $\frac{1}{3}$	$(v=3)$ $\frac{1}{6}$
$(v=1)$ $\frac{1}{6}$	$\frac{1}{6}\,(p=1)$	0	0
$(v=2)$ $\frac{1}{3}$	0	$\frac{1}{3}\,(p=2)$	0
$(v=3)$ $\frac{1}{2}$	$\frac{1}{3}\,(p=3)$	0	$\frac{1}{6}\,(p=3)$

(c) Optimal

Table 1: Non-Discriminatory Pricing by Matching

$1/3$, while consumers with value 3 retain surplus 1, so total consumer surplus is $1/3$.

The seller can do better while remaining non-discriminatory. Consider first the pricing rule described in Table 1a. This pricing rule is generated by first matching consumers from different groups (proportional to their group sizes) assortatively, and then charging each matched pair a common price. Since each matched pair contains the same mass of men and women, the induced price distribution is the same across groups. The resulting profit is $5/3 > 4/3$. The outcome is efficient, and male consumers retain positive surplus. Alternatively, using an anti-assortative matching scheme in Table 1b, and then charging each matched pair the same price, the resulting profit is also $5/3$. Under this pricing rule, all surplus is extracted from consumers who are served, but some consumers are excluded, creating a deadweight loss of $1/3$. Combining assortative matching and anti-assortative matching, Table 1c describes a profit-maximizing pricing rule, which yields a profit of $11/6 > 5/3$.

This example illustrates the key economic force in the paper. Under the non-discrimination constraint, the seller's problem is no longer simply to choose personalized prices, but to decide which consumers across groups are pooled together at a common price. Pooling consumers with similar values allows the seller to limit the rents left to the higher-value consumer when serving both. Pooling consumers with very different values allows the seller to limit the welfare losses when excluding some consumers. The optimal non-discriminatory pricing rule balances these two forces: it pools some consumers assortatively to economize on rents, and others anti-assortatively to economize on exclusion losses.

Related Literature The welfare effects of monopolistic price discrimination have been studied in a long-standing literature, which explores whether third degree price discrimination benefits consumers (see, e.g., Varian 1985; Aguirre, Cowan and Vickers 2010; Cowan 2016). Bergemann, Brooks and Morris (2015) show that any surplus division between the consumers and a monopolist can be achieved by some market segmentation.[4] In a model where the seller only observes protected characteristics, Cohen, Elmachtoub and Lei (2022) introduce multiple non-discrimination constraints and characterize the optimal prices. Similarly, Kallus and Zhou (2021) introduce several notions of non-discriminatory pricing, and characterize the optimal prices in a linear demand model. In contrast, this paper characterizes the optimal non-discriminatory pricing rules when the seller can engage in personalized pricing.

The personalized pricing model, where sellers are able to offer each consumer a price that depends on their values, has also been widely adopted in oligopoly models: Thisse and Vives (1988) show that in a Hotelling duopoly model consumer surplus can be higher under personalized pricing compared to uniform pricing. This framework is adopted by various papers that further investigate the effects of brand name (Shaffer and Zhang 2002), advertisement (Chen and Iyer 2002), or data sharing (Montes, Sand-Zantman and Valletti 2019). Rhodes and Zhou (2024) provide a comprehensive welfare analysis in a general oligopoly setting.

In the recent literature, Strack and Yang (2024) and He, Sandomirskiy and Tamuz (2024) study a notion of privacy, and characterize all signals that do not reveal certain information, which are referred to as privacy-preserving signals. A non-discriminatory pricing rule is mathematically equivalent to a privacy-preserving signal where the privacy sets are defined by the protected characteristics and pricing rules are interpreted as signals. Strack and Yang (2024) characterize the optimal privacy-preserving signals for *supermodular* decision problems. The pricing problem in this paper, in general, is not supermodular, and we develop new methods to solve this problem.[5]

Non-discriminatory pricing is also related to the notion of statistical parity in the algorithmic fairness literature (see, e.g., Darlington 1971; Calders and Verwer 2010; Hardt, Price and Srebro 2016).[6] These papers study the optimal fair algorithms for specific decision problems, typically with a binary state or a binary action.[7]

[4]See also: Haghpanah and Siegel (2022) and Haghpanah and Siegel (2023), who further consider segmentations in environments that feature nonlinear pricing; and Farboodi, Haghpanah and Shourideh (2025), who characterize when more information on consumers' characteristics leads to higher (lower) welfare.

[5]Section 5.3 in Strack and Yang (2024) illustrates the relation of privacy and non-discriminatory pricing through an example, but, importantly, does not provide a general characterization of optimal non-discriminatory pricing rules as it imposes an assumption on the support of the value distributions to make the profit function supermodular in that special case. Footnote 21 explains this distinction in further detail.

[6]Two other commonly adopted criteria are *separation* and *sufficiency*. It is well-known that no pair of these three common fairness criteria can be satisfied at the same time (see Barocas, Hardt and Narayanan (2019) and Carey and Wu (2023) for a comprehensive review of these criteria).

[7]In economics, Liang, Lu, Mu and Okumura (2024) and Doval and Smolin (2024) further characterize the

The rest of the paper is organized as follows: Section 2 introduces the model, Section 3 solves for the profit-maximizing non-discriminatory pricing rules, Section 4 discusses welfare implications and comparative statics. Section 6 presents extensions. Section 7 concludes.

2 Model

A monopolist sells a good or service to a continuum of consumers, each of whom demands one unit.[8] We normalize the total mass of consumers to one.

Consumers Each consumer is described by their value for the good $v \in V \subseteq \mathbb{R}_+$, the cost $c \in C \subseteq \mathbb{R}_+$ of being served, a *protected characteristic* $\theta \in \Theta := \{l, h\}$ (e.g., race or gender), and an auxiliary index $r \in [0,1]$ (for the purpose of randomization).[9] We denote by $\omega = (v, c, \theta, r) \in \Omega := V \times C \times \Theta \times [0,1]$ a consumer's type.

Let $\mathbb{P}$ be the product of a probability distribution on $V \times C \times \Theta$ and the Lebesgue measure on $[0,1]$. We denote by $G(\cdot) = \mathbb{P}[c \leq \cdot]$ the distribution of cost, by $\alpha_c = \mathbb{P}[\theta = h \mid c]$ the fraction of consumers with characteristic h conditional on having cost c,[10] and by $F_{c,\theta}(\cdot) := \mathbb{P}[v \leq \cdot \mid c, \theta]$ the distribution of values v of consumers of characteristic θ and cost c. We assume that $F_{c,\theta}$ admits a density $f_{c,\theta}$, has full support on an interval $(\underline{v}_c, \overline{v}_c)$ for some $0 \leq \underline{v}_c < \overline{v}_c \leq \infty$,[11] has finite mean $\mathbb{E}[v \mid c, \theta] < \infty$, and h-consumers have higher values for the product in the likelihood ratio order: $f_{c,h}(v)/f_{c,l}(v)$ is strictly increasing in v on $(\underline{v}_c, \overline{v}_c)$ for all c.

The likelihood-ratio assumption implies that, conditional on having the same cost, consumers with protected characteristic l have lower values (in first-order stochastic dominance) and react more strongly to price changes (i.e., are more elastic). One natural case captured by the above assumption is that of a normal good when h-consumers are richer. Alternatively, h-consumers could be the group with worse outside options. In particular, depending on the context, h-consumers could be either the advantaged group (e.g., rich consumers) or

entire Pareto frontier in terms of the payoffs of each protected group in a general setting.

[8]The methods we develop can also be applied to settings with multi-unit demand and vertically differentiated products. We focus on the case of unit demand mostly to simplify the exposition.

[9]All our results remain unchanged without r, as the optimal pricing rule we obtain turns out to be non-random.

[10]When there is no risk of confusion, we slightly abuse the notation and use v, c, θ, r to denote the random variable as well as a realization.

[11]In particular, $F_{c,l}$ and $F_{c,h}$ have common supports. This assumption is for the ease of exposition, and the result can be readily extended to distributions with different (interval) supports.

the disadvantaged group (e.g., those who have worse outside options).

Pricing Rules A *pricing rule* $p : \Omega \to \mathbb{R}_+ \cup \{+\infty\}$ is a random variable, where $p(\omega)$ is the price faced by consumers with type $\omega \in \Omega$, with $p(\omega) = +\infty$ representing exclusion (no transaction). In particular, a pricing rule p allows prices to be personalized, as different consumers could face different prices. The seller's profit under pricing rule p equals

$$\Pi(p) \coloneqq \mathbb{E}\left[(p(\omega) - c)\mathbf{1}\{v \geq p(\omega)\}\right].$$

For a pricing rule to be non-discriminatory, the distribution of prices consumers face can depend on the cost of serving them, but not on their protected characteristic (even if it correlates with their values).

Definition 1. A pricing rule p is *non-discriminatory*, denoted by $p \in \mathcal{D}$, if for all $c \in C$ and $M \subseteq \mathbb{R}_+$,

$$\mathbb{P}\left[p \in M \mid c, \theta = l\right] = \mathbb{P}\left[p \in M \mid c, \theta = h\right].$$

Equivalently, p is non-discriminatory if p is independent of θ conditional on c.

Non-discriminatory pricing rules exist, since uniform pricing is always non-discriminatory. As an example, U.S. fair lending laws require that in a loan market, black and white consumers with the same expected cost of default must be offered the same interest rates. This regulation is enforced: The Consumer Financial Protection Bureau (CFPB) launched 32 fair lending probes in 2022. For instance, the CFPB investigated Wells Fargo for "statistically significant disparities" in the rates at which the bank offered pricing exemptions (which correspond to $0.25\% - 0.75\%$ interest reductions relative to the rate calculated based on credit risk) to female and black loan applicants (CNBC 2023).

Remark 1 (The Role of Cost in Determining Non-Discriminatory Prices)**.** Non-discriminatory pricing requires that consumers with different protected characteristics face identical price distributions *conditional* on the cost of serving them. In disparate impact regulations, differences in outcomes are often permitted when they can be explained by characteristics that are materially relevant or justified by "business necessity." For example, in credit markets, firms are typically allowed to condition on the expected default risk even if it correlates with protected characteristics (e.g., *Anderson v. Wachovia Mortgage Corp.*). In insurance markets, firms are often allowed to condition premiums on the expected cost of treatments even though these costs correlate with age (which is typically a protected characteristic). Motivated by this practice, we use c to capture materially relevant characteristics that the seller is permitted to condition on. Thus, c serves a dual role in our model: it is the cost of serving a consumer and, at the same time, the characteristic on which price variation is permitted to depend.[12]

There are, however, markets in which costs are not considered as legitimate reasons to justify discriminatory pricing. A notable example is the EU/UK market for health insurance where equal premia for men and women are mandated since 2012. In Section 5.3, we discuss a model in which costs are not treated as materially relevant and which can thus be used to analyze such markets.

Remark 2 (Pricing Rules and Market Segmentations)**.** A pricing rule is closely related to a *market segmentation*, in the sense of Bergemann et al. (2015). A market segmentation $s : \Omega \to S$ is a random variable that maps consumers' types into some measurable space S. Each realization $s(\omega)$ corresponds to a *market segment*, so that $s(\omega) = s(\omega')$ means consumers with type ω and ω' belong to the same segment. In this regard, a pricing rule p itself is a market segmentation, where consumers who face the same price belong to the same segment. The converse is also true: given any market segmentation s, any pricing rule p that is measurable with respect to s can be interpreted as a rule that charges all consumers in the same segment the same price.

Consumer Surplus and Welfare Loss Finally, we denote by

$$CS(c, \theta; p) := \mathbb{E}\left[(v - p)^+ \mid c, \theta\right]$$

the average consumer surplus, and by

$$WL(c, \theta; p) := \mathbb{E}\left[\mathbf{1}\{p > v\}(v - c)^+ \mid c, \theta\right]$$

the welfare loss, of a θ-consumer with cost c under pricing rule p.

[12] The variable c is kept explicit even though the analysis is largely pointwise in c, because removing it would eliminate the distinction between prohibited and permitted sources of price variation that is central to these applications. Readers who prefer may interpret the results below as solving the pricing problem for a fixed c and then aggregating across cost classes. The same logic extends to settings with additional materially relevant characteristics.

3 Optimal Non-Discriminatory Pricing

We now maximize the seller's profit over non-discriminatory pricing rules. That is, we solve

$$\Pi^\star \coloneqq \sup_{p \in \mathcal{D}} \Pi(p)\,. \tag{1}$$

A non-discriminatory pricing rule p is *undominated* if there does not exist another non-discriminatory pricing rule such that both h-consumers and l-consumers have a higher average surplus, and the seller has a higher profit, with at least one of them being strictly higher. We focus on the undominated pricing rules among all profit-maximizing pricing rules.[13]

3.1 Optimal Pricing as Matching

We begin by showing that (1) can be decomposed into a two-stage problem: (i) the *matching stage*, where the seller determines how l- and h-consumers are paired across groups; and (ii) the *pricing stage*, where the seller chooses an optimal common price for each pair. Intuitively, since the non-discrimination constraint prevents the seller from tailoring prices independently to individual consumers, the seller must offer the same price to segments of consumers across the two protected classes. Since there are two protected groups, it is optimal that each of these segments consists only of a *pair* of values. For each $c \in C$, let $\mathcal{R}_c \subset \Delta(V^2)$ denote the set of all probability measures on V^2 with marginals $F_{c,l}$ and $F_{c,h}$. An element $\rho_c \in \mathcal{R}_c$ can be interpreted as a (possibly stochastic) matching scheme that, conditional on cost c, pairs l- and h-consumers by their values, proportional to their population sizes.

We first show that every non-discriminatory pricing rule $p \in \mathcal{D}$ induces a family of matching schemes $(\rho_c)_{c \in C}$. Fix $\tilde{c} \in C$. The pricing rule p segments the market, where all consumers in the same segment face the same price. Since there may be many l-consumers and many h-consumers in a given segment, a segment does not itself determine a pairing. A simple way to obtain one is therefore to pair l- and h-consumers randomly within each segment. Formally, for each $\tilde{c} \in C$, define the random matching $\rho_{\tilde{c}}$ by

$$\rho_{\tilde{c}}(V_l \times V_h) \coloneqq \mathbb{E}\left[\mathbb{P}\left[v \in V_l \mid p, c = \tilde{c}, \theta = l\right] \times \mathbb{P}\left[v \in V_h \mid p, c = \tilde{c}, \theta = h\right] \mid c = \tilde{c}\right], \tag{2}$$

where the expectation is taken over p. Since p is non-discriminatory, the distribution of

[13] We will further characterize the welfare outcomes of all profit-maximizing pricing rules later in Section 6.4.

prices conditional on $\tilde{c}$ is the same across groups. Therefore, averaging across price segments preserves the original conditional value distributions:

$$\mathbb{E}\left[\mathbb{P}[v \leq x \mid p, \theta = \tilde{\theta}, c = \tilde{c}] \mid c = \tilde{c}\right] = \mathbb{P}[v \leq x \mid \theta = \tilde{\theta}, c = \tilde{c}] = F_{\tilde{c},\tilde{\theta}}(x)$$

and hence $\rho_{\tilde{c}} \in \mathcal{R}_{\tilde{c}}$.

Conversely, every matching scheme $(\rho_c)_{c \in C}$ generates a feasible non-discriminatory pricing rule. Indeed, if each matched pair faces a common price, then the induced price distribution must be the same for l- and h-consumers conditional on cost c.[14] In fact, once a pair (v_l, v_h) is formed, the pricing stage problem is simple. The seller either does not sell, or charges a common price in $\{v_l, v_h\}$, leading to a profit of

$$\begin{aligned} \pi_c(v_l, v_h) &:= \max_{x \geq 0}(x - c)\left[(1 - \alpha_c)\mathbf{1}\{v_l \geq x\} + \alpha_c \mathbf{1}\{v_h \geq x\}\right] \\ &= \max\left\{\min\{v_l, v_h\} - c,\ \alpha_c(v_h - c)^+,\ (1 - \alpha_c)(v_l - c)^+\right\}. \end{aligned} \tag{3}$$

Thus, within a matched pair, the seller either sells to both consumers at the lower value, serves only one consumer and excludes the other, or excludes both. Since the seller can re-optimize independently within each pair, this weakly raises profit relative to the original rule p. We therefore obtain the following representation.[15]

Proposition 1 (Optimal Transport Representation)**.** *Let $\pi^\star$ be the value of the optimal transport problem*

$$\pi^\star := \int_C \left(\max_{\rho_c \in \mathcal{R}_c} \int_{V^2} \pi_c(v_l, v_h)\, \mathrm{d}\rho_c\right) G(\mathrm{d}c). \tag{4}$$

Then $\pi^\star = \Pi^\star$. Moreover, any solution of (4) induces a solution of (1), by pricing optimally within each pair, while any solution of (1) induces a solution of (4) via (2).

Proposition 1 leads to a simple interpretation of the seller's problem. For any cost c, consider the loss function $L_c : V^2 \to \mathbb{R}_+$ defined as subtracting the profit $\pi_c(v_l, v_h)$ from the total gains from trade $\alpha_c(v_h - c)^+ + (1 - \alpha_c)(v_l - c)^+$. That is,

$$L_c(v_l, v_h) := \min\{\alpha_c v_h + (1 - \alpha_c)v_l - \min\{v_l, v_h\}, \alpha_c(v_h - c)^+, (1 - \alpha_c)(v_l - c)^+\}. \tag{5}$$

[14] Formally, this uses the disintegration theorem; see the proof of Proposition 1 in Appendix A.1.

[15] Alternatively, Proposition 1 can be obtained from Proposition 2 of Strack and Yang (2024), which characterizes the optimal privacy-preserving signals for arbitrary decision problems.

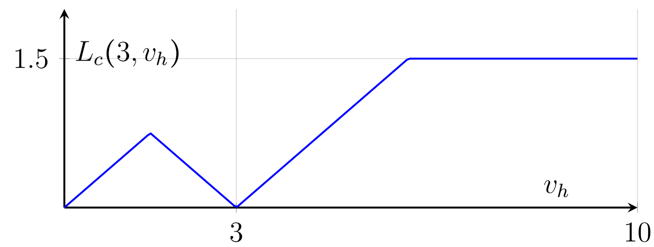


Figure 1: The cost function $L_c(v_l, v_h)$ for $c = 0$, $\alpha = 1/2$, and $v_l = 3$.

The loss function L_c highlights the key economic trade-off the seller faces at the matching stage. The first term of the minimum describes the rents left to consumers when both types are served, which decreases as the two matched types become more similar. Meanwhile, the second and the third terms of the minimum describe welfare loss when only one type is served, which decreases as the two matched types become further apart. Thus, at the matching stage, the seller effectively chooses a matching scheme to minimize total consumer rents and total welfare losses. As we will show below, the solution uses assortative matching to minimize consumer rents within the set of types where both types are served, and uses anti-assortative matching to minimize welfare losses within the set of types where only one type is served.

Relation to Classical Optimal Transport Problems The loss function L_c places (4) outside standard tractable optimal transport problems. For $c = 0$ and $\alpha_c = 1/2$,

$$L_c(v_l, v_h) = \frac{1}{2} \min\{|v_h - v_l|, v_h, v_l\}\,.$$

Classical transport objectives typically reward similarity through a cost $d(|v_h - v_l|)$, often making assortative matching optimal. Here, similarity is valuable only when both consumers are served, because it reduces rents. When one consumer is excluded, however, the seller instead benefits from matching low values with high values, which reduces welfare losses.

Thus, the objective L_c creates both assortative and anti-assortative forces. As demonstrated by Figure 1, it is neither supermodular nor submodular, is not translation invariant, and is non-monotone in $|v_h - v_l|$. These features place (4) outside the standard classes for which closed-form solutions are immediately available, and motivate the explicit duality construction below.

3.2 Profit-Maximizing Pricing Rules

We next derive an optimal pricing rule. For the ease of exposition, in the remainder of Section 3 and Section 4, we impose the following assumption on $F_{c,l}$ and $F_{c,h}$, which allows us to focus on the more economically interesting cases, where the seller cannot extract all gains from trade and faces a meaningful trade-off between surplus extraction and efficiency. The characterization of optimal pricing rules for general distributions is deferred to Section 6.3.

Assumption 1. *$\mathbb{P}[v \le c \mid c, \theta = l] < \|F_{c,l} - F_{c,h}\|$ for almost all $c \in C$.*

Here, $\|\cdot\|$ denotes the total variation distance.[16] Assumption 1 thus requires that, conditional on each cost, the distance between the value distribution of h-consumers and that of l-consumers is always greater than the share of l-consumers whose costs exceed their value. As we establish in Corollary 2, Assumption 1 is equivalent to no-full-surplus-extraction for all c. Intuitively, when $\|F_{c,l} - F_{c,h}\| = 0$, the value distributions of h and l consumers are the same, and thus the non-discrimination constraint does not have an effect. When c is large, there are many consumers who are inefficient to trade with, and thus the seller can fully extract from the rest of consumers by pairing them with those with values below c. In particular, if $c = 0$ almost surely, Assumption 1 is trivially satisfied as long as $\|F_{c,l} - F_{c,h}\| > 0$, since the only pricing rule that achieves full surplus extraction is to charge each consumer their value $p(v, c, \theta, r) = v$, which is discriminatory as $F_{c,l} \neq F_{c,h}$.

Before describing the optimal pricing rule, we introduce two simple exemplary non-discriminatory pricing rules motivated by the matching representation.

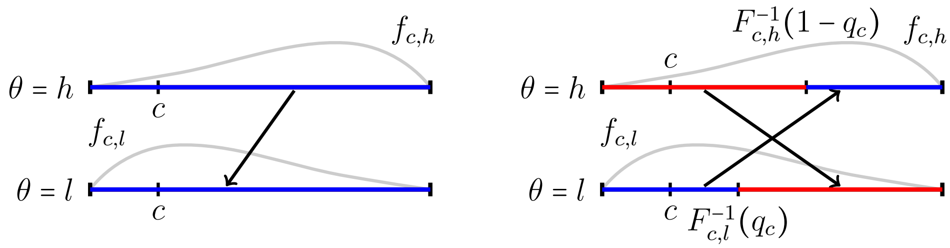
Definition 2 (Assortative Matching)**.** The assortative pricing rule p^{ass} is defined by

$$p^{ass}(v, c, \theta, r) := \begin{cases} v & \text{if } \theta = l \\ (F_{c,l}^{-1} \circ F_{c,h})(v) & \text{if } \theta = h \end{cases} \tag{6}$$

whenever the right-hand side is at least c, and $p^{ass}(v, c, \theta, r) := +\infty$ otherwise.

The pricing rule p^{ass} assortatively matches consumers conditional on c and charges each matched pair the lower of their values when that value is at least c; otherwise neither consumer transacts. Under p^{ass}, the seller's profit equals $\mathbb{E}[(v - c)^+ \mid \theta = l]$.

[16] Formally, $\|G - H\| = \sup_{A \subseteq \mathbb{R}_+} \left| \int_A \mathrm{d}G - \int_A \mathrm{d}H \right|$ for all CDFs G,H.



(a) Assortative. (b) Partly Anti-Assortative.

Figure 2: Assortative and Partly Anti-Assortative Pricing Rules

Regions of values matched are of the same color. The arrows illustrate a generic pair of values that are matched assortatively together, and the direction indicates the price each matched pair faces, conditional on being above c.

Definition 3 (Partly Anti-Assortative Matching)**.** A partly anti-assortative pricing rule $p^{anti} : \Omega \to V$ is defined by

$$p^{anti}(v, c, \theta, r) := \begin{cases} \max\{F^{-1}_{c,h}(F_{c,l}(v) + 1 - q_c),\, c\}, & \text{if } \theta = l \text{ and } v \le F^{-1}_{c,l}(q_c), \\ \max\{F^{-1}_{c,l}(F_{c,h}(v) + q_c),\, c\}, & \text{if } \theta = h \text{ and } v \le F^{-1}_{c,h}(1 - q_c), \\ \max\{v,\, c\}, & \text{otherwise} \end{cases}, \tag{7}$$

for some quantiles $(q_c)_{c\in C}$ with $\|F_{c,l} - F_{c,h}\| \le q_c \le 1$ for all $c \in C$.

Under p^{anti}, high-value consumers in one group are matched with low-value consumers in the other group, and the seller always excludes the low value in each pair.[17] Thus, consumers who are served always have their surplus extracted. Which consumers are served, on the other hand, depends on the choice of q_c. For example, if $q_c = 1$, then all l-consumers are excluded and the seller's profit equals $\mathbb{E}[\alpha_c(v - c)^+ \mid \theta = h]$. As q_c decreases, more high-value l-consumers are served and more low-value h-consumers are excluded. Denote by $(q^\star_c)_{c\in C}$ the quantiles that maximize the seller's profit among all partly anti-assortative pricing rules.[18]

The pricing rules p^{ass} and p^{anti} are separately tailored to the two competing forces identified by (5). Indeed, if the seller (suboptimally) always serves both types at the pricing stage, the assortative matching is optimal at the matching stage. In contrast, if the seller

[17]The constraint $q_c \ge \|F_{c,l} - F_{c,h}\|$ ensures that l-consumers with $v \ge F^{-1}_{c,l}(q_c)$ have higher values than h-consumers with $v < F^{-1}_{c,h}(1 - q_c)$ when matched.

[18]By standard Kuhn-Tucker arguments, $q^\star_c$ equals the maximum of the solution of $\alpha_c(F^{-1}_{c,h}(1 - q_c) - c)^+ = (1 - \alpha_c)(F^{-1}_{c,l}(q_c) - c)^+$ and $\|F_{c,l} - F_{c,h}\|$.

(suboptimally) always excludes the low type of each pair at the pricing stage, a partly anti-assortative matching is optimal at the matching stage. In general, however, since the seller can choose, pair by pair, whether to serve both consumers or exclude the lower-value consumer, the optimal pricing rule thus combines the features of p^{ass} and p^{anti}, as characterized below.

An Optimal Pricing Rule

We now describe an optimal pricing rule. For all $c \in C$, let $\Delta_c(v) := F_{c,l}(v) - F_{c,h}(v)$ and let $\overline{\Delta}_c^{-1}, \underline{\Delta}_c^{-1} : [0,1] \to V$ be the larger and smaller inverses of Δ_c, and let $v_c^\star$ be the unique solution to $f_{c,l}(v_c^\star) = f_{c,h}(v_c^\star)$.[19] The following lemma identifies some critical cutoffs.

Lemma 1. *For all $c \in C$, there exists a unique increasing vector $\kappa_c \in \mathbb{R}^5$ with $\kappa_c^4 < v_c^\star < \kappa_c^5$ such that*

$$\begin{aligned} \kappa_c^2 &= F_{c,l}^{-1}(\Delta_c(\kappa_c^3) + F_{c,h}(\kappa_c^1)) = F_{c,l}^{-1}(\Delta_c(\kappa_c^4)) = F_{c,l}^{-1}(\Delta_c(\kappa_c^5)) \\ \kappa_c^1 - c &= (1-\alpha_c) \cdot (\kappa_c^3 - c) = \alpha_c \cdot (\kappa_c^5 - \kappa_c^4). \end{aligned} \tag{8}$$

Henceforth, we will call κ_c the unique solution to (8) and define the pricing rule $p^\star$ as:

$$\begin{aligned} p^\star(v,c,l,r) &:= \begin{cases} \overline{\Delta}_c^{-1}(\Delta_c(\kappa_c^5) - F_{c,l}(v)), & \text{if } v < \kappa_c^2 \\ F_{c,h}^{-1}(F_{c,l}(v) - F_{c,l}(\kappa_c^2) + F_{c,h}(\kappa_c^1)), & \text{if } v \in [\kappa_c^2, \kappa_c^3) \\ v, & \text{if } v \geq \kappa_c^3 \end{cases} ; \\ p^\star(v,c,h,r) &:= \begin{cases} \underline{\Delta}_c^{-1}\left(F_{c,h}(v) + \Delta_c(\kappa_c^3)\right), & \text{if } v < \kappa_c^1 \\ F_{c,l}^{-1}(F_{c,h}(v) + \Delta_c(\kappa_c^4)), & \text{if } v \in [\kappa_c^4, \kappa_c^5) \\ v, & \text{if } v \in [\kappa_c^5, \infty) \cup (\kappa_c^1, \kappa_c^4) \end{cases} . \end{aligned} \tag{9}$$

Theorem 1 (Optimal Pricing)**.**

(i) $p^\star$ is a profit-maximizing non-discriminatory pricing rule. That is, $p^\star$ solves (1).

(ii) Every undominated profit-maximizing non-discriminatory pricing rule p induces the same average surplus for each protected characteristic conditional on cost. That is, $CS(c,\theta;p) = CS(c,\theta;p^\star)$ for all $c \in C$ and $\theta \in \{l,h\}$.

[19] Formally, since $F_{c,h}$ dominates $F_{c,l}$ in the likelihood ratio order, Δ_c is quasi-concave and is uniquely maximized at $v_c^\star$. Thus, for any $q \in [0, \Delta_c(v_c^\star)]$, there exists a unique pair $(\underline{\Delta}_c^{-1}(q), \overline{\Delta}_c^{-1}(q)) \in V^2$ such that $\underline{\Delta}_c^{-1}(q) \leq v_c^\star \leq \overline{\Delta}_c^{-1}(q)$ and $\Delta_c(\underline{\Delta}_c^{-1}(q)) = q = \Delta_c(\overline{\Delta}_c^{-1}(q))$.

Figure 3 plots the optimal pricing rule $p^\star$. Under $p^\star$, high-value consumers (i.e., $v \geq \kappa_c^3$ for $\theta = l$ and $v \geq \kappa_c^5$ for $\theta = h$) face a price equal to their values and thus are fully extracted; intermediate-value (i.e., $v \in (\kappa_c^2, \kappa_c^3)$ for $\theta = l$ and $v \in (\kappa_c^4, \kappa_c^5)$ for $\theta = h$) consumers face a price less than their values, thereby retaining rents; while low-value consumers (i.e., $v \leq \kappa_c^2$ for $\theta = l$ and $v \leq \kappa_c^1$ for $\theta = h$) face a price that exceeds their values and are excluded. Using (8), it can be verified that for all $c \in C$, the distributions of $p^\star$ conditional on (c, l) and on (c, h) are the same, and thus $p^\star$ is indeed non-discriminatory. Notably, the pricing rule $p^\star$ is non-monotone in consumers' values given c and θ; and does not depend on the randomization device r.

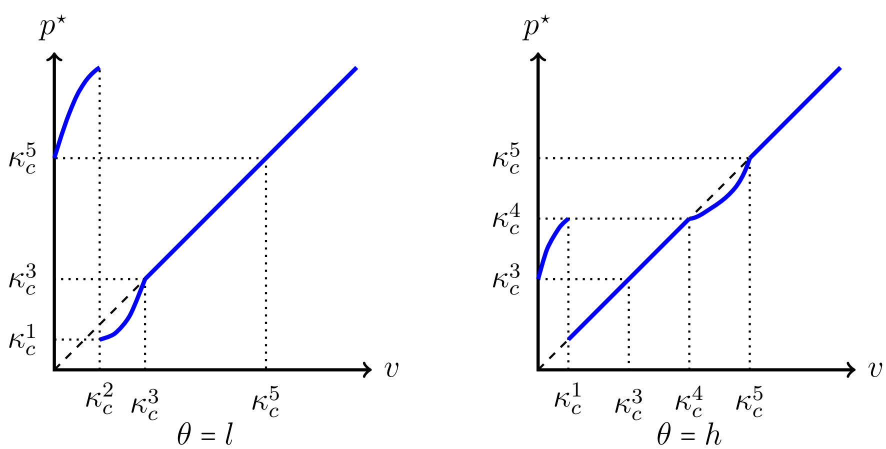


Figure 3: Optimal Pricing Rule $p^\star$.

From Proposition 1, the pricing rule $p^\star$ can alternatively be described by a family of matching schemes $(\rho_c^\star)_{c \in C}$ that solves (4). Figure 4 plots the matching scheme $\rho_c^\star$ for a given c. In Figure 4, the top interval depicts values of h-consumers, and the bottom interval depicts values of l-consumers. Subintervals with the same colors on each side are matched together: subintervals connected by solid arrows are matched positively assortatively, whereas subintervals connected by dashed arrows are matched by pairing consumers with the same values. The direction of the arrow indicates which value in a matched pair equals the price under $p^\star$. According to Figure 4, $\rho_c^\star$ matches h-consumers who have values $v \leq \kappa_c^1$ with l-consumers who have values $v \in (\kappa_c^3, \kappa_c^4]$. The seller's optimal price, by (8), for each of these matched pairs, equals the high value of the pair. Meanwhile, l-consumers with $v \leq \kappa_c^2$ are matched with an equal mass of h-consumers with $v > \kappa_c^5$, and the seller's optimal price for each of these matched pairs, by (8), equals the high value of the pair; l-consumers with

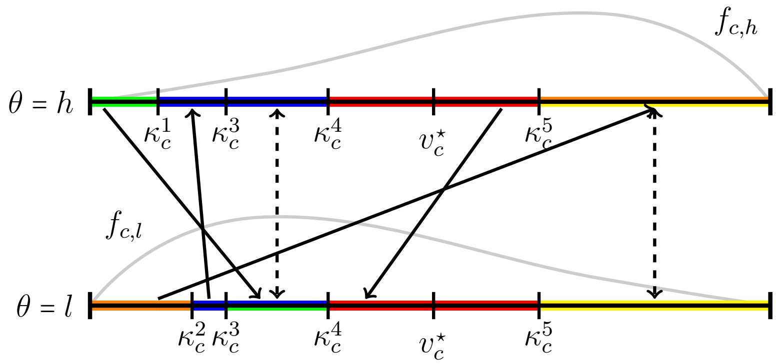


Figure 4: Matching Scheme $\rho_c^\star$.

$v \in (\kappa_c^2, \kappa_c^3]$ are matched assortatively with h-consumers with $v \in (\kappa_c^1, \kappa_c^3]$, and the seller's optimal price for each of these matched pairs, by (8), equals the low value of the pair; consumers with $v \in (\kappa_c^4, \kappa_c^5]$ are matched assortatively, and the seller's optimal price for each of these matched pairs, by (8), equals the low value of the pair. Lastly, all the remaining consumers are matched with those with the same values, and the seller's optimal price equals their values. By (8), each of these matching regions have equal mass of consumer values and thus the matching scheme $\rho_c^\star$ is well-defined.

In essence, $p^\star$ combines the features of p^{ass} and p^{anti}. In particular, it matches intermediate-value consumers assortatively and serves both types to minimize consumer rents, as the amount of rents left to consumers would be the smallest when the values paired together are similar. Meanwhile, it matches high- and low-value consumers (partly) anti-assortatively and excludes the low type to minimize welfare losses: the welfare loss is smaller when only low-value consumers are excluded. An immediate consequence of this matching pattern is that, in terms of consumer surplus, only consumers with intermediate values will retain rents, while low-value consumers are excluded and high-value consumers are fully extracted.

Remark 3. There are other pricing rules beyond $p^\star$ that maximize the seller's profit. For example, in the interval $[\kappa_c^4, \kappa_c^5]$ where the seller matches l and h consumers, any other matching that ensures every l-consumer is matched with an h-consumer of higher value yields the same expected profit of $\mathbb{E}[v \mid c, \theta = l, v \in [\kappa_c^4, \kappa_c^5]] - c$. As a result, for each individual consumer, their surplus might be different under different undominated profit-maximizing pricing rules. Nonetheless, Theorem 1 ensures that all undominated optimal pricing rules lead to the same *average* consumer surplus for each protected characteristic θ and cost c. In fact, the proof of Theorem 1 establishes a stronger sense of uniqueness: the average surplus *conditional* on each interval in the partition formed by $\kappa_c^1, \kappa_c^2, \kappa_c^3, \kappa_c^4, \kappa_c^5$ is the same under all undominated optimal pricing rules.

Remark 4. The characterization of the optimal matching $\rho_c^\star$ is also novel to the optimal transport literature. Closed-form solutions to concrete transport problems are rare and typically depend on supermodular or submodular objectives. The objective in (4) is neither supermodular nor submodular, and does not satisfy any of the commonly made assumptions like translation invariance or single quasi-convexity. The closed-form solution obtained in Theorem 1 thus contributes to the optimal transport literature by establishing a new class of tractable problems.[20]

There are some special cases where the optimal pricing rule $p^\star$ is simplified. First, we note that from Lemma 1 and Theorem 1, a simple support condition is necessary and sufficient for the assortative pricing rule to be optimal.

Corollary 1. *The assortative pricing rule p^{ass} is optimal if and only if $\alpha_c \cdot (\overline{v}_c - c) \le \underline{v}_c - c$.*

The sufficiency part of Corollary 1 follows immediately. Whenever $\alpha_c(\overline{v}_c - c) \le \underline{v}_c - c$, it is always optimal for the seller to serve both types in a matched pair, regardless of which pairs are matched. Thus, $\pi_c(v_l, v_h) = \min\{v_l, v_h\} - c$ is supermodular, and hence assortative matching is optimal. The proof of the necessity part relies on the proof of Theorem 1, and can be found in Appendix A.1.[21]

For specific distributions $F_{c,l}$ and $F_{c,h}$, the solutions to (8) have closed-form expressions, and therefore the optimal pricing rule can be written in closed form, as illustrated by the following example.

Example 1 (Beta Distributions)**.** Suppose that $c = 0$ almost surely, $\alpha_c = 1/2$, $V = [0,1]$, and $F_{c,l}(v) = v$, $F_{c,h}(v) = v^2$, $v \in [0,1]$.[22] The unique solution of (8) has a closed-form expression,

[20]To our knowledge, the only other papers that establish explicit properties of the solution for a concrete optimal transport problem without imposing these assumptions are (i) Boerma, Tsyvinski and Zimin (2025), who study the function $|v_l - v_h|^\beta$ for $\beta \in (0,1)$ and thus relax supermodularity/submodularity while keeping quasi-convexity and translation invariance, and (ii) Moscariello (2025), who considers S-shaped cost functions maintaining translation invariance.

[21]It is noteworthy that the sufficiency direction of Corollary 1 reproduces Proposition 10 of Strack and Yang (2024) when θ is binary. In this regard, Proposition 10 of Strack and Yang (2024), when restricted to binary θ, is a special case of Theorem 1. In fact, as we show in Section 4 below, the welfare implications change qualitatively once we relax this support condition.

[22]The likelihood-ratio-order assumption holds, as $\frac{f_{c,h}(v)}{f_{c,l}(v)} = \frac{2v}{1} = 2v$ is increasing in v. Assumption 1 also holds as $\|F_{c,l} - F_{c,h}\| = 1/4 > 0 = F_{c,l}(c)$. Moreover, $\Delta_c(v) := F_{c,l}(v) - F_{c,h}(v) = v - v^2$ is maximized at $v_c^\star = 1/2$.

given by

$$\kappa_c^1 = \frac{1}{2}\left(1 - \frac{1}{\sqrt{2}}\right), \qquad \kappa_c^2 = \frac{2\sqrt{2}-1}{8}, \qquad \kappa_c^3 = 1 - \frac{1}{\sqrt{2}}, \qquad \kappa_c^4 = \frac{1}{2\sqrt{2}}, \qquad \kappa_c^5 = 1 - \frac{1}{2\sqrt{2}}.$$

The optimal pricing rule in (9) can then be written in closed form:

$$p^\star(v,c,l,r) = \begin{cases} \frac{1+\sqrt{1-4(\kappa_c^2 - v)}}{2}, & \text{if } v < \kappa_c^2, \\ \sqrt{v - \kappa_c^2 + (\kappa_c^1)^2}, & \text{if } v \in [\kappa_c^2, \kappa_c^3), \\ v, & \text{if } v \geq \kappa_c^3. \end{cases}$$

$$p^\star(v,c,h,r) = \begin{cases} \frac{1-\sqrt{1-4(v^2 + \Delta_c(\kappa_c^3))}}{2}, & \text{if } v < \kappa_c^1, \\ v, & \text{if } v \in (\kappa_c^1, \kappa_c^4) \cup [\kappa_c^5, 1], \\ v^2 + \kappa_c^4 - (\kappa_c^4)^2, & \text{if } v \in [\kappa_c^4, \kappa_c^5). \end{cases}$$

3.3 Proof Sketch for Theorem 1

Fix $c \in C$. We solve the optimal transport problem

$$\pi^\star(c) \coloneqq \max_{\rho \in \mathcal{R}_c} \int_{V^2} \pi_c(v_l, v_h)\, \mathrm{d}\rho_c\,, \tag{10}$$

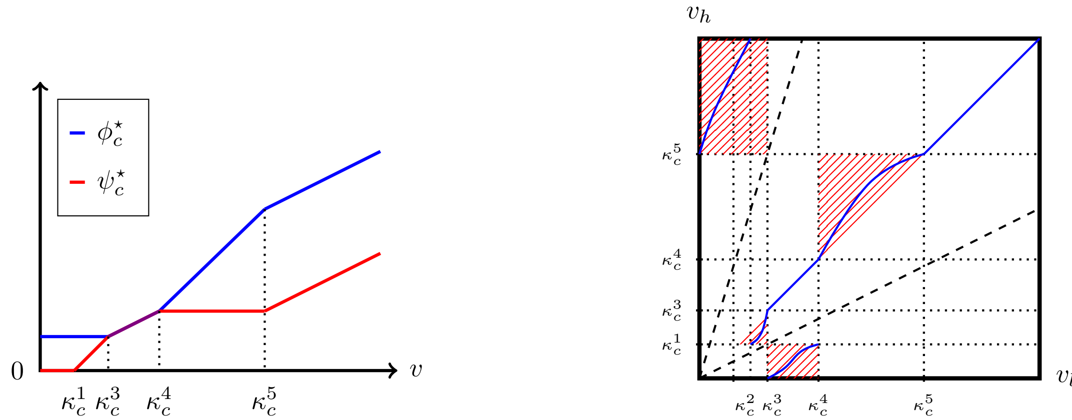
underlying Theorem 1 using duality arguments, and construct explicitly the multipliers that verifies the optimality of $\rho_c^\star$. The dual problem for each cost c is given by

$$\begin{aligned} \pi_\star(c) \coloneqq \inf_{\phi_c, \psi_c} & \left[\int_V \phi_c(v_l) F_{c,l}(\mathrm{d}v_l) + \int_V \psi_c(v_h) F_{c,h}(\mathrm{d}v_h) \right] \\ \text{s.t. } & \phi_c(v_l) + \psi_c(v_h) \geq \pi_c(v_l, v_h)\,, \end{aligned} \tag{11}$$

where the infimum is taken over all measurable functions $\phi_c, \psi_c : V \to \mathbb{R}$. Since π_c is continuous, the Kantorovich duality theorem holds (see, e.g., Villani 2009, Theorem 5.10).

Lemma 2 (Kantorovich Duality). *$\pi_\star(c) = \pi^\star(c)$ for all $c \in C$. Moreover, for any $c \in C$, for any measurable ϕ_c, ψ_c such that $\phi_c(v_l) + \psi_c(v_h) \geq \pi_c(v_l, v_h)$ for all $(v_l, v_h) \in V \times V$, and for any $\rho_c \in \mathcal{R}_c$, ρ_c is a solution of (10) and ϕ_c, ψ_c is a solution of (11) if and only if*

$$\phi_c(v_l) + \psi_c(v_h) = \pi_c(v_l, v_h) \tag{12}$$



(a) Optimal Duals $\phi^\star$, $\psi^\star$

(b) Complementary Slackness

Blue indicates the support of $\rho_c^\star$. Red indicates dual-binding regions. The dashed lines partition V^2 into regions where the seller gives rents or generates losses.

Figure 5: Optimal dual variables and complementary slackness conditions.

for all $(v_l, v_h) \in \text{supp}(\rho_c)$.

Therefore, it suffices to find, for each $c \in C$, a joint distribution $\rho_c^\star \in \mathcal{R}_c$ and a pair of functions $\phi_c^\star$ and $\psi_c^\star$ such that $(\phi_c^\star, \psi_c^\star)$ is feasible in the dual problem (11) and that the complementary slackness condition holds: $\phi_c^\star(v_l) + \psi_c^\star(v_h) = \pi_c(v_l, v_h)$ for all $(v_l, v_h) \in \text{supp}(\rho_c^\star)$. In Appendix A.1, we construct explicitly the optimal dual variables $(\phi_c^\star, \psi_c^\star)$. Figure 5a illustrates the functions $\phi_c^\star$ and $\psi_c^\star$ when $\alpha_c = 1/2$ and $c = 0$.

Then, we show that the complementary slackness condition (12) holds under the joint distribution $\rho_c^\star$ associated with the pricing rule $p^\star$. Figure 5b illustrates the support of the joint distribution of $\rho_c^\star$ when $c = 0$, where the blue region indicates the support of $\rho_c^\star$ and the dashed red region indicates the set of (v_l, v_h) at which $\phi_c^\star(v_l) + \psi_c^\star(v_h) = \pi_c(v_l, v_h)$.

As Figure 5b shows, the optimal match $\rho_c^\star$ is in fact a combination of assortative and (partly) anti-assortative matching. The region between the two dashed lines indicates value pairs where the seller would optimally serve both. In this region, $\rho_c^\star$ matches consumers assortatively. In contrast, the regions above and below the dashed lines indicate value pairs where the seller optimally excludes the low value. In this region, $\rho_c^\star$ matches consumers (partly) anti-assortatively.

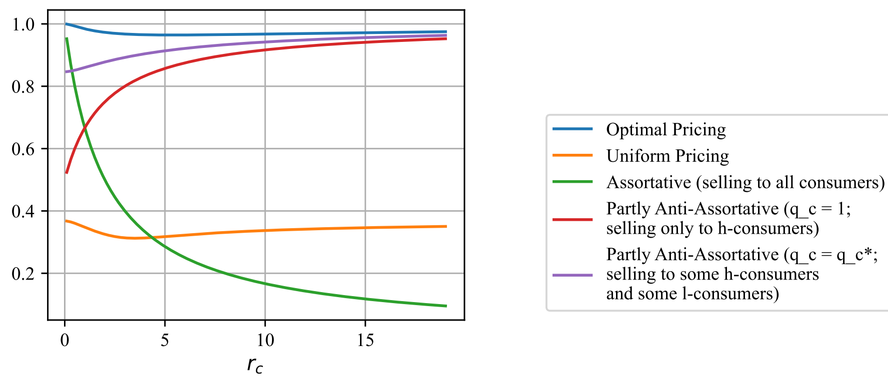


Figure 6: Seller profit divided by maximal social surplus for different pricing rules with $c = 0$, $\alpha_c = 1/2$, exponentially distributed values and $\mathbb{E}\left[v \mid c, \theta = l\right] = 1$.

3.4 Surplus Extraction by Non-Discriminatory Pricing

We now provide a general lower bound on the share of gains from trade the seller can guarantee under the optimal non-discriminatory pricing rule. For any cost $\tilde{c} \in C$, let $r_{\tilde{c}} + 1$ denote the ratio of gains from trade of h- and l-consumers:

$$r_{\tilde{c}} := \frac{\mathbb{E}\left[(v-c)^{+} \mid c = \tilde{c}, \theta = h\right]}{\mathbb{E}\left[(v-c)^{+} \mid c = \tilde{c}, \theta = l\right]} - 1\,.$$

The next proposition shows that the seller can guarantee a substantial fraction of gains from trade using non-discriminatory personalized pricing, and that this guarantee depends only on the relative gains from trade of the two groups.

Proposition 2. *For all $\tilde{c} \in C$,*

$$\frac{\mathbb{E}\left[(p^{\star} - c)\mathbf{1}\{v \geq p^{\star}\} \mid c = \tilde{c}\right]}{\mathbb{E}\left[(v-c)^{+} \mid c = \tilde{c}\right]} \geq \frac{\max\{1, \alpha_{\tilde{c}}(r_{\tilde{c}} + 1)\}}{\alpha_{\tilde{c}} r_{\tilde{c}} + 1} \geq \max\left\{\frac{1}{2-\alpha_{\tilde{c}}}, \frac{r_{\tilde{c}} + 1}{2 r_{\tilde{c}} + 1}\right\} > \frac{1}{2}\,.$$

The above bound is achieved by either p^{ass} or p^{anti} with $q_c = 1$ for all c.

According to Proposition 2, the seller can always guarantee more than half of gains from trade under the non-discrimination constraint. For example, if h-consumers have 40% higher gains from trade than l-consumers (that is, $r_c = 0.4$), then the seller can extract at least 77% of the total gains from trade under the optimal non-discriminatory personalized pricing rule. Similarly, if the groups are equal in size, $\alpha_c = 0.5$, the seller can guarantee at least 66% of the surplus. Thus, while non-discrimination regulations restrict the seller's ability to fully extract surplus, they do not eliminate the seller's ability to extract a large share of the gains from trade. Since p^{ass} and p^{anti} are generally not optimal, the bound in Proposition 2 is not tight in general, and the seller can often extract an even larger share of gains from trade under the optimal pricing rule $p^{\star}$, as illustrated by Figure 6, where groups are equally sized and values are exponentially distributed. In this case, the seller can always extract at least 90% of the gains from trade.

Comparison to Uniform Pricing Let p^u denote an optimal uniform pricing rule.[23] The guarantee in Proposition 2 is substantially stronger than what uniform pricing can provide. Indeed, while the personalized pricing rule $p^{\star}$ guarantees the seller strictly more than half of gains from trade, uniform pricing provides no comparable guarantees. The next proposition shows that, even under likelihood-ratio ordering, uniform pricing can extract an arbitrarily small share of total gains from trade.

Proposition 3. *For any $\tilde{c} \geq 0$, $\alpha_{\tilde{c}} \in (0,1)$, and $\varepsilon > 0$, there exists bounded support $F_{\tilde{c},l}, F_{\tilde{c},h}$ ordered in likelihood ratio, such that for any optimal uniform pricing rule p^u,*

$$\frac{\mathbb{E}\left[(p^u - c)\mathbf{1}\{v \geq p^u\} \mid c = \tilde{c}\right]}{\mathbb{E}\left[(v-c)^+ \mid c = \tilde{c}\right]} < \varepsilon \,.$$

Thus, banning personalized pricing altogether can sharply reduce the seller's profit relative to optimal non-discriminatory personalized pricing.[24]

4 Welfare Implications

In this section, we discuss the welfare implications of non-discrimination regulations using the characterization given by Theorem 1.

[23] That is, $p^u \in \operatorname{argmax}_{x\geq 0}(x-c)\left[\alpha_c(1-F_{c,h}(x)) + (1-\alpha_c)(1-F_{c,l}(x))\right]$ almost surely.

[24] Uniform pricing can guarantee half the surplus when the seller's profit, as a function of the uniform price, is concave, but not otherwise (see e.g. Bergemann, Castro and Weintraub 2022). In the example of Figure 6, the seller's profit function is not concave, and uniform pricing therefore yields less than 40% of gains from trade.

4.1 Allocation of Consumer Rents and Welfare Losses

An immediate implication of Theorem 1 is a characterization of consumer surplus and welfare losses under optimal non-discriminatory pricing. The non-discrimination constraint precludes fully personalized pricing, and thus may generate consumer rents and welfare losses. Proposition 4 shows exactly when these arise for each protected group and how they are allocated across the value distribution.

Proposition 4. *Under the optimal pricing rule $p^\star$, for all $c \in C$,*

(i) $CS(c, h; p^\star) > 0$*, while* $WL(c, h; p^\star) > 0$ *whenever* $\underline{v}_c < c$*;*

(ii) $CS(c, l; p^\star) > 0$ *and* $WL(c, l; p^\star) > 0$ *if and only if* $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$*;*

(iii) $WL(c, l; p^\star) \geq WL(c, h; p^\star)$*, with strict inequality if and only if* $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$*.*

Under any optimal pricing rule p every l-consumer with value $v \notin [\kappa_c^2, \kappa_c^3]$ retains zero surplus; whereas every h-consumer with value $v \notin [\kappa_c^4, \kappa_c^5]$ retains zero surplus.

In essence, Proposition 4 shows that consumer surplus and welfare losses are generally positive under $p^\star$, while l-consumers always have more welfare losses than h-consumers. Together with part (ii) of Theorem 1, the same is true for all undominated optimal pricing rules. More specifically, h-consumers always obtain positive surplus, and they incur welfare losses whenever the lowest valuation does not generate positive gains from trade (for example, when $\underline{v}_c = 0$). For l-consumers, surplus is positive if and only if $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$, which holds when the support $[\underline{v}_c, \overline{v}_c]$ is sufficiently wide, and in particular when $\underline{v}_c \leq c$. Thus, optimal non-discriminatory pricing typically leaves consumers with positive surplus, and does not in general induce efficient trades. Perhaps surprisingly, both protected groups can retain positive surplus and incur welfare losses even though l-consumers have lower values than h-consumers.

Since the non-discrimination constraint prohibits the seller from using any information θ conveys about v, consumers retain information rents, just as in standard screening problems. However, the allocation of consumer surplus and welfare losses differs sharply. In standard screening models, high values receive rents and some low values are excluded because of incentive compatibility, which typically translates to monotonicity of allocation rules. In contrast, according to Proposition 4, under any undominated optimal non-discriminatory pricing rule, intermediate-value consumers retain rents; high-value consumers are fully extracted, while low-value consumers are excluded. This difference is because the seller endogenously allocates consumer rents and welfare losses through the matching stage, and this allocation is governed

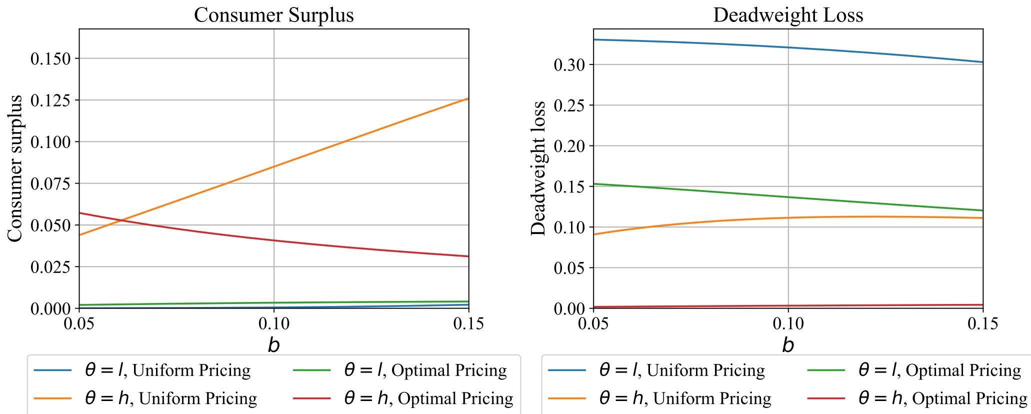


Figure 7: Consumer Surplus and Deadweight Losses

Comparisons between optimal non-discriminatory pricing and optimal uniform pricing. $c = 0$, $\alpha_c = {}^1\!/_2$, $F_{c,l} = \text{Beta}(1,2)$, $F_{c,h} = \text{Beta}(1,b)$, $b \in [{}^1\!/_{20}, {}^1\!/_5]$.

not by the elicitation of private information, but by the seller's choice of cross-group matching under the non-discrimination constraint. The seller leaves rents to intermediate-value consumers by pairing them with other intermediate-value consumers, which keeps total rents low when both types are served. Meanwhile, high-value consumers are paired with low-value consumers as this allows the seller to keep welfare losses low when one type is excluded.

Comparison to Uniform Pricing The consumer-surplus comparison to uniform pricing is ambiguous. Uniform pricing gives positive surplus to all consumers with $v \geq p^u$ and excludes the rest. Under $p^\star$, by contrast, rents are concentrated among intermediate-value consumers. Consequently, whether consumers retain more surplus, and whether welfare losses are larger, depends on the induced demand curvature and on the size of the assortative matching regions.

Figure 7 illustrates this: When h-consumer values are concentrated, the optimal uniform price is high and generates large deadweight losses, so both groups are better off under $p^\star$ than under p^u. When h-consumer values are more dispersed, the uniform price is lower and consumers are better off under p^u. Hence, banning personalized pricing may or may not benefit consumers.

4.2 Welfare Consequences across Groups

We next ask how the allocation of rents and losses characterized above maps into average surplus across protected groups. As Proposition 5 shows, the answer depends systematically on population shares: the larger the h-group is, the less surplus it retains under the optimal non-discriminatory pricing rule.

Proposition 5 (Effects of Population Sizes)**.** *Fix the value distributions $F_{c,l}, F_{c,h}$. Let $CS(c,\theta;p^\star,\alpha_c)$ denote the consumer surplus under pricing rule $p^\star$ when $\mathbb{P}[\theta = h \mid c] = \alpha_c$.*

(i) The surplus of h-consumers $CS(c,h;p^\star,\alpha_c)$ decreases in α_c.

(ii) For all $c \in C$, $\lim_{\alpha_c \to 1} CS(c,h;p^\star,\alpha_c) = \lim_{\alpha_c \to 0} CS(c,l;p^\star,\alpha_c) = 0$.

In particular, for all $c \in C$, there exists α_c, α_c' such that $CS(c,h;p^\star,\alpha_c) > CS(c,l;p^\star,\alpha_c)$ and $CS(c,h;p^\star,\alpha_c') < CS(c,l;p^\star,\alpha_c')$.

Figure 8 plots the surplus of h-consumers and l-consumers as functions of the share of h-types. As stated by Proposition 5, as α_c increases, $CS(c,h;p^\star,\alpha_c)$ decreases, while $CS(c,l;p^\star,\alpha_c)$ may increase or decrease.[25] Intuitively, as the h-group grows, the seller has higher incentives to tailor prices finely to the h-group, which reduces its surplus. Which protected group benefits more overall, however, depends on their population sizes and the shape of the underlying value distributions. Figure 8 illustrates this: On the left panel, h-consumers retain more surplus when in minority ($\alpha_c \approx 0$) than l-consumers do when in minority ($\alpha_c \approx 1$). On the right panel, the order reverses. In fact, even with fixed population sizes, different value distributions $F_{c,l}$ and $F_{c,h}$ can imply that either l-consumers or h-consumers benefit more from non-discriminatory pricing.[26]

These results, together with Proposition 4, therefore qualify the usual intuition that anti-discrimination regulation would automatically benefit primarily the socially disadvantaged group. Depending on the underlying value distributions and population sizes, either protected group may benefit more. At the same time, part of the surplus preserved by the non-discrimination constraint is sustained by the exclusion of some low-value consumers. In other words, intermediate-value consumers in both groups retain some rents at the expense

[25] A sufficient condition for $CS(c,l;p^\star,\alpha_c)$ to be nondecreasing in α_c is that $f_{c,h}$ is nondecreasing on $[0, v_c^\star]$; for instance, when $f_{c,h}$ is log-concave with mode at or above $v_c^\star$.

[26] As an example, fix $\alpha_c = 1/2$ for all c, and let $(F_{c,l}, F_{c,h})_{c \in C}$ belong to a scale family with base distributions $F_l = \mathrm{Beta}(a_l, b_l)$ and $F_h = \mathrm{Beta}(a_h, b_h)$ supported on $[0, K]$ for some $K \geq 1$. By Proposition A.1 (see Appendix A.3), Assumption 1 reduces to the single inequality $F_l(1) < \|F_l - F_h\|$. Under $(a_l, b_l) = (2, 4)$ and $(a_h, b_h) = (2, 3)$, l-consumers retain higher surplus; under $(a_l, b_l) = (2, 4)$ and $(a_h, b_h) = (4, 2)$, h-consumers do. We verify in Appendix A.3 that both pairs satisfy Assumption 1.

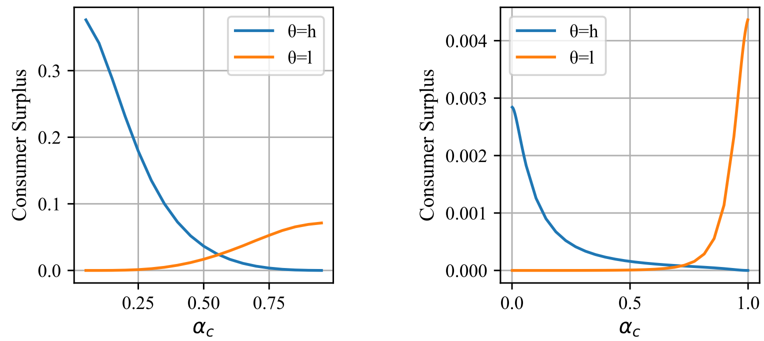


Figure 8: Consumer Surplus as functions of α_c.

The left panel has $c = 0$, $F_{c,l}$ given by Beta$(3, 3)$, $F_{c,h}$ given by Beta$(3.5, 3)$. The right panel has $c = 0$, $\underline{v}_c = 0$, $\overline{v}_c = 1$. $f_{c,l}(v) = 2v$, $f_{c,h}(v) = 1.98v$ for $v \leq 0.9$ and $f_{c,h}(v) \approx 2.0853v$ for $v > 0.9$ (with $b = (1 - 0.99 \cdot 0.9^2)/(1 - 0.9^2)$, so that $f_{c,h}$ integrates to one).

of welfare losses from low-value consumers, which are higher in the l-group than in the h-group according to part (iii) of Proposition 4.

5 Other Notions of Discrimination

5.1 Non-Discriminatory Outcomes

So far, we have studied a notion of non-discrimination that applies to the distribution of prices *offered* to consumers. Under this notion, consumers in different protected groups are given *equal opportunities*, in the sense that they face, overall, the same prices *before* deciding whether to purchase. We now consider a stronger notion of non-discrimination that instead requires *equal outcomes*: consumers in different protected groups must face the same transaction outcomes and transaction prices *after* purchase decisions are made. This comparison is useful because it shows how different regulations shape the seller's optimal allocation of surplus differently.[27] Changing the object constrained by regulation changes

[27] In practice, these differences are often driven by the availability of auditable data. Public guidance does not draw a sharp line between offered prices and transaction outcomes. Instead, regulators may use disparities in prices "quoted or charged," along with pricing discretion and other risk factors, as signals for further review, while relying on the subset of records that are systematically documented and comparable across applicants when conducting statistical analysis (Federal Financial Institutions Examination Council 2021).

how the seller can allocate rents and exclusion across groups.

Specifically, given a pricing rule p, denote by $y(\omega) = \mathbf{1}\{v \geq p(\omega)\}$ the random variable that indicates whether or not the product is sold to consumer ω.

Definition 4. A pricing rule p induces *non-discriminatory outcomes* if for all $c \geq 0$ and $M \subseteq \mathbb{R} \times \{0,1\}$,

$$\mathbb{P}[(p,y) \in M \mid c, \theta = l] = \mathbb{P}[(p,y) \in M \mid c, \theta = h].$$

In other words, a pricing rule induces non-discriminatory outcomes if the transaction event and the transaction price, (p,y), are independent of the protected characteristic θ conditional on cost c. Clearly, any pricing rule that induces non-discriminatory outcomes must also be non-discriminatory. Relative to the notion of non-discriminatory pricing based on offered prices, this stronger requirement restricts not only how prices are offered, but also how the seller may allocate exclusion across groups. Recall that in (6), we defined p^{ass} to be the pricing rule that matches h- and l-consumers assortatively and charges each pair the lower of their values whenever that value is at least c, and excludes the pair entirely otherwise.

Proposition 6 (Optimal Pricing Rule with Non-Discriminatory Outcomes)**.**

(i) p^{ass} *induces non-discriminatory outcomes and maximizes the seller's profit among all pricing rules that induce non-discriminatory outcomes.*

(ii) p^{ass} *yields a lower profit than the optimal non-discriminatory pricing rule:* $\Pi(p^\star) \geq \Pi(p^{ass})$*, and the inequality is strict if and only if* $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$ *for a positive measure of* $c \in C$*.*

(iii) The surplus of h*-consumers is higher under* p^{ass} *than under* $p^\star$*, while the surplus of* l*-consumers is lower. That is, for all* c*,*

$$CS(c,h;p^\star) \leq CS(c,h;p^{ass}) \quad \textit{and} \quad CS(c,l;p^\star) \geq CS(c,l;p^{ass}),$$

where the inequalities are strict if and only if $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$*.*

Proposition 6 shows that different standards of regulation qualitatively reshape the seller's optimal allocation of rents and exclusion. Relative to the notion of non-discriminatory pricing based on offered prices, requiring non-discriminatory outcomes removes the seller's ability to use exclusion asymmetrically across groups while still equalizing offer distributions. As a result, the optimal rule collapses to the assortative pricing rule p^{ass}. Thus, relative to the non-discriminatory pricing rule $p^\star$ based on offered prices, the stronger constraint transfers surplus from l-consumers to h-consumers: h-consumers are better off, while l-consumers are worse off. Thus, stricter fairness need not benefit the lower-value group more. In settings where the disadvantaged group has lower values—for example, because it is a poorer group—requiring non-discriminatory outcomes may therefore make the disadvantaged group worse off, even relative to the weaker notion of non-discriminatory pricing based on offers.

5.2 Disparate Impact and Disparate Treatment

As discussed in Section 1, the notion of non-discriminatory pricing studied in the previous sections is often referred to as *disparate impact* regulation. In legal terms, another natural notion of non-discrimination is *disparate treatment* regulation. Unlike disparate impact regulation, disparate treatment regulation does not impose restrictions on the output of pricing rules. Rather, it prohibits the use of protected characteristics as *inputs* into pricing decisions. Clearly, as noted in Section 1, in a data-rich environment where many proxies are available, disparate treatment regulation may have little effect. Nevertheless, when proxies are limited, it is less clear how disparate treatment and disparate impact differ in their welfare implications. The comparison between disparate treatment-based and disparate impact regulations would be policy-relevant, as which standard is applied by regulators depends on the underlying legal basis, regulations, and case law, which vary across different markets and across time.[28] Our framework allows us to compare these two notions directly.

Consider the following re-parameterization of the environment introduced in Section 2. A consumer's type is given by $\omega = (z, c, \theta, r) \in Z \times C \times \Theta \times [0, 1]$, where Z is an arbitrary measurable space. Unlike the model in Section 2, the value of a consumer is not part of their primitive type. Instead, the value of type $\omega = (z, c, \theta, r)$ is determined by $V_{c,\theta}(z)$, where $V_{c,\theta} : Z \to \mathbb{R}_+$ is a measurable function. With this parameterization, $(z, c) \in Z \times C$ can be interpreted as the consumer's *non-protected* characteristics, which, together with the protected characteristic θ, determines the consumer's value via $V_{c,\theta}(z)$. Throughout this section, we assume that $V_{c,h} \geq V_{c,l}$ for all $c \in C$, and maintain the assumption that the

[28] As noted in Section 1, housing and labor markets have long-standing legal bases for disparate impact claims, while in credit markets the basis has been more contested and has often depended on regulatory interpretation. The executive branch can also shape that regulatory basis. For instance, The White House (2015) emphasized the relevance of disparate impact reasoning in the era of big data, whereas the CFPB's 2026 final rule moved in a different direction by limiting disparate impact-based enforcement under ECOA and focusing instead on disparate treatment.

distributions of $V_{c,\theta}(z)$ conditional on (c,θ), denoted by $F_{c,\theta}$, are such that $F_{c,h}$ dominates $F_{c,l}$ in the likelihood ratio order for all $c \in C$ and satisfies Assumption 1.

Since a consumer's value is determined by (z,c,θ), and since $F_{c,h}$ dominates $F_{c,l}$ in the likelihood ratio order, the analysis above continues to apply and the optimal disparate impact-based non-discriminatory pricing rule is given by Theorem 1. With this alternative parameterization, however, we can also study disparate treatment regulation.

Definition 5. A pricing rule p is *disparate treatment-based non-discriminatory* if p does not depend on θ. That is, for any ω, ω' with $z = z'$, $c = c'$ and $r = r'$, we have $p(\omega) = p(\omega')$.

The implications of disparate treatment regulation depend on how the non-protected characteristics z are correlated with θ. If the non-protected characteristics perfectly identify θ—that is, if $\alpha_c(z) \in \{0,1\}$ almost surely—then prohibiting the direct use of θ has no bite: the seller can reconstruct the protected characteristic from proxies and effectively replicate any outcomes.[29] At the other extreme, if z is independent of θ conditional on c, then any disparate treatment-based rule is also non-discriminatory in the sense of disparate impact. In general, the strength of disparate treatment regulation depends on how informative the available proxies are about protected characteristics. Specifically, let

$$\alpha_c(z) := \mathbb{P}[\theta = h \mid z, c]$$

be the conditional distribution of θ given z and c.

Proposition 7 (Optimal Pricing under Disparate Treatment Regulation)**.** *Let $\overline{p}$ be an optimal disparate treatment-based non-discriminatory pricing rule. Then,*

(i) $\Pi(\overline{p}) = \mathbb{E}\left[\max\{0,\, V_{c,l}(z) - c,\, \alpha_c(z)(V_{c,h}(z) - c)\}\right]$.

(ii) $CS(c,l;\overline{p}) = 0$ *for all* $c \in C$.

(iii) Holding fixed the joint distribution of $(V_{c,l}(z), V_{c,h}(z), c)$, $\Pi(\overline{p})$ *is higher when the conditional distribution of* $\alpha_c(z)$ *given* $(V_{c,l}(z), V_{c,h}(z), c)$ *is more dispersed in the convex order for all* $c \in C$.

Proposition 7 shows that the welfare implications of disparate treatment regulation and disparate impact regulation are qualitatively different. First, l-consumers are always (weakly) worse off under disparate treatment regulation than under disparate impact regulation: they

[29] As a result, all previous results under the disparate impact-based constraint can also be made to satisfy the disparate treatment-based constraint if we assume that θ is measurable with respect to (z,c).

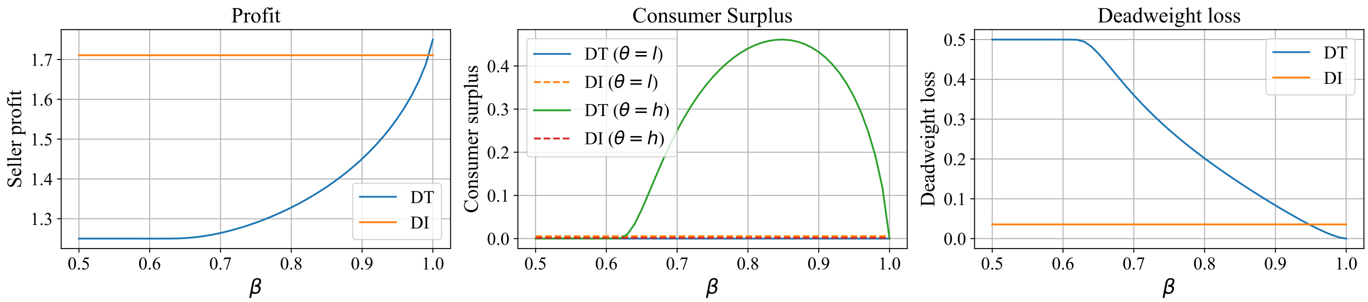


Figure 9: Disparate Treatment versus Disparate Impact

$\beta \in [1/2, 1]$ captures the informativeness of z about θ. $\alpha_c = 1/2$, $c = 0$, F_h is the exponential distribution with mean 3, and F_l is the exponential distribution with mean 1.

retain zero surplus under disparate treatment regulation; whereas they can retain a positive surplus under disparate impact regulation (by Proposition 4). Second, the seller benefits more from disparate treatment regulation when more informative proxies are available. If z is a perfect proxy of θ conditional on c, so that $\alpha_c(z) \in \{0, 1\}$ almost surely, then the seller can fully extract all gains from trade and profits would exceed those under disparate impact regulation. By contrast, when proxies are weak, disparate treatment regulation can constrain the seller more tightly and lead to lower profits. Thus, regulating inputs is not equivalent to regulating outcomes: the welfare comparison between disparate treatment and disparate impact depends on the availability of proxies, and the difference is especially stark for l-consumers.

An example Suppose that $c = 0$, $\alpha_c = 1/2$, and $Z = [0, 1]$. Fix $F_{c,l}, F_{c,h} \in \Delta(\mathbb{R}_+)$ that are likelihood-ratio ordered. Let $\overline{H}(z) := \max\{2z - 1, 0\}$ and $\underline{H}(z) := \min\{2z, 1\}$. For each $\beta \in [1/2, 1]$, let

$$H_h^\beta(z) := (1-\beta)\overline{H}(z) + \beta\underline{H}(z)$$
$$H_l^\beta(z) := \beta\overline{H}(z) + (1-\beta)\underline{H}(z)\,,$$

be the conditional distribution of z given $\theta = h$ and $\theta = l$, respectively. Moreover, let

$$V_{c,\theta}(z) := F_{c,\theta}^{-1}(H_\theta^\beta(z))\,.$$

By construction, z is uniformly distributed on $[0,1]$, the distribution of $V_{c,\theta}(z)$ conditional on (c,θ) is $F_{c,\theta}$, and $V_{c,h} \geq V_{c,l}$. As β increases, z and θ become more correlated, and hence the distribution of $\alpha_c(z)$ becomes more dispersed conditional on $(V_{c,h}(z), V_{c,l}(z), c)$.

Letting $F_{c,h}$ be the exponential distribution with mean 3 and $F_{c,l}$ be the exponential distribution with mean 1, Figure 9 plots the seller's profit, consumer surplus, and deadweight losses under the optimal pricing rule for the two notions of non-discrimination as functions of β. Compared to disparate impact regulation, the seller's optimal profit under disparate treatment regulation is lower when β is low and higher when β is high. Surplus for the l-group is always lower under disparate treatment regulation, whereas surplus for the h-group is non-monotone in β. Furthermore, when β is low, disparate impact regulation could benefit *both* groups of consumers *and* the seller relative to disparate treatment regulation.

5.3 Unconditional Non-Discriminatory Pricing

Thus far, the notion of non-discriminatory pricing only requires the price distribution to be independent of protected characteristics *conditional* on costs. That is, prices are allowed to vary across protected groups as long as the variation can be explained by differences in costs. In some settings, correlation between prices and protected characteristics is prohibited even if it is only through costs or other "business necessities." While such settings are relatively rare under the legal framework in the U.S. (see Remark 1), one such example would be the regulation of health insurance in many European countries, which requires insurance companies to charge the same prices to men and women even if their associated costs differ.[30] Our framework can be modified to accommodate such restriction as well. To demonstrate, suppose that v and c are perfectly correlated, where $v - c = s$ with probability 1. In the context of insurance with a risk-neutral insurance provider, s is the risk-premium consumers are willing to pay for insurance. The assumption that v and c are perfectly correlated then corresponds to assuming that consumers differ in their expected damages, but not the risk premium they are willing to pay. A pricing rule p is said to be non-discriminatory if for all $M \subseteq \mathbb{R}_+$

$$\mathbb{P}[p \in M \mid \theta = h] = \mathbb{P}[p \in M \mid \theta = l].$$

[30] After the Court of Justice of the European Union Test-Achats v. Belgium ruling, new insurance agreements signed since the 21st of December 2012 must use unisex tariffs.

That is, prices are required to be independent of protected characteristics *unconditionally*. By the same arguments as in the proof of Proposition 1, the optimal pricing problem can be represented by an optimal transport, where the seller first chooses a matching scheme ρ to match l- and h-consumers into pairs, and then sets an optimal price for each matched pair. In Appendix A.3.3, we solve this problem explicitly and show that the optimal pricing rule features similar qualitative patterns as $p^\star$: it mixes positive assortative matching with partly anti-assortative matching; low-value consumers are excluded, high-value consumers are extracted, and intermediate-value consumers retain some rents.

6 Extensions

6.1 Imperfect Price Discrimination

Thus far, we assumed that every pricing rule that satisfies the non-discrimination constraint is feasible. This requires the seller to know each consumer's type. In practice, sometimes the seller may not have access to enough data to perfectly estimate consumers' types, and can only obtain a noisy signal.

Our method can still be applied to characterize the profit-maximizing non-discriminatory pricing rule in this environment. Specifically, suppose now that consumers' true values are denoted by $w \geq 0$, whose distribution depends on an observable type v. The observable types v are distributed according to $F_{c,l}$, $F_{c,h}$ among consumers with protected characteristics l and h conditional on cost c, respectively. A pricing rule $p : V \times C \times \Theta \times [0,1] \to \mathbb{R}_+ \cup \{+\infty\}$ is defined as before, except that v does not stand for a consumer's true value but only provides noisy information about a consumer's value. Given a pricing rule p, the seller's profit is given by

$$\Pi(p) := \mathbb{E}\left[(p - c) \cdot \mathbf{1}\{w \geq p\}\right] .$$

By the same arguments as the proof of Proposition 1,[31] we may still recast the problem into an optimal transport:

$$\widetilde{\pi}^\star := \int_C \left(\max_{\rho_c \in \mathcal{R}_c} \int_{V^2} \widetilde{\pi}_c(v_l, v_h) \, \mathrm{d}\rho_c \right) G(\mathrm{d}c) \, , \tag{13}$$

[31] Alternatively, this can be derived from Lemma 3 of Strack and Yang (2024).

where[32]

$$\widetilde{\pi}_c(v_l, v_h) \coloneqq \max_{x \geq 0} \left[(x - c) \cdot (\alpha_c \mathbb{P}[w \geq x \mid v_h] + (1 - \alpha_c)\mathbb{P}[w \geq x \mid v_l]) \right] .$$

As a result, the profit-maximizing pricing rules can still be found by solving the optimal transport problem (13).

To illustrate, suppose that $c = 0$ almost surely, $\alpha_c = 1/2$, and that consumers' values w are distributed uniformly on $[0, 2v]$ conditional on v. It then follows that

$$\widetilde{\pi}_c(v_l, v_h) = \max \left\{ \frac{v_l}{4}, \frac{v_h}{4}, \frac{v_l v_h}{v_l + v_h} \right\} .$$

The objective $\widetilde{\pi}_c$ is qualitatively similar to π_c defined in (3). Indeed, just like π_c, $\widetilde{\pi}_c$ is supermodular on the region where $v_l \leq 3v_h$ and $v_h \leq 3v_l$; while it is linear whenever $3v_h < v_l$ or $3v_l < v_h$. As a result, the optimal matching scheme is qualitatively similar. Indeed, as we show in Appendix A.3.2, the corresponding optimal matching scheme is the same as $\rho_c^\star$, except that the cutoffs κ_c are defined by a different system of equations.

6.2 Implementable Welfare Outcomes

While we have so far focused on how the seller can maximize their profits using a non-discriminatory pricing rule, another natural question is what consumer welfare can be achieved. To explore this question, recall that from Remark 2, we may view pricing rules as price discriminating consumers based on a given market segmentation. In what follows, we explore the welfare outcomes (i.e., consumer surplus and seller profit) that can be induced by a non-discriminatory pricing rule that charges an optimal price in each market segment. That is, we calculate the average consumer surplus and the seller's profit that can be induced by some non-discriminatory pricing rule p that is (i) measurable with respect to some segmentation $s : \Omega \to S$, and (ii) is optimal given each segment for the seller.[33]

The set of feasible pairs of seller profit and average consumer surplus without the non-discrimination constraint is characterized by Bergemann et al. (2015) for $c = 0$, as the triangle spanned by the points $(\mathbb{E}[v], 0)$, $(r^\star, 0)$, $(r^\star, \mathbb{E}[v] - r^\star)$, where $r^\star = \max_p p \cdot \mathbb{P}[v \geq p]$ is the

[32] More precisely, $\widetilde{\pi}(v_l, v_h)$ is the optimal profit of the seller when selling to a consumer with value w, which is distributed according to $\alpha_c \mathbb{P}[w \in \cdot \mid v_h] + (1 - \alpha_c)\mathbb{P}[w \in \cdot \mid v_l]$. By standard arguments, a posted price mechanism is optimal. Therefore, it is without loss to restrict attention to (personalized) posted prices even if the seller cannot fully observe consumers' values.

[33] That is, the price $p(s)$ in each segment s satisfies $p(s) \in \operatorname{argmax}_{x \geq 0} \mathbb{E}[(x - c)\mathbf{1}\{v \geq x\} \mid s]$.

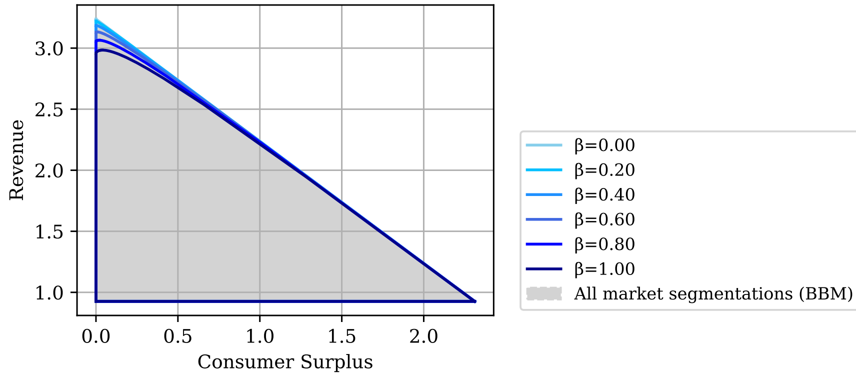


Figure 10: Feasible Welfare Outcomes. $c = 0$, $\alpha_c = 1/4$, $F_{c,h} = (\beta + \alpha(1-\beta))\overline{F} + (1 - (\beta + (1-\beta)\alpha))\underline{F}$ and $F_{c,l} = (1-\beta)\alpha\overline{F} + (1 - (1-\beta)\alpha)\underline{F}$.

optimal uniform pricing revenue.With the non-discrimination constraint, however, not every outcome in this triangle is feasible.

As an example, suppose that $c = 0$ and $\alpha_c = 1/4 =: \alpha$ of consumers are of protected characteristic h, and let $\overline{F}, \underline{F}$ be exponential distributions with means 10 and 1. Consider a parameterization of $F_{c,h}$ and $F_{c,l}$ where for some $\beta \in [0,1]$,

$$F_{c,h} := (\beta + \alpha(1-\beta))\overline{F} + (1 - (\beta + (1-\beta)\alpha))\underline{F}$$
$$F_{c,l} := (1-\beta)\alpha\overline{F} + (1 - (1-\beta)\alpha)\underline{F}\,.$$

By construction, $F_{c,h}$ dominates $F_{c,l}$ in the likelihood ratio order for all $\beta \in [0,1]$, and the total variation distance $\|F_{c,h} - F_{c,l}\|$ increases in β. Meanwhile, the overall distribution of values $\alpha F_{c,h} + (1-\alpha)F_{c,l}$ in the population remains unchanged in β. As a consequence, the surplus triangle of Bergemann et al. (2015) remains the same for all $\beta \in [0,1]$.

Figure 10 plots the set of feasible pairs of seller profit and average consumer surplus with the non-discrimination constraint for this parameterized family of type distributions. The shaded triangle corresponds to the feasible surplus division given by Bergemann et al. (2015), whereas the colored curves depict the boundary of the set of feasible welfare outcomes for different values of the parameter β.

As noted above, when $c = 0$ almost surely, Assumption 1 holds, and hence the highest feasible revenue is less than the total gains from trade $\mathbb{E}[v]$, due to the non-discrimination constraint. This is reflected in Figure 10 by the fact that the top-left corner of the triangle is not included in feasible surplus region. One notable feature of the example is the relatively minor restriction on implementable welfare outcomes even when the distributions of values become vastly different across consumer groups (recall that in this example, h-consumers value the good 10 times more than l-consumers at $\beta = 1$).

In this parametric example, there exists a segmentation that keeps the seller's revenue the same as the uniform pricing revenue, induces a non-discriminatory pricing rule in which all consumers buy, and thus the boundaries all reach the bottom-right corner of the triangle. However, this is not the case in general. Characterizing explicitly the feasible welfare outcomes in general, and in particular, when can the consumer-optimal outcome be attained under the non-discrimination constraint, is an exciting question for future research.

6.3 General Distributions

We now relax Assumption 1 and characterize the undominated profit-maximizing pricing rules for all distributions of values. According to Proposition 1, the optimal pricing rule can be found by solving an optimal transport problem for each $c \in C$. To this end, let

$$
\begin{aligned}
C_1 &:= \{c \in C : F_{c,l}(c) < \|F_{c,l} - F_{c,h}\|\} \\
C_2 &:= \{c \in C : F_{c,l}(c) \geq \|F_{c,l} - F_{c,h}\|, c < v_c^\star\} \\
C_3 &:= \{c \in C : F_{c,l}(c) \geq \|F_{c,l} - F_{c,h}\|, c \geq v_c^\star\}.
\end{aligned}
$$

By definition, C_1, C_2, C_3 partitions the set of possible costs C into three regions. Note that Assumption 1 imposes that $c \in C_1$ almost surely. We now define an optimal pricing rule $p^\star$ for general distributions conditioning on different realizations of c. When $c \in C_1$, let $p^\star$ be defined in (9).

When $c \in C_2$, since $F_{c,l}(c) \geq \Delta_c(v_c^\star)$, there exists a unique value $\eta_c^l \leq c$ such that $F_{c,l}(c) - F_{c,l}(\eta_c^l) = \Delta_c(v_c^\star)$. Let $\eta_c^h := F_{c,h}^{-1}(F_{c,l}(\eta_c^l))$. The pricing rule $p^\star$, when $c \in C_2$, is defined as follows:

$$p^\star(v,c,l,r) := \begin{cases} c, & \text{if } v < \eta_c^l \\ \overline{\Delta}_c^{-1}(\Delta_c(v_c^\star) + F_{c,l}(\eta_c^l) - F_{c,l}(v)), & \text{if } v \in [\eta_c^l, c) \\ v, & \text{if } v \geq c \end{cases}$$

$$p^\star(v,c,h,r) := \begin{cases} c, & \text{if } v < \eta_c^h \\ \underline{\Delta}_c^{-1}(F_{c,h}(v) - F_{c,h}(\eta_c^h) + \Delta_c(c)), & \text{if } v \in [\eta_c^h, c) \\ v, & \text{if } v \geq c \end{cases}$$

When $c \in C_3$, the pricing rule $p^\star$ is defined as follows:

$$p^\star(v,c,l,r) := \begin{cases} c, & \text{if } v < F_{c,l}^{-1}(F_{c,h}(c)) \\ \overline{\Delta}_c^{-1}(\Delta_c(c) + F_{c,h}(c) - F_{c,l}(v)), & \text{if } v \in [F_{c,l}^{-1}(F_{c,h}(c)), c) \\ v, & \text{if } v \geq c \end{cases}$$

$$p^\star(v,c,h,r) := \max\{v,c\}$$

Theorem 2 below establishes that $p^\star$ is an optimal non-discriminatory pricing rule.

Theorem 2 (Optimal Pricing for General Distributions)**.**

(i) $p^\star$ is a profit-maximizing non-discriminatory pricing rule, that is, $p^\star$ solves (1).

(ii) Every undominated profit-maximizing non-discriminatory pricing rule p induces the same average surplus for each protected characteristic conditional on cost. That is, $CS(c,\theta;p) = CS(c,\theta;p^\star)$ for all $c \in C$ and $\theta \in \{l,h\}$.

We note that Theorem 2 immediately implies Theorem 1, since $c \in C_1$ almost surely under Assumption 1. As another immediate consequence of Theorem 2, under the optimal pricing rule $p^\star$, the seller is able to fully extract all gains from trade whenever $c \in C_2 \cup C_3$.

Corollary 2. *For all $\tilde{c} \in C_2 \cup C_3$,*

$$\mathbb{E}[(p^\star - c)\mathbf{1}\{v \geq p^\star\} \mid c = \tilde{c}] = \mathbb{E}[(v-c)^+ \mid c = \tilde{c}]\,.$$

As a result,

$$CS(c,\theta;p^\star) = WL(c,\theta;p^\star) = 0\,,$$

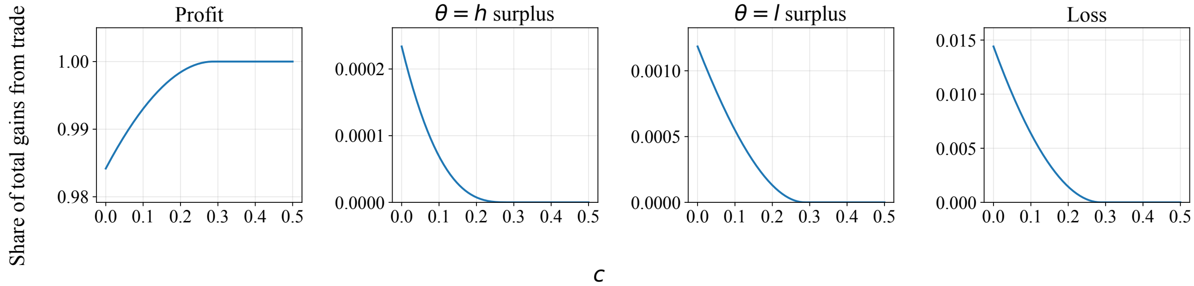


Figure 11: Welfare Outcomes as c increases. $\alpha = 1/2$, $F_h = \text{Exp}(2)$, $F_l = \text{Exp}(1)$.

for all $c \in C_2 \cup C_3$ *and* $\theta \in \{l, h\}$.

According to Corollary 2, the restrictions on the value distributions imposed by Assumption 1 are in fact equivalent to restricting attention to distributions where the seller cannot fully extract all gains from trade. Intuitively, this is because when $c \notin C_1$, there are few consumers who are efficient to trade with, and thus it is easier for the seller to pair them with high-value consumers at no cost.

With Theorem 2, we can also examine how welfare outcomes change as cost c increases. Specifically, suppose that $F_{c,\theta} = F_\theta$ and $\alpha_c = \alpha$ for all $c \in C$. In this case, C_1, C_2 and C_3 form a monotone partition of $\mathbb{R}_+$. Thus, as c increases, the welfare implications of the optimal pricing rule $p^\star$ would differ. In particular, when c is large enough, the seller would be able to fully extract all surplus from consumers. The welfare implications as c increases are illustrated by Figure 11.

6.4 All Profit-Maximizing Pricing Rules

So far, we have focused on revenue-maximizing pricing rules that are undominated, in the sense that there does not exist another revenue-maximizing pricing rule that generates a higher surplus for all consumer groups. We now explore the welfare outcomes of all optimal non-discriminatory pricing rules, including the dominated ones. In particular, we characterize the surplus of consumers with each protected characteristic under all optimal non-discriminatory pricing rules. To state our welfare results, for all $c \in C$, we say that $(\sigma_{c,l}, \sigma_{c,h})$ is a surplus outcome induced by an optimal non-discriminatory pricing rule if there exists a non-discriminatory pricing rule p such that $\Pi(p) = \Pi(p^\star)$ and that σ is the induced consumer surplus $CS(c, \theta; p) = \sigma_{c,\theta}$ for all $c \in C, \theta \in \{l, h\}$.

Proposition 8 (Welfare Outcomes). *$(\sigma_{c,l}, \sigma_{c,h})$ is a surplus outcome induced by an optimal non-discriminatory pricing rule if and only if*

$$0 \leq \sigma_{c,l} \leq CS(c, l; p^\star) \quad \textit{and} \quad \sigma_{c,h} = CS(c, h; p^\star)$$

According to Proposition 8, h-consumers retain the same amount of surplus under every optimal non-discriminatory pricing rule, whereas the average surplus of l-consumers ranges from zero to $\mathbb{E}[CS(c,\theta;p^\star) \mid \theta = l]$ across all optimal non-discriminatory pricing rules. Moreover, since profits are the same across all optimal non-discriminatory pricing rules, Proposition 8 in turn implies that h-consumers' deadweight losses are the same across all optimal non-discriminatory pricing rules, whereas the average deadweight loss of l-consumers ranges from $\mathbb{E}[WL(c,\theta;p^\star) \mid \theta = l]$ to $\mathbb{E}[WL(c,\theta;p^\star) + CS(c,\theta;p^\star) \mid \theta = l]$ across all optimal non-discriminatory pricing rules.

7 Conclusion

We characterize the profit-maximizing non-discriminatory pricing rules by showing that the pricing problem can be decomposed into a matching stage and a pricing stage, and explicitly solving the corresponding optimal transport problem. Under the optimal pricing rule, both groups of consumers typically retain a positive surplus but incur positive welfare losses. Moreover, unlike standard screening problems, surplus is allocated to intermediate-value consumers, while low-value consumers are excluded and high-value consumers are extracted. Relative to uniform pricing, the optimal non-discriminatory pricing rule extracts much more gains from trade for the seller, whereas the surplus implication for consumers depends on the underlying distributions. Strengthening non-discrimination by requiring both transaction outcomes and transaction prices to be the same across protected groups makes the low-value group worse off and the high-value group better off. Compared to disparate impact regulation, disparate treatment regulation is only effective when there are not enough non-protected proxies, and the low-value group is better off under regulation based on disparate impact than on disparate treatment. We also consider several extensions to the baseline model, including imperfect price discrimination, implementable outcomes, and non-discrimination constraints unconditional on costs.

References

Aguirre, I., S. Cowan, and J. Vickers (2010) "Monopoly Price Discrimination and Demand Curvature," *American Economic Review*, 100 (4), 1601–1615.

Barocas, S., M. Hardt, and A. Narayanan (2019) *Fairness and Machine Learning: Limitations and Opportunities*: MIT Press.

Bergemann, D., B. Brooks, and S. Morris (2015) "The Limits of Price Discrimination," *American Economic Review*, 105 (3), 921–957.

Bergemann, D., F. Castro, and G. Weintraub (2022) "Third-Degree Price Discrimination versus Uniform Pricing," *Games and Economic Behavior*, 131, 275–291.

Boerma, J., A. Tsyvinski, and A. P. Zimin (2025) "Sorting with Teams," *Journal of Political Economy*, 133 (2), 421–454.

Calders, T. and S. Verwer (2010) "Three Naive Bayes Approaches for Discrimination-Free Classification," *Data Mining and Knowledge Discovery*, 21, 277–292.

Carey, A. N. and X. Wu (2023) "The Statistical Fairness Field Guide: Perspectives from Social and Formal Sciences," *AI and Ethics*, 3, 1–23.

Chen, Y. and G. Iyer (2002) "Consumer Addressability and Customized Pricing," *Marketing Science*, 21 (2), 197–208.

Çinlar, E. (2010) *Probability and Stochastics*: Springer.

CNBC (2023) "Wells Fargo mortgage lenders probed over racial discrimination," December, https://www.cnbc.com/2023/12/11/wells-fargo-mortgage-lenders-probed-over-racial-discrimination.html, Accessed: 2024-12-25.

Cohen, M. C., A. N. Elmachtoub, and X. Lei (2022) "Price Discrimination with Fairness Constraints," *Management Science*, 68 (12), 8536–8552.

Cowan, S. (2016) "Welfare-Increasing Third-Degree Price Discrimination," *RAND Journal of Economics*, 47 (2), 326–340.

Darlington, R. B. (1971) "Another Look at Cultural Fairness," *Journal of Educational Measurement*, 8 (2), 71–82.

Doval, L. and A. Smolin (2024) "Persuasion and Welfare," *Journal of Political Economy*, 132 (7), 2451–2487.

Farboodi, M., N. Haghpanah, and A. Shourideh (2025) "Good Data and Bad Data: The Welfare Effects of Price Discrimination," Working Paper.

Federal Financial Institutions Examination Council (2021) "Interagency Fair

Lending Examination Procedures," https://www.ffiec.gov/sites/default/files/media/press-releases/2021/2021-june-17-fairlend.pdf, June, Released June 17, 2021.

Haghpanah, N. and R. Siegel (2022) "The Limits of Multi-Product Price Discrimination," *American Economic Review: Insight*, 4 (4), 443–458.

——— (2023) "Pareto Improving Segmentation of Multi-Product Markets," *Journal of Political Economy*, 131 (6), 1546–1575.

Hardt, M., E. Price, and N. Srebro (2016) "Equality of Opportunity in Supervised Learning," arXiv preprint arXiv:1610.02413.

He, K., F. Sandomirskiy, and O. Tamuz (2024) "Private Private Information," arXiv preprint arXiv:2112.14356.

Kallus, N. and A. Zhou (2021) "Fairness, Welfare, and Equity in Personalized Pricing," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, 296–314, New York, NY, USA: Association for Computing Machinery, 10.1145/3442188.3445895.

Liang, A., J. Lu, X. Mu, and K. Okumura (2024) "Algorithm Design: A Fairness-Accuracy Frontier," Working Paper.

Montes, R., W. Sand-Zantman, and T. Valletti (2019) "The Value of Personal Information in Online Markets with Endogenous Privacy," *Management Science*, 65 (3), 1342–1363.

Moscariello, P. (2025) "Redistricting with Endogenous Candidates,"Technical report, mimeo.

Rhodes, A. and J. Zhou (2024) "Personalization and Privacy Choice," Working Paper.

Shaffer, G. and Z. J. Zhang (2002) "Competitive One-to-One Promotions," *Management Science*, 48 (9), 1143–1160.

Strack, P. and K. H. Yang (2024) "Privacy-Preserving Signals," *Econometrica*, 92 (6), 1907–1938.

The White House (2015) "Big Data and Differential Pricing."

Thisse, J.-F. and X. Vives (1988) "On The Strategic Choice of Spatial Price Policy," *American Economic Review*, 78 (1), 128–137.

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

Villani, C. (2009) *Optimal Transport: Old and New*: Springer.

Appendix

In this appendix, Section A.1 contains the omitted proofs. Section A.2 contains the proofs of auxiliary lemmas. Section A.3 contains additional results and details for the extensions.

A.1 Omitted Proofs

Proof of Proposition 1. Consider any non-discriminatory pricing rule p, for each $c \in C$, let ρ_c be defined by (2). We first claim that

$$\Pi(p) \leq \int_C \left(\int_{V^2} \pi_c(v_l, v_h) \rho_c(\mathrm{d}v_l, \mathrm{d}v_h) \right) G(\mathrm{d}c) \,.$$

Indeed, note that, for all $\hat{c} \in C$, by the definition of $\pi_{\hat{c}}$,

$$\int_{V^2} \pi_{\hat{c}}(v_l, v_h) \rho_{\hat{c}}(\mathrm{d}v_l, \mathrm{d}v_h) = \int_{V^2} \max_{\tilde{p}} (\tilde{p} - \hat{c}) \left[(1 - \alpha_{\hat{c}}) \mathbf{1}\{v_l \geq \tilde{p}\} + \alpha_{\hat{c}} \mathbf{1}\{v_h \geq \tilde{p}\} \right] \rho_{\hat{c}}(\mathrm{d}v_l, \mathrm{d}v_h) \,.$$

Meanwhile, for any $\hat{c} \in C$ and $\hat{p} \in \mathbb{R}_+$ let

$$\mu^p_{\hat{c},l}(\cdot) := \mathbb{P}[v \in \cdot \mid p = \hat{p}, \theta = l, c = \hat{c}] \qquad \text{and} \qquad \mu^p_{\hat{c},h}(\cdot) := \mathbb{P}[v \in \cdot \mid p = \hat{p}, \theta = h, c = \hat{c}].$$

By the definition of $\rho_{\hat{c}}$, for all $\hat{c} \in C$, we have

$$\begin{aligned}
&\int_{V^2} \max_{\tilde{p}} (\tilde{p} - \hat{c}) \left[(1 - \alpha_{\hat{c}}) \mathbf{1}\{v_l \geq \tilde{p}\} + \alpha_{\hat{c}} \mathbf{1}\{v_h \geq \tilde{p}\} \right] \rho_{\hat{c}}(\mathrm{d}v_l, \mathrm{d}v_h) \\
=&\mathbb{E}\left[\int_{V^2} \max_{\tilde{p}} (\tilde{p} - \hat{c}) \left[(1 - \alpha_{\hat{c}}) \mathbf{1}\{v_l \geq \tilde{p}\} + \alpha_{\hat{c}} \mathbf{1}\{v_h \geq \tilde{p}\} \right] \mu^p_{\hat{c},l}(\mathrm{d}v_l) \mu^p_{\hat{c},h}(\mathrm{d}v_h) \mid c = \hat{c} \right] \\
\geq&\mathbb{E}\left[\max_{\tilde{p}} \int_{V^2} (\tilde{p} - \hat{c}) \left[(1 - \alpha_{\hat{c}}) \mathbf{1}\{v_l \geq \tilde{p}\} + \alpha_{\hat{c}} \mathbf{1}\{v_h \geq \tilde{p}\} \right] \mu^p_{\hat{c},l}(\mathrm{d}v_l) \mu^p_{\hat{c},h}(\mathrm{d}v_h) \mid c = \hat{c} \right] \\
=&\mathbb{E}\left[\max_{\tilde{p}} \left(\mathbb{E}\left[(\tilde{p} - c)(1 - \alpha_c) \mathbf{1}\{v \geq \tilde{p}\} \mid p, \theta = l, c = \hat{c} \right] + \mathbb{E}\left[(\tilde{p} - c) \alpha_c \mathbf{1}\{v \geq \tilde{p}\} \mid p, \theta = h, c = \hat{c} \right] \right) \mid c = \hat{c} \right] \\
=&\mathbb{E}\left[\max_{\tilde{p}} \mathbb{E}\left[(\tilde{p} - c) \mathbf{1}\{v \geq \tilde{p}\} \mid p, c = \hat{c} \right] \mid c = \hat{c} \right] \\
\geq&\mathbb{E}\left[\mathbb{E}\left[(p - c) \mathbf{1}\{v \geq p\} \mid p, c = \hat{c} \right] \mid c = \hat{c} \right] \\
=&\mathbb{E}\left[(p - c) \mathbf{1}\{v \geq p\} \mid c = \hat{c} \right],
\end{aligned} \tag{A.1}$$

Therefore,

$$\Pi(p) = \mathbb{E}[(p-c)\mathbf{1}\{v \geq p\}] \leq \int_C \left(\int_{V^2} \pi_c(v_l, v_h)\rho_c(\mathrm{d}v_l, \mathrm{d}v_h) \right) G(\mathrm{d}c)\,,$$

as desired.

Next, we show that

$$\sup_{p \in \mathcal{D}} \Pi(p) \geq \int_C \max_{\rho_c \in \mathcal{R}_c} \left(\int_{V^2} \pi_c(v_l, v_h)\rho_c(\mathrm{d}v_l, \mathrm{d}v_h) \right) G(\mathrm{d}c)\,.$$

To this end, we show that any $(\rho_c)_{c\in C}$ such that $\rho_c \in \mathcal{R}_c$ for all $c \in C$, we can construct a non-discriminatory pricing rule p such that

$$\Pi(p) = \int_C \left(\int_{V^2} \pi_c(v_l, v_h)\rho_c(\mathrm{d}v_l, \mathrm{d}v_h) \right) G(\mathrm{d}c)\,.$$

Indeed, consider any $(\rho_c)_{c\in C}$ such that $\rho_c \in \mathcal{R}_c$ for all $c \in C$. Since $V \subseteq \mathbb{R}_+$ is a standard Borel space, by the disintegration theorem (see, e.g., Çinlar 2010, Theorem 2.17, pp. 151), for each $c \in C$, there exists transition probabilities $\gamma_{c,l} : V \to \Delta(V)$ and $\gamma_{c,h} : V \to \Delta(V)$ such that for all measurable $V_l, V_h \subseteq V$

$$\int_{V_h} \gamma_{c,l}(V_l \mid v_h) F_{c,h}(\mathrm{d}v_h) = \rho_c(V_l \times V_h) = \int_{V_l} \gamma_{c,h}(V_h \mid v_l) F_{c,l}(\mathrm{d}v_l)\,. \tag{A.2}$$

Let $\Gamma_{c,\theta}(\cdot \mid v)$ be the CDF associated with $\gamma_{c,\theta}(\cdot \mid v)$, for all $v \in V$, $c \in C$ and $\theta \in \{l, h\}$. In the meantime, let $\xi_c : V^2 \to \mathbb{R}_+$ be a measurable selection of

$$\operatorname*{argmax}_{\tilde{p} \geq 0} (\tilde{p} - c)(\alpha_c \mathbf{1}\{v_h \geq \tilde{p}\} + (1-\alpha_c)\mathbf{1}\{v_l \geq \tilde{p}\})\,.$$

Then, let

$$p(v, c, \theta, r) := \begin{cases} \xi_c(v, \Gamma_{c,h}^{-1}(r \mid v)), & \text{if } \theta = l \\ \xi_c(\Gamma_{c,l}^{-1}(r \mid v), v), & \text{if } \theta = h \end{cases}, \tag{A.3}$$

for all $(v, c, \theta, r) \in \Omega$. By construction,

$$
\begin{aligned}
\Pi(p) =& \mathbb{E}[(p-c)\mathbf{1}\{v \geq p\}] \\
=& \mathbb{E}[\mathbb{E}[\mathbb{E}[(p-c)\mathbf{1}\{v \geq p\} \mid c, \theta] \mid c]] \\
=& \int_C \Bigg(\alpha_c \int_{V \times [0,1]} (\xi_c(\Gamma_{c,l}^{-1}(r \mid v_h), v_h) - c)\mathbf{1}\{v_h \geq \xi_c(\Gamma_{c,l}^{-1}(r \mid v_h), v_h)\} \, \mathrm{d}r \, F_{c,h}(\mathrm{d}v_h) \\
& + (1-\alpha_c) \int_{V \times [0,1]} (\xi_c(v_l, \Gamma_{c,h}^{-1}(r \mid v_l)) - c)\mathbf{1}\{v_l \geq \xi_c(v_l, \Gamma_{c,h}^{-1}(r \mid v_l))\} \, \mathrm{d}r \, F_{c,l}(\mathrm{d}v_l) \Bigg) G(\mathrm{d}c) \\
=& \int_C \Bigg(\alpha_c \int_{V^2} (\xi_c(v_l, v_h) - c)\mathbf{1}\{v_h \geq \xi_c(v_l, v_h)\} \gamma_{c,l}(\mathrm{d}v_l \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
& + (1-\alpha_c) \int_{V^2} (\xi_c(v_l, v_h) - c)\mathbf{1}\{v_l \geq \xi_c(v_l, v_h)\} \gamma_{c,h}(\mathrm{d}v_h \mid v_l) F_{c,l}(\mathrm{d}v_l) \Bigg) G(\mathrm{d}c) \\
=& \int_C \left(\int_{V^2} (\xi_c(v_l, v_h) - c)(\alpha_c \mathbf{1}\{v_h \geq \xi_c(v_l, v_h)\} + (1-\alpha_c)\mathbf{1}\{v_l \geq \xi_c(v_l, v_h)\}) \rho_c(\mathrm{d}v_l, \mathrm{d}v_h) \right) G(\mathrm{d}c) \\
=& \int_C \left(\int_{V^2} \pi_c(v_l, v_h) \rho_c(\mathrm{d}v_l, \mathrm{d}v_h) \right) G(\mathrm{d}c) \,,
\end{aligned}
\tag{A.4}
$$

where the second equality follows from the law of iterated expectation, the fourth equality follows from changing variables of the integration, the fifth equality follows from (A.2), and the last equality follows from the definition of ξ_c.

Moreover, for any $c \in C$ and for any measurable $M \subseteq V$,

$$
\begin{aligned}
\mathbb{P}[p \in M \mid c, \theta = l] =& \mathbb{P}[\xi_c(v, \Gamma_{c,h}^{-1}(r \mid v)) \in M \mid c, \theta = l] \\
=& \int_{V \times [0,1]} \mathbf{1}\{\xi_c(v_l, \Gamma_{c,h}^{-1}(r \mid v_l)) \in M\} \, \mathrm{d}r \, \mathrm{d}F_{c,l}(\mathrm{d}v_l) \\
=& \int_{V^2} \mathbf{1}\{\xi_c(v_l, v_h) \in M\} \gamma_{c,h}(\mathrm{d}v_h \mid v_l) F_{c,l}(\mathrm{d}v_l) \\
=& \int_{V^2} \mathbf{1}\{\xi_c(v_l, v_h) \in M\} \rho_c(\mathrm{d}v_l, \mathrm{d}v_h) \,,
\end{aligned}
$$

where the third equality again follows from changing variables of the integration, and the last equality follows from (A.2). Likewise,

$$\begin{aligned}\mathbb{P}[p \in M \mid c, \theta = h] &= \mathbb{P}[\xi_c(\Gamma_{c,l}^{-1}(r \mid v), v) \in M \mid c, \theta = h] \\ &= \int_{V \times [0,1]} \mathbf{1}\{\xi_c(\Gamma_{c,l}^{-1}(r \mid v_h), v_h) \in M\} \, \mathrm{d}r \, \mathrm{d}F_{c,h}(\mathrm{d}v_h) \\ &= \int_{V^2} \mathbf{1}\{\xi_c(v_l, v_h) \in M\} \gamma_{c,l}(\mathrm{d}v_l \mid v_h) F_{c,h}(\mathrm{d}v_h) \\ &= \int_{V^2} \mathbf{1}\{\xi_c(v_l, v_h) \in M\} \rho_c(\mathrm{d}v_l, \mathrm{d}v_h) \,.\end{aligned}$$

As a result, for all $c \in C$, and for all measurable $M \subseteq \mathbb{R}_+$,

$$\mathbb{P}[p \in M \mid c, \theta = l] = \int_{V^2} \mathbf{1}\{\xi_c(v_l, v_h) \in M\} \rho_c(\mathrm{d}v_l, \mathrm{d}v_h) = \mathbb{P}[p \in M \mid c, \theta = h] \,,$$

and thus p is indeed non-discriminatory.

Together, we have

$$\sup_{p \in \mathcal{D}} \Pi(p) = \int_C \left(\max_{\rho_c \in \mathcal{R}_c} \int_{V^2} \pi_c(v_l, v_h) \rho_c(\mathrm{d}v_l, \mathrm{d}v_h) \right) G(\mathrm{d}c) \,. \tag{A.5}$$

Furthermore, for any profit-maximizing non-discriminatory pricing rule p, let ρ_c be defined by (2), (A.1) and (A.5) then implies that $(\rho_c)_{c \in C}$ solves (4). Conversely, for any $(\rho_c)_{c \in C}$ that solves (4), let p be defined by (A.3). Then p solves (1) by (A.4) and (A.5). This completes the proof. □

Proof of Lemma 1. Since $F_{c,h}$ dominates $F_{c,l}$ in the likelihood ratio order, Δ_c is quasi-concave and has a unique maximizer for all c. Let $v_c^\star$ be the unique maximizer. Since $F_{c,l}$ and $F_{c,h}$ are CDFs on $\mathbb{R}_+$ that are absolutely continuous, $\Delta_c(0) = 0$ and $\lim_{v \to \infty} \Delta_c(v) = 0$, and

$$\|F_{c,l} - F_{c,h}\| = \max_{A \subseteq \mathbb{R}_+} \left| \int_A [f_{c,l}(v) - f_{c,h}(v)] \, \mathrm{d}v \right| = \max_v \Delta_c(v) = \Delta_c(v_c^\star) \,.$$

Moreover, since Δ_c is continuous and quasi-concave, for any $v \geq v_c^\star$, there exists a unique $g_c(v) \in [\underline{v}_c, v_c^\star]$ such that $\Delta_c(v) = \Delta_c(g_c(v))$. Moreover, the function $g_c : [v_c^\star, \infty) \to [\underline{v}_c, v_c^\star]$ is continuous and decreasing in v, with $g_c(v_c^\star) = v_c^\star$ and $\lim_{v \to \infty} g_c(v) = \underline{v}_c$. For any $v \geq v_c^\star$, let

$$h_c(v) := v - g_c(v) \,.$$

Note that h_c is increasing on $[v_c^\star, \overline{v}_c]$ and $h_c(v_c^\star) = 0$, $\lim_{v \to \infty} h_c(v) = \infty$. In particular, since

$F_{c,l}(c) < \|F_{c,l} - F_{c,h}\| = \Delta_c(v_c^\star)$, and thus $c < v_c^\star$, there exists a unique $\tilde{v}_c > v_c^\star$ such that $\alpha_c h_c(\tilde{v}_c) = (1-\alpha_c)(v_c^\star - c)$. Meanwhile let $\hat{v}_c := \inf\{v \geq v_c^\star : \alpha_c h_c(v) \geq (1-\alpha_c)(\underline{v}_c - c)\}$. Since h_c is nondecreasing, it must be that $\hat{v}_c \in [v_c^\star, \tilde{v}_c]$.

Note that, if $\hat{v}_c = v_c^\star$, then it must be that

$$\Delta_c\left(\frac{\alpha_c}{1-\alpha_c}h_c(\hat{v}_c) + c\right) + F_{c,h}(\alpha_c h_c(\hat{v}_c) + c) = F_{c,l}(c) < \|F_{c,l} - F_{c,h}\| = \Delta_c(v_c^\star) = \Delta_c(\hat{v}_c)\,;$$

If $\hat{v}_c \in (v_c^\star, \overline{v}_c)$, then it must be that $0 \leq \alpha_c h_c(\hat{v}_c) = (1-\alpha_c)(\underline{v}_c - c)$, and thus $\alpha_c h_c(\hat{v}_c) + c = (1-\alpha_c)\underline{v}_c + \alpha_c c \leq \underline{v}_c$. Therefore,

$$\Delta_c\left(\frac{\alpha_c}{1-\alpha_c}h_c(\hat{v}_c) + c\right) + F_{c,h}(\alpha_c h_c(\hat{v}_c) + c) = 0 < \Delta_c(\hat{v}_c)\,.$$

If $\hat{v}_c \geq \overline{v}_c$, then

$$\Delta_c\left(\frac{\alpha_c}{1-\alpha_c}h_c(\hat{v}_c) + c\right) + F_{c,h}(\alpha_c h_c(\hat{v}_c) + c) = 0 = \Delta_c(\hat{v}_c)\,.$$

Since Δ_c is quasi-concave, and hence is decreasing on $[v_c^\star, \overline{v}_c]$ while $\Delta_c(v) = 0$ for all $v \geq \overline{v}_c$. In the meantime, since $\alpha_c h_c(\tilde{v}_c)/(1-\alpha_c) + c = v_c^\star$ by definition,

$$\Delta_c\left(\frac{\alpha_c}{1-\alpha_c}h_c(\tilde{v}_c) + c\right) + F_{c,h}(\alpha_c h_c(\tilde{v}_c) + c) = \Delta_c(v_c^\star) + F_{c,h}(\alpha_c h_c(\tilde{v}_c) + c) \geq \Delta_c(v_c^\star) > \Delta_c(\tilde{v}_c)\,.$$

Together, since $\alpha_c h_c(v)/(1-\alpha_c) + c \leq v_c^\star$ for all $v \in [v_c^\star, \tilde{v}_c]$ and since h_c is increasing in v, the function

$$v \mapsto \Delta_c\left(\frac{\alpha_c}{1-\alpha_c}h_c(v) + c\right) + F_{c,h}(\alpha_c h_c(v) + c)$$

is increasing on $[v_c^\star, \tilde{v}_c]$. Together, there exists a unique $\kappa_c^5 \in [\hat{v}_c, \tilde{v}_c]$ such that

$$\Delta_c(\kappa_c^5) = \Delta_c\left(\frac{\alpha_c}{1-\alpha_c}h_c(\kappa_c^5) + c\right) + F_{c,h}(\alpha_c h_c(\kappa_c^5) + c)\,, \tag{A.6}$$

and that

$$\frac{\alpha_c}{1-\alpha_c}h_c(\kappa_c^5) \leq \frac{\alpha_c}{1-\alpha_c}h_c(\tilde{v}_c) = v_c^\star - c\,.$$

Let $\kappa_c^4 := g_c(\kappa_c^5)$ and $\kappa_c^1 := \alpha_c h_c(\kappa_c^5) + c$, $\kappa_c^3 := \alpha_c h_c(\kappa_c^5)/(1-\alpha_c) + c$, and $\kappa_c^2 := F_{c,l}^{-1}(\Delta_c(\kappa_c^5))$. By construction, $\kappa_c^5 \geq \hat{v}_c \geq v_c^\star > \kappa_c^4$, with at least one of the first two inequalities being strict, and $\kappa_c^3 \geq \kappa_c^1$. Moreover, since $\kappa_c^3 = \alpha_c h_c(\kappa_c^5)/(1-\alpha_c) + c \leq v_c^\star$ and since $\Delta_c(\kappa_c^3) \leq \Delta_c(\kappa_c^4) < \Delta_c(v_c^\star)$,

$\kappa_c^3 \leq \kappa_c^4$. In the meantime, since $F_{c,l}(\kappa_c^2) = \Delta_c(\kappa_c^3) + F_{c,h}(\kappa_c^1)$,

$$F_{c,l}(\kappa_c^3) - F_{c,l}(\kappa_c^2) = F_{c,h}(\kappa_c^3) - F_{c,h}(\kappa_c^1) \geq 0\,,$$

and hence $\kappa_c^2 \leq \kappa_c^3$. Lastly, since $\kappa_c^1 \leq \kappa_c^3 \leq v_c^\star$, it must be that $\Delta_c(\kappa_c^1) \leq \Delta_c(\kappa_c^3)$. Therefore,

$$F_{c,l}(\kappa_c^2) - F_{c,h}(\kappa_c^1) = \Delta_c(\kappa_c^3) \geq \Delta_c(\kappa_c^1) = F_{c,l}(\kappa_c^1) - F_{c,h}(\kappa_c^1)\,,$$

and hence $F_{c,l}(\kappa_c^2) \geq F_{c,l}(\kappa_c^1)$, which in turn implies $\kappa_c^2 \geq \kappa_c^1$. Together, it then follows that $\kappa_c^1 \leq \kappa_c^2 \leq \kappa_c^3 \leq \kappa_c^4 < v_c^\star < \kappa_c^5$.

For uniqueness, consider any increasing vector $\tilde{\kappa}_c = (\tilde{\kappa}_c^1, \tilde{\kappa}_c^2, \tilde{\kappa}_c^3, \tilde{\kappa}_c^4, \tilde{\kappa}_c^5)$ that solves (8) such that $\tilde{\kappa}_c^4 < v_c^\star < \tilde{\kappa}_c^5$. From (8), we have $\tilde{\kappa}_c^2 \geq \underline{v}_c$, and $\Delta_c(\tilde{\kappa}_c^4) = \Delta_c(\tilde{\kappa}_c^5)$, and hence $\tilde{\kappa}_c^4 = g_c(\tilde{\kappa}_c^5)$ and $h_c(\tilde{\kappa}_c^5) = \tilde{\kappa}_c^5 - \tilde{\kappa}_c^4$. The latter implies that $\tilde{\kappa}_c^3 = \alpha_c h_c(\tilde{\kappa}_c^5)/(1-\alpha_c) + c$. Together with (8), $\tilde{\kappa}_c^5$ must solve (A.6). As a result, since $\underline{v}_c \leq \tilde{\kappa}_c^2 \leq \tilde{\kappa}_c^3 \leq \tilde{\kappa}_c^4 < v_c^\star$, and since $\tilde{\kappa}_c^5 > v_c^\star$ and h_c is increasing on $[v_c^\star, \overline{v}_c]$, it follows that $\tilde{\kappa}_c^5 \in [\hat{v}_c, \tilde{v}_c]$. Therefore, since (A.6) has a unique solution on $[\hat{v}_c, \tilde{v}_c]$, we have $\tilde{\kappa}_c^5 = \kappa_c^5$.

Meanwhile, since $\tilde{\kappa}_c$ solves (8), it must be that $\tilde{\kappa}_c = \kappa_c$. Thus, κ_c is the unique increasing vector in $\mathbb{R}^5$ with $\kappa_c^4 < v_c^\star < \kappa_c^5$ that solves (8). ☐

Proof of Theorem 2. Note that Theorem 2 immediately implies Theorem 1, and therefore we prove Theorem 2 directly. For any $c \in C$, if $F_{c,l}(c) \geq \|F_{c,l} - F_{c,h}\|$, let

$$\phi_c^\star(v_l) := (1-\alpha_c)\cdot(v_l - c)^+ \quad \text{and} \quad \psi_c^\star(v_h) := \alpha_c \cdot (v_h - c)^+\,,$$

for all $(v_l, v_h) \in V \times V$. Meanwhile, if $F_{c,l}(c) < \|F_{c,l} - F_{c,h}\|$, let

$$\phi_c^\star(v_l) := \begin{cases} \kappa_c^1 - c, & \text{if } v_l \leq \kappa_c^3 \\ (1-\alpha_c)\cdot(v_l - c), & \text{if } v_l \in (\kappa_c^3, \kappa_c^4] \\ v_l - c - \alpha_c \cdot (\kappa_c^4 - c), & \text{if } v_l \in (\kappa_c^4, \kappa_c^5] \\ (1-\alpha_c)(v_l - c) + \kappa_c^1 - c, & \text{if } v_l > \kappa_c^5 \end{cases},$$

and let

$$\psi_c^\star(v_h) := \begin{cases} 0, & \text{if } v_h \le \kappa_c^1 \\ v_h - \kappa_c^1, & \text{if } v_h \in (\kappa_c^1, \kappa_c^3] \\ \alpha_c \cdot (v_h - c), & \text{if } v_h \in (\kappa_c^3, \kappa_c^4] \\ \alpha_c \cdot (\kappa_c^4 - c), & \text{if } v_h \in (\kappa_c^4, \kappa_c^5] \\ \alpha_c(v_h - c) - (\kappa_c^1 - c), & \text{if } v_h > \kappa_c^5 \end{cases}.$$

Lemma A.1. *For any $(v_l, v_h) \in V \times V$ and for any $c \in C$,*

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) \ge \pi_c(v_l, v_h).$$

The proof of Lemma A.1 is by inspection, using the system of equation (8) that defines κ_c. Details of the proof can be found in Section A.2. Next, we define a joint distribution $\rho_c^\star \in \Delta(V^2)$. When $F_{c,l}(c) < \|F_{c,l} - F_{c,h}\|$, define a transition probability $\gamma_c^\star : V \to \Delta(V)$ as follows:

$$\gamma_c^\star(v_l \le x \mid v_h) :=$$
$$\begin{cases} \mathbf{1}\{\underline{\Delta}_c^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^3)) \le x\}, & \text{if } v_h \le \kappa_c^1 \\ \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^3)) \le x\}, & \text{if } v_h \in (\kappa_c^1, \kappa_c^3] \\ \mathbf{1}\{v_h \le x\}, & \text{if } v_h \in (\kappa_c^3, \kappa_c^4] \\ \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^4)) \le x\}, & \text{if } v_h \in (\kappa_c^4, \kappa_c^5] \\ \frac{f_{c,l}(v_h)}{f_{c,h}(v_h)} \cdot \mathbf{1}\{v_h \le x\} + \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} \cdot \mathbf{1}\{F_{c,l}^{-1}(\Delta_c(\kappa_c^5) - \Delta_c(v_h)) \le x\}, & \text{if } v_h > \kappa_c^5, \end{cases},$$

for all $x \in V$ and for all $v_h \in V$. Meanwhile, when $F_{c,l}(c) \ge \|F_{c,l} - F_{c,h}\|$ and $c < v_c^\star$, define a transition probability $\gamma_c^\star : V \to \Delta(V)$ as:

$$\gamma_c^\star(v_l \le x \mid v_h) :=$$
$$\begin{cases} \mathbf{1}\left\{F_{c,l}^{-1}(F_{c,h}(v_h)) \le x\right\}, & \text{if } v_h \le \eta_c^h \\ \mathbf{1}\left\{\underline{\Delta}_c^{-1}(F_{c,h}(v_h) - F_{c,h}(\eta_c^h) + \Delta_c(c)) \le x\right\}, & \text{if } v_h \in (\eta_c^h, c] \\ \mathbf{1}\{v_h \le x\}, & \text{if } v_h \in (c, v_c^\star] \\ \frac{f_{c,l}(v_h)}{f_{c,h}(v_h)}\mathbf{1}\{v_h \le x\} + \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)}\mathbf{1}\{F_{c,l}^{-1}(\Delta_c(v_c^\star) - \Delta_c(v_h) + F_{c,l}(\eta_c^l)) \le x\}, & \text{if } v_h > v_c^\star \end{cases},$$

for all $x \in V$ and for all $v_h \in V$. When $F_{c,l}(c) \geq \|F_{c,l} - F_{c,h}\|$ and $c \geq v_c^\star$, define $\gamma_c^\star$ as:

$$\gamma_c^\star(v_l \leq x \mid v_h) :=$$
$$\begin{cases} \mathbf{1}\left\{F_{c,l}^{-1}(F_{c,h}(v_h)) \leq x\right\}, & \text{if } v_h \leq c \\ \frac{f_{c,l}(v_h)}{f_{c,h}(v_h)}\mathbf{1}\{v_h \leq x\} + \frac{f_{c,h}(v_h)-f_{c,l}(v_h)}{f_{c,h}(v_h)}\mathbf{1}\{F_{c,l}^{-1}(F_{c,l}(c) - \Delta_c(v_h)) \leq x\}, & \text{if } v_h > c \end{cases},$$

for all $x \in V$ and for all $v_h \in V$.

Then, for all $c \in C$, let $\rho_c^\star \in \Delta(V \times V)$ be defined as

$$\rho_c^\star(v_l \in A, v_h \in B) := \int_B \gamma_c^\star(A \mid v_h) F_{c,h}(\mathrm{d}v_h)\,, \tag{A.7}$$

for all measurable sets $A, B \subseteq V$. By construction, the marginals of $\rho_c^\star$ are exactly $F_{c,l}$ and $F_{c,h}$. That is,

Lemma A.2. *$\rho_c^\star \in \mathcal{R}_c$ for all $c \in C$.*

Combining Lemma 2, Lemma A.1 and Lemma A.2 with Lemma A.3 below, it then follows that $\rho_c^\star$ is a solution of (10).

Lemma A.3. *For any $c \in C$, $\phi_c^\star(v_l) + \psi_c^\star(v_h) = \pi_c(v_l, v_h)$ for all $(v_l, v_h) \in \operatorname{supp}(\rho_c^\star)$.*

Since $\rho_c^\star$ is a solution of (10) for all c, Proposition 1 implies that one can construct an optimal non-discriminatory pricing rule from $(\rho_c^\star)_{c\in C}$. To this end, for any $c \in C$, let $\beta_c^\star : V \to \Delta(V)$ be the conditional distribution of v_h given v_l implied by $\rho_c^\star$. That is, $\beta_c^\star$ is a version of the regular conditional probability defined by

$$\rho_c^\star(v_l \in A, v_h \in B) = \int_A \beta_c^\star(v_h \in B \mid v_l) F_{c,l}(\mathrm{d}v_l)\,, \tag{A.8}$$

for all measurable $A, B \subseteq V$. Next, let γ_c^{-1} and β_c^{-1} be the quantile function defined by $\gamma_c^\star$ and $\beta_c^\star$, respectively. That is,

$$\gamma_c^{-1}(r \mid v_h) := \inf\{v_l \in V : \gamma_c^\star([0, v_l] \mid v_h) \geq r\} \text{ and } \beta_c^{-1}(r \mid v_l) := \inf\{v_h \in V : \beta_c^\star([0, v_h] \mid v_l) \geq r\}$$

for all $r \in [0,1]$ and for all $(v_l, v_h) \in V^2$. Meanwhile, for any $(v_l, v_h) \in V^2$, let $\xi_c(v_l, v_h)$ be the minimum element of

$$\underset{\tilde{p}\geq 0}{\operatorname{argmax}}(\tilde{p}-c)(\alpha_c \mathbf{1}\{v_h \geq \tilde{p}\} + (1-\alpha_c)\mathbf{1}\{v_l \geq \tilde{p}\}) .$$

It then follows that $p^\star$ can be written as

$$p^\star(v,c,l,r) = \xi_c(v, \beta_c^{-1}(r \mid v)) \text{ and } p^\star(v,c,h,r) = \xi_c(\gamma_c^{-1}(r \mid v), v)$$

for all $v \in V$, $c \in C$ and $r \in [0,1]$. By construction,

$$\begin{aligned}
\mathbb{P}[p^\star \in M \mid c, \theta = l] =& \mathbb{P}[\xi_c(v, \beta_c^{-1}(r \mid v)) \in M \mid \theta = l, c] \\
=& \int_{V\times[0,1]} \mathbf{1}\{\xi_c(v_l, \beta_c^{-1}(r \mid v_l)) \in M\}\, \mathrm{d}r\, \mathrm{d}F_{c,l}(\mathrm{d}v_l) \\
=& \int_{V^2} \mathbf{1}\{\xi_c(v_l, v_h) \in M\} \beta_c^\star(\mathrm{d}v_h \mid v_l) F_{c,l}(\mathrm{d}v_l) \\
=& \int_{V^2} \mathbf{1}\{\xi_c(v_l, v_h) \in M\} \rho_c^\star(\mathrm{d}v_l, \mathrm{d}v_h) ,
\end{aligned}$$

where the third equality again follows from changing variables of the integration, and the last equality follows from (A.7) and (A.8). Likewise,

$$\begin{aligned}
\mathbb{P}[p^\star \in M \mid c, \theta = h] =& \mathbb{P}[\xi_c(\gamma_c^{-1}(r \mid v), v) \in M \mid \theta = h, c] \\
=& \int_{V\times[0,1]} \mathbf{1}\{\xi_c(\gamma_c^{-1}(r \mid v_h), v_h) \in M\}\, \mathrm{d}r\, \mathrm{d}F_{c,h}(\mathrm{d}v_h) \\
=& \int_{V^2} \mathbf{1}\{\xi_c(v_l, v_h) \in M\} \gamma_c^\star(\mathrm{d}v_l \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
=& \int_{V^2} \mathbf{1}\{\xi_c(v_l, v_h) \in M\} \rho_c^\star(\mathrm{d}v_l, \mathrm{d}v_h) .
\end{aligned}$$

As a result, for all $c \in C$, and for all measurable $M \subseteq \mathbb{R}_+$,

$$\mathbb{P}[p^\star \in M \mid c, \theta = l] = \int_{V^2} \mathbf{1}\{\xi_c(v_l, v_h) \in M\} \rho_c^\star(\mathrm{d}v_l, \mathrm{d}v_h) = \mathbb{P}[p^\star \in M \mid c, \theta = h] ,$$

and thus $p^\star$ is indeed non-discriminatory. Moreover,

$$
\begin{aligned}
\Pi(p^\star) =& \mathbb{E}[(p^\star - c)\mathbf{1}\{v \geq p^\star\}] \\
=& \mathbb{E}[\mathbb{E}[\mathbb{E}[(p^\star - c)\mathbf{1}\{v \geq p^\star\} \mid c, \theta] \mid c]] \\
=& \int_C \Bigg(\alpha_c \int_{V\times[0,1]} (\xi_c(\gamma_c^{-1}(r \mid v_h), v_h) - c)\mathbf{1}\{v_h \geq \xi_c(\gamma_c^{-1}(r \mid v_h), v_h)\} \, \mathrm{d}r \, F_{c,h}(\mathrm{d}v_h) \\
& + (1-\alpha_c) \int_{V\times[0,1]} (\xi_c(v_l, \beta_c^{-1}(r \mid v_l)) - c)\mathbf{1}\{v_l \geq \xi_c(v_l, \beta_c^{-1}(r \mid v_l))\} \, \mathrm{d}r \, F_{c,l}(\mathrm{d}v_l) \Bigg) G(\mathrm{d}c) \\
=& \int_C \Bigg(\alpha_c \int_{V^2} (\xi_c(v_l, v_h) - c)\mathbf{1}\{v_h \geq \xi_c(v_l, v_h)\} \gamma_c^\star(\mathrm{d}v_l \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
& + (1-\alpha_c) \int_{V^2} (\xi_c(v_l, v_h) - c)\mathbf{1}\{v_l \geq \xi_c(v_l, v_h)\} \beta_c^\star(\mathrm{d}v_h \mid v_l) F_{c,l}(\mathrm{d}v_l) \Bigg) G(\mathrm{d}c) \\
=& \int_C \left(\int_{V^2} (\xi_c(v_l, v_h) - c)(\alpha_c \mathbf{1}\{v_h \geq \xi_c(v_l, v_h)\} + (1-\alpha_c)\mathbf{1}\{v_l \geq \xi_c(v_l, v_h)\}) \rho_c^\star(\mathrm{d}v_l, \mathrm{d}v_h) \right) G(\mathrm{d}c) \\
=& \int_C \left(\int_{V^2} \pi_c(v_l, v_h) \rho_c^\star(\mathrm{d}v_l, \mathrm{d}v_h) \right) G(\mathrm{d}c) \, ,
\end{aligned}
$$

where the second equality follows from the law of iterated expectation, the fourth equality follows from changing variables of the integration, the fifth equality follows from (A.7) and (A.8), and the last equality follows from the definition of ξ_c. Thus, by Proposition 1, $p^\star$ is indeed an optimal non-discriminatory pricing rule. This completes the proof of (i).

Part (ii) follows from Proposition 8.[34] Indeed, Proposition 8 implies that any profit-maximizing non-discriminatory pricing rule p satisfies

$$
CS(c, h; p) = CS(c, h; p^\star) \quad \text{and} \quad 0 \leq CS(c, l; p) \leq CS(c, l; p^\star)
$$

for all $c \in C$. If $CS(c, l; p) < CS(c, l; p^\star)$ for some c, then p is dominated by $p^\star$, since both rules give the seller the same profit and the same surplus to h-consumers, while $p^\star$ gives strictly higher surplus to l-consumers. Therefore every undominated profit-maximizing non-discriminatory pricing rule must satisfy

$$
CS(c, \theta; p) = CS(c, \theta; p^\star)
$$

for all $c \in C$ and $\theta \in \{l, h\}$. $\square$

[34] Note that the proof of Proposition 8 depends on Proposition 1 and Lemma 2 but not on (ii) of Theorem 1.

Proof of Corollary 1. For the "if" part, suppose that

$$\underline{v}_c - c \geq \alpha_c(\overline{v}_c - c)$$

for almost all c. Then, by the proof of Proposition 4, $\kappa_c^1 \leq \underline{v}_c = \kappa_c^2 = \kappa_c^3 = \kappa_c^4$ and $\kappa_c^5 = \overline{v}_c$ for all c. Thus, $\Pi(p^\star) = \Pi(p^{ass})$. By Theorem 1, p^{ass} is optimal.

For the "only if" part, suppose that

$$\underline{v}_c - c < \alpha_c(\overline{v}_c - c)$$

for a positive measure of c. By the proof of Proposition 4, for any such c, $F_{c,l}(\kappa_c^3) > 0$, and thus $\kappa_c^3 > \underline{v}_c$. Therefore, $F_{c,h}(\kappa_c^4) \geq F_{c,h}(\kappa_c^3) > 0$ as well, which in turn implies that $F_{c,h}(\kappa_c^5) < 1$ and $\Delta_c(\kappa_c^4) > 0$, for all such c. In the meantime, under the assortative pricing rule p^{ass},

$$CS(c, h; p^{ass}) = \int_0^1 [F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q)]\, \mathrm{d}q\,,$$

whereas

$$CS(c, h; p^\star) = \int_{F_{c,h}(\kappa_c^4)}^{F_{c,h}(\kappa_c^5)} [F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q + \Delta_c(\kappa_c^4))]\, \mathrm{d}q < \int_0^1 [F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q)]\, \mathrm{d}q\,.$$

Thus, by Proposition 8, p^{ass} is not optimal. $\square$

Proof of Proposition 2. Recall that the assortative pricing rule p^{ass}, defined by (6) gives a profit of $\mathbb{E}[(v-c)^+ \mid c = \tilde{c}, \theta = l]$ for all $\tilde{c} \in C$. Since p^{ass} is non-discriminatory, and since

$$\mathbb{E}[(v-c)^+ \mid c = \tilde{c}] = \alpha_{\tilde{c}}\mathbb{E}[(v-c)^+ \mid c = \tilde{c}, \theta = h] + (1 - \alpha_{\tilde{c}})\mathbb{E}[(v-c)^+ \mid c = \tilde{c}, \theta = l]$$

we have

$$\frac{\Pi_{\tilde{c}}^\star}{\mathbb{E}[(v-c)^+ \mid c = \tilde{c}]} \geq \frac{\mathbb{E}[(v-c)^+ \mid c = \tilde{c}, \theta = l]}{\mathbb{E}[(v-c)^+ \mid c = \tilde{c}]} \geq \frac{1}{\alpha_{\tilde{c}} r_{\tilde{c}} + 1}\,, \tag{A.9}$$

for all $\tilde{c} \in C$. In the meantime, since the partly anti-assortative pricing rule p^{anti} defined by Definition 3, with $q_c = 1$ for all c, gives a profit of $\mathbb{E}[\alpha_c(v-c)^+ \mid c = \tilde{c}, \theta = h]$ conditional on every $\tilde{c} \in C$, and is also non-discriminatory, we have

$$\frac{\Pi_{\tilde{c}}^\star}{\mathbb{E}[(v-c)^+ \mid c = \tilde{c}]} \geq \frac{\alpha_{\tilde{c}}\mathbb{E}[(v-c)^+ \mid c = \tilde{c}, \theta = h]}{\mathbb{E}[(v-c)^+ \mid c = \tilde{c}]} = \frac{\alpha_{\tilde{c}}(r_{\tilde{c}} + 1)}{\alpha_{\tilde{c}} r_{\tilde{c}} + 1}\,, \tag{A.10}$$

for all $\tilde{c} \in C$. Since the right-hand side of (A.9) is decreasing in $\alpha_{\tilde{c}}$ and the right-hand side of (A.10) is increasing in $\alpha_{\tilde{c}}$, the maximum of the two is minimized at the unique crossing point. Holding $\alpha_{\tilde{c}}$ fixed and minimizing over $r_{\tilde{c}}$, the two right-hand sides coincide at $r_{\tilde{c}} = (1-\alpha_{\tilde{c}})/\alpha_{\tilde{c}}$ with common value $1/(2-\alpha_{\tilde{c}})$, so

$$\frac{\Pi^\star_{\tilde{c}}}{\mathbb{E}\left[(v-c)^+ \mid c=\tilde{c}\right]} \geq \max\left\{\frac{1}{\alpha_{\tilde{c}} r_{\tilde{c}}+1}, \frac{\alpha_{\tilde{c}}(r_{\tilde{c}}+1)}{\alpha_{\tilde{c}} r_{\tilde{c}}+1}\right\} \geq \frac{1}{2-\alpha_{\tilde{c}}}.$$

Symmetrically, holding $r_{\tilde{c}}$ fixed and minimizing over $\alpha_{\tilde{c}}$, the two right-hand sides coincide at $\alpha_{\tilde{c}} = 1/(r_{\tilde{c}}+1)$ with common value $(r_{\tilde{c}}+1)/(2r_{\tilde{c}}+1)$, so

$$\frac{\Pi^\star_{\tilde{c}}}{\mathbb{E}\left[(v-c)^+ \mid c=\tilde{c}\right]} \geq \max\left\{\frac{1}{\alpha_{\tilde{c}} r_{\tilde{c}}+1}, \frac{\alpha_{\tilde{c}}(r_{\tilde{c}}+1)}{\alpha_{\tilde{c}} r_{\tilde{c}}+1}\right\} \geq \frac{r_{\tilde{c}}+1}{2\,r_{\tilde{c}}+1}.$$

Combining the two yields the desired bound. □

Proof of Proposition 3. Fix $\tilde{c} \in C$ and $\varepsilon > 0$. Let $V_{\tilde{c}} := [0, 1+\tilde{c}]$, $T := 2 + e^{4/\varepsilon}$, and let

$$\eta := \frac{1}{T^2}, \qquad \xi := \frac{\eta}{4}(T-1).$$

Let $a := \tilde{c} + \eta$, and $b := \tilde{c} + \frac{\eta}{2}(T+1)$ and note that

$$b + \xi - \tilde{c} = \frac{\eta}{2}(T+1) + \frac{\eta}{4}(T-1) = \frac{\eta}{4}(3T+1) = \frac{1}{4T^2}(3T+1) < 1,$$

where the last inequality follows from

$$4T^2 - (3T+1) = (T-1)(4T+1) > 0.$$

Now, define $F_{\tilde{c}}$ by

$$F_{\tilde{c}}(x) := \begin{cases} 0, & \text{if } x \in [0, a), \\ 1 - \frac{\eta/2}{x-\tilde{c}-\eta/2}, & \text{if } x \in [a, b), \\ 1 - \frac{b+\xi-x}{T\xi}, & \text{if } x \in [b, b+\xi), \\ 1, & \text{if } x \in [b+\xi, 1+\tilde{c}]. \end{cases}$$

This distribution is absolutely continuous and compactly supported on $V_{\tilde{c}}$, with a density given by

$$f_{\tilde{c}}(x) = \begin{cases} \frac{\eta}{2(x-\tilde{c}-\eta/2)^2}, & \text{if } x \in (a,b), \\ \frac{1}{T\xi}, & \text{if } x \in (b, b+\xi) \\ 0, & \text{otherwise.} \end{cases}$$

Note that for any $x \geq \tilde{c}$, the profit under uniform price x when cost is $\tilde{c}$ is given by:

$$(x-\tilde{c})(1-F_{\tilde{c}}(x^-)) .$$

For $x < a$, we have $F_{\tilde{c}}(x^-) = 0$, and hence

$$(x-\tilde{c})(1-F_{\tilde{c}}(x^-)) = (x-\tilde{c}) .$$

For $x \in [a,b)$,

$$(x-\tilde{c})(1-F_{\tilde{c}}(x^-)) = (x-\tilde{c})\frac{\eta/2}{x-\tilde{c}-\eta/2} = \frac{\eta}{2}\frac{x-\tilde{c}}{x-\tilde{c}-\eta/2} ,$$

which is strictly decreasing on $[a,b)$, and hence is maximized at $x = a$, with the maximized value being η. For $x \in [b, b+\xi]$, since $1 - F_{\tilde{c}}(x^-) \leq 1/T$,

$$(x-\tilde{c})(1-F_{\tilde{c}}(x^-)) \leq \frac{x-\tilde{c}}{T} \leq \frac{b+\xi-\tilde{c}}{T} = \eta\frac{3T+1}{4T} < \eta .$$

Finally, for $x > b+\xi$, $F_{\tilde{c}}(x^-) = 1$. Therefore, the unique maximizer of

$$x \mapsto (x-\tilde{c})(1-F_{\tilde{c}}(x^-))$$

is $x = a = \tilde{c} + \eta$, with the maximized value being η.

Now, let $v^\star_{\tilde{c}} := \tilde{c} + \frac{3\eta}{2}$. Since $T > 2$, $v^\star_{\tilde{c}} \in (a,b)$. Moreover, $F_{\tilde{c}}(v^\star_{\tilde{c}}) = 1/2$. Define $f_{\tilde{c},l}$ and $f_{\tilde{c},h}$ as:

$$f_{\tilde{c},l}(x) := \begin{cases} (1+\alpha_{\tilde{c}})f_{\tilde{c}}(x), & \text{if } x < v^\star_{\tilde{c}}, \\ (1-\alpha_{\tilde{c}})f_{\tilde{c}}(x), & \text{if } x \geq v^\star_{\tilde{c}}, \end{cases} \quad \text{and} \quad f_{\tilde{c},h}(x) := \begin{cases} \alpha_{\tilde{c}} f_{\tilde{c}}(x), & \text{if } x < v^\star_{\tilde{c}}, \\ (2-\alpha_{\tilde{c}})f_{\tilde{c}}(x), & \text{if } x \geq v^\star_{\tilde{c}}. \end{cases} .$$

Note that by construction, since $F_{\tilde{c}}(v^\star_{\tilde{c}}) = 1/2$,

$$\int_0^\infty f_{\tilde{c},l}(x)\,\mathrm{d}x = F_{\tilde{c}}(v^\star_{\tilde{c}})(1+\alpha_{\tilde{c}}) + (1-F_{\tilde{c}}(v^\star_{\tilde{c}}))(1-\alpha_{\tilde{c}}) = 1 .$$

Likewise,
$$\int_0^\infty f_{\tilde{c},h}(x)\,\mathrm{d}x = F_{\tilde{c}}(v_{\tilde{c}}^\star)\alpha_{\tilde{c}} + (1 - F_{\tilde{c}}(v_{\tilde{c}}^\star))(2 - \alpha_{\tilde{c}}) = 1\,,$$
and hence $f_{\tilde{c},l}$ and $f_{\tilde{c},h}$ are densities supported on $V_{\tilde{c}}$. Let $F_{\tilde{c},l}$ and $F_{\tilde{c},h}$ be the CDFs of $f_{\tilde{c},l}$ and $f_{\tilde{c},h}$, respectively. Since

$$\begin{aligned}\alpha_{\tilde{c}} f_{\tilde{c},h}(x) + (1-\alpha_{\tilde{c}}) f_{\tilde{c},l}(x) =& \alpha_{\tilde{c}}^2 f_{\tilde{c}}(x) + (1-\alpha_{\tilde{c}})(1+\alpha_{\tilde{c}}) f_{\tilde{c}}(x)\\ =& (\alpha_{\tilde{c}}^2 + 1 - \alpha_{\tilde{c}}^2) f_{\tilde{c}}(x)\\ =& f_{\tilde{c}}(x)\,,\end{aligned}$$

for all $x < v_{\tilde{c}}^\star$. Likewise,

$$\begin{aligned}\alpha_{\tilde{c}} f_{\tilde{c},h}(x) + (1-\alpha_{\tilde{c}}) f_{\tilde{c},l}(x) =& \alpha_{\tilde{c}}(2-\alpha_{\tilde{c}}) f_{\tilde{c}}(x) + (1-\alpha_{\tilde{c}})^2 f_{\tilde{c}}(x)\\ =& (2\alpha_{\tilde{c}} - \alpha_{\tilde{c}}^2 + 1 - 2\alpha_{\tilde{c}} + \alpha_{\tilde{c}}^2) f_{\tilde{c}}(x)\\ =& f_{\tilde{c}}(x)\,,\end{aligned}$$

for all $x \geq v_{\tilde{c}}^\star$. As a result, $\alpha_{\tilde{c}} F_{\tilde{c},h} + (1-\alpha_{\tilde{c}}) F_{\tilde{c},l} = F_{\tilde{c}}$. Moreover, since

$$\frac{f_{\tilde{c},h}(x)}{f_{\tilde{c},l}(x)} = \begin{cases} \frac{\alpha_{\tilde{c}}}{1+\alpha_{\tilde{c}}}, & \text{if } x \leq v_{\tilde{c}}^\star \\ \frac{2-\alpha_{\tilde{c}}}{1-\alpha_{\tilde{c}}}, & \text{if } x > v_{\tilde{c}}^\star \end{cases}$$

for all $x \in [0, 1+\tilde{c}]$, and since
$$\frac{\alpha_{\tilde{c}}}{1+\alpha_{\tilde{c}}} < \frac{2-\alpha_{\tilde{c}}}{1-\alpha_{\tilde{c}}}\,,$$
$F_{\tilde{c},h}$ dominates $F_{\tilde{c},l}$ in the likelihood ratio order.

Together, under any optimal uniform pricing rule p^u, $p^u(\tilde{c}) = \tilde{c} + \eta$, and

$$\begin{aligned}\frac{\mathbb{E}[(p^u - \tilde{c})\mathbf{1}\{v \geq p^u\} \mid c = \tilde{c}]}{\mathbb{E}[(v-\tilde{c})^+ \mid c = \tilde{c}]} = \frac{\eta}{\eta + \frac{\eta}{2}\log T + \frac{\xi}{2T}} =& \frac{1}{1 + \frac{1}{2}\log T + \frac{\xi}{2T\eta}}\\ \leq& \frac{1}{1 + \frac{1}{2}\log T}\,.\end{aligned}$$

Since $T = 2 + e^{4/\varepsilon} > e^{4/\varepsilon}$,
$$\frac{1}{1 + \frac{1}{2}\log T} < \frac{1}{1 + 2/\varepsilon} = \frac{\varepsilon}{\varepsilon + 2} < \varepsilon\,,$$

as desired. $\square$

Proof of Proposition 4. For (i), by Theorem 1, to show that $CS(c,h;p^\star) > 0$, it suffices to show $F_{c,h}(\kappa_c^5) > F_{c,h}(\kappa_c^4)$. To see this, since $\kappa_c^4 < v_c^\star < \kappa_c^5$ and since $v_c^\star \in (\underline{v}_c, \overline{v}_c)$, it must be that $F_{c,h}(\kappa_c^5) > F_{c,h}(\kappa_c^4)$, as desired. In the meantime, to show that $WL(c,h;p^\star) > 0$, it suffices to show $F_{c,h}(\kappa_c^1) > 0$ by Theorem 1. According to (8), $\kappa_c^1 > c$ for all c. Therefore, whenever $\underline{v}_c < c$, $F_{c,h}(\kappa_c^1) > F_{c,h}(c) \geq F_{c,h}(\underline{v}_c) = 0$, as desired.

For (ii), suppose that $\alpha_c\overline{v}_c > \underline{v}_c - (1-\alpha_c)c$. By Theorem 1, it suffices to show that $F_{c,l}(\kappa_c^3) > F_{c,l}(\kappa_c^2) > 0$. We first claim that $\Delta_c(\kappa_c^4) > 0$. To see this, suppose the contrary. $\Delta_c(\kappa_c^4) = 0$. Then it must be that $\kappa_c^4 \leq \underline{v}_c$ and $\kappa_c^5 \geq \overline{v}_c$. Moreover, $\kappa_c^2 = F_{c,l}^{-1}(\Delta_c(\kappa_c^4)) = F_{c,l}^{-1}(0) = \underline{v}_c$, and $\Delta_c(\kappa_c^3) = F_{c,h}(\kappa_c^1) = 0$. Since κ_c is increasing, it must be that $\kappa_c^3 \leq \kappa_c^4 \leq \underline{v}_c$. Together,

$$\underline{v}_c = \kappa_c^2 \leq \kappa_c^3 \leq \kappa_c^4 \leq \underline{v}_c\,,$$

Therefore, $\kappa_c^3 = \kappa_c^4 = \underline{v}_c$. Thus,

$$(1-\alpha_c)(\kappa_c^3 - c) = (1-\alpha_c)(\underline{v}_c - c) = \alpha_c(\kappa_c^5 - \kappa_c^4) \geq \alpha_c(\overline{v}_c - \underline{v}_c)\,,$$

and hence

$$\alpha_c\overline{v}_c \leq \underline{v}_c - (1-\alpha_c)c\,,$$

a contradiction. As a result, $\Delta_c(\kappa_c^4) > 0$. This implies that $F_{c,l}(\kappa_c^2) > 0$. Together, we have that $CS(c,l;p^\star) > 0$ and $WL(c,l;p^\star) > 0$.

Conversely, suppose that $\alpha_c(\overline{v}_c - c) \leq \underline{v}_c - c$. From (8), $\kappa_c^1 - c = \alpha_c(\kappa_c^5 - \kappa_c^4) \geq 0$, and hence $\kappa_c^1 \geq c$; together with the fact that κ_c is increasing, this implies $\kappa_c^4 \geq \kappa_c^3 \geq c$ via $\kappa_c^1 - c = (1-\alpha_c)(\kappa_c^3 - c)$. We claim that $\kappa_c^1 \leq \underline{v}_c$. Suppose, by way of contradiction, that $\kappa_c^1 > \underline{v}_c$. Then $F_{c,h}(\kappa_c^1) > 0$, and (8) implies $\Delta_c(\kappa_c^4) = \Delta_c(\kappa_c^3) + F_{c,h}(\kappa_c^1) > 0$, so that $\Delta_c(\kappa_c^5) = \Delta_c(\kappa_c^4) > 0$ as well. Since Δ_c is quasi-concave with $\Delta_c(v) = 0$ for all $v \geq \overline{v}_c$, it follows that $\kappa_c^5 < \overline{v}_c$. Combining this with $\kappa_c^4 \geq c$, we obtain

$$\kappa_c^1 - c = \alpha_c(\kappa_c^5 - \kappa_c^4) < \alpha_c(\overline{v}_c - c) \leq \underline{v}_c - c\,,$$

contradicting $\kappa_c^1 > \underline{v}_c$. Hence $\kappa_c^1 \leq \underline{v}_c$, and (8) then implies that $0 = \Delta_c(\kappa_c^3) = \Delta_c(\kappa_c^4)$, and hence $\kappa_c^3 = \kappa_c^4 = \underline{v}_c$, which in turn implies that $\kappa_c^2 = \underline{v}_c$ and $\kappa_c^5 = \overline{v}_c$. Thus, by Theorem 1, $CS(c,l;p^\star) = WL(c,l;p^\star) = 0$, as desired.

For (iii), note that

$$WL(c,l;p^\star) = \int_0^{\kappa_c^2} (v-c)^+ f_{c,l}(v)\,\mathrm{d}v \quad \text{ and } \quad WL(c,h;p^\star) = \int_0^{\kappa_c^1} (v-c)^+ f_{c,h}(v)\,\mathrm{d}v\,.$$

Moreover, by Lemma 1, $\kappa_c^1 \le \kappa_c^2$. Since $F_{c,h}$ dominates $F_{c,l}$ in the likelihood ratio order, $f_{c,h}(v) \le f_{c,l}(v)$ for all $v \le \kappa_c^2 \le v_c^\star$. Together, we have $WL(c,l;p^\star) \ge WL(c,h;p^\star)$. Furthermore, suppose that $\alpha_c(\overline{v}_c - c) \le (\underline{v}_c - c)$. Then $WL(c,l;p^\star) = WL(c,h;p^\star) = 0$. Meanwhile, suppose $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$. By part (ii), $\kappa_c^2 > \underline{v}_c$; and by (8) together with $\kappa_c^5 > \kappa_c^4$ from Lemma 1, $\kappa_c^1 > c$. Thus, since

$$WL(c,l;p^\star) - WL(c,h;p^\star) = \int_0^{\kappa_c^2} (v-c)^+ (f_{c,l}(v) - f_{c,h}(v))\,\mathrm{d}v + \int_{\kappa_c^1}^{\kappa_c^2} (v-c)^+ f_{c,h}(v)\,\mathrm{d}v\,,$$

and since the integrands in both terms are non-negative by the likelihood ratio order (using $\kappa_c^1 \le \kappa_c^2 \le v_c^\star$), it suffices to show the first term is strictly positive. By the strict likelihood-ratio assumption, $f_{c,l} > f_{c,h}$ on $(\underline{v}_c, v_c^\star)$, and combined with $\kappa_c^2 > \max(c, \underline{v}_c)$, the first term is strictly positive. Thus $WL(c,l;p^\star) > WL(c,h;p^\star)$.

Lastly, from (A.12), (A.13), (A.14), and (A.15) in the proof of Proposition 8, the surplus of consumers with $\theta = l$ and $v \notin [\kappa_c^2, \kappa_c^3]$ and the surplus of consumers with $\theta = h$ and $v \notin [\kappa_c^4, \kappa_c^5]$ must be zero.

□

Proof of Proposition 5. For (i), consider each $c \in C$. Recall from the proof of Lemma 1 that $\tilde{v}_c$ is decreasing in α_c and that $\alpha_c h_c(v)/(1-\alpha_c) + c \le v_c^\star$ for all $v \in [v_c^\star, \tilde{v}_c]$. Since Δ_c therefore acts on its increasing side at this argument, the function

$$\alpha_c \mapsto \Delta_c\left(\frac{\alpha_c}{1-\alpha_c} h_c(v) + c\right) + F_{c,h}(\alpha_c h_c(v) + c)$$

is increasing in α_c for all $v \in [v_c^\star, \tilde{v}_c]$, and hence κ_c^5 defined in (A.6) is decreasing in α_c, which in turn implies that κ_c^4 is increasing in α_c. Therefore, since

$$CS(c,h;p^\star) = \int_{F_{c,h}(\kappa_c^4)}^{F_{c,h}(\kappa_c^5)} (F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q + \Delta_c(\kappa_c^4)))\,\mathrm{d}q,$$

$CS(c,h;p^\star)$ is decreasing in α_c.

For (ii), fix $c \in C$ and recall from (i) that κ_c^5 is decreasing and κ_c^4 is increasing in α_c on

$(0,1)$, while $\kappa_c^4 < v_c^\star < \kappa_c^5$ throughout.

Consider first $\alpha_c \to 1$. By monotonicity and the bound $\kappa_c^4 < v_c^\star < \kappa_c^5$, κ_c^4 converges to some limit in $[\underline{v}_c, v_c^\star]$ and κ_c^5 converges to some limit in $[v_c^\star, \overline{v}_c]$ as $\alpha_c \to 1$. Since $\Delta_c(\kappa_c^4) = \Delta_c(\kappa_c^5)$ by (8) and Δ_c is continuous, the two limits have the same Δ_c-value. Moreover, by (8), $(1-\alpha_c)(\kappa_c^3 - c) = \alpha_c(\kappa_c^5 - \kappa_c^4)$; since κ_c^3 is bounded, as $\alpha_c \to 1$ the LHS $\to 0$, so $\kappa_c^5 - \kappa_c^4 \to 0$. Combined with $\kappa_c^4 \leq v_c^\star \leq \kappa_c^5$, both limits equal $v_c^\star$. Hence $F_{c,h}(\kappa_c^4), F_{c,h}(\kappa_c^5) \to F_{c,h}(v_c^\star)$ and $\Delta_c(\kappa_c^4) \to 0$. The integrand of $CS(c,h;p^\star,\alpha_c)$ is bounded by $\overline{v}_c - \underline{v}_c$ while the interval of integration collapses to a point, so $CS(c,h;p^\star,\alpha_c) \to 0$.

Now consider $\alpha_c \to 0$. Since $\kappa_c^4, \kappa_c^5 \in [\underline{v}_c, \overline{v}_c]$, the difference $\kappa_c^5 - \kappa_c^4$ is bounded by $\overline{v}_c - \underline{v}_c$. By (8), $\kappa_c^1 - c = \alpha_c(\kappa_c^5 - \kappa_c^4) = (1-\alpha_c)(\kappa_c^3 - c)$. The first equality yields $\kappa_c^1 \to c$; the second then gives $\kappa_c^3 - c = (\kappa_c^1 - c)/(1-\alpha_c) \to 0$, i.e., $\kappa_c^3 \to c$. Since κ_c is increasing, $\kappa_c^2 \in [\kappa_c^1, \kappa_c^3]$, so $\kappa_c^2 \to c$ as well. The integrand of $CS(c,l;p^\star,\alpha_c)$ is bounded by $\overline{v}_c - \underline{v}_c$ and the integration interval shrinks to $\{c\}$, so $CS(c,l;p^\star,\alpha_c) \to 0$. This completes the proof. □

Proof of Proposition 6. To prove (i), consider any non-discriminatory pricing rule p that induces non-discriminatory outcomes. For each $c \in C$, define a matching scheme $\rho_c \in \Delta(V^2)$ as

$$\rho_c(V_l \times V_h) := \mathbb{E}\left[\mathbb{P}[v \in V_l \mid p, c, \theta = l, y] \times \mathbb{P}[v \in V_h \mid p, c, \theta = h, y] \mid c\right],$$

for all measurable $V_l, V_h \subseteq V$. Since p induces non-discriminatory outcomes,

$$\rho_c([0,z] \times V) = \mathbb{E}[\mathbb{P}[v \in [0,z] \mid p, \theta = l, c, y] \mid c] = F_{c,l}(z),$$

and

$$\rho_c(V \times [0,z]) = \mathbb{E}[\mathbb{P}[v \in [0,z] \mid p, \theta = h, c, y] \mid c] = F_{c,h}(z),$$

for all $z \geq 0$. Therefore, $\rho_c \in \mathcal{R}_c$.

For each c, given such matching scheme ρ_c, since (p,y) is independent of θ conditional on c, each matched pair $(v_l, v_h) \in \text{supp}(\rho_c)$ must face the same price under p; and either both purchase or both do not purchase. Therefore, the seller's profit under p must be weakly lower than selling to each matched pair $(v_l, v_h) \in \text{supp}(\rho_c)$ at a price $\min\{v_l, v_h\}$ whenever $\min\{v_l, v_h\} \geq c$, and not selling to the pair otherwise. That is,

$$\Pi(p) \leq \mathbb{E}\left[\int_{V^2} (\min\{v_l, v_h\} - c)^+ \, \mathrm{d}\rho_c\right].$$

As a result, for any pricing p that induces non-discriminatory outcomes,

$$\Pi(p) \leq \mathbb{E}\left[\max_{\rho_c \in \mathcal{R}_c} \int_{V^2} (\min\{v_l, v_h\} - c)^+ \, \mathrm{d}\rho_c\right] =: \overline{\pi} \, .$$

Since the objective $(\min\{v_l, v_h\} - c)^+$ of the optimal transport problem

$$\max_{\rho_c \in \mathcal{R}_c} \int_{V^2} (\min\{v_l, v_h\} - c)^+ \, \mathrm{d}\rho_c$$

is supermodular for all $c \in C$, the assortative matching must be a solution. Therefore,

$$\max_{\rho_c \in \mathcal{R}_c} \int_{V^2} (\min\{v_l, v_h\} - c)^+ \, \mathrm{d}\rho_c = \int_V (v - c)^+ F_{c,l}(\mathrm{d}v) \, .$$

Thus, by construction, under the pricing rule p^{ass},

$$\Pi(p^{ass}) = \mathbb{E}\left[\int_V (v - c)^+ F_{c,l}(\mathrm{d}v)\right] = \overline{\pi} \, .$$

Since p^{ass} also induces non-discriminatory outcomes, it is optimal. Furthermore, as p^{ass} is also non-discriminatory, $\Pi(p^\star) \geq \Pi(p^{ass})$. Lastly, since the solution of

$$\max_{\rho_c \in \mathcal{R}_c} \int_{V \times V} (\min\{v_l, v_h\} - c)^+ \rho_c(\mathrm{d}v_l, \mathrm{d}v_h)$$

must correspond to a pricing rule that is outcome-equivalent to the assortative matching for all $c \in C$, any profit-maximizing pricing rule that induces non-discriminatory outcomes must yield the same surplus to consumers.

For (iii), since $F_{c,h}$ dominates $F_{c,l}$ in the likelihood ratio order, under p^{ass}, each matched pair of consumers who are purchasing consumers must buy at the value of the consumer with $\theta = l$, and hence $0 = CS(c, l; p^{ass}) \leq CS(c, l; p^\star)$. Moreover, note that as established in the proof of (ii) of Proposition 4, $p^\star \neq p^{ass}$ if and only if $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$ for a positive measure of c. Thus, $CS(c, h; p^{ass}) > CS(c, h; p^\star)$ if and only if $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$ for a positive measure of c.

To see that $CS(c,h;p^{ass}) \geq CS(c,h;p^\star)$, first, note that

$$\begin{aligned} CS(c,h;p^\star) &= \int_{F_{c,h}(\kappa_c^4)}^{F_{c,h}(\kappa_c^5)} [F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q + \Delta_c(\kappa_c^4))]\,\mathrm{d}q \\ &= \int_{F_{c,h}(\kappa_c^4)}^{F_{c,h}(\kappa_c^5)} F_{c,h}^{-1}(q)\,\mathrm{d}q - \int_{F_{c,l}(\kappa_c^4)}^{F_{c,l}(\kappa_c^5)} F_{c,l}^{-1}(r)\,\mathrm{d}r \\ &\leq \int_{F_{c,l}(\kappa_c^4)}^{F_{c,l}(\kappa_c^5)} F_{c,h}^{-1}(q)\,\mathrm{d}q - \int_{F_{c,l}(\kappa_c^4)}^{F_{c,l}(\kappa_c^5)} F_{c,l}^{-1}(r)\,\mathrm{d}r\,, \end{aligned}$$

where the equality follows from changing variables and letting $r = q + \Delta_c(\kappa_c^4)$ and using $\Delta_c(\kappa_c^4) = \Delta_c(\kappa_c^5)$; and the inequality follows from the fact that $F_{c,l}(\kappa_c^5) - F_{c,l}(\kappa_c^4) = F_{c,h}(\kappa_c^5) - F_{c,h}(\kappa_c^4) + \Delta_c(\kappa_c^5) - \Delta_c(\kappa_c^4) = F_{c,h}(\kappa_c^5) - F_{c,h}(\kappa_c^4)$, that $F_{c,h}(\kappa_c^4) \leq F_{c,l}(\kappa_c^4)$ and $F_{c,h}(\kappa_c^5) \leq F_{c,l}(\kappa_c^5)$ due to the likelihood ratio assumption, and that $F_{c,h}^{-1}$ is nondecreasing.

In the meantime, since under p^{ass}, all h-consumers with $v \leq F_{c,h}^{-1}(F_{c,l}(c))$ are excluded, whereas all h-consumers with $v > F_{c,h}^{-1}(F_{c,l}(c))$ purchase at price $F_{c,l}^{-1}(F_{c,h}(v))$. Therefore,

$$CS(c,h;p^{ass}) = \int_{F_{c,l}(c)}^{1} [F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q)]\,\mathrm{d}q\,.$$

Now, note that by (8), $c \leq \kappa_c^4 \leq \kappa_c^5$, we have $F_{c,l}(c) \leq F_{c,l}(\kappa_c^4) \leq F_{c,l}(\kappa_c^5) \leq 1$. Moreover, since $F_{c,h}^{-1}(q) \geq F_{c,l}^{-1}(q)$ for all q by the likelihood-ratio order assumption, we have

$$\int_{F_{c,l}(c)}^{1} [F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q)]\,\mathrm{d}q \geq \int_{F_{c,l}(\kappa_c^4)}^{F_{c,l}(\kappa_c^5)} [F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q)]\,\mathrm{d}q\,.$$

As a result,

$$\begin{aligned} CS(c,h;p^{ass}) = \int_{F_{c,l}(c)}^{1} [F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q)]\,\mathrm{d}q &\geq \int_{F_{c,l}(\kappa_c^4)}^{F_{c,l}(\kappa_c^5)} [F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q)]\,\mathrm{d}q \\ &= \int_{F_{c,l}(\kappa_c^4)}^{F_{c,l}(\kappa_c^5)} F_{c,h}^{-1}(q)\,\mathrm{d}q - \int_{F_{c,l}(\kappa_c^4)}^{F_{c,l}(\kappa_c^5)} F_{c,l}^{-1}(r)\,\mathrm{d}r \\ &\geq CS(c,h;p^\star)\,, \end{aligned}$$

as desired.

For (ii), since p^{ass} is non-discriminatory, Theorem 1 implies that $\Pi(p^\star) \geq \Pi(p^{ass})$. Moreover, by (iii), since $CS(c,h;p^{ass}) > CS(c,h;p^\star)$ if and only if $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$, and since every undominated profit-maximizing rule p must yield $CS(c,\theta;p) = CS(c,\theta;p^\star)$ for all $c \in C$ and for all $\theta \in \{l, h\}$ by Theorem 1, it must be that p^{ass} is not optimal if and only if $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$ for a positive measure of $c \in C$. Therefore, $\Pi(p^\star) > \Pi(p^{ass})$ if and only if $\alpha_c(\overline{v}_c - c) > \underline{v}_c - c$. This completes the proof. □

Proof of Proposition 7. For (i), note that since $\overline{p}$ is a disparate treatment-based non-discriminatory pricing rule, $\overline{p}$ is only a function of (z, c, r). Moreover, since $\overline{p}$ is optimal, it must be that

$$\Pi(\overline{p}) = \mathbb{E}\left[\max_{p\geq 0}(p - c)\mathbb{E}[\mathbf{1}\{V_{c,\theta}(z) \geq p\} \mid z, c]\right] = \mathbb{E}\left[\max\{0,\ V_{c,l}(z) - c,\ \alpha_c(z)(V_{c,h}(z) - c)\}\right],$$

where the second equality follows from $V_{c,l} \leq V_{c,h}$ for all $c \in C$. Therefore, for almost all (z, c, r),

$$\overline{p}(z, c, r) \geq V_{c,l}(z),$$

which in turn implies that $\theta = l$ consumers always retain zero surplus, as they are either fully extracted or excluded.

For (iii), by (i), the seller's expected profit conditional on c is

$$\mathbb{E}\Big[g\big(V_{c,l}(z), V_{c,h}(z), \alpha_c(z)\big)\,\Big|\, c\Big],$$

where

$$g(v_l, v_h, \alpha) := \max\big\{0,\ v_l - c,\ \alpha(v_h - c)\big\}.$$

For each fixed (v_l, v_h, c) with $v_h \geq v_l$, $g(v_l, v_h, \cdot)$ is a maximum of three affine functions of α, hence convex in α. Consider two environments with the same joint distribution of $\big(V_{c,l}(z), V_{c,h}(z), c\big)$, such that $\alpha_c(z)$ in the first environment is more dispersed in the convex order than $\alpha_c(z')$ in the second, conditional on $\big(V_{c,l}(z), V_{c,h}(z)\big)$ and c. By convexity of g in α,

$$\mathbb{E}\Big[g\big(V_{c,l}(z), V_{c,h}(z), \alpha_c(z)\big)\,\Big|\, V_{c,l}(z), V_{c,h}(z), c\Big] \geq \mathbb{E}\Big[g\big(V_{c,l}(z'), V_{c,h}(z'), \alpha_c(z')\big)\,\Big|\, V_{c,l}(z'), V_{c,h}(z'), c\Big].$$

Since the joint distribution of $\big(V_{c,l}, V_{c,h}, c\big)$ coincides across the two environments, taking expectations over $\big(V_{c,l}, V_{c,h}, c\big)$ yields a weakly higher $\Pi(\overline{p})$ in the first environment, as desired. □

Proof of Proposition 8. By Proposition 1, any optimal non-discriminatory pricing rule p

can be identified by a family $(\rho_c)_{c\in C}$ of matching schemes, where $\rho_c \in \mathcal{R}_c$ is a solution of

$$\max_{\rho\in\mathcal{R}_c} \int_{V^2} \pi_c(v_l, v_h)\,\mathrm{d}\rho\,.$$

Thus, it suffices to consider the solutions of the optimal transport problem (10) for each c.

When $F_{c,l}(c) \geq \|F_{c,l} - F_{c,h}\|$, since $\pi^\star(c) = \mathbb{E}[(v-c)^+]$, it must be that $CS(c,h;p) = CS(c,l;p) = 0$. Now suppose that $F_{c,l}(c) < \|F_{c,l} - F_{c,h}\|$. Fix any such $c \in C$, consider any solution ρ_c of the optimal transport problem (10). By Lemma 2, for any solution ρ_c of (10), it must be that

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = \pi_c(v_l, v_h)\,,$$

for all $(v_l, v_h) \in \operatorname{supp}(\rho_c)$. Therefore, we have

$$\begin{aligned}
\operatorname{supp}(\rho_c) \subseteq [0,\kappa_c^3]\times[\kappa_c^5,\infty) &\cup \{(v_l,v_h)\in[\kappa_c^1,\kappa_c^3]\times[\kappa_c^1,\kappa_c^3] : v_l \geq v_h\} \\
&\cup [\kappa_c^3,\kappa_c^4]\times[0,\kappa_c^1] \\
&\cup \{(v_l,v_h)\in[\kappa_c^3,\kappa_c^4]\times[\kappa_c^3,\kappa_c^4] : v_l = v_h\} \\
&\cup \{(v_l,v_h)\in[\kappa_c^4,\kappa_c^5]\times[\kappa_c^4,\kappa_c^5] : v_l \leq v_h\} \\
&\cup \{(v_l,v_h)\in[\kappa_c^5,\infty)\times[\kappa_c^5,\infty) : v_l = v_h\}
\end{aligned} \tag{A.11}$$

Now consider any optimal non-discriminatory pricing rule p. By Proposition 1, there exists $(\rho_c)_{c\in C}$ such that ρ_c is a solution of (10) for all $c \in C$ and that for almost all $c \in C$ and for almost all matched pair $(v_l, v_h) \in \operatorname{supp}(\rho_c)$, these consumers face a price that equals

$$\operatorname*{argmax}_{x\in\{v_l,v_h\}} (x-c)[\alpha_c \mathbf{1}\{v_h \geq x\} + (1-\alpha_c)\mathbf{1}\{v_l \geq x\}]\,,$$

whenever $\max\{v_l, v_h\} \geq c$. Let $(\rho_c)_{c\in C}$ be the family of matching schemes associated with the non-discriminatory pricing rule p. Note that for any $(v_l, v_h) \in [\kappa_c^3, \kappa_c^4]\times[0,\kappa_c^1]$, $(1-\alpha_c)(v_l - c) \geq (1-\alpha_c)(\kappa_c^3 - c) = \kappa_c^1 - c \geq v_h - c$. Therefore, the optimal price for these matched pairs equals v_l, and hence l-consumers purchase at a price equal to their values, whereas h-consumers do not purchase. In particular, these consumers retain zero surplus, just as under $p^\star$. Likewise, for any $(v_l, v_h) \in [0,\kappa_c^3]\times[\kappa_c^5,\infty)$, $\alpha_c(v_h - c) \geq \alpha_c(\kappa_c^5 - c) = \alpha_c(\kappa_c^4 - c) + (1-\alpha_c)(\kappa_c^3 - c) \geq \alpha_c\kappa_c^3 - c + (1-\alpha_c)(\kappa_c^3 - c) = \kappa_c^3 - c \geq v_l - c$. Thus, the optimal price for these matched pairs (v_l, v_h) must equal v_h, and thus h-consumers purchase by paying their values, whereas l-consumers do not purchase, just as under $p^\star$. Furthermore, for any $(v_l, v_h) \in [\kappa_c^4, \kappa_c^5] \times [\kappa_c^4, \kappa_c^5]$ such that $v_l \le v_h$, $\alpha_c(v_h - c) \le \alpha_c(\kappa_c^5 - c) = \alpha_c(\kappa_c^4 - c) + (1-\alpha_c)(\kappa_c^3 - c) \le \kappa_c^4 - c \le v_l - c$. The optimal price for these matched pairs (v_l, v_h) must equal v_l, and hence both l- and h- consumers purchase by paying the value of l-consumers, just as under $p^\star$. Together, the pricing rule p must lead to the same outcomes as $p^\star$ for matched pairs (v_l, v_h) in $[0, \kappa_c^1] \times [\kappa_c^5, \infty)$, $\{(v_l, v_h) \in [\kappa_c^3, \kappa_c^4] \times [0, \kappa_c^1] : v_l \ge v_h\}$, and $\{(v_l, v_h) \in [\kappa_c^4, \kappa_c^5] \times [\kappa_c^4, \kappa_c^5] : v_l \le v_h\}$.

In the meantime, for any matched pair $(v_l, v_h) \in [\kappa_c^1, \kappa_c^3] \times [\kappa_c^1, \kappa_c^3]$, since $(1-\alpha_c)(v_l - c) \le (1-\alpha_c)(\kappa_c^3 - c) = \kappa_c^1 - c \le v_h - c$, the optimal price for these matched pairs must be v_h.

Together, we have

$$\begin{aligned}
\mathbb{E}[(v-p)^+ \mathbf{1}\{v \in [\kappa_c^4, \kappa_c^5]\} \mid c, \theta = h] =& CS(c,h;p) \\
=& CS(c,h;p^\star) \qquad\qquad (A.12)\\
=& \int_{F_{c,h}(\kappa_c^4)}^{F_{c,h}(\kappa_c^5)} (F_{c,h}^{-1}(q) - F_{c,l}^{-1}(q + F_{c,h}(\kappa_c^4) - F_{c,l}(\kappa_c^4)))\,\mathrm{d}q\,;
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[(v-c)^+ \mathbf{1}\{v \le \min\{p, \kappa_c^1\}\} \mid c, \theta = h] =& WL(c,h;p) \\
=& WL(c,h;p^\star) \qquad\qquad (A.13)\\
=& \int_0^{\kappa_c^1} (v-c)^+ F_{c,h}(\mathrm{d}v)\,;
\end{aligned}$$

$$\mathbb{E}[(v-p)^+ \mathbf{1}\{v \in [\kappa_c^1, \kappa_c^3]\} \mid c, \theta = l] = CS(c,l;p) = \int_{[\kappa_c^1, \kappa_c^3]^2} (v_l - v_h)\rho(\mathrm{d}v_l, \mathrm{d}v_h)\,; \qquad (A.14)$$

and

$$\begin{aligned}
\mathbb{E}[(v-c)^+ \mathbf{1}\{v \le \min\{p, \kappa_c^3\}\} \mid c, \theta = l] =& WL(c,l;p) \qquad\qquad (A.15)\\
=& \int_0^{\kappa_c^1} (v-c)^+ F_{c,l}(\mathrm{d}v) + \int_{[\kappa_c^1, \kappa_c^3] \times [\kappa_c^5, \infty)} (v_l - c)^+ \rho(\mathrm{d}v_l, \mathrm{d}v_h)
\end{aligned}$$

By (A.11), $\rho_c(v_l \notin [\kappa_c^1, \kappa_c^3], v_h \in [\kappa_c^1, x]) = 0$ and $\rho_c(v_l \in [\kappa_c^1, \kappa_c^3], v_h \notin [\kappa_c^1, \kappa_c^3]) = \rho_c(v_l \in [\kappa_c^1, x], v_h > \kappa_c^5)$, for all $x \in [\kappa_c^1, \kappa_c^3]$. Thus,

$$\begin{aligned}
\rho_c(v_l \in [\kappa_c^1, \kappa_c^3], v_h \in [\kappa_c^1, x]) =& \rho(v_l \in V, v_h \in [\kappa_c^1, x]) - \rho_c(v_l \notin [\kappa_c^1, \kappa_c^3], v_h \in [\kappa_c^1, x]) \\
=& F_{c,h}(x) - F_{c,h}(\kappa_c^1)\,,
\end{aligned}$$

and

$$\begin{aligned}\rho_c(v_l \in [\kappa_c^1, x], v_h \in [\kappa_c^1, \kappa_c^3]) =& \rho_c(v_l \in [\kappa_c^1, x], v_h \in V) - \rho_c(v_l \in [\kappa_c^1, x], v_h \notin [\kappa_c^1, \kappa_c^3]) \\ =& F_{c,l}(x) - F_{c,l}(\kappa_c^1) - \rho_c(v_l \in [\kappa_c^1, x], v_h > \kappa_c^5)\,,\end{aligned}$$

for all $x \in [\kappa_c^1, \kappa_c^3]$. Moreover, by (A.11), since $\rho_c \in \mathcal{R}_c$, it must be that

$$\rho_c(v_l \in [0, \kappa_c^3], v_h > \kappa_c^5) = \Delta_c(\kappa_c^5)\,.$$

Together with the fact that $F_{c,l}(\kappa_c^2) = \Delta_c(\kappa_c^5)$, and noting that the ρ_c-mass of $\{v_l \in [\kappa_c^1, x]\}$ cannot exceed the $F_{c,l}$-mass of the same interval, it follows that

$$\min\{F_{c,l}(x) - F_{c,l}(\kappa_c^1), F_{c,l}(\kappa_c^2)\} \geq \rho_c(v_l \in [\kappa_c^1, x], v_h > \kappa_c^5)\,. \tag{A.16}$$

As a result,

$$\begin{aligned}CS(c,l;p) =& \int_{[\kappa_c^1,\kappa_c^3]^2} (v_l - v_h)\rho_c(\mathrm{d}v_l, \mathrm{d}v_h) \\ =& \int_{\kappa_c^1}^{\kappa_c^3} v_l F_{c,l}(\mathrm{d}v_l) - \int_{\kappa_c^1}^{\kappa_c^3} v_l \rho_c(\mathrm{d}v_l, v_h > \kappa_c^5) - \int_{\kappa_c^1}^{\kappa_c^3} v_h F_{c,h}(\mathrm{d}v_h) \\ \leq& \int_{\kappa_c^1}^{\kappa_c^3} v F_{c,l}(\mathrm{d}v) - \int_{\kappa_c^1}^{\kappa_c^2} v F_{c,l}(\mathrm{d}v) - \int_{\kappa_c^1}^{\kappa_c^3} v F_{c,h}(\mathrm{d}v) \\ =& \int_{\kappa_c^2}^{\kappa_c^3} v F_{c,l}(\mathrm{d}v) - \int_{\kappa_c^1}^{\kappa_c^3} v F_{c,h}(\mathrm{d}v) \\ =& \int_{F_{c,l}(\kappa_c^2)}^{F_{c,l}(\kappa_c^3)} F_{c,l}^{-1}(q)\,\mathrm{d}q - \int_{F_{c,h}(\kappa_c^1)}^{F_{c,h}(\kappa_c^3)} F_{c,h}^{-1}(q)\,\mathrm{d}q \\ =& \int_{F_{c,l}(\kappa_c^2)}^{F_{c,l}(\kappa_c^3)} (F_{c,l}^{-1}(q) - F_{c,h}^{-1}(q + F_{c,h}(\kappa_c^3) - F_{c,l}(\kappa_c^3)))\,\mathrm{d}q \\ =& CS(c,l;p^\star)\,,\end{aligned}$$

where the inequality follows from (A.16) and the last equality follows from (8). Likewise,

$$\begin{aligned}
WL(c,l;p) &= \int_0^{\kappa_c^1} (v-c)^+ F_{c,l}(\mathrm{d}v) + \int_{[\kappa_c^1,\kappa_c^3]\times[\kappa_c^5,\infty)} (v_l-c)^+ \rho_c(\mathrm{d}v_l,\mathrm{d}v_h) \\
&\geq \int_0^{\kappa_c^1} (v-c)^+ F_{c,l}(\mathrm{d}v) + \int_{\kappa_c^1}^{\kappa_c^2} (v-c)^+ F_{c,l}(\mathrm{d}v) \\
&= \int_0^{\kappa_c^2} (v-c)^+ F_{c,l}(\mathrm{d}v)\,,
\end{aligned}$$

where the inequality follows from (A.16). Together, we have that

$$CS(c,h;p) = CS(c,h;p^\star)\,;\ \text{ and } WL(c,h;p) = WL(c,h;p^\star)\,,$$

while

$$0 \leq CS(c,l;p) \leq CS(c,l;p^\star)$$

for all $c \in C$. Since

$$\mathbb{E}[(v-c)^+ \mid c] = \pi^\star(c) + \alpha_c(CS(c,h;p) + WL(c,h;p)) + (1-\alpha_c)(CS(c,l;p) + WL(c,l;p))\,,$$

it then follows that

$$WL(c,l;p^\star) \leq WL(c,l;p) \leq \int_0^{\kappa_c^1} (v_l-c)^+ F_{c,l}(\mathrm{d}v_l) + \int_{\kappa_c^1}^{\kappa_c^3} (v-c)^+ (f_{c,l}(v) - f_{c,h}(v))\,\mathrm{d}v\,.$$

It now remains to show that for any c, whenever $CS(c,l;p^\star) > 0$, for all $\sigma_{c,l}$ such that

$$0 \leq \sigma_{c,l} \leq CS(c,l;p^\star)$$

there exists $\rho_c \in \mathcal{R}_c$ that solves (10) such that

$$\int_{[\kappa_c^1,\kappa_c^3]^2} (v_l - v_h)\rho_c(\mathrm{d}v_l,\mathrm{d}v_h) = \sigma_{c,l}$$

To this end, for each $c \in C$, let

$$\tilde{\gamma}_c(v_l \le x \mid v_h) := \begin{cases} \mathbf{1}\{\underline{\Delta}_c^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^3)) \le x\}, & \text{if } v_h \le \kappa_c^1 \\ \mathbf{1}\{v_h \le x\}, & \text{if } v_h \in (\kappa_c^1, \kappa_c^4] \\ \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^4)) \le x\}, & \text{if } v_h \in (\kappa_c^4, \kappa_c^5] \\ \frac{f_{c,l}(v_h)}{f_{c,h}(v_h)} \cdot \mathbf{1}\{v_h \le x\} + \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} \cdot \mathbf{1}\{J_c^{-1}(\Delta_c(\kappa_c^5) - \Delta_c(v_h)) \le x\}, & \text{if } v_h > \kappa_c^5, \end{cases},$$

where

$$J_c(v) := \begin{cases} \min\{F_{c,l}(v), \Delta_c(v) + F_{c,h}(\kappa_c^1)\}, & \text{if } v \le \kappa_c^3 \\ \Delta_c(\kappa_c^3) + F_{c,h}(\kappa_c^3), & \text{if } v > \kappa_c^3 \end{cases}$$

for all $v \in V$. By the same argument as the proof of Lemma A.2, $\tilde{\gamma}_c$ is indeed a transition probability. Then, let

$$\tilde{\rho}_c(v_l \in A, v_h \in B) := \int_B \tilde{\gamma}_c(A \mid v_h) F_{c,h}(\mathrm{d}v_h)\,.$$

Since $\rho_c^\star \in \mathcal{R}_c$ and since $F_{c,l}(\kappa_c^2) = \Delta_c(\kappa_c^5)$, it follows that $\tilde{\rho}_c \in \mathcal{R}_c$ as well. Moreover, for all $(v_l, v_h) \in \operatorname{supp}(\tilde{\rho}_c)$, $\phi_c^\star(v_l) + \psi_c^\star(v_h) = \pi_c(v_l, v_h)$. Thus, by Lemma 2, $\tilde{\rho}_c$ solves (10). In the meantime, by construction, $v_l = v_h$ for all $(v_l, v_h) \in \operatorname{supp}(\tilde{\rho}_c) \cap [\kappa_c^1, \kappa_c^3] \times [\kappa_c^1, \kappa_c^3]$. Therefore,

$$\int_{[\kappa_c^1, \kappa_c^3]^2} (v_l - v_h)\, \mathrm{d}\tilde{\rho}_c = 0\,.$$

Therefore, for any c such that $CS(c, l; p^\star) > 0$ and for any $\sigma_{c,l} \in [0, CS(c, l; p^\star)]$, let

$$\rho_c := \frac{\sigma_{c,l}}{CS(c, l; p^\star)} \rho_c^\star + \left(1 - \frac{\sigma_{c,l}}{CS(c, l; p^\star)}\right) \tilde{\rho}_c\,.$$

Since $\mathcal{R}_c$ is convex and since both $\rho_c^\star$ and $\tilde{\rho}_c$ are solutions of (10), ρ_c is in $\mathcal{R}_c$ solves (10) as well. Moreover,

$$\begin{aligned} \int_{[\kappa_c^1, \kappa_c^3]^2} (v_l - v_h)\, \mathrm{d}\rho_c =& \frac{\sigma_{c,l}}{CS(c, l; p^\star)} \int_{[\kappa_c^1, \kappa_c^3]^2} (v_l - v_h)\, \mathrm{d}\rho_c^\star + \left(1 - \frac{\sigma_{c,l}}{CS(c, l; p^\star)}\right) \int_{[\kappa_c^1, \kappa_c^3]^2} (v_l - v_h)\, \mathrm{d}\tilde{\rho}_c \\ =& CS(c, l; p^\star) \cdot \frac{\sigma_{c,l}}{CS(c, l; p^\star)} + 0 \cdot \left(1 - \frac{\sigma_{c,l}}{CS(c, l; p^\star)}\right) \\ =& \sigma_{c,l}\,, \end{aligned}$$

as desired. This completes the proof. $\square$

Proof of Corollary 2. For any $\tilde{c} \in C_2 \cup C_3$, $F_{\tilde{c},l}(\tilde{c}) \geq \|F_{\tilde{c},l} - F_{\tilde{c},h}\|$, so the dual variables in the proof of Theorem 2 are $\phi^\star_{\tilde{c}}(v_l) = (1 - \alpha_{\tilde{c}})(v_l - \tilde{c})^+$ and $\psi^\star_{\tilde{c}}(v_h) = \alpha_{\tilde{c}}(v_h - \tilde{c})^+$. Hence the dual value equals

$$\begin{aligned}\mathbb{E}[\phi^\star_{\tilde{c}}(v_l) + \psi^\star_{\tilde{c}}(v_h) \mid c = \tilde{c}] =& (1 - \alpha_{\tilde{c}})\mathbb{E}[(v - c)^+ \mid c = \tilde{c}, \theta = l] + \alpha_{\tilde{c}}\mathbb{E}[(v - c)^+ \mid c = \tilde{c}, \theta = h] \\ =& \mathbb{E}[(v - c)^+ \mid c = \tilde{c}] .\end{aligned}$$

By Lemma A.3 and weak duality, the seller's profit under $p^\star$ conditional on $c = \tilde{c}$ equals this dual value, yielding the first display. Since $CS(c, \theta; p^\star), WL(c, \theta; p^\star) \geq 0$ and the seller's profit cannot exceed total gains from trade, the second display follows. $\square$

A.2 Proofs of Auxiliary Lemmas

Proof of Lemma A.1. We show that $\phi^\star_c(v_l)+\psi^\star_c(v_h) \geq \pi_c(v_l, v_h)$ for all (v_l, v_h) by discussing all cases separately. Suppose first that $F_{c,l}(c) < \Delta_c(v^\star_c) = \|F_{c,l} - F_{c,h}\|$.

Case 1: $v_l \leq \kappa^3_c$.
When $v_h \leq \kappa^1_c$, $\pi_c(v_l, v_h)$ either equals $\min\{v_l, v_h\} - c \leq v_h - c \leq \kappa^1_c - c$, or $\alpha_c(v_h - c)^+ \leq \kappa^1_c - c$, or $(1 - \alpha_c)(v_l - c)^+ \leq (1 - \alpha_c)(\kappa^3_c - c) = \kappa^1_c - c$. Therefore,

$$\phi^\star_c(v_l) + \psi^\star_c(v_h) = \kappa^1_c - c \geq \pi_c(v_l, v_h) .$$

When $v_h \in (\kappa^1_c, \kappa^3_c]$, $v_h - c > \kappa^1_c - c = (1 - \alpha_c)(\kappa^3_c - c) \geq (1 - \alpha_c)(v_l - c)^+$ and hence $\min\{v_h, v_l\} - c \geq (1 - \alpha_c)(v_l - c)^+$. Therefore, $\pi_c(v_l, v_h)$ either equals $\min\{v_l, v_h\} - c \leq v_h - c$ or $\alpha_c(v_h - c) \leq v_h - c$. As a result,

$$\phi^\star_c(v_l) + \psi^\star_c(v_h) = v_h - c \geq \pi_c(v_l, v_h) .$$

When $v_h \in (\kappa^3_c, \kappa^4_c]$, $v_h > v_l$, and therefore $\pi_c(v_l, v_h) = \max\{v_l - c, \alpha_c(v_h - c)\}$. Moreover, since $(1 - \alpha_c)(\kappa^3_c - c) = \kappa^1_c - c$, we have $v_l - c \leq \kappa^3_c - c = \alpha_c(\kappa^3_c - c) + \kappa^1_c - c \leq \alpha_c(v_h - c) + \kappa^1_c - c$. Thus,

$$\phi^\star_c(v_l) + \psi^\star_c(v_h) = \alpha_c(v_h - c) + \kappa^1_c - c \geq \max\{v_l - c, \alpha_c(v_h - c)\} = \pi_c(v_l, v_h) .$$

When $v_h \in (\kappa^4_c, \kappa^5_c]$, $v_h > v_l$, and therefore $\pi_c(v_l, v_h) = \max\{v_l - c, \alpha_c(v_h - c)\}$. As argued above, we have $v_l - c \le \kappa_c^3 - c = \alpha_c(\kappa_c^3 - c) + \kappa_c^1 - c \le \alpha_c(\kappa_c^4 - c) + (\kappa_c^1 - c)$. Furthermore, by (8),

$$\alpha_c(\kappa_c^4 - c) + \kappa_c^1 - c = \alpha_c(\kappa_c^5 - c) \ge \alpha_c(v_h - c)\,.$$

Together,

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = \alpha_c(\kappa_c^4 - c) + \kappa_c^1 - c \ge \max\{v_l - c, \alpha_c(v_h - c)\} = \pi_c(v_l, v_h)\,.$$

When $v_h > \kappa_c^5$, note that since $\kappa_c^1 - c = (1-\alpha_c)(\kappa_c^3 - c) \le (1-\alpha_c)(\kappa_c^4 - c)$, and since $\alpha_c(\kappa_c^5 - \kappa_c^4) = \kappa_c^1 - c$, we have that $\alpha_c(v_h - c) \ge \alpha_c(\kappa_c^5 - c) = \alpha_c(\kappa_c^4 - c) + (1-\alpha_c)(\kappa_c^3 - c) \ge \kappa_c^3 - c \ge v_l - c$. Therefore, $\pi_c(v_l, v_h) = \alpha_c(v_h - c)$, and hence

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = \alpha_c(v_h - c) = \pi_c(v_l, v_h)\,.$$

Case 2: $v_l \in (\kappa_c^3, \kappa_c^4]$.
When $v_h \le \kappa_c^1$, $(1-\alpha_c)(v_l - c) > (1-\alpha_c)(\kappa_c^3 - c) = \kappa_c^1 - c \ge v_h - c$, and hence $\pi_c(v_l, v_h) = (1-\alpha_c)(v_l - c)$. Therefore,

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = (1-\alpha_c)(v_l - c) = \pi_c(v_l, v_h)\,.$$

When $v_h \in (\kappa_c^1, \kappa_c^3]$, $v_h \le \kappa_c^3 < v_l$ and hence $\min\{v_l, v_h\} - c = v_h - c$. Therefore, $\pi_c(v_l, v_h) = \max\{v_h - c, (1-\alpha_c)(v_l - c)\}$. Since $v_h - \kappa_c^1 + (1-\alpha_c)(v_l - c) > v_h - \kappa_c^1 + (1-\alpha_c)(\kappa_c^3 - c) = v_h - c$, we have

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = v_h - \kappa_c^1 + (1-\alpha_c)(v_l - c) \ge \max\{v_h - c, (1-\alpha_c)(v_l - c)\}\,.$$

When $v_h \in (\kappa_c^3, \kappa_c^4]$,

$$\begin{aligned}\phi_c^\star(v_l) + \psi_c^\star(v_h) =& \alpha_c(v_h - c) + (1-\alpha_c)(v_l - c)\\ &\ge \max\{\min\{v_l, v_h\} - c, (1-\alpha_c)(v_l - c), \alpha_c(v_h - c)\}\,.\end{aligned}$$

When $v_h \in (\kappa_c^4, \kappa_c^5]$, since $\kappa_c^1 - c = \alpha_c(\kappa_c^5 - \kappa_c^4)$ and since $(1-\alpha_c)(\kappa_c^3 - c) = \kappa_c^1 - c$, we have

$$\begin{aligned}\alpha_c(v_h - c) \leq \alpha_c(\kappa_c^5 - c) =& \alpha_c(\kappa_c^4 - c) + \kappa_c^1 - c \\ =& \alpha_c(\kappa_c^4 - c) + (1-\alpha_c)(\kappa_c^3 - c) \\ \leq& \alpha_c(\kappa_c^4 - c) + (1-\alpha_c)(v_l - c)\,.\end{aligned}$$

Together,

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = \alpha_c(\kappa_c^4 - c) + (1-\alpha_c)(v_l - c) \geq \max\{v_l - c, \alpha_c(v_h - c)\} = \pi_c(v_l, v_h)\,.$$

When $v_h > \kappa_c^5$, since $\kappa_c^1 - c = \alpha_c(\kappa_c^5 - \kappa_c^4)$, we have

$$\alpha_c(v_h - c) - (\kappa_c^1 - c) \geq \alpha_c(\kappa_c^5 - c) - (\kappa_c^1 - c) = \alpha_c(\kappa_c^4 - c) \geq \alpha_c(v_l - c)\,.$$

Therefore,

$$\alpha_c(v_h - c) - (\kappa_c^1 - c) + (1-\alpha_c)v_l \geq v_l - c\,.$$

Meanwhile, since $(1-\alpha_c)(\kappa_c^3 - c) = \kappa_c^1 - c$,

$$(1-\alpha_c)(v_l - c) \geq (1-\alpha_c)(\kappa_c^3 - c) = \kappa_c^1 - c\,,$$

and thus

$$\alpha_c(v_h - c) - (\kappa_c^1 - c) + (1-\alpha_c)(v_l - c) \geq \alpha_c(v_h - c)\,.$$

Together,

$$\begin{aligned}\phi_c^\star(v_l) + \psi_c^\star(v_h) =& \alpha_c(v_h - c) - (\kappa_c^1 - c) + (1-\alpha_c)(v_l - c) \\ \geq& \max\{v_l - c, \alpha_c(v_h - c)\} \\ =& \pi_c(v_l, v_h)\,.\end{aligned}$$

Case 3: $v_l \in (\kappa_c^4, \kappa_c^5]$.
When $v_h \leq \kappa_c^1$, we have

$$(1-\alpha_c)(v_l - c) > (1-\alpha_c)(\kappa_c^4 - c) \geq (1-\alpha_c)(\kappa_c^3 - c) = \kappa_c^1 - c \geq (v_h - c)^+\,.$$

Therefore, $\pi_c(v_l, v_h) = (1-\alpha_c)(v_l - c)$ and hence

$$\begin{aligned}\phi_c^\star(v_l) + \psi_c^\star(v_h) = v_l - c - \alpha_c(\kappa_c^4 - c) =& \alpha_c(v_l - \kappa_c^4) + (1-\alpha_c)(v_l - c) \\ >& (1-\alpha_c)(v_l - c) \\ =& \pi_c(v_l, v_h)\,.\end{aligned}$$

When $v_h \in (\kappa_c^1, \kappa_c^3]$,

$$v_l - c - \alpha_c(\kappa_c^4 - c) + v_h - \kappa_c^1 > v_l - c - \alpha_c(\kappa_c^4 - c) > (1-\alpha_c)(v_l - c)\,.$$

Moreover, since $(1-\alpha_c)(\kappa_c^3 - c) = \kappa_c^1 - c$ and since $\kappa_c^4 \geq \kappa_c^3$,

$$\begin{aligned}v_l - c - \alpha_c(\kappa_c^4 - c) + v_h - \kappa_c^1 =& v_l - c - \alpha_c(\kappa_c^4 - c) + v_h - c - (\kappa_c^1 - c) \\ >& (1-\alpha_c)(\kappa_c^4 - c) + v_h - c - (\kappa_c^1 - c) \\ \geq& (1-\alpha_c)(\kappa_c^3 - c) + v_h - c - (\kappa_c^1 - c) \\ =& v_h - c\,.\end{aligned}$$

Together,

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = v_l - c - \alpha_c(\kappa_c^4 - c) + v_h - \kappa_c^1 \geq \max\{v_h - c, (1-\alpha_c)(v_l - c)\} = \pi_c(v_l, v_h)\,.$$

When $v_h \in (\kappa_c^3, \kappa_c^4]$, since $v_l > \kappa_c^4 \geq v_h$,

$$v_l - c - \alpha_c(\kappa_c^4 - c) + \alpha_c(v_h - c) > (1-\alpha_c)(\kappa_c^4 - c) + \alpha_c(v_h - c) \geq v_h - c$$

In the meantime, since $v_h > \kappa_c^3 \geq \kappa_c^1 \geq c$,

$$v_l - c - \alpha_c(\kappa_c^4 - c) + \alpha_c(v_h - c) \geq v_l - c - \alpha_c(\kappa_c^4 - c) > (1-\alpha_c)(v_l - c)\,.$$

Therefore,

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = v_l - c - \alpha_c(\kappa_c^4 - c) + \alpha_c(v_h - c) \geq \max\{v_h - c, (1-\alpha_c)(v_l - c)\} = \pi_c(v_l, v_h)\,.$$

When $v_h \in (\kappa_c^4, \kappa_c^5]$, since $\kappa_c^1 - c = \alpha_c(\kappa_c^5 - \kappa_c^4)$ and since $\kappa_c^4 - c \geq \kappa_c^3 - c = {(\kappa_c^1 - c)}/{(1-\alpha_c)}$, we have

$$\alpha_c(\kappa_c^5 - c) = \alpha_c(\kappa_c^4 - c) + \kappa_c^1 - c = \alpha_c(\kappa_c^4 - c) + (1 - \alpha_c)(\kappa_c^3 - c) \leq \kappa_c^4 - c\,.$$

Therefore,

$$\alpha_c(v_h - c) \leq \alpha_c(\kappa_c^5 - c) \leq \kappa_c^4 - c \leq v_l - c\,,$$

and hence $\pi_c(v_l, v_h) = \max\{\min\{v_l, v_h\} - c, (1 - \alpha_c)(v_l - c)\}$. Together,

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = v_l - c \geq \max\{\min\{v_l, v_h\} - c, (1 - \alpha_c)(v_l - c)\} = \pi_c(v_l, v_h)\,.$$

When $v_h > \kappa_c^5$, since $\kappa_c^4 - c \geq \kappa_c^3 - c = {(\kappa_c^1 - c)}/{(1-\alpha_c)}$,

$$\begin{aligned}
&\alpha_c(v_h - c) + v_l - c - \alpha_c(\kappa_c^4 - c) - (\kappa_c^1 - c) \\
\geq &\alpha_c(v_h - c) + (1 - \alpha_c)(\kappa_c^4 - c) - (\kappa_c^1 - c) \\
\geq &\alpha_c(v_h - c) + (1 - \alpha_c)(\kappa_c^3 - c) - (\kappa_c^1 - c) \\
= &\alpha_c(v_h - c)\,.
\end{aligned}$$

Moreover, since $\kappa_c^1 - c = \alpha_c(\kappa_c^5 - \kappa_c^4)$,

$$\begin{aligned}
&\alpha_c(v_h - c) + v_l - c - (\kappa_c^1 - c) - \alpha_c(\kappa_c^4 - c) \\
> &\alpha_c(\kappa_c^5 - c) + v_l - c - (\kappa_c^1 - c) - \alpha_c(\kappa_c^4 - c) \\
= &\alpha_c(\kappa_c^5 - \kappa_c^4) - (\kappa_c^1 - c) + v_l - c \\
= &v_l - c\,.
\end{aligned}$$

Together,

$$\begin{aligned}
\phi_c^\star(v_l) + \psi_c^\star(v_h) = &\alpha_c(v_h - c) - (\kappa_c^1 - c) + v_l - c - \alpha_c(\kappa_c^4 - c) \\
\geq &\max\{v_l - c, \alpha_c(v_h - c)\} \\
= &\pi_c(v_l, v_h)\,.
\end{aligned}$$

Case 4: $v_l > \kappa_c^5$.

When $v_h \leq \kappa_c^1$, $(v_h - c)^+ \leq \kappa_c^1 - c = (1 - \alpha_c)(\kappa_c^3 - c) \leq (1 - \alpha_c)(v_l - c)$. Thus, $\pi_c(v_l, v_h) =$

$(1-\alpha_c)(v_l-c)$, and hence,

$$\phi_c^\star(v_l)+\psi_c^\star(v_h)=(1-\alpha_c)(v_l-c)+\kappa_c^1-c\geq(1-\alpha_c)(v_l-c)=\pi_c(v_l,v_h)\,.$$

When $v_h\in(\kappa_c^1,\kappa_c^3]$, $v_h\leq\kappa_c^3\leq\kappa_c^4<v_l$. Thus,

$$\phi_c^\star(v_l)+\psi_c^\star(v_h)=v_h-c+(1-\alpha_c)(v_l-c)\geq\max\{v_h-c,(1-\alpha_c)(v_l-c)\}=\pi_c(v_l,v_h)\,.$$

When $v_h\in(\kappa_c^3,\kappa_c^4]$, since $\kappa_c^1\geq c$,

$$\begin{aligned}\phi_c^\star(v_l)+\psi_c^\star(v_h)=&\alpha_c(v_h-c)+(1-\alpha_c)(v_l-c)+\kappa_c^1-c\\ \geq&\alpha_c(v_h-c)+(1-\alpha_c)(v_l-c)\\ \geq&\max\{\min\{v_l,v_h\}-c,\alpha_c(v_h-c),(1-\alpha_c)(v_l-c)\}\\ =&\pi_c(v_l,v_h)\,.\end{aligned}$$

When $v_h\in(\kappa_c^4,\kappa_c^5]$, since $\kappa_c^1-c=\alpha_c(\kappa_c^5-\kappa_c^4)$.

$$\begin{aligned}\phi_c^\star(v_l)+\psi_c^\star(v_h)=&\alpha_c(\kappa_c^4-c)+(1-\alpha_c)(v_l-c)+\kappa_c^1-c\\ =&\alpha_c(\kappa_c^5-c)-(\kappa_c^1-c)+(1-\alpha_c)(v_l-c)+\kappa_c^1-c\\ =&\alpha_c(\kappa_c^5-c)+(1-\alpha_c)(v_l-c)\\ \geq&\alpha_c(v_h-c)+(1-\alpha_c)(v_l-c)\\ \geq&\max\{\min\{v_l,v_h\}-c,\alpha_c(v_h-c),(1-\alpha_c)(v_l-c)\}\\ =&\pi_c(v_l,v_h)\,.\end{aligned}$$

When $v_h>\kappa_c^5$,

$$\begin{aligned}\phi_c^\star(v_l)+\psi_c^\star(v_h)=&\alpha_c(v_h-c)+(1-\alpha_c)(v_l-c)\\ \geq&\max\{\min\{v_l,v_h\}-c,\alpha_c(v_h-c),(1-\alpha_c)(v_l-c)\}\\ =&\pi_c(v_l,v_h)\,.\end{aligned}$$

Together, it follows that

$$\phi_c^\star(v_l)+\psi_c^\star(v_h)\geq\pi_c(v_l,v_h)\,,$$

for all $v_l, v_h \geq 0$, as desired.

Now suppose that $F_{c,l}(c) \geq \|F_{c,l} - F_{c,h}\|$. Then,

$$\begin{aligned}
\phi_c^\star(v_l) + \psi_c^\star(v_h) =& (1-\alpha_c)(v_l - c)^+ + \alpha_c(v_h - c)^+ \\
\geq& \max\{(1-\alpha_c)(v_l - c)^+, \alpha_c(v_h - c)^+, \min\{v_l, v_h\} - c\} \\
=& \pi_c(v_l, v_h)\,,
\end{aligned}$$

for all $v_l, v_h \in V$. This completes the proof. ☐

Proof of Lemma A.2. We first note that for all $c \in C$, $\gamma_c^\star(\cdot \mid v_h)$ is a probability measure for all $v_h \in V$. Indeed, for all $v_h \in V$ $\lim_{x\to\infty} \gamma_c^\star(v_l \leq x \mid v_h) = 1$ and $\gamma_c^\star(v_l < 0 \mid v_h) = 0$; $x \mapsto \gamma^\star(v_l \leq x \mid v_h)$ is right-continuous. Moreover, for any $c \in C$ and for any measurable set $A \subseteq V$, $\gamma_c^\star(A \mid \cdot)$ is a measurable function. Therefore, $\gamma_c^\star$ is a transition probability for all $c \in C$.

Next, we show that the marginals of $\rho_c^\star$ equal $F_{c,l}$ and $F_{c,h}$, respectively. By construction,

$$\rho_c^\star(v_l \in V, v_h \leq x) = \int_0^x F_{c,h}(\mathrm{d}v_h) = F_{c,h}(x)\,.$$

To show that $\rho_c^\star(v_l \leq x, v_h \in V) = F_{c,l}(x)$ for all $c \in C$ and for all $x \in V$, consider first the case when $F_{c,l}(c) < \Delta_c(v_c^\star)$. For all $x \leq \kappa_c^2$,

$$\gamma_c^\star(v_l \leq x \mid v_h) = \begin{cases} 0, & \text{if } v_h \leq \kappa_c^5 \\ \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} \mathbf{1}\{F_{c,l}^{-1}(\Delta_c(\kappa_c^5) - \Delta_c(v_h)) \leq x\}, & \text{if } v_h > \kappa_c^5 \end{cases}.$$

Note that the derivative of $v \mapsto \Delta_c(\kappa_c^5) - \Delta_c(v)$ equals $f_{c,h} - f_{c,l}$. Therefore,

$$\begin{aligned}
\int_0^\infty \gamma_c^\star(v_l \leq x \mid v_h) F_{c,h}(\mathrm{d}v_h) &= \int_{\kappa_c^5}^\infty \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} \mathbf{1}\{F_{c,l}^{-1}(\Delta_c(\kappa_c^5) - \Delta_c(v_h)) \leq x\} F_{c,h}(\mathrm{d}v_h) \\
&= \int_{\kappa_c^5}^\infty (f_{c,h}(v_h) - f_{c,l}(v_h)) \mathbf{1}\{\Delta_c(\kappa_c^5) - \Delta_c(v_h) \leq F_{c,l}(x)\}\,\mathrm{d}v_h \\
&= \int_0^{\Delta_c(\kappa_c^5)} \mathbf{1}\{z \leq F_{c,l}(x)\}\,\mathrm{d}z \\
&= \min\{\Delta_c(\kappa_c^5), F_{c,l}(x)\} \\
&= F_{c,l}(x)\,,
\end{aligned}$$

where the third equality follows from changing variables of integration, and the last equality follows from $F_{c,l}(x) \le F_{c,l}(\kappa_c^2) = \Delta_c(\kappa_c^5)$, which in turn follows from (8).

For all $x \in (\kappa_c^2, \kappa_c^3]$,

$$\gamma_c^\star(v_l \le x \mid v_h) = \begin{cases} 0, & \text{if } v_h \in [0, \kappa_c^1] \cup (\kappa_c^3, \kappa_c^5] \\ \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^3)) \le x\}, & \text{if } v_h \in (\kappa_c^1, \kappa_c^3] \\ \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)}, & \text{if } v_h > \kappa_c^5 \end{cases}.$$

Therefore,

$$\begin{aligned}
&\int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
&= \int_{\kappa_c^1}^{\kappa_c^3} \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^3)) \le x\} F_{c,h}(\mathrm{d}v_h) + \int_{\kappa_c^5}^\infty \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} F_{c,h}(\mathrm{d}v_h) \\
&= \int_{\kappa_c^1}^{\kappa_c^3} \mathbf{1}\{F_{c,h}(v_h) \le F_{c,l}(x) - \Delta_c(\kappa_c^3)\} F_{c,h}(\mathrm{d}v_h) + F_{c,l}(\kappa_c^5) - F_{c,h}(\kappa_c^5) \\
&= \int_{F_{c,h}(\kappa_c^1)}^{F_{c,h}(\kappa_c^3)} \mathbf{1}\{z \le F_{c,l}(x) - \Delta_c(\kappa_c^3)\} \,\mathrm{d}z + \Delta_c(\kappa_c^5) \\
&= \min\{F_{c,l}(x) - \Delta_c(\kappa_c^3), F_{c,h}(\kappa_c^3)\} - F_{c,h}(\kappa_c^1) + \Delta_c(\kappa_c^5) \\
&= F_{c,l}(x) - \Delta_c(\kappa_c^3) - F_{c,h}(\kappa_c^1) + \Delta_c(\kappa_c^5) \\
&= F_{c,l}(x)\,,
\end{aligned}$$

where the third equality follows from changing variables for integration, the fifth equality follows from $F_{c,l}(x) - \Delta_c(\kappa_c^3) \le F_{c,l}(\kappa_c^3) - \Delta_c(\kappa_c^3) = F_{c,h}(\kappa_c^3)$, and the last equality also follows from (8).

For all $x \in (\kappa_c^3, \kappa_c^4]$,

$$\gamma_c^\star(v_l \le x \mid v_h) = \begin{cases} \mathbf{1}\{\underline{\Delta}_c^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^3)) \le x\}, & \text{if } v_h \le \kappa_c^1 \\ 1, & \text{if } v_h \in (\kappa_c^1, \kappa_c^3] \\ \mathbf{1}\{v_h \le x\}, & \text{if } v_h \in (\kappa_c^3, \kappa_c^4] \\ 0, & \text{if } v_h \in (\kappa_c^4, \kappa_c^5] \\ \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)}, & \text{if } v_h > \kappa_c^5 \end{cases}.$$

Thus,

$$
\begin{aligned}
&\int_0^{\infty} \gamma_c^{\star}(v_l \leq x \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
=&\int_0^{\kappa_c^1} \mathbf{1}\{\underline{\Delta}_c^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^3)) \leq x\} F_{c,h}(\mathrm{d}v_h) + F_{c,h}(\kappa_c^3) - F_{c,h}(\kappa_c^1) \\
&+ \int_{\kappa_c^3}^{\kappa_c^4} \mathbf{1}\{v_h \leq x\} F_{c,h}(\mathrm{d}v_h) + \int_{\kappa_c^5}^{\infty} \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} F_{c,h}(\mathrm{d}v_h) \\
=&\int_0^{\kappa_c^1} \mathbf{1}\{F_{c,h}(v_h) + \Delta_c(\kappa_c^3) \leq \Delta_c(x)\} F_{c,h}(\mathrm{d}v_h) \\
&+ F_{c,h}(\kappa_c^3) - F_{c,h}(\kappa_c^1) + F_{c,h}(x) - F_{c,h}(\kappa_c^3) + F_{c,l}(\kappa_c^5) - F_{c,h}(\kappa_c^5) \\
=&\int_0^{F_{c,h}(\kappa_c^1)} \mathbf{1}\{z \leq \Delta_c(x) - \Delta_c(\kappa_c^3)\} \,\mathrm{d}z + F_{c,h}(x) - F_{c,h}(\kappa_c^1) + \Delta_c(\kappa_c^5) \\
=&\min\{\Delta_c(x) - \Delta_c(\kappa_c^3), F_{c,h}(\kappa_c^1)\} + F_{c,h}(x) - F_{c,h}(\kappa_c^1) + \Delta_c(\kappa_c^5) \\
=&\Delta_c(x) - \Delta_c(\kappa_c^3) + F_{c,h}(x) - F_{c,h}(\kappa_c^1) + \Delta_c(\kappa_c^5) \\
=&F_{c,l}(x) - (\Delta_c(\kappa_c^3) + F_{c,h}(\kappa_c^1)) + \Delta_c(\kappa_c^5)\,,
\end{aligned}
$$

where the third equality follows from changing variables of the integration, the fifth equality follows from $\Delta_c(x) - \Delta_c(\kappa_c^3) \leq \Delta_c(\kappa_c^4) - \Delta_c(\kappa_c^3) = F_{c,h}(\kappa_c^1)$, which in turn follows from (8); whereas the last equality also follows from (8).

For all $x \in (\kappa_c^4, \kappa_c^5]$,

$$
\gamma_c^{\star}(v_l \leq x \mid v_h) = \begin{cases} 1, & \text{if } v_h \leq \kappa_c^4 \\ \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^4)) \leq x\}, & \text{if } v_h \in (\kappa_c^4, \kappa_c^5] \\ \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)}, & \text{if } v_h > \kappa_c^5 \end{cases}.
$$

Thus,

$$
\begin{aligned}
&\int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
=&\int_0^{\kappa_c^4} 1 F_{c,h}(\mathrm{d}v) + \int_{\kappa_c^4}^{\kappa_c^5} \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^4)) \le x\} F_{c,h}(\mathrm{d}v_h) + \int_{\kappa_c^5}^\infty \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} F_{c,h}(\mathrm{d}v_h) \\
=&F_{c,h}(\kappa_c^4) + \int_{\kappa_c^4}^{\kappa_c^5} \mathbf{1}\{F_{c,h}(v_h) \le F_{c,l}(x) - \Delta_c(\kappa_c^4)\} F_{c,h}(\mathrm{d}v_h) + \int_{\kappa_c^5}^\infty [f_{c,h}(v_h) - f_{c,l}(v_h)] \, \mathrm{d}v_h \\
=&F_{c,h}(\kappa_c^4) + \min\{F_{c,l}(x) - \Delta_c(\kappa_c^4), F_{c,h}(\kappa_c^5)\} - F_{c,h}(\kappa_c^4) + \Delta_c(\kappa_c^5) \\
=&F_{c,h}(\kappa_c^4) + F_{c,l}(x) - \Delta_c(\kappa_c^4) - F_{c,h}(\kappa_c^4) + \Delta_c(\kappa_c^5) \\
=&F_{c,l}(x) \, ,
\end{aligned}
$$

where the third equality follows from changing variables of the integration, the fourth equality follows from $F_{c,l}(x) - \Delta_c(\kappa_c^4) \le F_{c,l}(\kappa_c^5) - \Delta_c(\kappa_c^4) = F_{c,l}(\kappa_c^5) - \Delta_c(\kappa_c^5) = F_{c,h}(\kappa_c^5)$, which in turn follows from (8); and the last equality follows from (8).

For all $x > \kappa_c^5$,

$$
\begin{cases}
1, & \text{if } v_h \le \kappa_c^5 \\
\frac{f_{c,l}(v_h)}{f_{c,h}(v_h)} \mathbf{1}\{v_h \le x\} + \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)}, & \text{if } v_h > \kappa_c^5
\end{cases} .
$$

Therefore,

$$
\begin{aligned}
&\int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
=&\int_0^{\kappa_c^5} 1 F_{c,h}(\mathrm{d}v_h) + \int_{\kappa_c^5}^\infty \left(\frac{f_{c,l}(v_h)}{f_{c,h}(v_h)} \mathbf{1}\{v_h \le x\} + \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} \right) F_{c,h}(\mathrm{d}v_h) \\
=&F_{c,h}(\kappa_c^5) + \int_{\kappa_c^5}^x f_{c,l}(v_h) \, \mathrm{d}v_h + \int_{\kappa_c^5}^\infty [f_{c,h}(v_h) - f_{c,l}(v_h)] \, \mathrm{d}v_h \\
=&F_{c,h}(\kappa_c^5) + F_{c,l}(x) - F_{c,l}(\kappa_c^5) + F_{c,l}(\kappa_c^5) - F_{c,h}(\kappa_c^5) \\
=&F_{c,l}(x) \, .
\end{aligned}
$$

Together, we have that

$$
\rho_c^\star(v_l \le x, v_h \in V) = \int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) = F_{c,l}(x) \, ,
$$

for all $x \in V$ and hence $\rho_c^\star \in \mathcal{R}_c$ for all $c \in C$ such that $F_{c,l}(c) < \|F_{c,l} - F_{c,h}\|$.

Now consider the case when $F_{c,l}(c) \geq \|F_{c,l} - F_{c,h}\|$ and $c \leq v_c^\star$. If $x \leq \eta_c^l$, then

$$\gamma_c^\star(v_l \leq x \mid v_h) = \begin{cases} 0, & \text{if } v_h > \eta_c^h \\ \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h)) \leq x\}, & \text{if } v_h \leq \eta_c^h \end{cases}.$$

Thus,

$$\begin{aligned}
\int_0^\infty \gamma_c^\star(v_l \leq x \mid v_h) F_{c,h}(\mathrm{d}v_h) &= \int_0^{\eta_c^h} \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h)) \leq x\} F_{c,h}(\mathrm{d}v_h) \\
&= \int_0^{\eta_c^h} \mathbf{1}\{F_{c,h}(v_h) \leq F_{c,l}(x)\} F_{c,h}(\mathrm{d}v_h) \\
&= \int_0^{F_{c,h}(\eta_c^h)} \mathbf{1}\{z \leq F_{c,l}(x)\}\,\mathrm{d}z \\
&= \min\{F_{c,h}(\eta_c^h), F_{c,l}(x)\} \\
&= F_{c,l}(x)\,,
\end{aligned}$$

where the third equality follows from changing variables of the integration, and the last equality follows from $F_{c,l}(x) \leq F_{c,l}(\eta_c^l) = F_{c,h}(\eta_c^h)$.

If $x \in (\eta_c^l, c]$, then

$$\gamma_c^\star(v_l \leq x \mid v_h) = \begin{cases} 1, & \text{if } v_h \leq \eta_c^h \\ \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} \mathbf{1}\{F_{c,l}^{-1}(\Delta_c(v_c^\star) - \Delta_c(v_h) + F_{c,l}(\eta_c^l)) \leq x\}, & \text{if } v_h > v_c^\star \\ 0, & \text{otherwise} \end{cases}.$$

Thus,

$$
\begin{aligned}
&\int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
=&F_{c,h}(\eta_c^h) + \int_{v_c^\star}^\infty \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} \mathbf{1}\{F_{c,l}^{-1}(\Delta_c(v_c^\star) - \Delta_c(v_h) + F_{c,l}(\eta_c^l)) \le x\} F_{c,h}(\mathrm{d}v_h) \\
=&F_{c,h}(\eta_c^h) + \int_{v_c^\star}^\infty (f_{c,h}(v_h) - f_{c,l}(v_h)) \mathbf{1}\{\Delta_c(v_c^\star) - \Delta_c(v_h) \le F_{c,l}(x) - F_{c,l}(\eta_c^l)\} \,\mathrm{d}v_h \\
=&F_{c,h}(\eta_c^h) + \int_0^{\Delta_c(v_c^\star)} \mathbf{1}\{z \le F_{c,l}(x) - F_{c,l}(\eta_c^l)\} \,\mathrm{d}z \\
=&F_{c,h}(\eta_c^h) + \min\{\Delta_c(v_c^\star), F_{c,l}(x) - F_{c,l}(\eta_c^l)\} \\
=&F_{c,h}(\eta_c^h) + F_{c,l}(x) - F_{c,l}(\eta_c^l) \\
=&F_{c,l}(x)\,,
\end{aligned}
$$

where the third equality follows from changing variables of the integration, the fifth equality follows from $F_{c,l}(x) - F_{c,l}(\eta_c^l) \le F_{c,l}(c) - F_{c,l}(\eta_c^l) = \Delta_c(v_c^\star)$, which in turn follows from the definition of η_c^l; and the last equality follows from $F_{c,h}(\eta_c^h) = F_{c,l}(\eta_c^l)$.

If $x \in (c, v_c^\star]$, then

$$
\gamma_c^\star(v_l \le x \mid v_h) = \begin{cases} 1, & \text{if } v_h \le \eta_c^h \\ \mathbf{1}\{\underline{\Delta}_c^{-1}(F_{c,h}(v_h) - F_{c,h}(\eta_c^h) + \Delta_c(c)) \le x\}, & \text{if } v_h \in (\eta_c^h, c] \\ \mathbf{1}\{v_h \le x\}, & \text{if } v_h \in (c, v_c^\star] \\ \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)}, & \text{if } v_h > v_c^\star \end{cases}.
$$

Thus,

$$
\begin{aligned}
&\int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
=&\int_0^{\eta_c^h} F_{c,h}(\mathrm{d}v_h) + \int_{\eta_c^h} \mathbf{1}\{\underline{\Delta}_c^{-1}(F_{c,h}(v_h) - F_{c,h}(\eta_c^h) + \Delta_c(c)) \le x\} F_{c,h}(\mathrm{d}v_h) \\
&+ \int_c^{v_c^\star} \mathbf{1}\{v_h \le x\} F_{c,h}(\mathrm{d}v_h) + \int_{v_c^\star}^\infty \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} F_{c,h}(\mathrm{d}v_h) \\
=&F_{c,h}(\eta_c^h) + \int_{\eta_c^h}^c \mathbf{1}\{F_{c,h}(v_h) - F_{c,h}(\eta_c^h) \le \Delta_c(x) - \Delta_c(c))\} F_{c,h}(\mathrm{d}v_h) + F_{c,h}(x) - F_{c,h}(c) + \Delta_c(v_c^\star) \\
=&F_{c,h}(\eta_c^h) + \int_0^{F_{c,h}(c) - F_{c,h}(\eta_c^h)} \mathbf{1}\{z \le \Delta_c(x) - \Delta_c(c)\}\, \mathrm{d}z + F_{c,h}(x) - F_{c,h}(c) + \Delta_c(v_c^\star) \\
=&F_{c,h}(\eta_c^h) + \min\{\Delta_c(x) - \Delta_c(c), F_{c,h}(c) - F_{c,h}(\eta_c^h)\} + F_{c,h}(x) - F_{c,h}(c) + \Delta_c(v_c^\star) \\
=&F_{c,h}(\eta_c^h) + \Delta_c(x) - \Delta_c(c) + F_{c,h}(x) - F_{c,h}(c) + \Delta_c(v_c^\star) \\
=&F_{c,h}(\eta_c^h) + F_{c,l}(x) + \Delta_c(v_c^\star) - F_{c,l}(c) \\
=&F_{c,h}(\eta_c^h) + F_{c,l}(x) - F_{c,l}(\eta_c^l) \\
=&F_{c,l}(x)\,,
\end{aligned}
$$

where the third equality follows from changing variables of the integration, the fifth equality follow from $\Delta_c(x) - \Delta_c(c) \le \Delta_c(v_c^\star) - \Delta_c(c) = F_{c,h}(c) - F_{c,h}(\eta_c^h)$, which in turn follows from the definition of η_c^h; and the last equality also follows from the definition of η_c^h and η_c^l.

If $x > v_c^\star$, then

$$
\gamma_c^\star(v_l \le x \mid v_h) = \begin{cases} 1, & \text{if } v_h \le v_c^\star \\ \frac{f_{c,l}(v_h)}{f_{c,h}(v_h)} \mathbf{1}\{v_h \le x\} + \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)}, & \text{if } v_h > v_c^\star. \end{cases} .
$$

Thus,

$$
\begin{aligned}
&\int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
=&\int_0^{v_c^\star} 1 F_{c,h}(\mathrm{d}v_h) + \int_{v_c^\star}^\infty \left(\frac{f_{c,l}(v_h)}{f_{c,h}(v_h)} \mathbf{1}\{v_h \le x\} + \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} \right) F_{c,h}(\mathrm{d}v_h) \\
=&F_{c,h}(v_c^\star) + \int_{v_c^\star}^x f_{c,l}(v_h)\,\mathrm{d}v_h + \int_{v_c^\star}^\infty [f_{c,h}(v_h) - f_{c,l}(v_h)]\,\mathrm{d}v_h \\
=&F_{c,h}(v_c^\star) + F_{c,l}(x) - F_{c,l}(v_c^\star) + F_{c,l}(v_c^\star) - F_{c,h}(v_c^\star) \\
=&F_{c,l}(x)\,.
\end{aligned}
$$

Together, whenever $F_{c,l}(c) \ge \|F_{c,l} - F_{c,h}\|$ and $c \le v_c^\star$,

$$
\rho_c^\star(v_l \le x, v_h \in V) = \int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) = F_{c,l}(x)\,,
$$

for all $x \in V$ and hence $\rho_c^\star \in \mathcal{R}_c$.

Lastly, consider the case when $F_{c,l}(c) \ge \|F_{c,l} - F_{c,h}\|$ and $c > v_c^\star$. If $x \le F_{c,l}^{-1}(F_{c,h}(c))$, then

$$
\gamma_c^\star(v_l \le x \mid v_h) = \begin{cases} \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h)) \le x\}, & \text{if } v_h \le c \\ 0, & \text{if } v_h > c \end{cases}.
$$

Thus,

$$
\begin{aligned}
\int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) &= \int_0^c \mathbf{1}\{F_{c,l}^{-1}(F_{c,h}(v_h)) \le x\} F_{c,h}(\mathrm{d}v_h) \\
&= \int_0^c \mathbf{1}\{F_{c,h}(v_h) \le F_{c,l}(x)\}\,\mathrm{d}z \\
&= \int_0^{F_{c,h}(c)} \mathbf{1}\{z \le F_{c,l}(x)\}\,\mathrm{d}z \\
&= \min\{F_{c,h}(c), F_{c,l}(x)\} \\
&=F_{c,l}(x)\,,
\end{aligned}
$$

where the third equality follows from changing variables of the integration.

If $x > F_{c,l}^{-1}(F_{c,h}(c))$, then

$$\gamma_c^\star(v_l \le x \mid v_h) = \begin{cases} 1, & \text{if } v_h \le c \\ \frac{f_{c,l}(v_h)}{f_{c,h}(v_h)}\mathbf{1}\{v_h \le x\} + \frac{f_{c,h}(v_h)-f_{c,l}(v_h)}{f_{c,h}(v_h)}, & \text{if } v_h > c \end{cases}.$$

Thus,

$$\begin{aligned}
&\int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) \\
=&\int_0^c F_{c,h}(\mathrm{d}v_h) + \int_c^\infty \frac{f_{c,l}(v_h)}{f_{c,h}(v_h)}\mathbf{1}\{v_h \le x\} F_{c,h}(\mathrm{d}v_h) + \int_c^\infty \frac{f_{c,h}(v_h) - f_{c,l}(v_h)}{f_{c,h}(v_h)} F_{c,h}(\mathrm{d}v_h) \\
=&F_{c,h}(c) + \int_c^\infty \mathbf{1}\{v_h \le x\} f_{c,l}(v_h)\,\mathrm{d}v_h + \int_c^\infty (f_{c,h}(v_h) - f_{c,l}(v_h))\,\mathrm{d}v_h \\
=&F_{c,h}(c) + F_{c,l}(x) - F_{c,l}(c) + F_{c,l}(c) - F_{c,h}(c) \\
=&F_{c,l}(x)\,.
\end{aligned}$$

Together, whenever $F_{c,l}(c) \ge \|F_{c,l} - F_{c,h}\|$ and $c > v_c^\star$,

$$\rho_c^\star(v_l \le x, v_h \in V) = \int_0^\infty \gamma_c^\star(v_l \le x \mid v_h) F_{c,h}(\mathrm{d}v_h) = F_{c,l}(x)\,,$$

for all $x \in V$, and hence $\rho_c^\star \in \mathcal{R}_c$. This completes the proof. □

Proof of Lemma A.3. Consider first the case when $F_{c,l}(c) < \|F_{c,l} - F_{c,h}\|$. By construction,

$$\begin{aligned}
\operatorname{supp}(\rho_c^\star) \subseteq &[0, \kappa_c^2] \times [\kappa_c^5, \infty) \cup [\kappa_c^3, \kappa_c^4] \times [0, \kappa_c^1] \\
&\cup \{(v_l, v_h) : v_l = F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^3))\} \\
&\cup \{(v_l, v_h) : v_l = v_h\,, v_l, v_h \in [\kappa_c^3, \kappa_c^4] \cup [\kappa_c^5, \infty)\} \\
&\cup \{(v_l, v_h) : v_l = F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^4))\,, v_l, v_h \in [\kappa_c^4, \kappa_c^5]\}\,.
\end{aligned}$$

For all $(v_l, v_h) \in \text{supp}(\rho_c^\star) \cap [0, \kappa_c^2] \times [\kappa_c^5, \infty)$, since

$$\begin{aligned} \alpha_c(v_h - c) \geq \alpha_c(\kappa_c^5 - c) =& \alpha_c(\kappa_c^4 - c) + \kappa_c^1 - c \\ =& \alpha_c(\kappa_c^4 - c) + (1 - \alpha_c)(\kappa_c^3 - c) \\ \geq& \alpha_c(\kappa_c^3 - c) + (1 - \alpha_c)(\kappa_c^3 - c) \\ =& \kappa_c^3 - c \\ \geq& \kappa_c^2 - c \\ \geq& v_l - c\,, \end{aligned}$$

$\pi_c(v_l, v_h) = \alpha_c(v_h - c)$. Thus,

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = \alpha_c(v_h - c) = \pi_c(v_l, v_h)\,.$$

For all $(v_l, v_h) \in \text{supp}(\rho_c^\star) \cap [\kappa_c^2, \kappa_c^3] \times [\kappa_c^1, \kappa_c^3]$, it must be that $v_l = F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^3))$. Moreover, since $v \mapsto \Delta_c(v)$ is increasing on $[0, v_c^\star]$, $\Delta_c(v) \leq \Delta_c(\kappa_c^3)$. Therefore,

$$F_{c,l}(v) \leq F_{c,h}(v) + \Delta_c(\kappa_c^3)\,,$$

for all $v \in [\kappa_c^1, \kappa_c^3]$, and hence

$$v \leq F_{c,l}^{-1}(F_{c,h}(v) + \Delta_c(\kappa_c^3))$$

for all $v \in [\kappa_c^1, \kappa_c^3]$.

Therefore, for all $(v_l, v_h) \in \text{supp}(\rho_c^\star) \cap [\kappa_c^2, \kappa_c^3] \times [\kappa_c^1, \kappa_c^3]$, it must be that

$$v_l = F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^3)) \geq v_h\,,$$

Together with the fact that

$$(1 - \alpha_c)(v_l - c) \leq (1 - \alpha_c)(\kappa_c^3 - c) = \kappa_c^1 - c \leq v_h - c\,,$$

which follows from (8), it must be that

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = v_h - c = \pi_c(v_l, v_h)\,.$$

For all $(v_l, v_h) \in \operatorname{supp}(\rho^\star) \cap [\kappa_c^3, \kappa_c^4] \times [0, \kappa_c^1]$,

$$v_h - c \le \kappa_c^1 - c = (1-\alpha_c)(\kappa_c^3 - c) \le (1-\alpha_c)(v_l - c)\,.$$

Therefore, $\pi_c(v_l, v_h) = (1-\alpha_c)(v_l - c)$. As a result,

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = (1-\alpha_c)(v_l - c) = \pi_c(v_l, v_h)\,.$$

For all $(v_l, v_h) \in \operatorname{supp}(\rho_c^\star) \cap \{(v_l, v_h) : v_l = v_h\,, v_l, v_h \in [\kappa_c^3, \kappa_c^4] \cup [\kappa_c^5, \infty)\}$, we have $\pi_c(v_l, v_h) = v_l - c = v_h - c$. Therefore,

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = (1-\alpha_c)(v_l - c) + \alpha_c(v_h - c) = v_h - c = v_l - c = \pi_c(v_l, v_h)\,.$$

For all $(v_l, v_h) \in \operatorname{supp}(\rho_c^\star) \cap \{(v_l, v_h) : v_l = F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^4))\,, v_l, v_h \in [\kappa_c^4, \kappa_c^5]\}$, it must be that $v_l \le v_h$. Indeed, since Δ_c is quasi-concave, $\Delta_c(v) \ge \Delta_c(\kappa_c^4) = \Delta_c(\kappa_c^5)$ for all $v \in [\kappa_c^4, \kappa_c^5]$. As a result,

$$F_{c,h}(v) + \Delta_c(\kappa_c^4) \le F_{c,l}(v)\,,$$

and hence

$$v_l = F_{c,l}^{-1}(F_{c,h}(v_h) + \Delta_c(\kappa_c^4)) \le v_h\,.$$

Therefore, $\pi_c(v_l, v_h) = v_l - c$, and hence

$$\phi_c^\star(v_l) + \psi_c^\star(v_h) = v_l - c = \pi_c(v_l, v_h)\,.$$

Together, it follows that

$$\phi^\star(v_l) + \psi^\star(v_h) = \pi_c(v_l, v_h)$$

for all $(v_l, v_h) \in \operatorname{supp}(\rho^\star)$, as desired.

Now consider the case when $F_{c,l}(c) \ge \|F_{c,l} - F_{c,h}\|$. Note that for all $(v_l, v_h) \in \operatorname{supp}(\rho_c^\star)$, either $v_l = v_h$ or $\min\{v_l, v_h\} \le c$. In both cases, we have

$$\pi_c(v_l, v_h) = \alpha_c(v_h - c)^+ + (1-\alpha_c)(v_l - c)^+ = \phi_c^\star(v_l) + \psi_c^\star(v_h)\,,$$

as desired. $\square$

A.3 Additional Results

A.3.1 Scale Family Distributions

In this section, we show that Assumption 1, as well as the solutions implied by Lemma 1, have a simple structure when the value distributions are in a scaled family. Say that $(F_{c,l}, F_{c,h})_{c\in C}$ belongs to a scale family if there exists F_h, F_l such that $F_{c,h}(x) = F_h(x/c)$ and $F_{c,l}(x) = F_l(x/c)$ for all c and for all x.

Proposition A.1. *Suppose that $(F_{c,l}, F_{c,h})_{c\in C}$ belongs to a scale family, then Assumption 1 is equivalent to*

$$F_l(1) < \|F_l - F_h\|.$$

Proof. For any $c \in C$,

$$\begin{aligned}
\|F_{c,l} - F_{c,h}\| &= \max_v \Delta_c(v) = \max_v [F_{c,l}(v) - F_{c,h}(v)] \\
&= \max_v \left[F_l\left(\frac{v}{c}\right) - F_h\left(\frac{v}{c}\right) \right] \\
&= \max_{\tilde{v}} [F_l(\tilde{v}) - F_h(\tilde{v})] && (\tilde{v} = \tfrac{v}{c}) \\
&= \|F_l - F_h\|.
\end{aligned}$$

Meanwhile, for any $c \in C$, $F_{c,l}(c) = F_l(1)$ by definition. This completes the proof. □

While Assumption 1—or, more generally, the division of C_1, C_2 and C_3—is defined by pointwise inequalities for all c, Proposition A.1 shows that when $(F_{c,l}, F_{c,h})_{c\in C}$ belongs to a scale family, Assumption 1 simplifies to a single inequality, in which case exactly one of C_1, C_2, C_3 is nonempty and equals C. In particular, the optimal pricing rule $p^\star$ always has the same structure.

For example, suppose that $F_{c,l}(v) = 1 - e^{-v/\lambda_l c}$ and $F_{c,h}(v) = 1 - e^{-v/\lambda_h c}$ for all $v \geq 0$ and for some $0 < \lambda_l < \lambda_h$. Let $\gamma := \lambda_h/\lambda_l$. Then, $\Delta_c(v) = e^{-v/\lambda_h c} - e^{-v/\lambda_l c}$ for all $v \geq 0$. Moreover, since $v_c^\star$ is the unique maximize of Δ_c, by the first order condition, $\Delta_c'(v_c^\star) = f_{c,l}(v_c^\star) - f_{c,h}(v_c^\star) = 0$. Therefore,

$$\gamma = \frac{\lambda_h}{\lambda_l} = \exp\left(\frac{-v_c^\star}{\lambda_h c} + \frac{v_c^\star}{\lambda_l c}\right) = \exp\left(-\frac{v_c^\star}{\lambda_h c}(1-\gamma)\right).$$

Therefore,

$$\exp\left(-\frac{v_c^\star}{\lambda_h c}\right) = \gamma^{-\frac{1}{\gamma-1}}$$

and

$$\exp\left(-\frac{v_c^\star}{\lambda_l c}\right) = \left(\exp\left(-\frac{v_c^\star}{\lambda_h c}\right)\right)^{\frac{\lambda_h}{\lambda_l}} = \gamma^{-\frac{\gamma}{\gamma-1}}\,.$$

As a result,

$$\Delta_c(v_c^\star) = \exp\left(-\frac{v_c^\star}{\lambda_h c}\right) - \exp\left(-\frac{v_c^\star}{\lambda_l c}\right) = \gamma^{-\frac{1}{\gamma-1}} - \gamma^{-\frac{\gamma}{\gamma-1}} = \gamma^{-\frac{\gamma}{\gamma-1}}(\gamma-1)\,,$$

and hence Assumption 1 simplifies to

$$F_{c,l}(c) = 1 - e^{\frac{-1}{\lambda_l}} < \gamma^{-\frac{\gamma}{\gamma-1}}(\gamma-1) = \Delta_c(v_c^\star)\,.$$

As a second illustration, suppose $F_l = \mathrm{Beta}(a_l, b_l)$ and $F_h = \mathrm{Beta}(a_h, b_h)$ are beta distributions supported on $[0, K]$ for some $K \geq 1$, so that $F_{c,\theta}(v) = F_\theta(v/c)$ defines the scale family. By Proposition A.1, Assumption 1 reduces to $F_l(1) < \|F_l - F_h\|$. At $K = 10$, the left-hand side equals ≈ 0.082. Under $(a_l, b_l) = (2, 4)$ and $(a_h, b_h) = (2, 3)$, $\|F_l - F_h\| \approx 0.138$, and numerical computation yields $CS(c, l; p^\star) \approx 0.0013 > CS(c, h; p^\star) \approx 0.00035$, so l-consumers retain higher surplus. Under $(a_l, b_l) = (2, 4)$ and $(a_h, b_h) = (4, 2)$, $\|F_l - F_h\| \approx 0.625$, and $CS(c, h; p^\star) \approx 0.141 > CS(c, l; p^\star) \approx 0.020$, so h-consumers do. Both pairs thus satisfy Assumption 1.

In fact, when $(F_{c,l}, F_{c,h})_{c\in C}$ belongs to a scale family, and when α_c is constant, then the solutions to (8) are linear in c.

Proposition A.2. *Suppose that $\alpha_c = \alpha$ for all $c \in C$ and that $(F_{c,l}, F_{c,h})_{c\in C}$ belongs to a scale family. Then, $\kappa_c = c \cdot \kappa_1$ for all $c \in C$.*

Proof. Let κ_1 be the solution to (8) when $c = 1$, and let $\kappa_c := c \cdot \kappa_1$. Then, for all $c \in C$, $F_{c,l}(\kappa_c^j) = F_{1,l}(\kappa_1^j)$ and $F_{c,h}(\kappa_c^j) = F_{1,h}(\kappa_1^j)$ for all $j \in \{1, 2, 3, 4, 5\}$, and

$$\kappa_c^1 - c = c(\kappa_c^1 - 1) = c\alpha(\kappa_1^3 - 1) = (1-\alpha)(\kappa_c^3 - c)$$

while

$$(1-\alpha)(\kappa_c^3 - c) = c(1-\alpha)(\kappa_1^3 - 1) = c\alpha(\kappa_1^5 - \kappa_1^4) = \alpha(\kappa_c^5 - \kappa_c^4)\,.$$

Therefore, $\kappa_c = c \cdot \kappa_1$ must solve (8). $\square$

A.3.2 Imperfect Price Discrimination

Throughout this subsection, we focus on the case $c = 0$ and suppress the subscript c on F_l, F_h, f_l, f_h, Δ, $\underline{\Delta}^{-1}$, $\overline{\Delta}^{-1}$, $v^\star$, $\overline{v}$, $\underline{\beta}$, $\overline{\beta}$, $\tilde{\kappa}$, $\tilde{\gamma}^\star$, $\tilde{\rho}^\star$, and $\mathcal{R}$ for notational convenience.

We now solve the optimal transport problem (13), and show that the solution is qualitatively similar to the solution in the baseline model. To describe the solution, let $\tilde{\kappa} = (\tilde{\kappa}^1, \dots, \tilde{\kappa}^5)$ be an increasing vector with $\tilde{\kappa}^4 < v^\star < \tilde{\kappa}^5$ that solves the following system of equations:

$$\begin{aligned}
&\tilde{\kappa}^2 = F_l^{-1}(\Delta(\tilde{\kappa}^3) + F_h(\tilde{\kappa}^1)) = F_l^{-1}(\Delta(\tilde{\kappa}^4)) = F_l^{-1}(\Delta(\tilde{\kappa}^5)) \\
&\frac{\tilde{\kappa}^1\tilde{\kappa}^2}{\tilde{\kappa}^1 + \tilde{\kappa}^2} = \frac{\tilde{\kappa}^3}{4} - \int_{\tilde{\kappa}^2}^{\tilde{\kappa}^3} \left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)}\right)^2 \mathrm{d}z = \int_{\tilde{\kappa}^4}^{\tilde{\kappa}^5} \left(\frac{\overline{\beta}(z)}{z + \overline{\beta}(z)}\right)^2 \mathrm{d}z - \frac{\tilde{\kappa}^5 - \tilde{\kappa}^4}{4},
\end{aligned} \tag{A.17}$$

where

$$\underline{\beta}(z) \coloneqq F_h^{-1}(F_l(z) - \Delta(\tilde{\kappa}^3)), \qquad z \in [\tilde{\kappa}^2, \tilde{\kappa}^3],$$

and

$$\overline{\beta}(z) \coloneqq F_h^{-1}(F_l(z) - \Delta(\tilde{\kappa}^4)), \qquad z \in [\tilde{\kappa}^4, \tilde{\kappa}^5].$$

The existence of such a vector $\tilde{\kappa}$ is established by the following lemma.

Lemma A.4. *The system* (A.17) *admits an increasing solution*

$$0 \le \tilde{\kappa}^1 \le \tilde{\kappa}^2 \le \tilde{\kappa}^3 \le \tilde{\kappa}^4 < v^\star < \tilde{\kappa}^5 \le \overline{v}.$$

Proof. We work in quantile space. Let $\Delta^\star \coloneqq \Delta(v^\star)$. For each $q \in (0, \Delta^\star)$, define

$$k_2(q) \coloneqq F_l^{-1}(q), \qquad k_4(q) \coloneqq \underline{\Delta}^{-1}(q), \qquad k_5(q) \coloneqq \overline{\Delta}^{-1}(q).$$

Note that $k_2(q) \le k_4(q) < v^\star < k_5(q)$. For each $t \in [k_2(q), k_4(q)]$, define

$$k_1(q,t) \coloneqq F_h^{-1}(q - \Delta(t))$$

and

$$\beta_{q,t}(z) \coloneqq F_h^{-1}(F_l(z) - \Delta(t)), \qquad z \in [k_2(q), t].$$

Consider the function

$$\Gamma_q(t) := \frac{k_1(q,t)k_2(q)}{k_1(q,t)+k_2(q)} - \left(\frac{t}{4} - \int_{k_2(q)}^{t} \left(\frac{\beta_{q,t}(z)}{z+\beta_{q,t}(z)}\right)^2 \mathrm{d}z\right). \tag{A.18}$$

Fix any $q \in (0, \Delta^\star)$. Clearly, at $t = k_2(q)$, $\Gamma_q(t) = k_2(q)/4 > 0$. We will now show that $\Gamma_q(t) < 0$ at $t = k_4(q)$. To this end, we first claim that $\beta_{q,t}(z) \le z$ for all $z \in [k_2(q), t]$ and all $t \in [k_2(q), k_4(q)]$. Indeed, by definition, $\beta_{q,t}(z) \le z$ is equivalent to $F_l(z) - \Delta(t) \le F_h(z)$, which in turn is equivalent to $\Delta(z) \le \Delta(t)$. Since $z \le t \le k_4(q) < v^\star$ and Δ is strictly increasing on $[0, v^\star]$, we have $\Delta(z) \le \Delta(t)$, as desired. Meanwhile, for any $z \in [k_2(q), t]$ and for all $t \in [k_2(q), k_4(q)]$, note that $F_l(z) \ge F_l(k_2(q)) = q \ge \Delta(t)$, where the last inequality uses $\Delta(t) \le \Delta(k_4(q)) = q$.

Together with the fact that $k_1(q,t) = 0$ at $t = k_4(q)$, we have

$$-\Gamma_q(t) = \frac{k_4(q)}{4} - \int_{k_2(q)}^{k_4(q)} \left(\frac{\beta_{q,k_4(q)}(z)}{z+\beta_{q,k_4(q)}(z)}\right)^2 \mathrm{d}z \ge \frac{k_2(q)}{4} > 0,$$

and hence $\Gamma_q(t) < 0$ at $t = k_4(q)$. Since $t \mapsto \Gamma_q(t)$ is continuous on $[k_2(q), k_4(q)]$ for all $q \in (0, \Delta^\star)$, by the intermediate value theorem, for all $q \in (0, \Delta^\star)$, there exists $t \in [k_2(q), k_4(q)]$ such that $\Gamma_q(t) = 0$. Moreover, this solution is unique. Indeed, $t \mapsto k_1(q,t)$ is decreasing in t, while $t \mapsto \frac{t}{4} - \int_{k_2(q)}^{t} \left(\frac{\beta_{q,t}(z)}{z+\beta_{q,t}(z)}\right)^2 \mathrm{d}z$ is increasing: $t/4$ is increasing, the additional mass of the integral has integrand at most $1/4$, and $\left(\frac{\beta_{q,t}(z)}{z+\beta_{q,t}(z)}\right)^2$ is decreasing in t since $\Delta(t)$ increases. Denote this unique solution by $t(q)$. Standard continuity arguments then imply that $t(q)$ is continuous in q.

Now define

$$A(q) := \frac{k_1(q,t(q))k_2(q)}{k_1(q,t(q))+k_2(q)}$$

and

$$J(q) := \int_{k_4(q)}^{k_5(q)} \left(\frac{\overline{\beta}_q(z)}{z+\overline{\beta}_q(z)}\right)^2 \mathrm{d}z - \frac{k_5(q)-k_4(q)}{4},$$

where

$$\overline{\beta}_q(z) := F_h^{-1}(F_l(z) - q), \qquad z \in [k_4(q), k_5(q)].$$

Both $A(q)$ and $J(q)$ are continuous in q.

First, we show that $A(q) - J(q) < 0$ for all sufficiently small q. Indeed, as $q \downarrow 0$, $k_2(q) =$

$F_l^{-1}(q) \to 0$, and hence $A(q) \le k_2(q) \to 0$. Meanwhile, we claim that $\lim_{q\downarrow 0} J(q) > 0$. To see this, note that as $q \downarrow 0$, $k_4(q) = \underline{\Delta}^{-1}(q) \downarrow 0$ and $k_5(q) = \overline{\Delta}^{-1}(q) \uparrow \overline{v}$, since $\Delta(0) = \Delta(\overline{v}) = 0$ and Δ is strictly increasing on $[0, v^\star]$ and strictly decreasing on $[v^\star, \overline{v}]$. Define the extended integrand

$$g_q(z) := \left[\left(\frac{\overline{\beta}_q(z)}{z + \overline{\beta}_q(z)}\right)^2 - \frac{1}{4}\right]\mathbf{1}_{[k_4(q), k_5(q)]}(z)\,, \qquad z \in (0, \overline{v})\,,$$

so that $J(q) = \int_0^{\overline{v}} g_q(z)\,\mathrm{d}z$. Since $\overline{\beta}_q(z)/(z + \overline{\beta}_q(z)) \in [0,1]$, $|g_q(z)| \le 3/4$ uniformly, and the constant $3/4$ on $(0, \overline{v})$ is an integrable dominator. For each fixed $z \in (0, \overline{v})$, $k_4(q) < z < k_5(q)$ for all sufficiently small q, and the continuity of F_h^{-1} implies $\overline{\beta}_q(z) = F_h^{-1}(F_l(z) - q) \to F_h^{-1}(F_l(z)) =: \beta_0(z)$. Therefore, $g_q(z) \to (\beta_0(z)/(z + \beta_0(z)))^2 - 1/4$. By the dominated convergence theorem,

$$\lim_{q\downarrow 0} J(q) = \int_0^{\overline{v}} \left[\left(\frac{\beta_0(z)}{z + \beta_0(z)}\right)^2 - \frac{1}{4}\right]\mathrm{d}z\,.$$

Since $F_h \le F_l$ pointwise by the monotone likelihood-ratio dominance, $\beta_0(z) \ge z$, and hence $\beta_0(z)/(z + \beta_0(z)) \ge 1/2$. Therefore, the integrand is nonnegative on $(0, \overline{v})$, and is strictly positive whenever $\beta_0(z) > z$, that is, whenever $\Delta(z) > 0$. Since Δ is strictly increasing on $[0, v^\star]$ and strictly decreasing on $[v^\star, \overline{v}]$ with $\Delta(0) = \Delta(\overline{v}) = 0$ and $\Delta(v^\star) > 0$, $\Delta > 0$ on $(0, \overline{v})$, and hence the integrand is strictly positive almost everywhere. Together, $\lim_{q\downarrow 0} J(q) > 0$, and therefore $A(q) - J(q) < 0$ for all sufficiently small q.

Next, we show that $A(q) - J(q) > 0$ for q close to $\Delta^\star$. As $q \uparrow \Delta^\star$, $k_4(q), k_5(q) \to v^\star$, and hence $J(q) \to 0$. Meanwhile, by (A.18),

$$A(q) = \frac{t(q)}{4} - \int_{k_2(q)}^{t(q)} \left(\frac{\beta_{q,t(q)}(z)}{z + \beta_{q,t(q)}(z)}\right)^2 \mathrm{d}z \ge \frac{k_2(q)}{4}\,,$$

and hence $A(q)$ is bounded away from zero near $\Delta^\star$. Therefore, $A(q) - J(q) > 0$ for q close to $\Delta^\star$.

Together, by the intermediate value theorem, there exists $q^\star \in (0, \Delta^\star)$ such that $A(q^\star) = J(q^\star)$. Setting

$$\tilde{\kappa}^1 := k_1(q^\star, t(q^\star))\,, \quad \tilde{\kappa}^2 := k_2(q^\star)\,, \quad \tilde{\kappa}^3 := t(q^\star)\,, \quad \tilde{\kappa}^4 := k_4(q^\star)\,, \quad \tilde{\kappa}^5 := k_5(q^\star)$$

gives an increasing solution of (A.17), as desired. □

Given such a solution $\tilde{\kappa}$, define the within-pair optimal posted price

$$\xi(v_l, v_h) := \begin{cases} v_l \,, & \text{if } v_l \geq 3v_h \\ v_h \,, & \text{if } v_h \geq 3v_l \\ \frac{2v_l v_h}{v_l + v_h} \,, & \text{otherwise,} \end{cases}$$

which is the maximizer of $x \mapsto x \cdot (\alpha_c \mathbb{P}[w \geq x \mid v_h] + (1 - \alpha_c)\mathbb{P}[w \geq x \mid v_l])$ at $c = 0$, $\alpha_c = 1/2$ under the maintained assumption that $w \mid v \sim U[0, 2v]$. Define the pricing rule

$$\tilde{p}^{\star}(v, l, r) := \begin{cases} \xi\Big(v, \overline{\Delta}^{-1}(\Delta(\tilde{\kappa}^5) - F_l(v))\Big) \,, & \text{if } v < \tilde{\kappa}^2 \\ \xi\Big(v, F_h^{-1}(F_l(v) - F_l(\tilde{\kappa}^2) + F_h(\tilde{\kappa}^1))\Big) \,, & \text{if } v \in [\tilde{\kappa}^2, \tilde{\kappa}^3) \\ v \,, & \text{if } v \in [\tilde{\kappa}^3, \tilde{\kappa}^4] \\ \xi\Big(v, F_h^{-1}(F_l(v) - \Delta(\tilde{\kappa}^4))\Big) \,, & \text{if } v \in (\tilde{\kappa}^4, \tilde{\kappa}^5] \\ v \,, & \text{if } v > \tilde{\kappa}^5 \end{cases} ;$$

$$\tilde{p}^{\star}(v, h, r) := \begin{cases} \xi\Big(\underline{\Delta}^{-1}(F_h(v) + \Delta(\tilde{\kappa}^3)), v\Big) \,, & \text{if } v < \tilde{\kappa}^1 \\ \xi\Big(F_l^{-1}(F_h(v) + \Delta(\tilde{\kappa}^3)), v\Big) \,, & \text{if } v \in [\tilde{\kappa}^1, \tilde{\kappa}^3) \\ v \,, & \text{if } v \in [\tilde{\kappa}^3, \tilde{\kappa}^4] \\ \xi\Big(F_l^{-1}(F_h(v) + \Delta(\tilde{\kappa}^4)), v\Big) \,, & \text{if } v \in (\tilde{\kappa}^4, \tilde{\kappa}^5] \\ v \,, & \text{if } v > \tilde{\kappa}^5 \end{cases} ,$$

for all $v \in V$ and $r \in [0, 1]$. As Proposition A.3 below shows, $\tilde{p}^{\star}$ is optimal. The corresponding matching scheme is identical to the one given in the baseline model, with the only difference being how the thresholds $\tilde{\kappa}$ are defined; each matched pair is charged the within-pair optimal price $\xi(v_l, v_h)$.

Proposition A.3. *$\tilde{p}^{\star}$ is an optimal non-discriminatory pricing rule.*

Proof. Consider the following pair of functions $\tilde{\phi}$ and $\tilde{\psi}$:

$$\tilde{\phi}(v_l) := \begin{cases} \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2+\tilde{\kappa}^1}\,, & \text{if } v_l \le \tilde{\kappa}^2 \\ \frac{\tilde{\kappa}^1\tilde{\kappa}^2}{\tilde{\kappa}^1+\tilde{\kappa}^2} + \int_{\tilde{\kappa}^2}^{v_l} \left(\frac{\underline{\beta}(z)}{z+\underline{\beta}(z)} \right)^2 \mathrm{d}z\,, & \text{if } v_l \in (\tilde{\kappa}^2, \tilde{\kappa}^3] \\ \frac{v_l}{4}\,, & \text{if } v_l \in (\tilde{\kappa}^3, \tilde{\kappa}^4] \\ \frac{\tilde{\kappa}^4}{4} + \int_{\tilde{\kappa}^4}^{v_l} \left(\frac{\overline{\beta}(z)}{z+\overline{\beta}(z)} \right)^2 \mathrm{d}z\,, & \text{if } v_l \in (\tilde{\kappa}^4, \tilde{\kappa}^5] \\ \frac{v_l}{4} + \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2+\tilde{\kappa}^1}\,, & \text{if } v_l > \tilde{\kappa}^5 \end{cases}$$

and

$$\tilde{\psi}(v_h) := \begin{cases} 0\,, & \text{if } v_h \le \tilde{\kappa}^1 \\ \frac{\underline{\beta}^{-1}(v_h)v_h}{v_h+\underline{\beta}^{-1}(v_h)} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2+\tilde{\kappa}^1} - \int_{\tilde{\kappa}^2}^{\underline{\beta}^{-1}(v_h)} \left(\frac{\underline{\beta}(z)}{z+\underline{\beta}(z)} \right)^2 \mathrm{d}z\,, & \text{if } v_h \in (\tilde{\kappa}^1, \tilde{\kappa}^3] \\ \frac{v_h}{4}\,, & \text{if } v_h \in (\tilde{\kappa}^3, \tilde{\kappa}^4] \\ \frac{v_h\overline{\beta}^{-1}(v_h)}{v_h+\overline{\beta}^{-1}(v_h)} - \frac{\tilde{\kappa}^4}{4} - \int_{\tilde{\kappa}^4}^{\overline{\beta}^{-1}(v_h)} \left(\frac{\overline{\beta}(z)}{z+\overline{\beta}(z)} \right)^2 \mathrm{d}z\,, & \text{if } v_h \in (\tilde{\kappa}^4, \tilde{\kappa}^5] \\ \frac{v_h}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2+\tilde{\kappa}^1}\,, & \text{if } v_h > \tilde{\kappa}^5 \end{cases}.$$

By construction, $\tilde{\phi}$ and $\tilde{\psi}$ are continuous at all cutoff points. Moreover, the last equality of (A.17) gives

$$\tilde{\phi}(\tilde{\kappa}^5) = \frac{\tilde{\kappa}^5}{4} + \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2+\tilde{\kappa}^1}\,, \qquad \tilde{\psi}(\tilde{\kappa}^5) = \frac{\tilde{\kappa}^5}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2+\tilde{\kappa}^1}\,,$$

which is the boundary identity needed for feasibility at $(\tilde{\kappa}^5, \tilde{\kappa}^5)$.

We now show that

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \ge \widetilde{\pi}(v_l, v_h) \tag{A.19}$$

for all v_l, v_h. To this end, we first establish two inequalities that follow from (A.17). First, note that the function

$$v \mapsto \frac{\tilde{\kappa}^1 v}{\tilde{\kappa}^1 + v} + \int_v^{\tilde{\kappa}^3} \left(\frac{\underline{\beta}(z)}{z+\underline{\beta}(z)} \right)^2 \mathrm{d}z$$

is decreasing in v. This, together with (A.17), implies that

$$\frac{\tilde{\kappa}^1\tilde{\kappa}^3}{\tilde{\kappa}^1+\tilde{\kappa}^3} \le \frac{\tilde{\kappa}^1\tilde{\kappa}^2}{\tilde{\kappa}^1+\tilde{\kappa}^2} + \int_{\tilde{\kappa}^2}^{\tilde{\kappa}^3} \left(\frac{\underline{\beta}(z)}{z+\underline{\beta}(z)} \right)^2 \mathrm{d}z = \frac{\tilde{\kappa}^3}{4}\,.$$

Meanwhile, since the function

$$v \mapsto \frac{v}{4} - \int_{\tilde{\kappa}^2}^{v} \left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)} \right)^2 \mathrm{d}z$$

is increasing in v, it follows that

$$\frac{\tilde{\kappa}^1 \tilde{\kappa}^2}{\tilde{\kappa}^1 + \tilde{\kappa}^2} = \frac{\tilde{\kappa}^3}{4} - \int_{\tilde{\kappa}^2}^{\tilde{\kappa}^3} \left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)} \right)^2 \mathrm{d}z \geq \frac{\tilde{\kappa}^2}{4} \,.$$

Together, the above two inequalities imply

$$\tilde{\kappa}^2 \leq 3\tilde{\kappa}^1 \leq \tilde{\kappa}^3 \,. \tag{A.20}$$

Next, by the last equality of (A.17) and the fact that $\overline{\beta}(z) \leq \tilde{\kappa}^5$ for all $z \in [\tilde{\kappa}^4, \tilde{\kappa}^5]$,

$$\frac{\tilde{\kappa}^2}{4} \leq \frac{\tilde{\kappa}^1 \tilde{\kappa}^2}{\tilde{\kappa}^1 + \tilde{\kappa}^2} = \int_{\tilde{\kappa}^4}^{\tilde{\kappa}^5} \left[\left(\frac{\overline{\beta}(z)}{z + \overline{\beta}(z)} \right)^2 - \frac{1}{4} \right] \mathrm{d}z \leq \int_{\tilde{\kappa}^4}^{\tilde{\kappa}^5} \left[\left(\frac{\tilde{\kappa}^5}{z + \tilde{\kappa}^5} \right)^2 - \frac{1}{4} \right] \mathrm{d}z = \frac{(\tilde{\kappa}^5 - \tilde{\kappa}^4)^2}{4(\tilde{\kappa}^5 + \tilde{\kappa}^4)} \leq \frac{(\tilde{\kappa}^5 - \tilde{\kappa}^2)^2}{4(\tilde{\kappa}^5 + \tilde{\kappa}^2)} \,.$$

Therefore,

$$\tilde{\kappa}^5 \geq 3\tilde{\kappa}^2 \,. \tag{A.21}$$

Then, we discuss all the cases.

Case 1: $v_l \leq \tilde{\kappa}^2$.

When $v_h \leq \tilde{\kappa}^1$,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{\tilde{\kappa}^2 \tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} \geq \frac{v_l v_h}{v_l + v_h} \,.$$

Moreover, since $\tilde{\kappa}^1 \leq \tilde{\kappa}^2$,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{\tilde{\kappa}^2 \tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} \geq \frac{\tilde{\kappa}^1}{4} \geq \frac{v_h}{4} \,.$$

Furthermore, (A.20) implies that

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{\tilde{\kappa}^2 \tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} \geq \frac{\tilde{\kappa}^2}{4} \geq \frac{v_l}{4} \,.$$

Together,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_h}{4}, \frac{v_l}{4}\right\} = \widetilde{\pi}(v_l, v_h)\,,$$

as desired.

When $v_h \in (\tilde{\kappa}^1, \tilde{\kappa}^3]$, by (A.20), $3v_h \geq 3\tilde{\kappa}^1 \geq \tilde{\kappa}^2 \geq v_l$, and hence

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_h}{4}\right\}\,.$$

Note that the function

$$v_h \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_h}{4}$$

is increasing in v_h for all $v_l \leq \tilde{\kappa}^2$. Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_h}{4} \geq \tilde{\phi}(v_l) + \tilde{\psi}(\tilde{\kappa}^1) - \frac{\tilde{\kappa}^1}{4} = \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} - \frac{\tilde{\kappa}^1}{4} \geq 0\,,$$

where the last inequality follows from the cutoff monotonicity $\tilde{\kappa}^1 \leq \tilde{\kappa}^2$. Meanwhile, since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is decreasing in v_l and increasing in v_h, it follows that

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq \tilde{\phi}(\tilde{\kappa}^2) + \tilde{\psi}(\tilde{\kappa}^1) - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} = 0\,.$$

Together,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)\,,$$

as desired.

When $v_h \in (\tilde{\kappa}^3, \tilde{\kappa}^4]$,

$$\begin{aligned}
\tilde{\phi}(v_l) + \tilde{\psi}(v_h) =& \frac{\tilde{\kappa}^3}{4} - \int_{\tilde{\kappa}^2}^{\tilde{\kappa}^3} \left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)} \right)^2 \mathrm{d}z + \frac{v_h}{4} \\
=& \int_{\tilde{\kappa}^2}^{\tilde{\kappa}^3} \left[\frac{1}{4} - \left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)} \right)^2 \right] \mathrm{d}z + \frac{v_h}{4} + \frac{\tilde{\kappa}^2}{4} \\
\geq& \frac{v_h}{4} + \frac{\tilde{\kappa}^2}{4} \\
\geq& \max\left\{ \frac{v_l v_h}{v_l + v_h}, \frac{v_l}{4}, \frac{v_h}{4} \right\} = \widetilde{\pi}(v_l, v_h)\,,
\end{aligned}$$

as desired.

When $v_h \in (\tilde{\kappa}^4, \tilde{\kappa}^5]$, since $v_l \leq \tilde{\kappa}^2 \leq \tilde{\kappa}^4 < v_h$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{ \frac{v_l v_h}{v_l + v_h}, \frac{v_h}{4} \right\}.$$

Note that the function

$$v_h \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_h}{4}$$

is decreasing in v_h for all $v_l \leq \tilde{\kappa}^2$. Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_h}{4} \geq \tilde{\phi}(v_l) + \tilde{\psi}(\tilde{\kappa}^5) - \frac{\tilde{\kappa}^5}{4} = \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} + \frac{\tilde{\kappa}^5}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} - \frac{\tilde{\kappa}^5}{4} = 0\,.$$

Meanwhile, since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is decreasing in v_l and increasing in v_h,

$$\begin{aligned}
\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq& \tilde{\phi}(\tilde{\kappa}^2) + \tilde{\psi}(\tilde{\kappa}^4) - \frac{\tilde{\kappa}^2\tilde{\kappa}^4}{\tilde{\kappa}^2 + \tilde{\kappa}^4} \\
=& \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} + \frac{\tilde{\kappa}^4}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^4}{\tilde{\kappa}^2 + \tilde{\kappa}^4} \\
\geq& \frac{\tilde{\kappa}^2}{4} + \frac{\tilde{\kappa}^4}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^4}{\tilde{\kappa}^2 + \tilde{\kappa}^4} \geq 0\,,
\end{aligned}$$

where the second inequality uses (A.20). Together,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)\,,$$

as desired.

Lastly, when $v_h > \tilde{\kappa}^5$, by (A.21), $v_h > \tilde{\kappa}^5 \geq 3\tilde{\kappa}^2 \geq 3v_l$. Thus, $\widetilde{\pi}(v_l, v_h) = v_h/4$, and hence

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} + \frac{v_h}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} = \frac{v_h}{4} = \widetilde{\pi}(v_l, v_h)\,,$$

as desired.

Case 2: $v_l \in (\tilde{\kappa}^2, \tilde{\kappa}^3]$.

When $v_h \leq \tilde{\kappa}^1$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_l}{4}\right\}.$$

Note that the function

$$v_l \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is increasing in v_l, and the function

$$v_l \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l}{4}$$

is decreasing in v_l. Indeed, the latter follows from $\underline{\beta}(v_l) \leq v_l$ for $v_l \in [\tilde{\kappa}^2, \tilde{\kappa}^3]$, which in turn is because $\Delta(v_l) \leq \Delta(\tilde{\kappa}^3)$ on this interval. Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq \tilde{\phi}(\tilde{\kappa}^2) + \tilde{\psi}(v_h) - \frac{\tilde{\kappa}^2 v_h}{\tilde{\kappa}^2 + v_h} = \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} - \frac{\tilde{\kappa}^2 v_h}{\tilde{\kappa}^2 + v_h} \geq 0\,,$$

where the last inequality follows from $v_h \leq \tilde{\kappa}^1$. Moreover,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l}{4} \geq \tilde{\phi}(\tilde{\kappa}^3) + \tilde{\psi}(v_h) - \frac{\tilde{\kappa}^3}{4} = 0\,.$$

Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)$, as desired.

When $v_h \in (\tilde{\kappa}^1, \tilde{\kappa}^3]$, by (A.20),

$$\widetilde{\pi}(v_l, v_h) = \frac{v_l v_h}{v_l + v_h}\,.$$

In particular, $\widetilde{\pi}$ is supermodular on $(\tilde{\kappa}^2, \tilde{\kappa}^3] \times (\tilde{\kappa}^1, \tilde{\kappa}^3]$. Therefore, since $\underline{\beta}$ is increasing, by construction,

$$\tilde{\phi}(v_l) = \max_{v_h' \in (\tilde{\kappa}^1, \tilde{\kappa}^3]} \left[\widetilde{\pi}(v_l, v_h') - \tilde{\psi}(v_h')\right] \geq \widetilde{\pi}(v_l, v_h) - \tilde{\psi}(v_h)\,,$$

as desired.

When $v_h \in (\tilde{\kappa}^3, \tilde{\kappa}^4]$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_h}{4}\right\}.$$

Note that

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{\tilde{\kappa}^1 \tilde{\kappa}^2}{\tilde{\kappa}^1 + \tilde{\kappa}^2} + \int_{\tilde{\kappa}^2}^{v_l} \left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)}\right)^2 \mathrm{d}z + \frac{v_h}{4} \geq \frac{v_h}{4}\,.$$

Moreover, since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is increasing in v_h and decreasing in v_l,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq \tilde{\phi}(\tilde{\kappa}^3) + \tilde{\psi}(\tilde{\kappa}^3) - \frac{\tilde{\kappa}^3}{2} = 0\,.$$

Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)$, as desired.

When $v_h \in (\tilde{\kappa}^4, \tilde{\kappa}^5]$, note that the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_h}{4}$$

is increasing in v_l and decreasing in v_h. Also, the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is decreasing in v_l and increasing in v_h. Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_h}{4} \geq \tilde{\phi}(\tilde{\kappa}^2) + \tilde{\psi}(\tilde{\kappa}^5) - \frac{\tilde{\kappa}^5}{4} = \frac{\tilde{\kappa}^2 \tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} - \frac{\tilde{\kappa}^2 \tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} = 0\,,$$

and

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \ge \tilde{\phi}(\tilde{\kappa}^3) + \tilde{\psi}(\tilde{\kappa}^4) - \frac{\tilde{\kappa}^3\tilde{\kappa}^4}{\tilde{\kappa}^3 + \tilde{\kappa}^4} = \frac{\tilde{\kappa}^3}{4} + \frac{\tilde{\kappa}^4}{4} - \frac{\tilde{\kappa}^3\tilde{\kappa}^4}{\tilde{\kappa}^3 + \tilde{\kappa}^4} \ge 0\,.$$

Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \ge \widetilde{\pi}(v_l, v_h)$, as desired.

When $v_h > \tilde{\kappa}^5$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_h}{4}\right\}\,.$$

Note that

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{\tilde{\kappa}^1\tilde{\kappa}^2}{\tilde{\kappa}^1 + \tilde{\kappa}^2} + \int_{\tilde{\kappa}^2}^{v_l}\left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)}\right)^2 \mathrm{d}z + \frac{v_h}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} = \int_{\tilde{\kappa}^2}^{v_l}\left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)}\right)^2 \mathrm{d}z + \frac{v_h}{4} \ge \frac{v_h}{4}\,.$$

Meanwhile, since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is decreasing in v_l and increasing in v_h,

$$\begin{aligned}
\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \ge& \tilde{\phi}(\tilde{\kappa}^3) + \tilde{\psi}(\tilde{\kappa}^5) - \frac{\tilde{\kappa}^3\tilde{\kappa}^5}{\tilde{\kappa}^3 + \tilde{\kappa}^5} \\
=& \frac{\tilde{\kappa}^3}{4} + \frac{\tilde{\kappa}^5}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} - \frac{\tilde{\kappa}^3\tilde{\kappa}^5}{\tilde{\kappa}^3 + \tilde{\kappa}^5} \\
=& \frac{\tilde{\kappa}^3}{4} + \frac{\tilde{\kappa}^5}{2} - \frac{\tilde{\kappa}^4}{4} - \int_{\tilde{\kappa}^4}^{\tilde{\kappa}^5}\left(\frac{\overline{\beta}(z)}{z + \overline{\beta}(z)}\right)^2 \mathrm{d}z - \frac{\tilde{\kappa}^3\tilde{\kappa}^5}{\tilde{\kappa}^3 + \tilde{\kappa}^5} \\
\ge& \frac{\tilde{\kappa}^3}{4} + \frac{\tilde{\kappa}^5}{2} - \frac{\tilde{\kappa}^4}{4} - \int_{\tilde{\kappa}^4}^{\tilde{\kappa}^5}\left(\frac{\tilde{\kappa}^5}{z + \tilde{\kappa}^5}\right)^2 \mathrm{d}z - \frac{\tilde{\kappa}^3\tilde{\kappa}^5}{\tilde{\kappa}^3 + \tilde{\kappa}^5} \\
=& \left[\frac{\tilde{\kappa}^4\tilde{\kappa}^5}{\tilde{\kappa}^4 + \tilde{\kappa}^5} - \frac{\tilde{\kappa}^4}{4}\right] - \left[\frac{\tilde{\kappa}^3\tilde{\kappa}^5}{\tilde{\kappa}^3 + \tilde{\kappa}^5} - \frac{\tilde{\kappa}^3}{4}\right] \\
\ge& 0\,,
\end{aligned}$$

where the third equality uses the last equality of (A.17), the second inequality follows from $\overline{\beta}(z) \le \tilde{\kappa}^5$ for all $z \in [\tilde{\kappa}^4, \tilde{\kappa}^5]$, and the last inequality follows from the fact that the function

$$t \mapsto \frac{t\tilde{\kappa}^5}{t + \tilde{\kappa}^5} - \frac{t}{4}$$

is increasing for $t \le \tilde{\kappa}^5$ and from $\tilde{\kappa}^4 \ge \tilde{\kappa}^3$. Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \ge \widetilde{\pi}(v_l, v_h)$, as desired.

Case 3: $v_l \in (\tilde{\kappa}^3, \tilde{\kappa}^4]$.

When $v_h \le \tilde{\kappa}^1$, by (A.20), $v_l > \tilde{\kappa}^3 \ge 3\tilde{\kappa}^1 \ge 3v_h$. Thus, $\widetilde{\pi}(v_l, v_h) = v_l/4$, and hence

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{v_l}{4} + 0 = \frac{v_l}{4} = \widetilde{\pi}(v_l, v_h)\,,$$

as desired.

When $v_h \in (\tilde{\kappa}^1, \tilde{\kappa}^3]$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_l}{4}\right\}.$$

Note that

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{v_l}{4} + \tilde{\psi}(v_h) \ge \frac{v_l}{4}\,.$$

Moreover, since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is increasing in v_l and decreasing in v_h,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \ge \tilde{\phi}(\tilde{\kappa}^3) + \tilde{\psi}(\tilde{\kappa}^3) - \frac{\tilde{\kappa}^3}{2} = 0\,.$$

Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \ge \widetilde{\pi}(v_l, v_h)$, as desired.

When $v_h \in (\tilde{\kappa}^3, \tilde{\kappa}^4]$,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{v_l}{4} + \frac{v_h}{4} \ge \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_l}{4}, \frac{v_h}{4}\right\} = \widetilde{\pi}(v_l, v_h)\,,$$

as desired.

When $v_h \in (\tilde{\kappa}^4, \tilde{\kappa}^5]$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_h}{4}\right\}.$$

First, note that the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_h}{4}$$

is increasing in v_l and decreasing in v_h. Therefore,

$$\begin{aligned}\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_h}{4} \geq& \tilde{\phi}(\tilde{\kappa}^3) + \tilde{\psi}(\tilde{\kappa}^5) - \frac{\tilde{\kappa}^5}{4} \\ =& \frac{\tilde{\kappa}^3}{4} + \frac{\tilde{\kappa}^5}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} - \frac{\tilde{\kappa}^5}{4} \\ =& \frac{\tilde{\kappa}^3}{4} - \frac{\tilde{\kappa}^3}{4} + \int_{\tilde{\kappa}^2}^{\tilde{\kappa}^3} \left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)}\right)^2 \mathrm{d}z \\ \geq& 0\,.\end{aligned}$$

Moreover, the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is decreasing in v_l and increasing in v_h. Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq \tilde{\phi}(\tilde{\kappa}^4) + \tilde{\psi}(\tilde{\kappa}^4) - \frac{\tilde{\kappa}^4}{2} = 0\,.$$

Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)$, as desired.

When $v_h > \tilde{\kappa}^5$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_h}{4}\right\}\,.$$

First, note that

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{v_l}{4} + \frac{v_h}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} \geq \frac{\tilde{\kappa}^3}{4} + \frac{v_h}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} = \int_{\tilde{\kappa}^2}^{\tilde{\kappa}^3} \left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)}\right)^2 \mathrm{d}z + \frac{v_h}{4} \geq \frac{v_h}{4}\,.$$

Moreover, since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is decreasing in v_l and increasing in v_h,

$$
\begin{aligned}
\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq & \tilde{\phi}(\tilde{\kappa}^4) + \tilde{\psi}(\tilde{\kappa}^5) - \frac{\tilde{\kappa}^4 \tilde{\kappa}^5}{\tilde{\kappa}^4 + \tilde{\kappa}^5} \\
= & \frac{\tilde{\kappa}^4}{4} + \frac{\tilde{\kappa}^5}{4} - \frac{\tilde{\kappa}^2 \tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} - \frac{\tilde{\kappa}^4 \tilde{\kappa}^5}{\tilde{\kappa}^4 + \tilde{\kappa}^5} \\
= & \frac{\tilde{\kappa}^4}{4} + \frac{\tilde{\kappa}^5}{4} + \frac{\tilde{\kappa}^5 - \tilde{\kappa}^4}{4} - \int_{\tilde{\kappa}^4}^{\tilde{\kappa}^5} \left(\frac{\overline{\beta}(z)}{z + \overline{\beta}(z)} \right)^2 \mathrm{d}z - \frac{\tilde{\kappa}^4 \tilde{\kappa}^5}{\tilde{\kappa}^4 + \tilde{\kappa}^5} \\
= & \frac{\tilde{\kappa}^5}{2} - \frac{\tilde{\kappa}^4 \tilde{\kappa}^5}{\tilde{\kappa}^4 + \tilde{\kappa}^5} - \int_{\tilde{\kappa}^4}^{\tilde{\kappa}^5} \left(\frac{\overline{\beta}(z)}{z + \overline{\beta}(z)} \right)^2 \mathrm{d}z \\
= & \int_{\tilde{\kappa}^4}^{\tilde{\kappa}^5} \left[\left(\frac{\tilde{\kappa}^5}{z + \tilde{\kappa}^5} \right)^2 - \left(\frac{\overline{\beta}(z)}{z + \overline{\beta}(z)} \right)^2 \right] \mathrm{d}z \\
\geq & 0\,,
\end{aligned}
$$

where the second equality follows from the last equality of (A.17), and the last inequality follows from $\overline{\beta}(z) \leq \tilde{\kappa}^5$ for all $z \in [\tilde{\kappa}^4, \tilde{\kappa}^5]$. Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)$, as desired.

Case 4: $v_l \in (\tilde{\kappa}^4, \tilde{\kappa}^5]$.

When $v_h \leq \tilde{\kappa}^1$, by (A.20), $v_l > \tilde{\kappa}^4 \geq \tilde{\kappa}^3 \geq 3\tilde{\kappa}^1 \geq 3v_h$, and hence $\widetilde{\pi}(v_l, v_h) = v_l/4$. Note that, since $\tilde{\kappa}^4 \leq v^\star \leq \tilde{\kappa}^5$ and $\Delta(\tilde{\kappa}^4) = \Delta(\tilde{\kappa}^5)$, $\Delta(z) \geq \Delta(\tilde{\kappa}^4)$ for all $z \in [\tilde{\kappa}^4, \tilde{\kappa}^5]$, and hence $\overline{\beta}(z) \geq z$ on this interval. Therefore,

$$
\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{\tilde{\kappa}^4}{4} + \int_{\tilde{\kappa}^4}^{v_l} \left(\frac{\overline{\beta}(z)}{z + \overline{\beta}(z)} \right)^2 \mathrm{d}z \geq \frac{\tilde{\kappa}^4}{4} + \frac{1}{4}(v_l - \tilde{\kappa}^4) = \frac{v_l}{4} = \widetilde{\pi}(v_l, v_h)\,,
$$

as desired.

When $v_h \in (\tilde{\kappa}^1, \tilde{\kappa}^3]$,

$$
\widetilde{\pi}(v_l, v_h) = \max\left\{ \frac{v_l v_h}{v_l + v_h}, \frac{v_l}{4} \right\}.
$$

As argued above,

$$
\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \tilde{\phi}(v_l) = \frac{\tilde{\kappa}^4}{4} + \int_{\tilde{\kappa}^4}^{v_l} \left(\frac{\overline{\beta}(z)}{z + \overline{\beta}(z)} \right)^2 \mathrm{d}z \geq \frac{v_l}{4}.
$$

Moreover, since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is increasing in v_l and decreasing in v_h,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq \tilde{\phi}(\tilde{\kappa}^4) + \tilde{\psi}(\tilde{\kappa}^3) - \frac{\tilde{\kappa}^4 \tilde{\kappa}^3}{\tilde{\kappa}^4 + \tilde{\kappa}^3} = \frac{\tilde{\kappa}^4}{4} + \frac{\tilde{\kappa}^3}{4} - \frac{\tilde{\kappa}^4 \tilde{\kappa}^3}{\tilde{\kappa}^4 + \tilde{\kappa}^3} \geq 0\,,$$

as desired.

When $v_h \in (\tilde{\kappa}^3, \tilde{\kappa}^4]$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_l}{4}\right\}.$$

Again, as argued above,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \tilde{\phi}(v_l) = \frac{\tilde{\kappa}^4}{4} + \int_{\tilde{\kappa}^4}^{v_l} \left(\frac{\overline{\beta}(z)}{z + \overline{\beta}(z)}\right)^2 \mathrm{d}z \geq \frac{v_l}{4}\,.$$

Meanwhile, note that the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is increasing in v_l and decreasing in v_h. Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq \tilde{\phi}(\tilde{\kappa}^4) + \tilde{\psi}(\tilde{\kappa}^4) - \frac{\tilde{\kappa}^4}{2} = 0\,.$$

Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)$, as desired.

When $v_h \in (\tilde{\kappa}^4, \tilde{\kappa}^5]$, first note that since the function

$$(v_l, v_h) \mapsto \frac{v_l v_h}{v_l + v_h}$$

is supermodular and since $\overline{\beta}$ is increasing, by construction,

$$\tilde{\phi}(v_l) = \max_{v_h' \in [\tilde{\kappa}^4, \tilde{\kappa}^5]} \left[\frac{v_l v_h'}{v_l + v_h'} - \tilde{\psi}(v_h')\right] \geq \frac{v_l v_h}{v_l + v_h} - \tilde{\psi}(v_h)\,,$$

and thus

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \frac{v_l v_h}{v_l + v_h}\,.$$

Next, note that since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_h}{4}$$

is increasing in v_l and decreasing in v_h,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_h}{4} \geq \tilde{\phi}(\tilde{\kappa}^4) + \tilde{\psi}(\tilde{\kappa}^5) - \frac{\tilde{\kappa}^5}{4} = \frac{\tilde{\kappa}^4}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} = \frac{\tilde{\kappa}^4}{4} - \frac{\tilde{\kappa}^3}{4} + \int_{\tilde{\kappa}^2}^{\tilde{\kappa}^3} \left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)}\right)^2 \mathrm{d}z \geq 0\,.$$

Lastly, since $\overline{\beta}(z) \geq z$ on $[\tilde{\kappa}^4, \tilde{\kappa}^5]$, the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l}{4} \tag{A.22}$$

is increasing in both v_l and v_h on this rectangle: the derivative in v_l equals $(\overline{\beta}(v_l)/(v_l + \overline{\beta}(v_l)))^2 - 1/4 \geq 0$, and $\tilde{\psi}$ is increasing on $[\tilde{\kappa}^4, \tilde{\kappa}^5]$. Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l}{4} \geq \tilde{\phi}(\tilde{\kappa}^4) + \tilde{\psi}(\tilde{\kappa}^4) - \frac{\tilde{\kappa}^4}{4} = \frac{\tilde{\kappa}^4}{4} \geq 0\,. \tag{A.23}$$

Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)$, as desired.

When $v_h > \tilde{\kappa}^5$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_h}{4}\right\}\,.$$

Note that

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \frac{\tilde{\kappa}^4}{4} + \frac{v_h}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} \geq \frac{\tilde{\kappa}^3}{4} + \frac{v_h}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} = \int_{\tilde{\kappa}^2}^{\tilde{\kappa}^3} \left(\frac{\underline{\beta}(z)}{z + \underline{\beta}(z)}\right)^2 \mathrm{d}z + \frac{v_h}{4} \geq \frac{v_h}{4}\,.$$

Moreover, since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is decreasing in v_l and increasing in v_h,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq \tilde{\phi}(\tilde{\kappa}^5) + \tilde{\psi}(\tilde{\kappa}^5) - \frac{\tilde{\kappa}^5}{2} = 0\,.$$

Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)$, as desired.

Case 5: $v_l > \tilde{\kappa}^5$.

When $v_h \leq \tilde{\kappa}^1$, by (A.20), $v_l > \tilde{\kappa}^5 \geq \tilde{\kappa}^3 \geq 3\tilde{\kappa}^1 \geq 3v_h$, and hence $\widetilde{\pi}(v_l, v_h) = v_l/4$. Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{v_l}{4} + \frac{\tilde{\kappa}^2 \tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} \geq \frac{v_l}{4} = \widetilde{\pi}(v_l, v_h)\,,$$

as desired.

When $v_h \in (\tilde{\kappa}^1, \tilde{\kappa}^3]$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_l}{4}\right\}.$$

Note that

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \tilde{\phi}(v_l) = \frac{v_l}{4} + \frac{\tilde{\kappa}^2 \tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} \geq \frac{v_l}{4}\,.$$

Moreover, since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is increasing in v_l and decreasing in v_h,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq \tilde{\phi}(\tilde{\kappa}^5) + \tilde{\psi}(\tilde{\kappa}^3) - \frac{\tilde{\kappa}^5 \tilde{\kappa}^3}{\tilde{\kappa}^5 + \tilde{\kappa}^3} = \frac{\tilde{\kappa}^5}{4} + \frac{\tilde{\kappa}^3}{4} - \frac{\tilde{\kappa}^5 \tilde{\kappa}^3}{\tilde{\kappa}^5 + \tilde{\kappa}^3} \geq 0\,.$$

Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)$, as desired.

When $v_h \in (\tilde{\kappa}^3, \tilde{\kappa}^4]$,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{v_l}{4} + \frac{v_h}{4} + \frac{\tilde{\kappa}^2 \tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} \geq \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_l}{4}, \frac{v_h}{4}\right\} = \widetilde{\pi}(v_l, v_h)\,,$$

as desired.

When $v_h \in (\tilde{\kappa}^4, \tilde{\kappa}^5]$,

$$\widetilde{\pi}(v_l, v_h) = \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_l}{4}\right\}.$$

Note that

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \tilde{\phi}(v_l) = \frac{v_l}{4} + \frac{\tilde{\kappa}^2 \tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} \geq \frac{v_l}{4}.$$

Moreover, since the function

$$(v_l, v_h) \mapsto \tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h}$$

is increasing in v_l and decreasing in v_h,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) - \frac{v_l v_h}{v_l + v_h} \geq \tilde{\phi}(\tilde{\kappa}^5) + \tilde{\psi}(\tilde{\kappa}^5) - \frac{\tilde{\kappa}^5}{2} = 0.$$

Together, $\tilde{\phi}(v_l) + \tilde{\psi}(v_h) \geq \widetilde{\pi}(v_l, v_h)$, as desired.

When $v_h > \tilde{\kappa}^5$,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{v_l}{4} + \frac{v_h}{4} \geq \max\left\{\frac{v_l v_h}{v_l + v_h}, \frac{v_l}{4}, \frac{v_h}{4}\right\} = \widetilde{\pi}(v_l, v_h),$$

as desired.

This completes the verification of (A.19) for all (v_l, v_h).

We now construct a feasible primal coupling. To this end, define a transition probability $\tilde{\gamma}^\star : V \to \Delta(V)$ from h-values to l-values as follows:

$$\tilde{\gamma}^\star(v_l \leq x \mid v_h) :=$$
$$\begin{cases} \mathbf{1}\{\underline{\Delta}^{-1}(F_h(v_h) + \Delta(\tilde{\kappa}^3)) \leq x\}, & \text{if } v_h \leq \tilde{\kappa}^1 \\ \mathbf{1}\{F_l^{-1}(F_h(v_h) + \Delta(\tilde{\kappa}^3)) \leq x\}, & \text{if } v_h \in (\tilde{\kappa}^1, \tilde{\kappa}^3] \\ \mathbf{1}\{v_h \leq x\}, & \text{if } v_h \in (\tilde{\kappa}^3, \tilde{\kappa}^4] \\ \mathbf{1}\{F_l^{-1}(F_h(v_h) + \Delta(\tilde{\kappa}^4)) \leq x\}, & \text{if } v_h \in (\tilde{\kappa}^4, \tilde{\kappa}^5] \\ \frac{f_l(v_h)}{f_h(v_h)}\mathbf{1}\{v_h \leq x\} + \frac{f_h(v_h) - f_l(v_h)}{f_h(v_h)}\mathbf{1}\{F_l^{-1}(\Delta(\tilde{\kappa}^5) - \Delta(v_h)) \leq x\}, & \text{if } v_h > \tilde{\kappa}^5, \end{cases}$$

for all $x \in V$ and for all $v_h \in V$. Then, let $\tilde{\rho}^\star \in \Delta(V \times V)$ be defined as

$$\tilde{\rho}^\star(v_l \in A, v_h \in B) \coloneqq \int_B \tilde{\gamma}^\star(A \mid v_h) F_h(\mathrm{d}v_h)\,, \tag{A.24}$$

for all measurable sets $A, B \subseteq V$. By construction, the v_h-marginal of $\tilde{\rho}^\star$ is F_h. The mixing weights on $v_h > \tilde{\kappa}^5$ lie in $[0,1]$, since $\tilde{\kappa}^5 > v^\star$ and $f_l \le f_h$ on $[v^\star, \overline{v}]$ by the monotone likelihood-ratio dominance. To verify the v_l-marginal, a change of variables on each piece of $\tilde{\gamma}^\star$, using the identities $f_h(\underline{\beta}(y))\underline{\beta}'(y) = f_l(y)$, $f_h(\overline{\beta}(y))\overline{\beta}'(y) = f_l(y)$, and $\Delta' = f_l - f_h$, yields the v_l-density contributions f_l on $[0, \tilde{\kappa}^2]$, f_l on $(\tilde{\kappa}^2, \tilde{\kappa}^3]$, $f_l - f_h$ plus f_h on $(\tilde{\kappa}^3, \tilde{\kappa}^4]$, f_l on $(\tilde{\kappa}^4, \tilde{\kappa}^5]$, and f_l on $(\tilde{\kappa}^5, \overline{v}]$. Summing yields f_l on V, and hence $\tilde{\rho}^\star \in \mathcal{R}$.

It remains to verify complementary slackness. To this end, consider any $(v_l, v_h) \in \operatorname{supp}(\tilde{\rho}^\star)$. If $v_h \le \tilde{\kappa}^1$, then $v_l = \underline{\Delta}^{-1}(F_h(v_h) + \Delta(\tilde{\kappa}^3)) \in [\tilde{\kappa}^3, \tilde{\kappa}^4]$, and hence $v_l \ge 3v_h$ by (A.20). Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{v_l}{4} = \widetilde{\pi}(v_l, v_h)\,,$$

as desired. If $v_h \in (\tilde{\kappa}^1, \tilde{\kappa}^3]$, then $v_l = \underline{\beta}^{-1}(v_h)$, and by construction,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \tilde{\phi}(v_l) + \tilde{\psi}(\underline{\beta}(v_l)) = \frac{v_l \underline{\beta}(v_l)}{v_l + \underline{\beta}(v_l)} = \widetilde{\pi}(v_l, \underline{\beta}(v_l)) = \widetilde{\pi}(v_l, v_h)\,,$$

as desired. If $v_h \in (\tilde{\kappa}^3, \tilde{\kappa}^4]$, then $v_l = v_h$, and hence

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{v_h}{2} = \widetilde{\pi}(v_h, v_h) = \widetilde{\pi}(v_l, v_h)\,,$$

as desired. If $v_h \in (\tilde{\kappa}^4, \tilde{\kappa}^5]$, then $v_l = \overline{\beta}^{-1}(v_h)$, and hence

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \tilde{\phi}(v_l) + \tilde{\psi}(\overline{\beta}(v_l)) = \frac{v_l \overline{\beta}(v_l)}{v_l + \overline{\beta}(v_l)}\,.$$

Moreover, by (A.22) and (A.23),

$$\frac{v_l \overline{\beta}(v_l)}{v_l + \overline{\beta}(v_l)} = \tilde{\phi}(v_l) + \tilde{\psi}(\overline{\beta}(v_l)) \ge \max\left\{\frac{v_l}{4}, \frac{\overline{\beta}(v_l)}{4}\right\}\,,$$

and hence $\widetilde{\pi}(v_l, \overline{\beta}(v_l)) = v_l\overline{\beta}(v_l)/(v_l + \overline{\beta}(v_l))$. Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \widetilde{\pi}(v_l, v_h)\,,$$

as desired. Finally, if $v_h > \tilde{\kappa}^5$, the support of $\tilde{\rho}^\star$ either has $v_l = v_h$, in which case

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{v_h}{2} = \widetilde{\pi}(v_h, v_h) = \widetilde{\pi}(v_l, v_h)\,,$$

or it has $v_l = F_l^{-1}(\Delta(\tilde{\kappa}^5) - \Delta(v_h)) \leq \tilde{\kappa}^2$. In the latter case, by (A.21), $v_h \geq 3v_l$, and hence $\widetilde{\pi}(v_l, v_h) = v_h/4$. Therefore,

$$\tilde{\phi}(v_l) + \tilde{\psi}(v_h) = \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} + \frac{v_h}{4} - \frac{\tilde{\kappa}^2\tilde{\kappa}^1}{\tilde{\kappa}^2 + \tilde{\kappa}^1} = \frac{v_h}{4} = \widetilde{\pi}(v_l, v_h)\,,$$

as desired.

Together, complementary slackness holds on $\mathrm{supp}(\tilde{\rho}^\star)$. By Kantorovich duality, $\tilde{\rho}^\star$ solves the transport problem. As a result, the pricing rule $\tilde{p}^\star$ induced by $\tilde{\rho}^\star$ is an optimal non-discriminatory pricing rule. □

A.3.3 Unconditional Non-Discriminatory Pricing

We now solve the problem introduced in Section 5.3. For any $v_l, v_h \in V^2$ let $c_h := v_h - s$ and $c_l := v_l - s$, and let

$$\begin{aligned}\pi(v_l, v_h) &:= \max\left\{\min\{v_l, v_h\} - \frac{c_l + c_h}{2}, \frac{1}{2}(v_l - c_l), \frac{1}{2}(v_h - c_h)\right\}\\ &= \max\left\{\min\{v_l, v_h\} - \frac{v_l + v_h}{2} + s, \frac{1}{2}s\right\}\\ &= \max\left\{s - \frac{1}{2}|v_h - v_l|, \frac{1}{2}s\right\}.\end{aligned}$$

The seller's problem can be represented by

$$\max_{\rho\in\mathcal{R}} \int \pi(v_l, v_h)\,\mathrm{d}\rho\,. \tag{A.25}$$

The solution of the problem can be described as follows: Let $\underline{\kappa}$ and $\overline{\kappa}$ be the unique vector with $\underline{\kappa} \le \overline{\kappa}$ such that

$$\begin{aligned} \Delta(\underline{\kappa}) &= \Delta(\overline{\kappa}) \\ \overline{\kappa} - \underline{\kappa} &= s \end{aligned}$$

Note that by definition, since Δ is quasi-concave, $\underline{\kappa} \le v^\star \le \overline{\kappa}$, and then $f_h(v) \ge f_l(v)$ for all $v \ge \overline{\kappa}$. Define a transition probability $\gamma^\star : V \to \Delta(V)$ as follows:

$$\gamma^\star(v_l \le x \mid v_h) := \begin{cases} \mathbf{1}\{v_h \le x\}, & \text{if } v_h \le \underline{\kappa} \\ \mathbf{1}\{F_l^{-1}(F_h(v_h) + \Delta(\underline{\kappa})) \le x\}, & \text{if } v_h \in (\underline{\kappa}, \overline{\kappa}] \\ \frac{f_l(v_h)}{f_h(v_h)}\mathbf{1}\{v_h \le x\} + \frac{f_h(v_h)-f_l(v_h)}{f_h(v_h)}\mathbf{1}\{F_l^{-1}(\Delta(\overline{\kappa}) - \Delta(v_h)) \le x\}, & \text{if } v_h > \overline{\kappa} \end{cases}.$$

Then, let

$$\rho^\star(v_l \in A, v_h \in B) := \int_B \gamma^\star(A \mid v_h) F_h(\mathrm{d}v_h),$$

for all measurable $A, B \subseteq V$. It can be verified that the marginals of $\rho^\star$ are F_l and F_h, respectively, and thus $\rho^\star \in \mathcal{R}$.

Proposition A.4. *$\rho^\star$ solves the optimal transport problem* (A.25)*.*

Proof. We prove this by constructing the dual variables $\phi^\star$ and $\psi^\star$ explicitly. Let

$$\phi^\star(v_l) := \begin{cases} 0, & \text{if } v_l \le \underline{\kappa} \\ \frac{1}{2}(v_l - \underline{\kappa}), & \text{if } v_l \in (\underline{\kappa}, \overline{\kappa}] \\ \frac{1}{2}s, & \text{if } v_l > \overline{\kappa} \end{cases}$$

and let

$$\psi^\star(v_h) := \begin{cases} s, & \text{if } v_h \le \underline{\kappa} \\ \frac{1}{2}s + \frac{1}{2}(\overline{\kappa} - v_h), & \text{if } v_h \in (\underline{\kappa}, \overline{\kappa}] \\ \frac{1}{2}s, & \text{if } v_h > \overline{\kappa} \end{cases}.$$

We first verify that $\phi^\star$ and $\psi^\star$ is feasible given π. To see this, suppose first that $v_l \le \underline{\kappa}$. Then, for $v_h \le \underline{\kappa}$,

$$\phi^\star(v_l) + \psi^\star(v_h) = s \ge \max\left\{s - |v_l - v_h|, \frac{1}{2}s\right\} = \pi(v_l, v_h).$$

For $v_h \in (\underline{\kappa}, \overline{\kappa}]$,

$$\begin{aligned}\phi^\star(v_l) + \psi^\star(v_h) = \frac{1}{2}s + \frac{1}{2}(\overline{\kappa} - v_h) =& \frac{1}{2}s + \frac{1}{2}(\overline{\kappa} - \underline{\kappa}) - \frac{1}{2}(v_h - \underline{\kappa}) \\ =& s - \frac{1}{2}(v_h - \underline{\kappa}) \\ \geq s - \frac{1}{2}(v_h - v_l)\,.\end{aligned}$$

Therefore,

$$\phi^\star(v_l) + \psi^\star(v_h) = \frac{1}{2}s + \frac{1}{2}(\overline{\kappa} - v_h) \geq \max\left\{s - \frac{1}{2}(v_h - v_l), \frac{1}{2}s\right\} = \pi(v_l, v_h)\,.$$

For $v_h > \overline{\kappa}$, $v_h - v_l \geq \overline{\kappa} - \underline{\kappa} = s$ and thus

$$\phi^\star(v_l) + \psi^\star(v_h) = \frac{1}{2}s = \max\left\{s - \frac{1}{2}(v_h - v_l), \frac{1}{2}s\right\} = \pi(v_l, v_h)\,.$$

Next, suppose that $v_l \in (\underline{\kappa}, \overline{\kappa}]$. Then, for $v_h \leq \underline{\kappa}$,

$$\phi^\star(v_l) + \psi^\star(v_h) = s + \frac{1}{2}(v_l - \underline{\kappa}) \geq s \geq \pi(v_l, v_h)\,.$$

For $v_h \in (\underline{\kappa}, \overline{\kappa}]$, since $|v_h - v_l| \leq \overline{\kappa} - \underline{\kappa}$ and since $\overline{\kappa} - \underline{\kappa} = s$,

$$\begin{aligned}\phi^\star(v_l) + \psi^\star(v_h) =& \frac{1}{2}(s + (\overline{\kappa} - \underline{\kappa}) - (v_h - v_l)) \\ =& s - \frac{1}{2}(v_h - v_l) \\ \geq& s - \frac{1}{2}|v_h - v_l| \\ =& \pi(v_l, v_h)\,.\end{aligned}$$

For $v_h > \overline{\kappa}$,

$$\begin{aligned}\phi^\star(v_l) + \psi^\star(v_h) =& \frac{1}{2}s + \frac{1}{2}(v_l - \underline{\kappa}) \\ =& \frac{1}{2}s + \frac{1}{2}(\overline{\kappa} - \underline{\kappa}) - \frac{1}{2}(\overline{\kappa} - v_l) \\ \geq& s - \frac{1}{2}(v_h - v_l)\end{aligned}$$

Therefore,

$$\phi^\star(v_l) + \psi^\star(v_h) = \frac{1}{2}s + \frac{1}{2}(v_l - \underline{\kappa}) \geq \max\left\{s - \frac{1}{2}(v_h - v_l), \frac{1}{2}s\right\} = \pi(v_l, v_h).$$

Lastly, suppose that $v_l > \overline{\kappa}$. Then, for $v_h \leq \underline{\kappa}$,

$$\phi^\star(v_l) + \psi^\star(v_h) = \frac{3}{2}s > s \geq \pi(v_l, v_h).$$

For $v_h \in (\underline{\kappa}, \overline{\kappa}]$,

$$\phi^\star(v_l) + \psi^\star(v_h) = s + \frac{1}{2}(\overline{\kappa} - v_h) \geq s \geq \pi(v_l, v_h).$$

For $v_h > \overline{\kappa}$,

$$\phi^\star(v_l) + \psi^\star(v_h) = s \geq \pi(v_l, v_h).$$

Together,

$$\phi^\star(v_l) + \psi^\star(v_h) \geq \pi(v_l, v_h)$$

for all $(v_l, v_h) \in V^2$.

By Lemma 2, it remains to show that $\phi^\star(v_l) + \psi^\star(v_h) = \pi(v_l, v_h)$ for all $(v_l, v_h) \in \text{supp}(\rho^\star)$. Note that, since $\overline{\kappa} - \underline{\kappa} = s$,

$$\phi^\star(v) + \psi^\star(v) = s = \pi(v, v)$$

for all $v \in V$. Moreover, for all $v_l, v_h \in (\underline{\kappa}, \overline{\kappa}]^2$ such that $v_h \geq v_l$,

$$\phi^\star(v_l) + \psi^\star(v_h) = s - \frac{1}{2}(v_h - v_l) = \pi(v_l, v_h).$$

Lastly, for all $v_l \leq \underline{\kappa}$ and for all $v_h > \overline{\kappa}$, since $v_h - v_l > \overline{\kappa} - \underline{\kappa} = s$,

$$\phi^\star(v_l) + \psi^\star(v_h) = \frac{1}{2}s = \pi(v_l, v_h).$$

In the meantime, note that since Δ is quasi-concave, $\Delta(v) \geq \Delta(\underline{\kappa})$ for all $v \in [\underline{\kappa}, \overline{\kappa}]$. Therefore, $F_l^{-1}(F_h(v_h) + \Delta(\underline{\kappa})) \leq v_h$. Together, it follows that

$$\phi^\star(v_l) + \psi^\star(v_h) = \pi(v_l, v_h)$$

for all $(v_l, v_h) \in \text{supp}(\rho^\star)$. This completes the proof. □

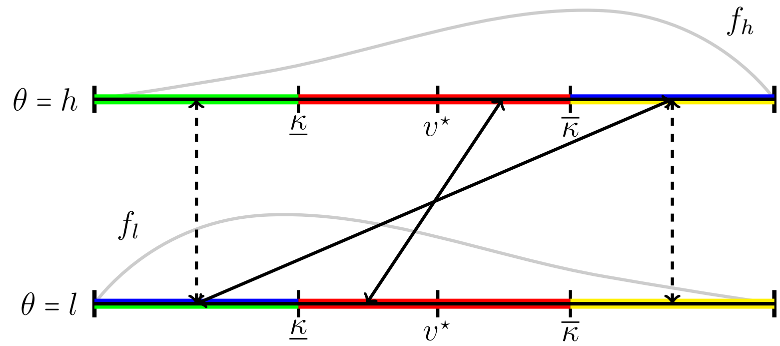


Figure 12: Optimal matching scheme $\rho^\star$

The matching $\rho^\star$ matches all consumers in $[\underline{\kappa}, \overline{\kappa}]$ assortatively (solid arrow). Then matches all the remain consumers with the same value until the masses are exhausted in one group (dashed arrows), and matches all the remaining consumers assortatively (solid arrow). Consumers labeled in the same colors are matched together.

This optimal matching scheme, while distinct from the optimal matching scheme described in Section 3, is qualitatively similar. In particular, the optimal matching here also features a mix of assortative and anti-assortative match. Consumers with intermediate values are matched together, whereas consumers with extreme values are matched.