

CAMBRIDGE WORKING PAPERS IN ECONOMICS
JANEWAY INSTITUTE WORKING PAPERS

Market segmentation through information

Matthew
Elliott
University of
Cambridge

Andrea
Galeotti
London
Business School

Andrew
Koh
Massachusetts
Institute of
Technology

Wenhao
Li
Pennsylvania
State University

Abstract

We explore the power that precise information about consumers' preferences grants an intermediary in shaping competition. We think of an intermediary as an information designer who chooses what information to reveal to firms, which then compete à la Bertrand in a differentiated product market. We characterize the information designs that maximize consumer and producer surplus, showing how information can be used to segment markets to intensify or soften competition. We also show how the power of the intermediary is further enhanced when it has some control over which products consumers are aware of.

Reference Details

2105 Cambridge Working Papers in Economics
2114 Janeway Institute Working Paper Series

Published 14 January 2021
Updated 5 May 2024

Key Words Information Design, Market Segmentation, Price Discrimination
JEL Codes D43, D83, L13

Websites www.econ.cam.ac.uk/cwpe
www.janeway.econ.cam.ac.uk/working-papers

Market Segmentation through Information*

Matthew Elliott
Cambridge

Andrea Galeotti
London Business School

Andrew Koh
MIT

Wenhao Li
Penn State

This version: May 5, 2024

Abstract

We explore the power that precise information about consumers' preferences grants an intermediary in shaping competition. We think of an intermediary as an information designer who chooses what information to reveal to firms, which then compete à la Bertrand in a differentiated product market. We characterize the information designs that maximize consumer and producer surplus, showing how information can be used to segment markets to intensify or soften competition. We also show how the power of the intermediary is further enhanced when it has some control over which products consumers are aware of.

*First version: August 2019.

This paper subsumes "Matching and Information Design in Marketplaces" by the same authors and "Using Information to Amplify Competition" by Wenhao Li. We thank Nageeb Ali, Ben Golub, Konstantin Guryev, Sergiu Hart, Navin Kartik, Stephen Morris, Ludovic Renou, and Alex Wolitzky. Jörg Kalbfuss and Alastair Langtry provided excellent research assistance. Elliott acknowledges funding from the European Research Council under the grant EMBED #757229 and the JM Keynes Fellowships Fund. Galeotti acknowledges funding from the European Research Council under the grant #724356.

1 INTRODUCTION

The last two decades have witnessed the emergence of a new internet based business model whereby revenue streams emanate from collecting and using information about users to target advertisements. Concerns about competition and users' privacy issues have attracted the attention of antitrust authorities around the world. We explore the power that information grants internet intermediaries in shaping market competition when firms are willing to offer targeted discounts. We do so by extending the information design problem with a monopolist considered by [Bergemann, Brooks, and Morris \(2015\)](#) to an oligopoly setting.

Antitrust authorities often seem to lean on two benchmarks to guide their thinking towards possible economic harms in downstream markets—complete information, in which all firms know all consumers' preferences, and no information, in which all firms know only the aggregate distribution of consumers' preferences. Comparing these cases reveals that the use of information that permits price discrimination is typically welfare enhancing and can sometimes increase consumer surplus.¹ This provides a salient, cautionary note for regulations that limit the use of information about consumer preferences.²

An intermediary who commands access to consumers' data has more options available than just choosing between either withholding all information, or disclosing all information to all firms. For example, in response to privacy concerns, Google has attempted to replace the use of third-party tracking cookies on its Chrome web browser with its "Privacy Sandbox". The "Privacy Sandbox" groups users into "cohorts" based on their browsing behaviour and targets firms' ads and promotions to these cohorts rather than to individuals. Technologies like this can package information about consumers and disclose it to firms in a fairly complicated way. The aim of this paper is to shed new light on the effect of such technologies on price competition.

We consider an information designer who chooses what information about consumers to reveal to competing firms who, then, play a simultaneous pricing game. The information designer can be thought of as an intermediary whose objective is increasing in consumer surplus and producer surplus; we study the maximal combinations of producer and consumer surplus such an intermediary can achieve.

We illustrate the main ideas with an example. There are two single product firms, A and B , with a zero marginal cost of production. A single consumer has unit demand and her type is identified by her valuations for the two products (θ_A, θ_B) . Both firms know that it is equally likely that the consumer has one of four valuations: $(1, 1/2)$, $(1/2, 1)$, $(3/8, 3/16)$ and $(3/16, 3/8)$. In the efficient allocation both types $(1, 1/2)$

¹While this debate been reopened by the possibility of data-driven price discrimination, its provenance dates back at least to [Pigou \(1920\)](#) and [Robinson \(1933\)](#). See also [Rhodes and Zhou \(2022\)](#) for a modern comparison of personalized pricing vs uniform pricing.

²For instance, the Council of Economic Advisors (CEA) report on big data and price discrimination observes that "Economic reasoning suggests that differential pricing, whether online or offline, can benefit both buyers and sellers," and goes on to conclude that "we should be cautious about proposals to regulate online pricing." ([Council of Economic Advisors, 2015](#)).

and $(3/8, 3/16)$ consume product A , and the other types consume product B ; the total available surplus is $TS^* = 11/16$. A platform knows the valuations of consumers and can disclose information to firms. Three benchmark information structures are illustrative.

Case 1: Full disclosure. The platform tells both firms the exact realised consumer's type. The equilibrium outcome is efficient, firm A sets prices $p_A(\theta_A, \theta_B) = \max\{\theta_A - \theta_B, 0\}$ and firm B sets prices $p_B(\theta_A, \theta_B) = \max\{\theta_B - \theta_A, 0\}$. The producers and consumers split TS^* equally.

Case 2: Consumer-optimal information design. The consumer optimal information design implements an equilibrium of the induced pricing game that yields more consumer surplus than can be obtained in any other equilibrium given any other information design. We consider the consumer types with a higher value for A 's product (a symmetric information structure holds for the other types.) The designer sends two messages: when the type is $(3/8, 3/16)$ firms receive message $m = (3/16, 0)$, when the type is $(1, 1/2)$ they receive m with probability $3/5$ and $m' = (1/2, 0)$ with the remaining probability. The messages are price recommendations for the two firms. For example, message $m = (3/16, 0)$ recommends firm A to set a price $3/16$ and firm B to set a price 0 . It can be easily checked that no firm wants to deviate from these recommendations, the equilibrium outcome is efficient, the producer surplus is roughly 36% of TS^* and the consumer surplus is roughly 64% of TS^* .

To see that this outcome is consumer-optimal, consider an arbitrary information structure. An option available to firm 1 is to ignore any message it receives and to set the same price to all consumers. Under this strategy, the worst case scenario for firm 1 is when firm 2 sets a price of 0 to all consumers. In this case, firm A can only hope to sell to types $(1, 1/2)$ and $(3/8, 3/16)$, and each of them is willing to pay at most $\theta_A - \theta_B$ for product A . We note that the profit maximizing uniform price for firm 1 is $1/2$ yielding it profits equal to $1/4$, roughly 18% of TS^* . This is a lower bound on firm 1's profits (and, by symmetry, firm 2's) yielding an overall lower bound on producer surplus of 36% of TS^* . As the information structure proposed above achieves this bound and all remaining surplus goes to consumers, consumer surplus is maximized.

This insight is generalized in Theorem 1. It extends to price competition the consumer optimal information structure for the monopoly problem in [Bergemann, Brooks, and Morris \(2015\)](#) and shares a similar economic logic to the construction of revenue maximizing information structure in the auction setting of [Bergemann, Brooks, and Morris \(2017\)](#) when bidders know their own values. In this example, and in general, the consumer-optimal outcome is obtained by grouping together only consumers who like the same product the most and, then, segmenting this group in a way such that the most preferred firm does not exclude any consumer. A recent paper by [Bergemann, Brooks, and Morris \(2023\)](#) generalizes this logic to incorporate the case where there is uncertainty about costs.

Case 3: Producer-optimal outcome. We now construct an information design that implements an equilibrium where all available surplus is extracted by the firms. The designer sends price-recommendation $m = (1, 1)$ if the consumer's type is either $(1, 1/2)$

or $(1/2, 1)$, and sends price-recommendation $m' = (3/8, 3/8)$ otherwise. It can be checked that firms have an incentive to follow these pricing recommendations, so TS^* is entirely captured as producer surplus.

The market has been perfectly segmented through information. A segment groups consumers who like product A the most with other consumers who value product A less and like product B the most. The exact mix of such consumers is set to create an incentive for firms to engage in a niche market strategy and price to extract all surplus from the consumers who value their product the most, while excluding the other consumers.

In general, an information structure that implements an outcome in which all surplus is extracted by the producers, which we call a *producer-perfect outcome*, does not always exist. Theorem 2 provides a necessary and sufficient condition for its existence. The condition can be interpreted as an aggregate incentive compatibility constraint for the firms. For firm i it requires that consumers can be segmented so that, aggregate infra-marginal losses from i deviating downwards to a price p whenever i is meant to set a higher price, are larger than the maximum additional profits i could gain from the deviation (i.e., those additional profits it would obtain by stealing *all* the customers from other firms that value i 's product above p). It is not obvious that aggregate incentive compatibility is necessary for a producer-perfect outcome to exist because it is optimistic about the profitability of the contemplated deviations; and it is not obvious it is sufficient because there are many other deviations available to the firms. We show that it is both necessary and sufficient. The aggregate incentive compatibility condition is easier to satisfy when consumers have a strong taste for their most preferred product (Proposition 2).

Even when the aggregate incentive compatibility constraint is not satisfied, an intermediary can still have considerable power to extract producer surplus (without foregoing efficiency). We show this in a simple but rich setting exhibiting vertical differentiation (high and low value consumers) as well as horizontal differentiation (consumers are willing to pay extra for their preferred product) by characterizing the form and value of producer-optimal segmentations (Proposition 3). The intermediary's power is gradually ameliorated as products become less horizontally differentiated and only in the extreme case as products become homogeneous does information become ineffective.

Finally, we show that in the cases when the aggregate incentive compatibility constraint fails, granting the intermediary the additional power to control which consumers can access which products³ fully restores its ability to soften competition. However, its ability to intensify competition is unaffected. We characterize all combinations of producer and consumer surplus that can be obtained when the intermediary controls access (Theorem 3), and show that these outcomes remain obtainable even when the intermediary's power to control access to products is substantially curtailed.

The main message of our analysis is that information is a powerful tool for softening and for intensifying price competition. But there is a substantial difference on the way

³See Bergemann and Bonatti (2019) for a discussion on the distinction between information and access design.

in which information is used to soften competition relative to intensify competition. A general principle underlying the producer optimal information design is that the information designer segments the market (creating possibly different segments for different firms) so that when pricing to a segment the firm's demand consists of both those consumers that value its product most highly, and those that prefer another product. This is done to induce the firm to charge a price that excludes those consumers who most prefer another product, thereby granting other firms market power over the excluded consumers and softening competition. In contrast, in the consumer-optimal information design the same segments are created for all firms, and each segment contains only consumers with the same most preferred product. This intensifies competition. Moreover, the producer of this most preferred product for a given segment is induced to price sufficiently low that all consumers in the segment want to buy their product.

Although the debate about privacy and the use of consumer data is certainly more complex and multifaceted than our analysis captures, it may nevertheless be of interest to antitrust authorities mandated to protect consumer surplus. The information structures that maximize producer and consumer surplus are relatively easy to implement—the intermediary just has to use consumer information to create (possibly firm specific) market segments and then suggest targeted discounts that differ across these segments. Moreover, both the producer-optimal and consumer-optimal information designs are consistent with privacy enhancing technologies. Hence, enhancing users' privacy need be no impediment to extracting consumer surplus—to the contrary, it may facilitate it.⁴

1.1 Related literature. Our paper contributes to a recent literature studying how information shapes consumer and producer surplus. [Bergemann, Brooks, and Morris \(2015\)](#) characterizes the consumer and producer surplus outcomes attainable when a designer can provide different information on consumer valuations to a monopolist able to price discriminate. We extend the analysis to an oligopoly setting—the introduction of competition poses additional technical challenges, but also leads to new economic insights which can be related to contemporary regulatory debates.⁵

Two papers that also consider how information shapes market competition are [Bergemann, Brooks, and Morris \(2017\)](#) and [Bergemann, Brooks, and Morris \(2021\)](#). [Bergemann, Brooks, and Morris \(2017\)](#) study an information designer that, in a first price auction, discloses to bidders information about their valuations. Bidders have different valuations and the costs for the auctioneer to serve bidders are homogeneous; this

⁴It may also be of interest to antitrust authorities that in our setting firms can sometimes collectively benefit from practices that reduce the mass of high value consumers in the market. For example, by coordinating on inefficient industry standards that make all products worse for some consumers, it may be possible for all firms to extract higher profits if the producer optimal information design is being implemented (see Section 5).

⁵A literature studied how firms choose which information about an aggregate parameter (e.g., demand shock) to share when competing, see, among others, [Novshek and Sonnenschein \(1982\)](#), [Vives \(1988\)](#), [Raith \(1996\)](#). Recent papers have taken a design approach and studied how equilibrium varies in the information structure ([Bergemann and Morris, 2013](#); [Bimpikis, Crapis, and Tahbaz-Salehi, 2019](#)).

corresponds to a Bertrand oligopoly where firms produce homogeneous products at different costs. We focus on product differentiation, whereas sellers' production costs are known; this corresponds, in the auction setting, to complete information of all bidders values and variation in the auctioneer's cost of supplying the good to different bidders. [Bergemann et al. \(2021\)](#) consider a homogeneous Bertrand oligopoly where consumers have the same valuation but they may access only a subset of price quotations. These search frictions create de-facto product differentiation because a consumer who does not observe the price quotation of a firm is effectively a consumer with zero valuation for that product. In this sense, the environment of [Bergemann, Brooks, and Morris \(2021\)](#) is a special case of our model in which product differentiation is derived via search frictions.⁶

Throughout the paper we clarify the connection of our results with these two papers. Here, we point out one substantial difference, which also highlights a methodological contribution. In [Bergemann, Brooks, and Morris \(2017\)](#) to characterise the bidder-surplus maximizing outcome (the producer-optimal outcome in our setting) is sufficient to check deviations in which bidders increase their bids (downward price-deviations in our setting).⁷ Similarly, in [Bergemann, Brooks, and Morris \(2021\)](#) only downward deviations bind at the producer-optimal outcome. We show more generally that in a Bertrand model with product differentiation, checking downward deviations is sufficient to characterise producer surplus only when firms are able to extract all available surplus. When such an outcome is not implementable, deviations in which firms increase their prices can become binding at the producer optimal outcome.

We investigate what outcomes an intermediary with exogenous consumer data can achieve by sharing the data with firms. Complementary to this, [Ali, Lewis, and Vasser-man \(2020\)](#) consider a disclosure game in which a consumer chooses some verifiable information about her preferences to convey to firms. They show that the ability to reveal only partial information can play firms against each other and intensify competition.

Our focus is on a setting in which firms are uncertain about consumer valuations, while [Roesler and Szentes \(2017\)](#) study the problem in which consumers have uncertain valuation and face a monopolist which prices uniformly; they characterise the signal structure which is best for consumers. [Armstrong and Zhou \(2019\)](#) extend this setting to the duopoly case with uniform pricing, and characterise both firm-optimal and consumer-optimal signal structures.

Our paper also relates to a burgeoning literature on markets for information broadly conceived—the transaction, pricing, and design of information (see, e.g., [Admati and Pfleiderer \(1986\)](#), [Armstrong and Vickers \(2019\)](#) [Lizzeri \(1999\)](#), [Taylor \(2004\)](#) [Calzolari and Pavan \(2006\)](#), [Bergemann and Bonatti \(2015\)](#), [Bergemann et al. \(2018\)](#), [Acemoglu et al. \(2019\)](#), [Bergemann et al. \(2019\)](#), [Fainmesser and Galeotti \(2019\)](#), [Kehoe et al. \(2018\)](#)

⁶We thank a referee for drawing our attention to the connections between our framework and these papers, and to Stephen Morris for detailed discussions on the connection.

⁷In particular, uniform upward deviations which were also employed by [Feldman et al. \(2016\)](#) to study correlated equilibria in auctions with complete information.

Montes et al. (2019), Jones and Tonetti (2020), Bounie, Dubus, and Waelbroeck (2020); also see Bergemann and Bonatti (2019) for a summary). Perhaps the closest paper to ours is Bounie, Dubus, and Waelbroeck (2021). Like us, they consider an intermediary choosing what information to reveal to firms about consumer valuations. Their paper focuses on an intermediary who can share information to a single firm or both and conduct their analysis within a Hotelling model with linear transportation costs. We abstract away from the way industry profits are shared between firms and the intermediary and study our information design problem in a general oligopoly model with differentiated products and arbitrary information structures.

2 MODEL

There is a finite set of firms, indexed $\mathcal{N} = \{1, \dots, n\}$ each of which produces a single product at constant marginal cost, which we normalize to zero for each firm.⁸ There is a continuum of consumers with unit mass each of whom demands a single unit inelastically.⁹ Consumers have different valuations for different firms: type $\theta \in \Theta = [0, 1]^n$ has value θ_1 for firm 1, θ_2 for firm 2 and so on. The distribution of consumers over Θ is given by $\mu \in \Delta(\Theta)$.¹⁰ Note that μ could have atoms and, therefore, this formulation nests both discrete and continuous types.

Define the types that value product i the most by $E_i := \{\theta \in \Theta : \theta_i > \max_{j \neq i} \theta_j\}$. We assume that consumers have strict preferences so that

$$\mu\left(\{\theta \in \Theta : |\operatorname{argmax}_j \theta_j| > 1\}\right) = 0 \quad \text{and hence} \quad \mu\left(\bigcup_j E_j\right) = 1.$$

An information designer, knowing the valuation of each consumer for each product, commits to an information structure which specifies a joint distribution over types and messages

$$\psi \in \Delta\left(\Theta \times M\right) \quad \text{such that} \quad \operatorname{marg}_M \psi = \mu$$

where $M := \prod_i M_i := [0, 1]^n$ is an n -dimensional message space. Let Ψ be the set of all information structures. Write $\psi(\cdot | m_i) \in \Delta(\Theta \times M_{-i})$ to denote the conditional joint distribution over types and other firms' messages, and $\psi_i(\cdot | \theta) \in \Delta(M_i)$ as the distribution of messages received by firm i conditional on θ .

Call $m_i \in M_i$ a message realisation for firm i . Given the messages received, firms play a simultaneous move pricing game. A pure strategy for firm $i \in \{1, \dots, n\}$ is $p_i : M_i \rightarrow [0, 1]$.¹¹ A mixed strategy for firm i is $\sigma_i : M_i \rightarrow \Delta([0, 1])$. Each con-

⁸In Online Appendix II we show that an environment with positive marginal costs is in effect identical to one with zero marginal costs under a suitable transformation of consumer valuations.

⁹All results translate into an alternate setting with a single consumer of uncertain type.

¹⁰We make the mild assumption that μ can be decomposed into a discrete measure and an absolutely continuous measure (which thus admits a density).

¹¹Given that we have normalized each firm's marginal cost to zero, the restriction on non-negative prices should be interpreted as a condition that prevents firms from pricing below their marginal costs. We could dispense of this restriction and, instead, refine away equilibria in which firms price below marginal costs (see, for example, Bergemann and Välimäki (1996)).

sumer observes the prices she is being offered by the different firms and chooses to either purchase a product which maximizes her surplus given these prices or to not purchase any product and obtain zero surplus. From standard arguments (Bergemann and Morris, 2016) we will focus, without loss, on the interpretation that messages are price recommendations.

The information designer can be thought of as an intermediary that has detailed information about consumer preferences, and chooses how to segment the market for each firm.¹²

2.1 Switching cost environment. Throughout the paper we will illustrate key ideas in a duopoly market (firm 1 and 2) with four types of consumers. A mass $\mu \in (0, 1)$ of consumers have a high valuation, normalised to 1 for their preferred product; half of them prefer product i and incur an utility loss of γ_H if they buy product $j \neq i$. The remaining mass of consumers $1 - \mu$ have a low valuation equal to $1 - v$, with $v \geq 0$; half of them prefer product i and incur an utility loss of γ_L if they buy product $j \neq i$. We assume throughout that $\min\{1 - v - \gamma_L, 1 - \gamma_H\} > 0$ so valuations are always positive. Table 1 summarizes consumer valuations and Table 2 shows the mass of each consumer type.

The parameters γ_H and γ_L capture the level of horizontal differentiation in the market and can be interpreted as switching costs. The parameters v and μ capture vertical heterogeneity in consumer preferences. By changing these parameters we alter the level of price competition and we will explore how these interact with information design in shaping market outcomes.

Table 1: Valuations

	Preferred Product	
	Firm 1	Firm 2
High	$(1, 1 - \gamma_H)$	$(1 - \gamma_H, 1)$
Low	$(1 - v, 1 - v - \gamma_L)$	$(1 - v - \gamma_L, 1 - v)$

Table 2: Distribution

	Preferred Product	
	Firm 1	Firm 2
High	$\mu/2$	$\mu/2$
Low	$(1 - \mu)/2$	$(1 - \mu)/2$

3 CONSUMER-OPTIMAL INFORMATION DESIGNS

We first establish an upper bound on consumer surplus and then construct an information structure that implements it. Firm i can always ignore any information received

¹²Bergemann and Morris (2013, 2016) consider many-player settings and examine how the informational environment maps to resultant equilibria. With a single receiver, Kamenica and Gentzkow (2011) show that concavification of the designer's payoff as a function of receiver's posteriors binds the designer's maximum attainable utility (see also Kamenica (2019)). However, there are well-known difficulties applying such techniques when the type space is large, multidimensional, with many receivers. We show that it is helpful to reframe certain information design problems as matching problems (see also Dworzak and Kolotilin (2022); Kolotilin, Corrao, and Wolitzky (2022)). More recently, Smolin and Yamashita (2023) develops a verification approach for information design in concave games with many players; because ours is a model of price competition, payoffs in our setting are non-smooth (and not necessarily concave) in prices.

by the designer and set a uniform price p_i . Even if all other firms charge a uniform price of 0, which is the worst case scenario for firm i , firm i can guarantee itself profits of

$$\underline{\Pi}^* = \max_{p_i \in [0,1]} p_i \int_{\Theta: \theta_i - p_i \geq \max_{j \neq i} \theta_j} d\mu.$$

Profits $\sum_i \underline{\Pi}_i^*$ are therefore a lower bound on the producer surplus that obtains for any information design. Let $TS^* = \sum_{i=1}^n \int_{\theta \in E_i} \theta_i d\mu$ be the total surplus available in the economy. An upper bound on consumer surplus over all information structures and all equilibria is

$$CS^* = TS^* - \sum_{i=1}^n \underline{\Pi}_i^*.$$

A starting point will be to suppose that i is a monopolist so an information structure specifies a joint distribution over valuations for i (θ_i), and price recommendations (m_i). Let $\sigma_{\mu_i}^*$ be the consumer-optimal uniform-profit preserving information structure developed in [Bergemann, Brooks, and Morris \(2015\)](#) when the distribution of valuations for the monopolist is given by $\mu_i \in \Delta([0, 1])$. Let $\sigma_{\mu_i}^*(\cdot | \theta_i) \in \Delta(M_i)$ denote the distribution over price recommendations given the valuation θ_i .

We now construct the consumer-optimal information structure ψ^* for the oligopoly case. Define the projection $\Lambda_i(\theta) = \theta_i - \max_{j \neq i} \theta_j$ which sends each type $\theta \in E_i$ to her residual valuation for i i.e., the difference between that consumer's valuation for i , and her favourite product. Define $\mu_i^* := \frac{1}{\mu(E_i)} \mu \circ \Lambda_i^{-1}$ as the distribution of residual valuations for firm i .¹³ For each firm i and each $\theta \in E_i$,

$$\psi^* \left((\cdot, m_{-i} = 0) | \theta \right) = \sigma_{\mu_i^*}^* \left(\cdot | \Lambda_i(\theta) \right).$$

In words, for each $\theta \in E_i$, the conditional distribution over firms' messages is such that (i) the distribution over firm i 's messages is the same as that under the consumer-optimal information structure of [Bergemann, Brooks, and Morris \(2015\)](#) facing the distribution of i 's residual valuations (μ_i^*); and (ii) all firms other than i receive the message 0. We note that the design ψ^* is not equivalent to a design in which firms are given all information about all consumers' valuations, which we refer as to the full information design.

Theorem 1. The consumer-optimal surplus is CS^* and the design ψ^* is consumer-optimal.¹⁴ The full information design is consumer-optimal if and only if for all firms i , all consumers in E_i have the same residual valuation.

Our construction of the consumer-optimal information structure shares a similar economic logic to the construction of revenue-maximizing (bidder surplus-minimizing) information structure in [Bergemann, Brooks, and Morris \(2017\)](#) when bidders know their own value. In both cases the information designer publicly reveals the identity of

¹³That is $\Lambda_i^{-1}(E) := \{\theta \in \Theta : \Lambda_i(\theta) = E\}$ for all $E \in \mathcal{B}([0, 1])$.

¹⁴We observe that, by construction, the producer-surplus implemented under ψ^* is the lowest producer surplus that can be implemented in any equilibrium for any information structure.

the highest value player. This is the highest value bidder in the auction; in our setting, it is the firm that produces the consumer's ideal product. By disclosing this information, the other players learn their comparative disadvantages which, in turn, intensify competition: in the auction the non-highest value bidders bid their value and in our setting the firms offering a non-ideal match to a consumer charge a price which is equal to their marginal cost.

3.1 Consumer-optimal in the switching cost environment. We illustrate Theorem 1 for the environment of Section 2.1. Supposing firm j sets a price 0 to all consumers, firm i will face demand $1/2$ at price γ_L and demand $\mu/2$ at price γ_H . Hence, a lower bound of firm i 's profit is given by

$$\underline{\Pi}^* = \max\left(\frac{\mu\gamma_H}{2}, \frac{\gamma_L}{2}\right),$$

and the consumer-optimal surplus is $CS^* = TS^* - \max(\mu\gamma_H, \gamma_L)$, where $TS^* = 1 - v(1 - \mu)$.

We now construct the consumer-optimal design ψ^* . Consider first the case $\gamma_H\mu \geq \gamma_L$ and let $x = [\gamma_L(1 - \mu)]/[2(\gamma_H - \gamma_L)]$. The designer assigns a mass $\mu/2 - x$ of high-value consumers in E_1 to message $m_1 = (\gamma_H, 0)$ and all the remaining consumers in E_1 to message $m_2 = (\gamma_L, 0)$. Firms have incentives to follow the price-recommendations. Note that the mass of high-value consumers x has been chosen so that firm 1 is indifferent between charging γ_H and γ_L upon receiving the price recommendation γ_L .¹⁵ Under ψ^* , firm 1 achieves the lower bound profit $\underline{\Pi}^*$ and all consumers in E_1 buy from firm 1. In the second case, when $\gamma_H\mu < \gamma_L$, the designer assigns all consumers in E_1 to a single pair of recommendations $m = (\gamma_L, 0)$.

In both cases the consumer-optimal information design differs from the full information design. Under the full information design each firm would get $\mu\gamma_H + (1 - \mu)\gamma_L$ which is strictly higher than $\underline{\Pi}^*$ if and only if $\gamma_H \neq \gamma_L$. It is only when the residual valuation of high and low types are the same, i.e., $\gamma_L = \gamma_H$, that the full information design is consumer-optimal.¹⁶

4 PRODUCER-PERFECT INFORMATION DESIGNS

We first characterize conditions under which there exist information structures such that, in an equilibrium of the resultant pricing game, the following property holds:

P (Producer-perfect outcome) almost all consumers pay their max valuation i.e., type θ pays $\max_i \theta_i$.

¹⁵The assumption that $\gamma_H\mu \geq \gamma_L$ assures that this mass is feasible, i.e., $x \leq \mu/2$.

¹⁶The consumer optimal information design is based on firms being unable (or unwilling) to charge prices below their marginal costs. If firms could set such prices then there would exist an equilibrium, following the full information design, in which all potential surplus is extracted as consumer surplus. For example, a consumer that has value 10 for product 1 and value 8 for product 2 could be charged a price of 0 by firm 1, a price of -2 by firm 2, and resolve indifference in favor of buying from firm 1.

Condition P requires that all available surplus is extracted by the firms as producer surplus such that the outcome is efficient and there is no consumer surplus. It is the outcome that would obtain under perfect collusion when transfers are possible, and is also efficient. Let $\Gamma(\psi)$ denote the Bayesian pricing game induced by the information structure ψ . Let Γ^* denote the set of induced games in which there exists an equilibrium satisfying condition P, and let $\Psi^* := \{\psi : \Gamma(\psi) \in \Gamma^*\}$ be the set of information structures that can be used to fulfil condition P. We refer to $\psi \in \Psi^*$ as a producer-perfect information structure and to the induced outcome as the producer-perfect outcome. We say that a producer-perfect information structure exists whenever $\Psi^* \neq \emptyset$.

Suppose an information structure induces a producer-perfect outcome. Then consumers of type $\theta \in E_i$ must buy from firm i at a price $p_i = \theta_i$. A possible deviation available to firm i is to then deviate downwards to a price \hat{p}_i whenever it is supposed to set a price above \hat{p}_i to all consumers of types $\theta \in E_i$ such that $\theta_i > \hat{p}_i$. At this lower price firm i will continue to sell to all these consumers and might be able to make some additional sales to consumer types $\theta' \notin E_i$. Indeed, there is an upper bound on the additional sales firm i can possibly make via such a deviation. At best, firm i can make additional sales to all those consumer types $\theta' \notin E_i$ who value i 's product weakly above \hat{p}_i . Thus a sufficient condition for no firm to ever want to deviate downwards like this is

$$\int_{\theta \in E_i: \theta_i > \hat{p}_i} (\theta_i - \hat{p}_i) d\mu \geq \hat{p}_i \int_{\theta \in \Theta \setminus E_i: \theta_i \geq \hat{p}_i} d\mu$$

for all $\hat{p}_i < \sup\{\theta_i : \theta \in E_i\}$ and all firms i . (AIC)

where AIC abbreviates aggregate incentive compatibility for reasons which will soon be apparent. The left-hand side of this inequality is firm i 's aggregate infra-marginal losses from setting price $\hat{p}_i \leq \theta_i$ instead of θ_i to all consumers in E_i with valuations above \hat{p}_i , and the right-hand side is the maximum business stealing profit that firm i can hope to obtain from such a deviation.

It is not obvious that condition AIC needs to be satisfied in a producer-perfect design, or that satisfying it is sufficient to achieve the producer-perfect outcome—it only considers some very particular deviations and, for those deviations, it may be overly optimistic about the profitability of them. Our next result shows that AIC is exactly what is required to implement the producer-perfect outcome.

Theorem 2. A producer-perfect information structure exists if and only if the aggregate incentive compatibility condition (AIC) holds.

We outline the proof of Theorem 2 in Section 4.1. In Section 4.2 we illustrate when the AIC condition holds; Section 4.3 showcases AIC in the switching cost environment; Section 4.4 shows that when AIC holds all efficient outcomes are implementable by some information design.¹⁷

¹⁷In Bergemann, Brooks, and Morris (2017), the analogous outcome to our producer-perfect outcome

4.1 Outline of proof of Theorem 2. The first steps are to show that a producer-perfect information design must satisfy several conditions. A basic condition is that ψ must agree with the actual distribution of consumers:

$$\text{marg}_{\Theta}\psi = \mu \quad \text{(Consistency)}$$

Furthermore, all consumers must buy their most preferred product and pay their full valuation for it. This implies that the messages firm i receives must perfectly separate consumers $\theta, \theta' \in E_i$ with different values $\theta_i \neq \theta'_i$ for product i :

$$\int_{E_i \times M: \theta_i = m_i} d\psi = \mu(E_i) \quad \text{for all firms } i \in \mathcal{N} \quad \text{(Separation)}$$

We have argued that each firm i charges types $\theta \in E_i$ a price of θ_i . It remains to define how to assign types not in E_i to price recommendations M_i in a way that firm i follows the price recommendations (**Firm IC**) and consumers in E_i buys product i (**Consumer IC**).

We start with **Consumer IC**. Consider a consumer of type $\theta \in E_j$ and suppose all firms follow the price recommendations. By **Separation**, firm j charges θ_j to this consumer. If firm i receives message m_i about this consumer firm i will set a price equal to m_i and hence the consumer can buy product i at a price m_i . So, for the consumer to instead buy firm j 's product at price θ_j , we need that θ_j is lower than m_i . This is what **Consumer IC** states:

$$\int_{\Theta \setminus E_i \times M: m_i \leq \theta_i} d\psi = 0 \quad \text{for all firms } i \in \mathcal{N} \quad \text{(Consumer IC)}$$

We finally consider **Firm IC**. **Separation** and **Consumer IC** implies that a firm never wishes to charge a price above the price recommendation (as demand will be zero). Hence, we only need to prevent that, upon receiving message m_i , undercutting deviations to $\hat{p}_i < m_i$ are not profitable: the infra-marginal losses for consumers in E_i (now being charged a price less than their valuations) must be greater than the extra profits made via any additional sales to consumers not in E_i :

$$\underbrace{(m_i - \hat{p}_i) \int_{E_i \times M_{-i}: \theta_i = m_i} d\psi(\cdot | m_i)}_{\text{Infra-marginal losses}} \geq \underbrace{\hat{p}_i \int_{\Theta \setminus E_i \times M_{-i}: \theta_i \geq \hat{p}_i} d\psi(\cdot | m_i)}_{\text{Business-stealing gains}} \quad \text{(Firm IC)}$$

for all firms $i \in \mathcal{N}$, all $m_i \in M_i$ ¹⁸ and all $\hat{p}_i < m_i$. Note that this implies that for almost all consumers, firm i receives a message $m_i \in \{\theta_i : \theta \in E_i\}$. Thus, our information design problem can be recast as a problem of matching types $\theta \notin E_i$ to messages $\{\theta_i : \theta \in E_i\}$ for each firm i .

is an outcome in which the winning bidder pays 0 for the item. There is no information structure which implements this outcome.

¹⁸Except for possibly a zero (Lebesgue) measure set $M \subset [0, 1]$ where for each $m \in M$, $\mu(\{E_i : \theta_i = m\}) = 0$.

Lemma 1 summarizes the properties of a producer-perfect information design.

Lemma 1. A producer-perfect information design exists if and only if there exists an information structure ψ which, for all firms $i \in \mathcal{N}$, satisfies **Separation**, **Consistency**, **Consumer IC** and **Firm IC**.

Let ψ satisfy **Separation**, **Consistency** and **Consumer IC**. The next step in proving Theorem 2 is to determine the maximum mass of types not in E_i that can be matched to each of firm i 's message $m_i \in M_i$ without violating one of firm i 's incentive compatibility conditions. Thus, we consider the mass of types not in E_i that can be assigned to a given message m_i for firm i that makes firm i indifferent between following the recommendation and deviating to *any* price $\hat{p}_i \leq m_i$. For all $m_i \in E_i$, this matching capacity is given by tightening and rearranging **Firm IC**:

$$\int_{\Theta \setminus E_i \times M_{-i}; \theta_i \geq \hat{p}_i} d\psi(\cdot | m_i) = \frac{(m_i - \hat{p}_i)}{\hat{p}_i} \int_{E_i \times M_{-i}; \theta_i = m_i} d\psi(\cdot | m_i),$$

where the right-hand-side is the exact measure of consumers not in E_i with valuation for i 's product in $[\hat{p}_i, m_i)$ so that, if matched to m_i , makes firm i indifferent between following the recommendation and charging m_i and deviating down to \hat{p}_i . In other words, the maximal matching capacity for each message m_i and each deviation $\hat{p}_i < m_i$. We define this as:

$$G_i(\hat{p}_i | m_i) := \frac{(m_i - \hat{p}_i)}{\hat{p}_i} \int_{E_i \times M_{-i}; \theta_i = m_i} d\psi(\cdot | m_i).$$

To find the overall capacity for matching consumer types $\theta \notin E_i$ to messages M_i' we integrate $G_i(\hat{p}_i | m_i)$ across all possible price recommendations greater than \hat{p}_i that firm i can receive. Formally, we define the function

$$H_i(\hat{p}_i) := \int_{m_i > \hat{p}_i} G(\hat{p}_i | m_i) d\psi_i = \int_{E_i; \theta_i > \hat{p}_i} \frac{(\theta_i - \hat{p}_i)}{\hat{p}_i} d\mu,$$

where $\psi_i := \text{marg}_{M_i} \psi$ is the marginal distribution over firm i 's messages. The second equality follows from noting that after integrating over all messages greater than \hat{p}_i , the integral against ψ restricted to types E_i concentrates on the messages $[\hat{p}_i, 1]$ for firm i (by **Separation**) so we can replace ψ with μ .

The value of $H_i(\hat{p}_i)$ gives us a maximum measure of types not in E_i with a value for product i larger than \hat{p}_i that can be matched to messages $m_i > \hat{p}_i$ if we wish to bind all firm i 's IC constraints. But the available mass of consumers not in E_i which have at least value \hat{p}_i for i 's product is

$$\int_{\Theta \setminus E_i; \theta_i \geq \hat{p}_i} d\mu.$$

Hence, if this mass of consumers is greater than $H_i(\hat{p}_i)$ then we cannot construct a producer-perfect structure: there is no way to assign all these consumers messages

$m_i \in \{M_i : m_i > \hat{p}_i\}$ without firm i sometimes having a profitable deviation to capture some of these consumers. On the other hand, if this mass of consumers is weakly less than $H_i(\hat{p}_i)$ for all \hat{p}_i , then there is a way of assigning the consumers not in E_i messages $m_i \in M_i$ such that firm i wants to follow the price recommendation m_i . These steps, as well as the prior ones, are formalized in Appendix A.

We have shown that a producer-perfect information design exists if and only if for all firms i and all prices \hat{p}_i ,

$$H_i(\hat{p}_i) \geq \int_{\Theta \setminus E_i : \theta_i \geq \hat{p}_i} d\mu,$$

which is exactly the AIC condition.

4.2 The Aggregate Incentive Compatibility Condition. We first observe that the AIC condition fails when firms offer products which are highly substitutable and, therefore, price competition is intense.¹⁹ The extreme case is one where firms offer homogeneous products; that is, for each consumer type θ , $\theta_i = \theta_j$ for all products i and j , but some consumer types may have higher valuation than others.²⁰

Proposition 1 (Product differentiation is necessary for producer profits). Suppose firms offer homogeneous products. Then for any $\psi \in \Psi$ and any equilibrium induced by ψ , each consumer buys from some firm at a price of zero and all firms make zero profits.

When products are homogeneous, if all firms have no further information beyond their common prior μ , standard undercutting arguments imply producer surplus must be zero and consumers are charged 0. If all firms observe a common public signal prior to making the pricing decision then the same logic applies to the posterior distribution induced by each message realization. The merit of Proposition 1 is to show that this undercutting logic also applies when the designer sends private price recommendations to firms. In particular, consider the highest transaction price \bar{p} that is ever paid in equilibrium to any firm. Consider now the different prices charged by the firms to these consumers. As products are homogeneous, these prices must be weakly higher than \bar{p} . It turns out that even with private information, at least one firm must have a profitable deviation to just undercut unless $\bar{p} = 0$.

We now show that an increase in the polarization of consumer preferences with respect to product offering relaxes the AIC condition.

Definition 1. Consumers' preferences are more polarized under distribution $\hat{\mu}$ relative to distribution μ whenever, for each firm i , (i) the distribution of valuations for firm i among consumers in E_i under $\hat{\mu}$ dominates that under μ such that for all x ,

$$\hat{\mu}(\{\theta \in E_i : \theta_i \geq x\}) \geq \mu(\{\theta \in E_i : \theta_i \geq x\});$$

¹⁹There exists $\epsilon > 0$ such that if, for each $\theta \in \Theta$, $|\theta_i - \theta_j| < \epsilon$ for all products i and j , then the producer-perfect outcome is not feasible.

²⁰More formally, $\mu(\{\theta \in \Theta : \theta_1 = \theta_2 = \dots, \theta_n\}) = 1$. Note that we are temporarily relaxing the assumption that almost all consumers have strict preferences.

and (ii) the distribution of valuations for firm i among consumers not in E_i under $\hat{\mu}$ is dominated by that under μ such that for all x ,

$$\hat{\mu}(\{\theta \notin E_i : \theta_i \geq x\}) \leq \mu(\{\theta \notin E_i : \theta_i \geq x\}).$$

Proposition 2 (Polarization aids segmentation). If a producer-perfect information structure exists under μ , and $\hat{\mu}$ is more polarized than μ , then a producer-perfect information structure exists under $\hat{\mu}$.

There are various ways in which consumers' preference can become more polarised. An obvious avenue is for firms to make their products more differentiated.²¹ We refer to [Johnson and Myatt \(2006\)](#) for examples on how firms can use product design and advertising to shape the distribution of consumers' preferences. As these managerial options increase polarization they also aid the feasibility of the producer-perfect outcome.

On the other hand, firm actions that uniformly increase the value consumers place on one product relative to another, thereby skewing the mass of consumer valuations towards a particular firm can inhibit the ability to achieve the producer-perfect outcome. This is because the firm with a reduced consumer base has stronger incentives to undercut other firms. Hence, imbalanced competition in which some firms have a much smaller market share than others can severely inhibit an intermediary from implementing a producer-perfect outcome.²²

Furthermore, it can also be seen that a merger weakens the AIC condition. Consider, for instance, a merger between two firms i and j into a single firm k . Clearly, this slackens downward deviations for the new firm k : since there are fewer consumers firm k must not sell to in a producer-perfect equilibrium, downward deviations are easier to deter. At the same time, the downward deviations for all other firms remain as before so AIC is slackened.

4.3 Producer-perfection in the switching cost environment. We apply [Theorem 2](#) to the switching cost environment ([Section 2.1](#)) for the case $\gamma_H = \gamma_L = \gamma \geq v$. It suffices to check the AIC condition for one firm (by the symmetry of the type distribution). AIC requires two aggregate downward pricing deviations to be unprofitable. First to the price $1 - \gamma$ to steal high value consumers who prefer the other firm:

$$\underbrace{\gamma \cdot \frac{\mu}{2}}_{\text{inframarginal losses high value consumers}} + \underbrace{(\gamma - v) \cdot \frac{1 - \mu}{2}}_{\text{inframarginal losses low value consumers}} \geq \underbrace{(1 - \gamma) \frac{\mu}{2}}_{\text{potential consumer stealing gains}} \quad (\text{AIC-High})$$

²¹In the switching-cost environment increasing polarization is equivalent to increase the switching costs, see [Section 4.3](#).

²²This is similar to the argument that, in the presence of switching costs, firms with a smaller customer base can be a stronger competitive constraint on market behaviour than more established firms ([Klemperer, 1995](#)). Based in part on this logic the UK antitrust authorities prohibited the acquisition of Abbey National by Lloyds TSB Group in 2001.

and, second, to the price $1 - \gamma - v$ to steal low value consumers who prefer the other firm:

$$\underbrace{(\gamma + v) \cdot \frac{\mu}{2}}_{\substack{\text{inframarginal losses} \\ \text{high value consumers}}} + \underbrace{\gamma \cdot \frac{1 - \mu}{2}}_{\substack{\text{inframarginal losses} \\ \text{low value consumers}}} \geq \underbrace{(1 - \gamma - v) \frac{1}{2}}_{\substack{\text{potential consumer} \\ \text{stealing gains}}}. \quad (\text{AIC-Low})$$

AIC-High and **AIC-Low** simplify to

$$\text{either (i) } \gamma \geq \frac{1}{2} \quad \text{or} \quad \text{(ii) } \underline{\mu}(\gamma, v) := \frac{1 - v - 2\gamma}{v} \leq \mu \leq \frac{\gamma - v}{1 - \gamma - v} := \bar{\mu}(\gamma, v).$$

When product differentiation is sufficiently large (condition (i) holds) a producer-perfect outcome can be implemented via public information: firms receive a public message revealing whether the consumer has a high or low value, but they don't learn which product the consumer prefers. As product differentiation is sufficiently large, each firm prefers to charge a high price and extract the surplus from the consumers who like their product the most, instead of stealing the consumers who value their rival's product more.

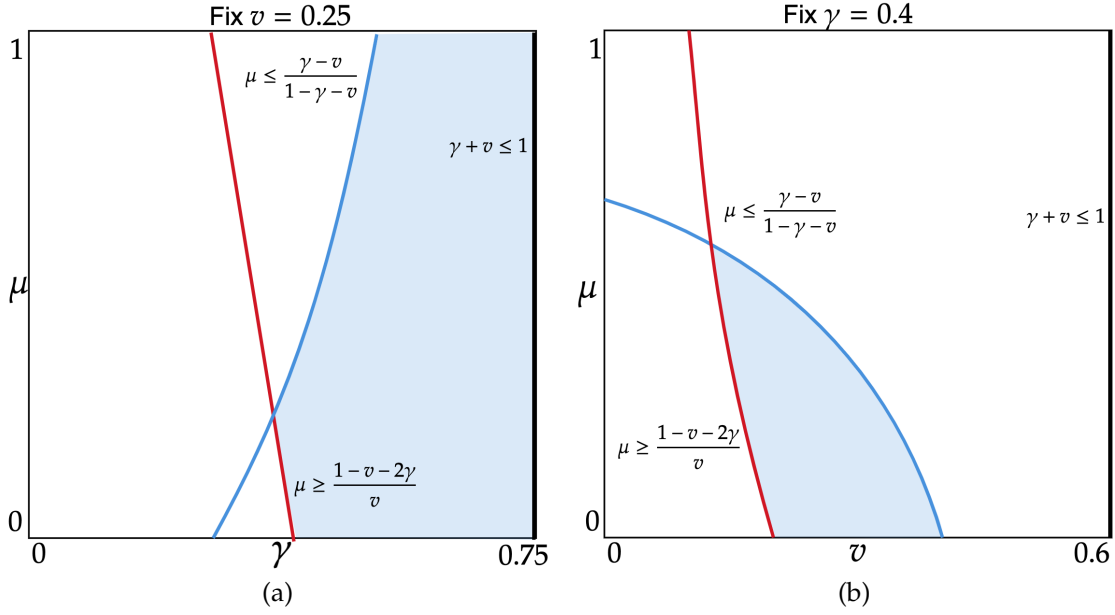
When product differentiation is not sufficiently large (condition (i) fails), the business stealing effect is too strong and, under the above public information design, firms profitably undercut the recommended price. To sustain a producer-perfect outcome the designer must now send two private messages to each firm. One message tells firm 1 to charge 1 to a segment that contains all high types who prefer product 1, and a fraction α of high types, and a fraction β of low types who prefer product 2. The other message tells firm 1 to charge $1 - v$ to a segment that contains all low types who prefer product 1, and all the remaining consumers who prefer product 2.

A producer-perfect outcome exists when we can find α and β so that it is profitable for firm 1 to follow these price recommendations. This requires that the fraction of high types μ takes an intermediate value (condition (ii)). For intermediate values of μ , we can find α and β so that firm 1 finds downward pricing deviations to steal firm 2's high value consumers unprofitable, while at the same time finding downward deviations to steal firm 2's low value consumers unprofitable; When the relative mass of high or low value consumers is too high, one of these deviations is always profitable.

Figure 1 illustrates the region of the parameter space under which AIC holds (shaded light-blue area). If $\gamma = 0$ then the maximum amount of producer surplus which can be extracted via information is 0 (Proposition 1) and indeed AIC is never fulfilled. An increase in γ makes consumers' preferences more polarized (as in Definition 1) and so it facilitates AIC (Proposition 2).²³ Further note that the effect of v on AIC is ambiguous (Panel (b)): increasing v decreases the inframarginal losses from low-types to deviate to $1 - \gamma$ and so it tightens **AIC-High**, however it increases the inframarginal losses

²³For each firm i , an increase in γ keeps the distribution of valuations for i in E_i unchanged, but decreases the valuations for i in E_j

Figure 1: Parameters under which a producer-perfect outcome exists



from high-value consumers and decreases the potential gains from stealing consumers to deviate to $1 - \gamma - v$ and so it relaxes **AIC-Low**.

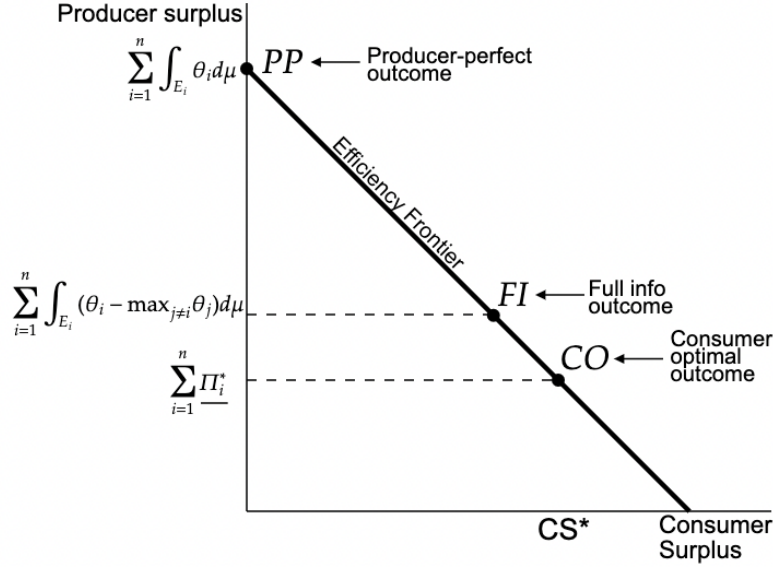
In Online Appendix **IV** we consider some other canonical duopoly examples. A first benchmark is when consumers' valuations are uniformly distributed over the unit square; in this case AIC always holds. A second benchmark is the Hotelling model when the valuations for the two products (θ_1, θ_2) are anti-correlated, i.e., $\theta_2 = 1 - \theta_1$ and θ_1 is uniformly distributed over the unit interval. Also in this case the AIC holds. When, instead, θ_1 is drawn from a truncated (at 0 and 1) normal distribution with mean $1/2$ and variance σ^2 , the AIC condition holds for $\sigma \geq 0.15$. As the variance increases, consumers preferences become more polarised which slackens constraints posed by **Firm IC**, consistent with Proposition 2.

4.4 Efficient Frontier under Aggregate Incentive Compatibility. By combining Theorem 1 and Theorem 2 we obtain a characterization of the efficient outcomes that are implementable when the AIC condition holds which is illustrated in Figure 2.

Theorem 1 established that the consumer-optimal outcome is efficient; this is point CO in Figure 2. It also established that the full information design implements an efficient outcome, but often is not consumer-optimal; this is point FI in Figure 2. These two outcomes can always be implemented, and, therefore we can always implement any other efficient outcome between them. Theorem 2 established that, when the AIC condition is met, the designer can allocate all available economic surplus to producers; this is illustrated by point PP in Figure 2.

We note that whenever AIC holds, all points between PP and CO are obtainable. To see this, suppose we wish to obtain a point $X = \lambda PP + (1 - \lambda)CO$ for some $\lambda \in (0, 1)$. We can partition the distribution of consumers μ into $\mu_{PP} = \lambda \cdot \mu$ and $\mu_{CO} = (1 - \lambda) \cdot \mu$ and then apply the producer-perfect information design to μ_{PP} and the consumer-optimal

Figure 2: Efficient Information Structures under AIC



design to μ_{CO} . Since AIC condition holds for μ it also holds for μ_{PP} because this is simply a renormalization of measure.

5 EFFICIENT FRONTIER WHEN AGGREGATE INCENTIVE COMPATIBILITY FAILS

In this section we provide full characterization of the the efficient frontier within the switching-cost environment. This amounts to characterizing the maximum amount of producer surplus that can be obtained across any efficient outcome. It showcases the potential for information to relax price competition when the level of product differentiation is low enough that AIC fails. At the same time, this makes a methodological point by illustrating the essential role that upward deviations play in constraining the amount of producer surplus which can be extracted.

Recall that total available surplus is $TS^*(v, \mu) = 1 - v(1 - \mu)$. Let $PS^*(\mu, \gamma, v)$ denote the highest producer surplus which can be implemented through information, conditional on efficiency. We refer to $PS^*(\mu, \gamma, v)$ as the producer-optimal surplus, and the information structure which implements $PS^*(\mu, \gamma, v)$ as the producer-optimal information structure.²⁴

We set $\gamma_H = \gamma_L = \gamma \in [v, 1/2)$. We have already discussed that when $\gamma \geq 1/2$ the producer-perfect outcome is implementable. Here we focus on the case where $\gamma \geq v$; we develop the case of $\gamma < v$ in Appendix B.4 as the main ideas are similar. Let

$$\hat{\mu}(\gamma, v) := 1 - v(1 - \bar{\mu}(\gamma, v)) \quad \text{and} \quad \underline{\gamma}(v) := \begin{cases} \frac{1-v}{2} - \frac{\sqrt{\frac{1}{4} + v - 3v^2} - \frac{1}{2}}{2} & \text{if } v \leq 1/3 \\ \frac{1-v}{2} & \text{if } v > 1/3. \end{cases}$$

²⁴If (i) the designer is constrained to send a finite number of messages and (ii) firms, upon receipt of these messages play pure strategies, then $PS^*(\mu, \gamma, v)$ is also the highest producer surplus implementable across all equilibria—efficient or not. We prove this in Online Appendix III.

Proposition 3. Suppose $v \leq \gamma < 1/2$. The following characterizes the producer surplus obtained at the producer-optimal efficient outcome:

(i) If $\underline{\gamma}(v) \leq \gamma$:

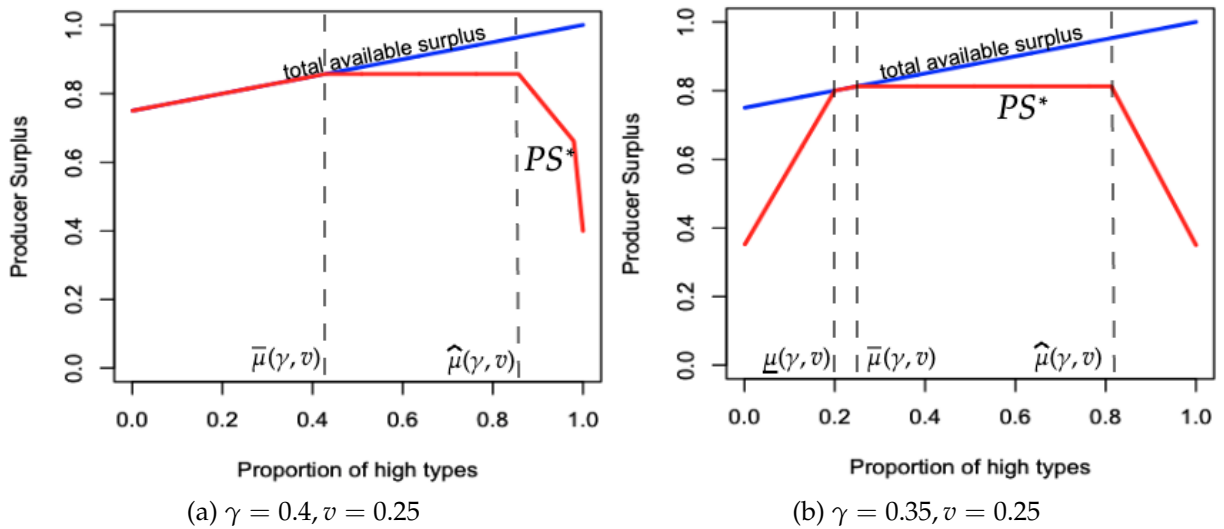
$$PS^*(\mu, \gamma, v) = \begin{cases} \gamma + \mu \cdot \frac{TS^*(v, \underline{\mu}(\gamma, v)) - \gamma}{\underline{\mu}(\gamma, v)} & \text{if } \mu < \underline{\mu}(\gamma, v) \\ TS^*(v, \mu) & \text{if } \underline{\mu}(\gamma, v) \leq \mu \leq \bar{\mu}(\gamma, v) \\ TS^*(v, \mu) - v(\mu - \bar{\mu}(\gamma, v)) & \text{if } \bar{\mu}(\gamma, v) \leq \mu \leq \hat{\mu}(\gamma, v) \\ g(\mu, \gamma, v) & \text{if } \hat{\mu}(\gamma, v) \leq \mu, \end{cases}$$

where g is continuous, strictly decreasing in μ , and $g(1, \gamma, v) = \gamma$.

(ii) If $\underline{\gamma}(v) > \gamma$:

$$PS^*(\mu, \gamma, v) = \gamma.$$

Figure 3: Illustration of optimal producer surplus as distribution varies ($\gamma \geq v$)



It turns out that if ψ implements an efficient equilibria, the support of price recommendations must be in the set $\{0, \gamma, 1 - v, 1\}$.²⁵ We split our discussion of Proposition 3 into three cases. We first consider case A with $\gamma \geq \frac{1-v}{2}$ so that downward deviations to $1 - v - \gamma$ are always unprofitable. We then consider case B with $\underline{\gamma}(v) \leq \gamma < \frac{1-v}{2}$ so that downward deviations to $1 - v - \gamma$ are potentially profitable and must be deterred by the information design. We finally consider case C with $\gamma < \underline{\gamma}(v)$.

Case A. $PS^*(\mu, \gamma, v)$ is depicted in Panel (a) of Figure 6. Note that $\underline{\mu}(\gamma, v) \leq 0$ and so, when $\mu \leq \bar{\mu}(\gamma, v)$, the producer-perfect outcome obtains (see Section 4.3); in the figure, $PS^*(\mu, \gamma, v)$ (red line) coincides with $TS^*(\mu, v)$ (blue line).

²⁵We show this in Appendix B and the proof adapts a simple "contagion argument" to the Bertrand price setting environment.

Consider the boundary case in which $\mu = \bar{\mu}(\gamma, v)$; our discussion in Section 4.3 implies that AIC just holds and, in particular, **AIC-High** holds with equality:

$$\gamma \cdot \frac{\bar{\mu}(\gamma, v)}{2} + (\gamma - v) \cdot \frac{1 - \bar{\mu}(\gamma, v)}{2} = (1 - \gamma) \cdot \frac{\bar{\mu}(\gamma, v)}{2}.$$

Suppose we take a small measure of low-value consumers and make them high-value (i.e., we increase μ above $\bar{\mu}(\gamma, v)$). Condition **AIC-High** now fails and the producer-perfect outcome is not implementable. What then is the best for producers that information design can implement?

In response to this change in distribution we modify the producer-perfect information design such that some high-value consumers are now assigned to the low-value message. Specifically, the new measure $(\mu - \bar{\mu}(\gamma, v))/2$ of high-value consumers for, say, firm 1 are assigned to the message $1 - v$ for firm 1. That is, we implement the producer-perfect information design as if the distribution of high types is $\bar{\mu}(\gamma, v)$ instead of the real μ .

Under this design the incentive of firm 1 to undercut to $1 - \gamma$ when she receives price recommendation 1 or $1 - v$ are exactly as when μ was equal to $\bar{\mu}(\gamma, v)$. This deviation can only steal the high value consumers for firm 2 who are being charged a price of 1 (and hence have zero consumer surplus), and this mass of consumers is exactly the same as when the mass of high value consumers was equal to $\bar{\mu}(\gamma, v)$. Likewise, the incentives to deviate downwards to a price of $1 - \gamma - v$ are also exactly the same as before. As before, this deviation is just sufficient to steal all firm 2's consumers; Firm 2's high value consumers who are being charged the lower price of $1 - v$ are just willing to buy at this price as well as firm 2's low value consumers.

However, whereas checking downward deviations were both necessary and sufficient to implement a perfect-producer outcome when $\mu = \bar{\mu}(\gamma, v)$, this is no longer the case. Under $\mu > \bar{\mu}(\gamma, v)$ and the modified information design, firm 1 knows that by following recommendation $1 - v$ it leaves a rent v to the mass of high types $(\mu - \bar{\mu}(\gamma, v))/2$. By raising the price to 1, firm 1 will lose the low-value consumers in this segment, but will extract the extra amount v from these high-value consumers.²⁶ Thus, the segmentation must ensure that firms do not want to deviate upwards to the price 1.²⁷ This requires:

$$(1 - v) \cdot \left(\frac{1 - \bar{\mu}(\gamma, v)}{2} \right) \geq 1 \cdot \frac{\mu - \bar{\mu}(\gamma, v)}{2}.$$

For sufficiently small $\mu - \bar{\mu}(\gamma, v)$, these upwards deviations are slack but, as the mass of high types increases the deviation becomes more and more attractive and it binds when $\mu = \hat{\mu}(\gamma, v)$.

As we vary μ from $\bar{\mu}(\gamma, v)$ to $\hat{\mu}(\gamma, v)$ the maximum amount of producer surplus that can be obtained is constant and it coincides with the producer-perfect outcome in an

²⁶Such types are charged either price 1 or price $1 - v$ by firm 2, so firm 1 makes the sale.

²⁷Note that if both firms charge price 1, this is not an equilibrium since one firm would have incentive to deviate downwards.

economy with a mass of high type consumers is $\bar{\mu}(\gamma, v)$ (red flat line in panel (a)). Analogously, producers extract all available surplus $TS^*(v, \mu)$ except the information rent enjoyed by the high-type consumers treated as low-types, which is equal to $v(\mu - \bar{\mu}(\gamma, v))$.

Beyond $\mu > \hat{\mu}(\gamma, v)$, to cope with the upward deviations, the designer has to give up some more producer surplus. In particular, it turns out to be optimal for the designer to create a new segment containing the mass of high type consumers in excess of $\hat{\mu}(\gamma, v)$ and firms compete for these consumers under full information, while the information design for the remaining consumers is the same as before. This causes the maximum attainable producer surplus to strictly decrease as μ increases. As all consumers become the high-type, the best the information designer can do is to provide full information about consumer preferences, which in this case is the same as the consumer-optimal information design.

Case B. Panel (b) of Figure 6 depicts $PS^*(\mu, \gamma, v)$ for an example of Case B. In this case we have $\underline{\mu}(\gamma, v) > 0$. When $\mu \geq \underline{\mu}(\gamma, v)$, the discussion is identical to Case A and is omitted. When $\mu < \underline{\mu}(\gamma, v)$ AIC is violated as discussed in Section 4.3; we now illustrate the producer surplus that we can be attained.

We first note that when the market is entirely composed by low-types, and so $\mu = 0$, the highest producer surplus is attained by informing firms about the type of each consumer and having them compete a la Bertrand. Firms then charge a price of γ to the consumers who prefer their product the most and 0 to the other consumers and producer surplus is γ .

Next, suppose that the fraction of low types is $\mu \in (0, \underline{\mu}(\gamma, v))$. The following design is producer-optimal: construct two distributions (μ^1, μ^2) where the the fraction of high types in μ^1 is $\underline{\mu}(\gamma, \mu)$ and the fraction of high types in μ^2 is zero. By construction, the AIC holds for distribution μ^1 , and so, by Theorem 2, the designer can extract all surplus as producer surplus. From the remaining distribution μ^2 , the best the designer can do is to extract γ . Thus, this attains

$$\begin{aligned} PS^*(\mu, \gamma, v) &= \frac{\mu}{\underline{\mu}(\gamma, v)} \cdot TS^*(v, \underline{\mu}(\gamma, \mu)) + \frac{\underline{\mu}(\gamma, v) - \mu}{\underline{\mu}(\gamma, v)} \gamma \\ &= \gamma + \frac{\mu}{\underline{\mu}(\gamma, v)} \left(TS^*(v, \underline{\mu}(\gamma, \mu)) - \gamma \right), \end{aligned}$$

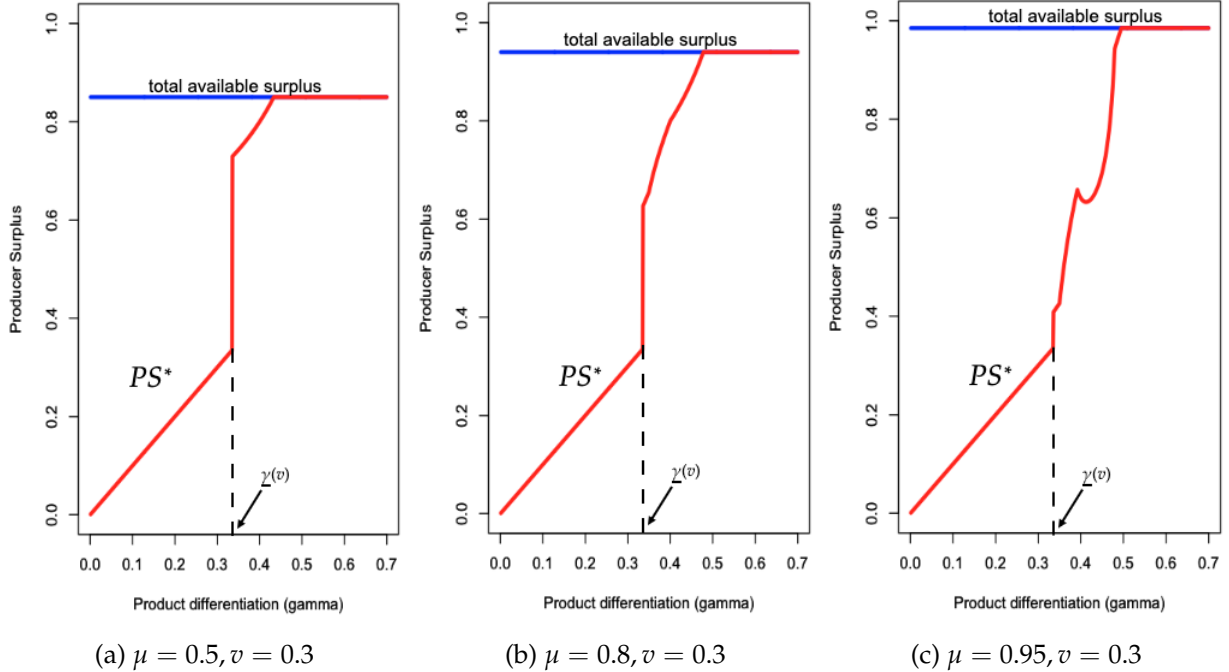
which is increasing in $\mu \leq \underline{\mu}(\gamma, v)$ (panel (b) increasing red line).

Case C. Note that $\underline{\gamma}(v)$ is the minimum value of γ such that $\hat{\mu}(\gamma, v) \leq \bar{\mu}(\gamma, v)$. Hence, whenever $\gamma < \underline{\gamma}(v)$ there never exists a distribution of types under which AIC holds. Thus, we cannot perform the same partition of distribution μ that we did in case B. Indeed, in this case horizontal differentiation is sufficiently low that information design has no bite: the producer-optimal information design is to simply provide full information, which, in this case, coincides with the consumer-optimal information design.

We conclude by elaborating how the producer-optimal outcome varies as we fix the proportion of high-types μ and vary the degree of product differentiation γ ; Figure 4

illustrates this. For γ sufficiently high, AIC is satisfied and all surplus can be extracted. For intermediate values of γ , the optimal information design treats high-value consumers as if they are low-value. And finally, for lower values of γ upward deviation constraints start to bind and the information design must be modified accordingly.

Figure 4: Illustration of optimal producer surplus as differentiation varies



We make two observations. First, the change in producer surplus is discontinuous in γ . For instance, as γ increases past $\underline{\gamma}(v)$, by our discussion above, there now exists *some* proportion of high and low types which can be combined such that AIC holds. For this subset of the market, when γ was just below $\underline{\gamma}(v)$ we extracted γ , whereas when γ was just above $\underline{\gamma}(v)$ we extract all surplus, a discrete jump in PS^* . As γ increases further, we can increase continuously the size of this submarket and, therefore, producer surplus increases continuously until AIC holds. More generally, if we start from an environment (μ, v, γ) in which only downward deviations bind, then increasing γ at the margin slackens the binding constraints and hence PS^* increases.

However, this is not true when the environment is such that upward deviation constraints bind. Our second observation is that PS^* is not necessarily monotone in γ , as illustrated in Panel (c) where we see that PS^* can decrease in γ at around 0.4 before increasing again. The reason is that when there are many high types ($\mu = 0.95$), two upward deviation constraints bind. Upon receipt of $1 - v$ the firm is tempted to deviate up to price 1 to steal high-types which also face price $1 - v$ from the other firm. In addition, upon receipt of $1 - v$, the firm is also tempted to deviate up to price $2\gamma > 1 - v$ to steal high types which face price γ from the other firm.²⁸ As such,

²⁸The optimal information structure assigns some high types the messages $(1 - v, \gamma)$ for firm 1 and 2 respectively. This takes pressure off firm 1's downward deviation constraint from $1 - v$ to the price $1 - v - \gamma$ which, in turn, means that more high-types can be assigned the messages $(1 - v, 1)$ in an

increasing γ can intensify the incentive to deviate upwards to 2γ which explains the non-monotonicity in PS^* .

Our results characterizing producer-optimal information structures when AIC is violated relates to the elegant paper of [Bergemann, Brooks, and Morris \(2021\)](#) who study producer-optimal information in search markets in which there is uncertainty about how many firms the consumer has access to. Translated into our setting, their model (for $n = 2$) corresponds to measure $1 - \mu$ on valuation $(1, 1)$, and measure $\mu/2$ on each of $(1, 0)$ and $(0, 1)$. That is, consumers are either contested in which case they have identical valuations, or captive in which case they have only positive valuation for one firm.²⁹

Our switching cost setting corresponds to an environment in which consumers are aware of the existence and offerings of both firms. Different consumers might have different valuations—as captured by the degree of vertical heterogeneity v —and products might be more or less differentiated—as captured by the degree of horizontal differentiation γ . This yields different and complementary insights relative to the search environment of [Bergemann, Brooks, and Morris \(2021\)](#). In particular, our analysis allow us to understand the limits of information to extract product surplus as we vary both dimensions of heterogeneity smoothly from the region in which AIC holds to the region in which it fails and, finally, to the region where there is no product differentiation and the maximum amount of extractable surplus is zero. Furthermore, our environment sheds light on the possibility that upward deviations pose a constraint on surplus extraction; this is a methodological point which, to our knowledge, has not been made in the literature.

6 MATCHING AND INFORMATION DESIGN

So far the only role of the intermediary is to disclose information to firms about consumers' valuations. In practice, however, the intermediary might also control what price offers each consumer can evaluate by withholding firms' access to certain consumers.³⁰ We now enrich our model with this possibility.

A joint matching and information design (henceforth, just design) is a map from the consumers' types to a joint distribution over the firms that the consumers receive an offer from, and the information each firm receives about the consumers' valuations. Hence, a design is a joint distribution

$$\lambda \in \Delta\left(\Theta \times M \times 2^N\right)$$

incentive-compatible manner. This, in turn, takes pressure off firm 2's downward deviation constraint from 1 to $1 - v$ and allows the price of 1 to be extracted from more high-types.

²⁹See, e.g., [Armstrong and Vickers \(2019\)](#).

³⁰See [Bergemann and Bonatti \(2019\)](#) for a discussion on the distinction between information and access design. [Bergemann, Brooks, and Morris \(2021\)](#) (Section V) analyze a model in which consumers' valuations are homogeneous across products, but consideration sets are endogeneously generated via sequential search. Here we let the consideration sets be chosen by a platform.

where recall $\mathcal{M} := \prod_{i \in \mathcal{N}} M_i := [0, 1]^n$ is a message space, and $2^{\mathcal{N}}$ is the power set of the set of firms \mathcal{N} . Let Λ be the space of all designs which fulfil the analogous consistency requirement that $\text{marg}_{\Theta} \lambda = \mu$.

For a given type $\theta \in [0, 1]^n$, the design λ induces a measure $\lambda(\cdot | \theta) \in \Delta([0, 1]^n \times 2^{\mathcal{N}})$ over both vector-valued messages $\mathbf{m} \in [0, 1]^n$, and sets $S \in 2^{\mathcal{N}}$, where set S contains the firms that have been matched to (and hence have access to) the consumer. Sometimes we refer to S as the consideration set of consumer type θ . Consumers observe offers only from firms in their consideration set.

The design λ induces a matching scheme

$$\text{marg}_{\Theta \times 2^{\mathcal{N}}} \lambda := \phi_{\lambda} \in \Delta(\Theta \times 2^{\mathcal{N}}),$$

which is a joint distribution over consumer types and consideration sets: for each type θ and each set of firms $S \subseteq \mathcal{N}$, this gives the probability that firms S are the ones with access to a consumer of the type θ . It will sometimes be helpful to start with a fixed matching scheme ϕ , in which case we are restricting the space of designs to

$$\Lambda_{\phi} := \left\{ \lambda \in \Lambda : \text{marg}_{\Theta \times 2^{\mathcal{N}}} \lambda = \phi \right\},$$

i.e., the designs which induce the same joint distribution over types and consideration sets as the matching scheme ϕ .

As before, the design induces a simultaneous price setting game among firms and we are interested in Bayes-Correlated Equilibria (Bergemann and Morris, 2016), henceforth equilibria. We wish to characterise the feasible surplus set defined as follows:

Definition 2. The feasible surplus set $SUR \subset \mathbb{R}_{\geq 0}^2$ comprises the pairs of producer surplus (PS) and consumer surplus (CS) that can be implemented as an equilibrium outcome of some design $\lambda \in \Lambda$. The lower envelope of SUR , denoted by LE , is the set of all pairs $(CS, PS) \in SUR$ with the property that if $CS' = CS$ and $PS' < PS$ then $(CS', PS') \notin SUR$.

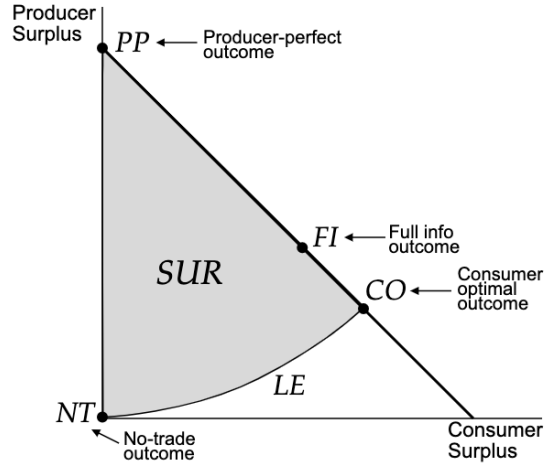
Definition 3. The consumer optimal point (CO) is a point in set SUR with the highest consumer surplus among all points in SUR . If there are multiple such points, we choose the one with highest producer surplus. The ϕ -consumer optimal design is a design which implements the highest consumer surplus among all designs in Λ_{ϕ} .

Theorem 3. The following characterizes SUR :

- (i) The producer-perfect point (PP) is obtained by matching each consumer in E_i only to firm i and by fully revealing the consumer type to firm i .
- (ii) Each point in the lower envelope of SUR (LE) can be implemented through a ϕ -consumer-optimal design for some ϕ .
- (iii) The consumer-optimal point (CO) is obtained through the unrestricted matching denoted $\phi^*(\cdot | \theta) = \delta_{\mathcal{N}}$ for all $\theta \in \Theta$, paired with the ϕ^* -consumer optimal design.

- (iv) The feasible surplus set is the convex hull generated by the producer-optimal point PP and the lower envelope: $SUR = \text{conv}(PP \cup LE)$.

Figure 5: Illustration of SUR



The proof of Theorem 3 is developed in Appendix C. Notice that the producer-perfect point PP can be trivially achieved by simply making the consumer's favourite firm a monopolist (by now allowing the consumer access any other firm) and providing perfect information about that consumer's valuation. The point NT can be achieved via the empty matching such that there is no trade. The key step is to show that any point in the lower envelope of SUR is implemented with designs which are consumer-optimal for *some* matching design.

Theorem 3 shows that augmenting the platform's ability to design both access and information cannot deliver more consumer surplus than simply choosing the consumer-optimal information design under complete matching. Intuitively, competition is intensified when all firms are included in each consideration set of all consumers. Following the same logic, every point in SUR can be seen as a consumer-optimal outcome constrained to a specific matching.

We have granted the intermediary full flexibility to design matchings. However, as we explain next, Theorem 3 is robust to situations in which intermediaries face some constraints on the matching they can design. First, we observe that every point in SUR can be achieved via matching consumers to at most two firms. This follows from the nature of price competition, in which the consumer's second favourite firm among her consideration set poses a necessary and sufficient constraint on her favourite firm's pricing strategy.

Second, the PP outcome can be approximately implemented in large markets (with many firms) even when the platform is constrained to show each consumer the offerings of at least $K > 1$ firms. Suppose, for instance, that valuations for each firm is iid and drawn from some atomless distribution. Then a point close to PP can be implemented through the following design: for each consumer with valuation θ

- (i) Matching: match the consumer to her favourite firm, as well as her $K - 1$ least favorite firms; and

- (ii) Information: publicly announce the consumer’s highest valuation $\max_i \theta_i$ and nothing else.

As the number of firms grows large, under this design almost all surplus is extracted as producer surplus (it approximately implements the point PP). Both matching and information play essential roles. Suppose that all firms learn the consumer’s highest valuation $\max_i \theta_i$. If all firms set this price, one of them will be lucky and sell to the consumer at a price equal to her valuation. To minimize the incentives of the firm to deviate and set a lower price, the design matches the consumer to her $K - 1$ least favourite products which in effect polarizes her preferences. Thus, since $\max_i \theta_i$ is of order 1 and each firm has $1/n$ chance of selling to the consumer at her highest valuation, this delivers expected profits of order $1/n$. Conversely, the consumer’s $K - 1$ th least preferred product concentrates exponentially around 0 so the requisite downward deviation to make the sale with high probability decays exponentially. Thus, for large markets, such a deviation is not profitable.³¹

7 CONCLUDING REMARKS

7.1 Takeaways for regulators. We have explored how platforms can use information on market participants to shape price competition. We showed how, by packaging information about consumers preferences in different ways, the platform can relax or intensify market competition to obtain different ratio of consumer to producer surplus on the efficient frontier.

From the perspective of an antitrust authority mandated with protecting consumer surplus, this raises a delicate problem. Outrightly preventing the use of information will typically sacrifice efficiency while, without regulation on information disclosure, a platform that wishes to increase consumer surplus can intensify price competition well beyond the complete information case (hence achieving greater consumer surplus). At the same time, an intermediary with a revenue model based on monetizing consumer information via charges levied on firms may design information structures where consumers gets much less than under complete information; if there is enough polarization in consumers’ preferences the intermediary can implement the same outcome perfect collusion would yield in an otherwise competitive downstream market.

Our analysis shows that there are distinct principles on how information is disclosed which matter for attendant market outcomes. We therefore suggest that regulators might formulate guidelines or rules of conduct that ensure that such groups of consumers are formed in line with the principles characterizing the consumer-optimal information structure—i.e., only consumers with similar preferences (and hence the same most preferred product) should be grouped together and this information should be disclosed publicly to firms.

7.2 Price discrimination in practice. Our analysis is based on the assumption that firms will price discriminate if they can. If firms expect consumers to become aware of

³¹Online Appendix I formalizes these arguments.

differential pricing based on consumers' willingness to pay, the ensuing reputational damage may deter the implementation of these practices. In this case, information design is irrelevant since firms must charge uniform prices.

There are, however, ways in which price discrimination can be concealed. First, a 2019 report by the UK's Digital Competition Experts Panel writes that if firms can "send secret deals to consumers, for example by directly offering discounts via email, the price discrimination becomes entirely opaque." The use of discount codes is widespread and encouraged by internet intermediaries.³² In fact, when firms attempt to conceal price discrimination from consumers in this way it will be relatively challenging to detect it empirically. A web-scraping 'robot,' used in experiments like that run by [Cavallo and Rigobon \(2016\)](#) to compare online and offline prices, does not have the same web-surfing or purchase history as real profiles. As such, firms do not have the opportunity to target them with discount codes (for instance, through social media feeds). Second, in industries where the cost of providing the service being sold depends on the characteristics of the individual (e.g., insurance and credit markets), and in industries that use dynamic demand-based pricing (e.g., flights and ride-hailing), it is hard for consumers to understand what underlies price differences.³³ Again, in such cases, it is challenging for empirical work using publicly available data to identify price discrimination.

All this points to a lack of strong evidence for widespread price discrimination not necessarily implying that such practices are not taking place, albeit in more subtle ways. And this is why the possibility that consumer data are used to facilitate discriminatory pricing has drawn regulatory interest. China's new anti-monopoly guidelines—tailored exclusively to reigning in tech firms—explicitly outline the phenomena for data being used to "achieve coordinated behaviour" ([State Administration for Market Regulation, 2021](#)).³⁴ In a similar vein, a recent report by the Competition and Markets Authority in the UK reported that "even if there is limited evidence for personalized pricing, this could change quickly" ([Competition & Markets Authority, 2021](#)). Similar issues are highlighted in regulatory documents from the EU, US and Canada.³⁵

REFERENCES

ACEMOGLU, D., A. MAKHDOUMI, A. MALEKIAN, AND A. OZDAGLAR (2019): "Too much data: Prices and inefficiencies in data markets," Tech. rep., National Bureau of

³²See Google's marketer playbook and Facebook's webpage for small businesses. Targeted discounts are also ubiquitous in the grocery market; supermarkets collect detailed data on consumers and price discriminates using coupons. [Hannak et al. \(2014\)](#) compare the prices charged to real consumer profiles obtained via Amazon Mechanical Turk. They find evidence that Home Depot, Sears, many travel sites (e.g. Cheaptickets, Orbitz, Priceline etc.) price discriminate.

³³A 2018 report by the Competition and Markets Authority, the UK's competition regulator found that some home and motor insurance firms use complex and opaque pricing techniques to charge consumers with a higher willingness to pay markedly higher prices ([Competition & Markets Authority, 2018](#)).

³⁴There is considerable anecdotal evidence for widespread price discrimination occurring in China. A survey conducted in 2019 by the Beijing Consumer Association finds that 88% of consumers believe that the practice of big data-enabled price discrimination is significant, and 57% have personally experienced this.

³⁵See [European Commission \(2019\)](#), [Council of Economic Advisors \(2015\)](#), [Competition Bureau Canada \(2018\)](#).

Economic Research.

- ADMATI, A. R. AND P. PFLEIDERER (1986): "A monopolistic market for information," *Journal of Economic Theory*, 39, 400–438.
- ALI, S. N., G. LEWIS, AND S. VASSERMAN (2020): "Voluntary disclosure and personalized pricing," in *Proceedings of the 21st ACM Conference on Economics and Computation*, 537–538.
- ARMSTRONG, M. AND J. VICKERS (2019): "Discriminating against captive customers," *American Economic Review: Insights*, 1, 257–72.
- ARMSTRONG, M. AND J. ZHOU (2019): "Consumer information and the limits to competition," .
- BERGEMANN, D. AND A. BONATTI (2015): "Selling cookies," *American Economic Journal: Microeconomics*, 7, 259–94.
- (2019): "Markets for information: An introduction," *Annual Review of Economics*, 11, 85–107.
- BERGEMANN, D., A. BONATTI, AND T. GAN (2019): "The economics of social data," .
- BERGEMANN, D., A. BONATTI, AND A. SMOLIN (2018): "The design and price of information," *American economic review*, 108, 1–48.
- BERGEMANN, D., B. BROOKS, AND S. MORRIS (2015): "The limits of price discrimination," *American Economic Review*, 105, 921–57.
- (2017): "First-price auctions with general information structures: Implications for bidding and revenue," *Econometrica*, 85, 107–143.
- (2021): "Search, information and prices," *Journal of Political Economy*, 129.
- (2023): "On the Alignment of Consumer Surplus and Total Surplus Under Competitive Price Discrimination," .
- BERGEMANN, D. AND S. MORRIS (2013): "Robust predictions in games with incomplete information," *Econometrica*, 81, 1251–1308.
- (2016): "Bayes correlated equilibrium and the comparison of information structures in games," *Theoretical Economics*, 11, 487–522.
- BERGEMANN, D. AND J. VÄLIMÄKI (1996): "Learning and strategic pricing," *Econometrica: Journal of the Econometric Society*, 1125–1149.
- BIMPIKIS, K., D. CRAPIS, AND A. TAHBAZ-SALEHI (2019): "Information sale and competition," *Management Science*, 65, 2646–2664.
- BOUNIE, D., A. DUBUS, AND P. WAELBROECK (2020): "Market for Information and Selling Mechanisms," Available at SSRN 3454193.
- (2021): "Selling strategic information in digital competitive markets," *The RAND Journal of Economics*, 52, 283–313.
- CALZOLARI, G. AND A. PAVAN (2006): "On the optimality of privacy in sequential contracting," *Journal of Economic Theory*, 130, 168–204.
- CAVALLO, A. AND R. RIGOBON (2016): "The billion prices project: Using online prices for measurement and research," *Journal of Economic Perspectives*, 30, 151–78.
- COMPETITION & MARKETS AUTHORITY (2018): "Pricing algorithms: Economic working paper on the use of algorithms to facilitate collusion and personalised pricing," https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/746353/Algorithms_econ_report.pdf.
- (2021): "Algorithms: How they can reduce competition and harm consumers," <https://www.gov.uk/government/publications/algorithms-how-they-can-reduce-competition-and-harm-consumers/algorithms-how-they-can-reduce-competition-and-harm-consumers>.
- COMPETITION BUREAU CANADA (2018): "Big data and innovation: key themes

- for competition policy in Canada," <https://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/04342.html>.
- COUNCIL OF ECONOMIC ADVISORS (2015): "Big Data and Differential Pricing," https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/docs/Big_Data_Report_Nonembargo_v2.pdf.
- DWORCZAK, P. AND A. KOLOTILIN (2022): "The persuasion duality," *arXiv preprint arXiv:1910.11392*.
- EUROPEAN COMMISSION (2019): "Competition policy for the digital era," <https://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf>.
- FAINMESSER, I. P. AND A. GALEOTTI (2019): "Pricing network effects: Competition," *Johns Hopkins Carey Business School Research Paper*.
- FELDMAN, M., B. LUCIER, AND N. NISAN (2016): "Correlated and coarse equilibria of single-item auctions," in *International Conference on Web and Internet Economics*, Springer, 131–144.
- HANNAK, A., G. SOELLER, D. LAZER, A. MISLOVE, AND C. WILSON (2014): "Measuring price discrimination and steering on e-commerce web sites," in *Proceedings of the 2014 conference on internet measurement conference*, 305–318.
- JOHNSON, J. P. AND D. P. MYATT (2006): "On the simple economics of advertising, marketing, and product design," *American Economic Review*, 96, 756–784.
- JONES, C. I. AND C. TONETTI (2020): "Nonrivalry and the Economics of Data," *American Economic Review*, 110, 2819–58.
- KAMENICA, E. (2019): "Bayesian persuasion and information design," *Annual Review of Economics*, 11, 249–272.
- KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.
- KEHOE, P. J., B. J. LARSEN, AND E. PASTORINO (2018): "Dynamic competition in the era of big data," Tech. rep., Working paper, Stanford University and Federal Reserve Bank of Minneapolis.
- KLEMPERER, P. (1995): "Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade," *The Review of Economic Studies*, 62, 515–539.
- KOLOTILIN, A., R. CORRAO, AND A. WOLITZKY (2022): "Persuasion with Non-Linear Preferences," .
- LIZZERI, A. (1999): "Information revelation and certification intermediaries," *The RAND Journal of Economics*, 214–231.
- MONTES, R., W. SAND-ZANTMAN, AND T. VALLETTI (2019): "The value of personal information in online markets with endogenous privacy," *Management Science*, 65, 1342–1362.
- NOVSHEK, W. AND H. SONNENSCHNEIN (1982): "Fulfilled expectations Cournot duopoly with information acquisition and release," *The Bell Journal of Economics*, 214–218.
- PIGOU, A. (1920): "The Economics of Welfare (London, 1920)," 12.
- RAITH, M. (1996): "A general model of information sharing in oligopoly," *Journal of economic theory*, 71, 260–288.
- RHODES, A. AND J. ZHOU (2022): "Personalized pricing and competition," Available at SSRN 4103763.
- ROBINSON, J. (1933): *The economics of imperfect competition*, London: Macmillan.
- ROESLER, A.-K. AND B. SZENTES (2017): "Buyer-optimal learning and monopoly pricing," *American Economic Review*, 107, 2072–80.

- SMOLIN, A. AND T. YAMASHITA (2023): "Information Design in Games: Certification Approach," Available at arXiv: <https://arxiv.org/abs/2202.10883>.
- STATE ADMINISTRATION FOR MARKET REGULATION (2021): "Anti-Monopoly Guidelines for the Platform Economy Sector," http://gkml.samr.gov.cn/nsjg/fldj/202102/t20210207_325967.html.
- TAYLOR, C. R. (2004): "Consumer privacy and the market for customer information," *RAND Journal of Economics*, 631–650.
- VIVES, X. (1988): "Aggregation of information in large Cournot markets," *Econometrica: Journal of the Econometric Society*, 851–876.

APPENDIX TO MARKET SEGMENTATION THROUGH INFORMATION

Appendix **A** collects the proofs of Theorem 1 (consumer-optimal) and Theorem 2 (producer-perfect) designs, as well as comparative statics on AIC (Propositions 1 and 2). Appendix **B** collects proofs of results on producer-optimal designs in the switching cost environment, as well as characterizes the producer-optimal design when $\gamma < v$. Appendix **C** collects the proofs of the characterization of welfare outcomes when the designer controls both information and matching (Theorem 3).

A: PROOFS FOR SECTIONS 3 (CONSUMER OPTIMAL) AND 4 (PRODUCER PERFECT)

A.1 Proof of Theorem 1. We first establish that $\underline{\Pi}_i^*$ is a lower bound on firm i 's profit. Given information structure ψ and opponent firms follow strategy profile σ , firm i can always ignore any information from information structure ψ . In particular, fix any joint distribution over types and other firms' prices $\psi_{-i} \in \Delta(\Theta \times M_{-i})$. The firm can always obtain a profit

$$\Pi_i^*(\psi_{-i}) := \max_{p_i \geq 0} p_i \cdot \mathbb{P}_{(\theta, p_{-i}) \sim \psi_{-i}} \left(\theta_i \geq p_i, \theta_i - p_i \geq \max_{j \neq i} (\theta_j - p_j) \right)$$

and the probability of making a sale when firm i charges p_i is minimized when all other firms charge all consumers a price of zero, so we have the lower bound

$$\Pi_i^*(\psi_{-i}) \geq \max_{p_i \geq 0} p_i \cdot \mathbb{P}_{\theta \sim \mu} \left(\theta_i - \max_{j \neq i} \theta_j \geq p_i \right) =: \underline{\Pi}_i^*$$

We are ready to prove Theorem 1. When the consumer is in set E_i , ψ^* will recommend firms other than i to price at zero. Given that opponent firms follow the price recommendation, type $\theta \in E_i$ will have a residual value of $\Lambda_i(\theta)$ for product i . In words, type θ purchase product i at price p_i if $\Lambda_i(\theta) \geq p_i$. If firm i only knows the consumer is in set E_i , firm i will believe that the consumer's residual value is distributed as μ_i^* . Firm i will exactly obtain its lower bound profit $\underline{\Pi}_i^*$:

$$\max_{p_i \in [0,1]} p_i \mu(E_i) \mu_i^*([p_i, 1]) = \max_{p_i \in [0,1]} p_i \mathbb{P}(\Lambda_i(\theta) \geq p_i) = \max_{p_i \in [0,1]} p_i \mathbb{P}(\theta_i - \max_{j \neq i} \theta_j \geq p_i) = \underline{\Pi}_i^*.$$

Note that we don't need to put in the condition $\theta \in E_i$ because it will be implied by $\Lambda_i(\theta) \geq p_i$.

ψ^* then applies the uniform profit preserving extremal information structure (Bergemann, Brooks, and Morris, 2015) to μ_i^* . Upon receipt of each message m_i , by the property of uniform profit preserving extremal information structures, each residual value for product i in support of the posterior is an optimal price for firm i . Suppose that firm i price at the lowest residual value for product i in support of each posterior.³⁶ Firm i 's overall profit remains $\underline{\Pi}_i^*$ and all consumers in E_i will purchase product i . Thus, consumer surplus is exactly at the upper bound CS^* . It remains to show that each firm $j \neq i$ will follow price recommendation to price at zero. Since firm i prices at the lowest

³⁶Since we focus only on partial implementation.

residual value for product i in support of the posterior induced by each message m_i , firm j will make zero sales at price 0 so finds it (weakly) unprofitable to deviate to any positive price.

A.2 Proof of Theorem 2. We start by completing the proof of Lemma 1.

Proof of Lemma 1. From the discussion in the main text, it follows that **Consistency**, **Separation** and **Consumer-IC** are necessary if ψ is an producer-perfect information structure. To see that **Firm IC** is necessary, recall that **Firm IC** is:

$$\underbrace{(m_i - \hat{p}_i)\psi(\{\theta \in E_i : \theta_i = m_i\} | m_i)}_{\text{Inframarginal losses}} \geq \underbrace{\hat{p}_i\psi(\{\theta \notin E_i : \theta_i \geq \hat{p}_i\} | m_i)}_{\text{Business-stealing gains}} \quad (\text{Firm IC})$$

For type $\theta \notin E_i$, firms other than i will charge type θ at least her value for their product (by **Separation** and **Consumer-IC**). Therefore, each type $\theta \notin E_i$ with $\theta_i > \hat{p}_i$ will strictly prefer to purchase product i at price \hat{p}_i . Note we include types with $\theta_i = \hat{p}_i$ when calculate business-stealing gains. It does not matter if there is no atom at $\theta_i = \hat{p}_i$. When there is atom at $\theta_i = \hat{p}_i$, we will arrive at the same formula by considering deviating to a price $p < \hat{p}_i$ and taking p to \hat{p}_i .

To see these four properties are also sufficient, note that **Consistency** ensure that ψ is a valid information structure. **Separation**, **Consistency** ensure that each type $\theta \in E_i$ is willing to purchase product i at her value. Thus property P follows. Finally **Firm IC** ensures that firm i will follow the price recommendations. \square

Proof of Theorem 2. We first show that AIC is necessary.

Step 1: Necessity. By Lemma 1, if ψ is a producer-perfect information structure, then **Firm IC** must hold. We get AIC by aggregating **Firm IC** over messages $m_i \in [\hat{p}_i, 1]$:

$$\begin{aligned} \int_{m_i \geq \hat{p}_i} (m_i - \hat{p}_i)\psi(\{\theta \in E_i : \theta_i = m_i\} | m_i)\psi_i(dm_i) \\ \geq \int_{m_i \geq \hat{p}_i} \hat{p}_i\psi(\{\theta \notin E_i : \theta_i \geq \hat{p}_i\} | m_i)\psi_i(dm_i). \end{aligned}$$

where we let $\psi_i := \text{marg}_{M_i}\psi$.

$$\begin{aligned} LHS &= \mathbb{P}(m_i \geq \hat{p}_i, \theta_i \in E_i, \theta_i = m_i)\mathbb{E}(m_i - \hat{p}_i | m_i \geq \hat{p}_i, \theta_i \in E_i, \theta_i = m_i) \\ &= \mathbb{P}(\theta_i \geq \hat{p}_i, \theta_i \in E_i, \theta_i = m_i)\mathbb{E}(\theta_i - \hat{p}_i | \theta_i \geq \hat{p}_i, \theta_i \in E_i, \theta_i = m_i) \\ &= \mathbb{P}(\theta_i \geq \hat{p}_i, \theta_i \in E_i)\mathbb{E}(\theta_i - \hat{p}_i | \theta_i \geq \hat{p}_i, \theta_i \in E_i) \quad (\text{by separation}) \\ &= \int_{\theta \in E_i: \theta_i \geq \hat{p}_i} (\theta_i - \hat{p}_i)d\mu. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} RHS &= \hat{p}_i\mathbb{P}(m_i \geq \hat{p}_i, \theta \notin E_i : \theta_i \geq \hat{p}_i) \\ &= \hat{p}_i\mu(\{\theta \notin E_i | \theta_i \geq \hat{p}_i\}) \quad (\text{by Consumer-IC}) \end{aligned}$$

So whenever AIC fails, **Firm IC** also fails.

Step 2: Sufficiency. To show AIC is also sufficient. We construct an information structure ψ and satisfies **Consistency**, **Separation**, **Consumer IC** and **Firm IC**. The proof then follows from Lemma 1.

Recall that we assumed that μ can be decomposed into a discrete and absolutely continuous measure. Let f denote the density of μ . and V denote the the set of atoms. For each Borel set $B \in \mathcal{B}(\Theta)$, we have the following decomposition:

$$\mu(B) = \sum_{\theta \in B \cap V} \mu(\theta) + \int_{\theta \in B} f(\theta) d\theta.$$

We construct ψ in such a way that, conditional on each θ , different firms' messages are independently distributed. For each $i \in \mathcal{N}$ and $\theta \in \Theta$, let $\psi_i(m_i|\theta)$ denote the conditional probability density of firm i receiving message m_i conditional on the realized type is θ . We use $\psi_i(m_i|\theta)$ to denote the probability mass function when the conditional distribution has atom at m_i . It suffices to construct $\psi_i(\cdot|\theta)$ for each firm i and each $\theta \in \Theta$. Note that this construction automatically fulfills **Consistency** as long as $\psi_i(\cdot|\theta)$ is valid density or probability mass function.

For each firm i and each $m_i \in [0, 1]$, define $G_i(\cdot|m_i)$ as follows:

$$G_i(\hat{p}_i|m_i) := \begin{cases} \frac{m_i - \hat{p}_i}{\hat{p}_i} \mu(\{\theta \in E_i : \theta_i = m_i\}) & \text{marg}_{\Theta_i} \mu \text{ has an atom at } m_i \\ \frac{m_i - \hat{p}_i}{\hat{p}_i} f_i(m_i) & \text{otherwise.} \end{cases}$$

where

$$f_i(m_i) := \int_{\theta_{-i}: (m_i, \theta_{-i}) \in E_i} f(m_i, \theta_{-i}) d\theta_{-i}.$$

is the marginal density at m_i . For $\hat{p}_i \geq m_i$, $G_i(\hat{p}_i|m_i)$ takes value zero.

Let $\psi_i(m_i)$ denote the probability density (mass) of firm i receiving message m_i whenever the distribution is atomless (has an atom) at m_i . By multiplying $\psi_i(m_i)$ and divide \hat{p}_i on both side of **Firm IC**, we will have the following version of **Firm IC** which is easier to verify:

$$G(\hat{p}_i|m_i) \geq \int_{\theta \notin E_i: \theta_i \geq \hat{p}_i} \psi_i(m_i|\theta) d\mu.$$

Let $V_i := \{\theta_i | \theta \in V \cap E_i\}$ be the set of atoms of the distribution of θ_i for $\theta \in E_i$.

We then have:

$$H_i(\hat{p}_i) = \int_{E_i: \theta_i \geq \hat{p}_i} \frac{(\theta_i - \hat{p}_i)}{\hat{p}_i} d\mu = \int_{m_i \notin V_i: m_i \geq \hat{p}_i} G(\hat{p}_i|m_i) dm_i + \sum_{m_i \in V_i: m_i \geq \hat{p}_i} G_i(\hat{p}_i|m_i).$$

Hence AIC implies

$$H_i(\hat{p}_i) \geq Q_i(\hat{p}_i) := \mu(\{\theta \notin E_i | \theta_i \geq \hat{p}_i\}) \text{ for all } i \text{ and all } \hat{p}_i$$

For each $c \in [0, 1]$, let $\gamma(c)$ denote the cutoff that satisfies

$$H_i(\gamma(c)) = Q_i(c).$$

Intuitively, $\gamma(c)$ is the cutoff above which the capacity is just enough to accommodate the total mass of types not in E_i with value for product i above c . γ is well defined since $H_i(c) \geq Q_i(c)$, $H_i(c)$ is continuous and strictly decreasing in c and $H_i(\bar{c}) = 0$, where \bar{c} is the largest possible realization of θ_i in E_i . Further note that γ is weakly increasing and $\gamma(c) \geq c$.

For all $\theta \in E_i$, probability mass $\psi_i(m_i = \theta_i | \theta) = 1$ fulfilling **separation**. Recall that we assumed μ can be decomposed into a discrete measure and a measure which is absolutely continuous hence admits a density. Order the elements of the following set in an increasing order:

$$V_i^- := \{\theta_i | \theta \notin E_i, \theta \in V\} := \{0 \leq v_1 < \dots < v_k < \dots < v_K < 1\}.$$

For each $\theta \notin E_i$ and $m_i \in M_i$, we construct $\psi_i(m_i | \theta)$ as follows:

If $\theta_i = v_k$ for some k , then

$$\psi_i(m_i | \theta) := \begin{cases} \frac{G_i(\gamma(v_k) | m_i) - \lim_{v \downarrow v_k} G_i(\gamma(v) | m_i)}{H_i(\gamma(v_k)) - \lim_{v \downarrow v_k} H_i(\gamma(v))} & \text{if } m_i > \gamma(v_k); \\ 0 & \text{otherwise.} \end{cases}$$

Note the denominator is positive since Q_i has atom at v_k . In addition, $\lim_{v \downarrow v_k} \gamma(v) > \gamma(v_k)$.

If $\theta_i \notin V_i^-$, then

$$\psi_i(m_i | \theta) := \begin{cases} \left(\frac{\partial G_i(\hat{p}_i | m_i)}{\partial \hat{p}_i} / \frac{dH_i(\hat{p}_i)}{d\hat{p}_i} \right) \Big|_{\hat{p}_i = \gamma(\theta_i)} & \text{if } m_i > \gamma(\theta_i); \\ 0 & \text{otherwise.} \end{cases}$$

Note H_i is always differentiable even if μ has atoms.

When m_i is an atom ($m_i \in V_i$), $\psi_i(m_i | \theta)$ is interpreted as conditional probability mass function; otherwise it is interpreted as conditional probability density function. It is valid probability distribution since it integrates over m_i to 1: if θ is such that $\theta_i \notin V_i^-$ then

$$\begin{aligned} \int_{m_i} \psi_i(m_i | \theta) dm_i &= \int_{m_i > \gamma(\theta_i)} \left(\frac{\partial G_i(\hat{p}_i | m_i)}{\partial \hat{p}_i} / \frac{dH_i(\hat{p}_i)}{d\hat{p}_i} \right) \Big|_{\hat{p}_i = \gamma(\theta_i)} dm_i \\ &= 1. \end{aligned}$$

Similarly, if $\theta_i = v_k$ for some k then

$$\begin{aligned} \int_{m_i} \psi_i(m_i | \theta) dm_i &= \int_{m_i > \gamma(v_k)} \frac{G_i(\gamma(v_k) | m_i) - \lim_{v \downarrow v_k} G_i(\gamma(v) | m_i)}{H_i(\gamma(v_k)) - \lim_{v \downarrow v_k} H_i(\gamma(v))} dm_i \\ &= 1 \end{aligned}$$

where both equalities follow from the observation

$$\frac{dH_i(\hat{p}_i)}{d\hat{p}_i} = \int_{m_i \geq \hat{p}_i} \frac{\partial G_i(\hat{p}_i, m_i)}{\partial \hat{p}_i} dm_i + \sum_{m_i \in V_i: m_i > \hat{p}_i} \frac{\partial G_i(\hat{p}_i, m_i)}{\partial \hat{p}_i}.$$

Therefore, the construction fulfills **Consistency**.

Since $\gamma(\theta_i) \geq \theta_i$, only consumers not in E_i with valuations less than θ_i are matched to the message $m_i = \theta_i$. Hence, the construction fulfills **Consumer IC**.

We are left to verify **Firm IC**: For $m_i \in [0, 1]$ and $\hat{p}_i < m_i$,

$$\begin{aligned} \int_{\theta \notin E_i, \theta_i \geq \hat{p}_i} \psi_i(m_i | \theta) d\mu &\leq \int_{\theta \notin E_i, \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i | \theta) d\mu && \text{(by } \gamma(\theta_i) \geq \theta_i) \\ &= \int_{\theta \notin E_i, m_i > \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i | \theta) d\mu && \text{(by } \psi_i(m_i | \theta) = 0 \text{ if } m_i \leq \gamma(\theta_i)) \\ &= \sum_{\theta \notin E_i, \theta \in V, m_i > \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i | \theta) \mu(\theta) \\ &\quad + \int_{\theta \notin E_i, m_i > \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i | \theta) f(\theta) d\theta \\ &= \sum_{k: m_i > \gamma(v_k) \geq \hat{p}_i} \psi_i(m_i | v_k) \mu(\theta \notin E_i, \theta \in V, \theta_i = v_k) + \\ &\quad \int_{\theta_i \notin V_i^-: m_i > \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i | \theta_i) \left(\int_{\theta_{-i}: (\theta_i, \theta_{-i}) \notin E_i} f(\theta_i, \theta_{-i}) d\theta_{-i} \right) d\theta_i \\ &\quad \text{(by } \psi_i(m_i | \theta) \text{ only depending on } \theta_i) \\ &= \sum_{k: m_i > \gamma(v_k) \geq \hat{p}_i} \psi_i(m_i | v_k) (Q_i(v_k) - \lim_{v \downarrow v_k} Q_i(v)) \\ &\quad + \int_{\theta_i \notin V_i^-: m_i > \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i | \theta_i) (-1) Q_i'(\theta_i) d\theta_i \\ &\quad \text{(by definition of } Q_i \text{ and since } Q_i \text{ is differentiable at } c \notin V_i^-) \\ &= \sum_{k: m_i > \gamma(v_k) \geq \hat{p}_i} \psi_i(m_i | v_k) (H_i(\gamma(v_k)) - \lim_{v \downarrow v_k} H_i(\gamma(v))) + \\ &\quad + \int_{\theta_i \notin V_i^-: m_i > \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i | \theta_i) (-1) H_i'(\gamma(\theta_i)) d\gamma(\theta_i) \\ &\quad \text{(by } H_i(\gamma(c)) = Q_i(c)) \\ &= \sum_{k: m_i > \gamma(v_k) \geq \hat{p}_i} (G_i(\gamma(v_k), m_i) - \lim_{v \downarrow v_k} G_i(\gamma(v), m_i)) + \\ &\quad + \int_{\theta_i \notin V_i^-: m_i > \gamma(\theta_i) \geq \hat{p}_i} \frac{\partial G_i(c, m_i)}{\partial c} \Big|_{c=\gamma(\theta_i)} (-1) d\gamma(\theta_i) \\ &\quad \text{(by the construction of } \psi_i(m_i | \theta)) \\ &= \sum_{k: m_i > \gamma(v_k) \geq \hat{p}_i} (G_i(\gamma(v_k), m_i) - \lim_{v \downarrow v_k} G_i(\gamma(v), m_i)) + \\ &\quad + \int_{\theta_i \notin V_i^-: m_i > \gamma(\theta_i) \geq \hat{p}_i} \frac{\partial G_i(\gamma(\theta_i), m_i)}{\partial \theta_i} (-1) d\theta_i \\ &\leq G_i(\hat{p}_i, m_i) \end{aligned}$$

The last step is by noticing that $\gamma(v)$ is an increasing function of v with an upward jump at each $v \in V_i^-$ and is differentiable at each $v \notin V_i^-$. Therefore, $G_i(\gamma(v), m_i)$ is a decreasing function in v and has a downward jump at each $v \in V_i^-$ and is differentiable at each $v \notin V_i^-$. Hence **Firm IC** is fulfilled. \square

A.3 Proof of Propositions 1 and 2.

Proof of Proposition 1. Let $\bar{\theta}$ be the highest possible value for any product, i.e., $\bar{\theta} = \inf\{\theta \in [0, 1] \mid \mu([0, \theta]^n) = 1\}$. If $\bar{\theta} = 0$, Proposition 1 directly follows. For the rest of the proof we suppose $\bar{\theta} > 0$. Fix an arbitrary information structure ψ and first observe that in any equilibrium induced by ψ , the consumer must buy from some firm with strictly positive probability otherwise all firms make zero profits and it is strictly profitable for any firm to deviate uniformly to the price $\bar{\theta} - \epsilon$ for some $\epsilon > 0$.

Next define the following equilibrium object

$$F^\psi(x) := \mathbb{P}(\text{The consumer pays price } p \leq x \mid \text{the consumer buys one of the products}).$$

We will now show that the highest price in the support of F^ψ , $\bar{p} := \inf\{p \in [0, 1] : F^\psi(p) = 1\}$ must be zero. To this end, suppose, towards a contradiction, that $\bar{p} > 0$. Now define

$$F_i^\psi := \mathbb{P}\left(\begin{array}{l} \text{The consumer pays price } p \leq x \\ \text{and buys from } i \end{array} \mid \text{the consumer buys one of the products}\right),$$

noting that for any $p \in [0, 1]$, $F^\psi(p) = \sum_{i=1}^n F_i^\psi(p)$. Since \bar{p} was defined as the highest point in the support of F^ψ , choose $\epsilon > 0$ sufficiently small such that

$$F^\psi(\bar{p}) - F^\psi(\bar{p} - \epsilon) = \sum_{i=1}^n [F_i^\psi(\bar{p}) - F_i^\psi(\bar{p} - \epsilon)] > 0.$$

There must then exist some firm i such that

$$F_i^\psi(\bar{p}) - F_i^\psi(\bar{p} - \epsilon) \leq \frac{1}{n} [F^\psi(\bar{p}) - F^\psi(\bar{p} - \epsilon)].$$

Now consider the following uniform downward deviation for i :³⁷ whenever it would have chosen price $p \in (\bar{p} - \epsilon, 1]$, charge price $\bar{p} - \epsilon$ instead; if it would have chosen price $p \leq \bar{p} - \epsilon$, leaves prices unchanged. We conclude by showing that this deviation is strictly profitable.

Note that i 's loss on the extensive margin—the reduction in price charged to consumers she previously sold to—is at most

$$\epsilon [F_i^\psi(\bar{p}) - F_i^\psi(\bar{p} - \epsilon)];$$

³⁷Feldman, Lucier, and Nisan (2016); Bergemann, Brooks, and Morris (2017) also consider uniform upward deviations in auction settings.

on the other hand, the business stealing gain is at least

$$(\bar{p} - \epsilon) \sum_{j \neq i} [F_j^\psi(\bar{p}) - F_j^\psi(\bar{p} - \epsilon)]$$

since by deviating to $\bar{p} - \epsilon$, firm i now poaches all consumers who were previously buying from some firm $j \neq i$ at prices strictly greater than $\bar{p} - \epsilon$. For this to be an equilibrium induced by ψ , a necessary condition is that for all uniform downward deviations to $\bar{p} - \epsilon$ for $\epsilon > 0$,

$$\epsilon [F_i^\psi(\bar{p}) - F_i^\psi(\bar{p} - \epsilon)] \geq (\bar{p} - \epsilon) \sum_{j \neq i} [F_j^\psi(\bar{p}) - F_j^\psi(\bar{p} - \epsilon)].$$

But this implies

$$\frac{\bar{p} - \epsilon}{\epsilon} \leq \frac{F_i^\psi(\bar{p}) - F_i^\psi(\bar{p} - \epsilon)}{\sum_{j \neq i} [F_j^\psi(\bar{p}) - F_j^\psi(\bar{p} - \epsilon)]} \leq \frac{1}{n-1}$$

which is a contradiction for sufficiently small $\epsilon > 0$. \square

Proof of Proposition 2. Since producer-perfect information structure exists under μ , by Theorem 2, AIC holds:

$$\int_{\theta \in E_i: \theta_i \geq \hat{p}_i} \left(\frac{\theta_i - \hat{p}_i}{\hat{p}_i} \right) d\mu \geq \mu(\{\theta \notin E_i \mid \theta_i \geq \hat{p}_i\}), \text{ for all } i \text{ and all } \hat{p}_i > 0.$$

Define

$$\tilde{F}_{E_i}(c) := \hat{\mu}(E_i) - \lim_{x \downarrow c} \hat{\mu}(\{\theta \in E_i \mid \theta_i \geq x\}) \quad \text{and} \quad F_{E_i}(c) := \mu(E_i) - \lim_{x \downarrow c} \mu(\{\theta \in E_i \mid \theta_i \geq x\}).$$

Part (i) of definition 1 implies $\tilde{F}_{E_i} \leq F_{E_i}$.

We have

$$\begin{aligned} \int_{\theta \in E_i: \theta_i \geq \hat{p}_i} \left(\frac{\theta_i - \hat{p}_i}{\hat{p}_i} \right) d\hat{\mu} &= \int_{\theta_i \in [0,1]} \max \left\{ \frac{\theta_i - \hat{p}_i}{\hat{p}_i}, 0 \right\} d\tilde{F}_{E_i} \\ &\geq \int_{\theta_i \in [0,1]} \max \left\{ \frac{\theta_i - \hat{p}_i}{\hat{p}_i}, 0 \right\} dF_{E_i} \\ &= \int_{\theta \in E_i: \theta_i \geq \hat{p}_i} \left(\frac{\theta_i - \hat{p}_i}{\hat{p}_i} \right) d\mu \\ &\geq \mu(\{\theta \notin E_i \mid \theta_i \geq \hat{p}_i\}) \\ &\geq \hat{\mu}(\{\theta \notin E_i \mid \theta_i \geq \hat{p}_i\}) \end{aligned}$$

The first step replaces the restriction $\theta_i \geq \hat{p}_i$ with the max function and replacing measure $\hat{\mu}$ with un-normalized CDF \tilde{F}_{E_i} ; the second step is because $\max\{\frac{\theta_i - \hat{p}_i}{\hat{p}_i}, 0\}$ is an increasing function in θ_i and \tilde{F}_{E_i} first order stochastically dominates F_{E_i} ; the third step is by noticing that the distribution of θ_i induced by measure μ conditional on E_i is the same as the distribution of θ_i induced by F_{E_i} except that F_{E_i} put extra mass at $\theta_i = 0$.

The fourth step is by AIC. The last step is by (ii) of definition 1. Therefore, AIC also holds under $\hat{\mu}$. By Theorem 2, producer-perfect information structure exists under $\hat{\mu}$. \square

B: PROOFS FOR SECTION 5 (PRODUCER OPTIMAL IN SWITCHING COST MODEL)

B.1 Preliminaries for analyzing the switching cost environment. We first set out some notation before stating several lemmas which will be useful for analyzing the switching cost environment.

Notation. For compactness, we write $H1$ to represent the type $(1, 1 - \gamma)$, $H2$ to represent $(1 - \gamma, 1)$, $L1$ to represent $(1 - v, 1 - v - \gamma)$ and $L2$ to represent $(1 - v - \gamma, 1 - v)$. Under this labelling, our type space is $\Theta = \{H1, H2, L1, L2\}$. We sometimes use $H := 1$ and $L := 1 - v$ to represent the high and low prices.

Recall that an information structure is a joint distribution $\psi \in (\Theta \times M^2)$, and furthermore since prices are supported on a finite set, it is a probability mass function. We use $\psi_\theta(\cdot) := \psi(\cdot|\theta)$ to represent the joint distribution over firm 1 and 2's messages conditional on type θ . For instance, $\psi_{H1}(H, L)$ represents the mass of type $H1$ assigned the message H for firm 1, and L for firm 2. Since there are just two firms, we suppress the bold font for types.

Lemma 2. AIC holds in the switching cost model if and only if

$$\mu \geq \frac{1 - v - 2\gamma}{v} =: \underline{\mu}(\gamma, v) \quad \text{and} \quad \mu \leq \frac{\max\{\gamma - v, 0\}}{1 - \gamma - v} =: \bar{\mu}(\gamma, v).$$

Proof of Lemma 2.

$$G(p|1) = \frac{1 - p}{p} \cdot \frac{\mu}{2} \quad p < 1$$

and

$$G(p|1 - v) = \frac{1 - v - p}{p} \cdot \frac{1 - \mu}{2} \quad p < 1 - v$$

so

$$H(p) = \begin{cases} \frac{1 - p}{p} \cdot \frac{\mu}{2} & \text{if } p < 1 \\ \frac{1 - v - p}{2p}(1 - \mu) + \frac{\mu}{2} \cdot \frac{1 - p}{p} & \text{if } p < 1 - v \end{cases}$$

There are two prices to check: $1 - \gamma$ and $1 - v - \gamma$. If $1 - \gamma \leq 1 - v \iff v \leq \gamma$ then we have

$$\begin{aligned} H(1 - \gamma) &= \frac{\gamma - v}{2(1 - \gamma)}(1 - \mu) + \frac{\mu}{2} \cdot \frac{\gamma}{1 - \gamma} \geq \frac{\mu}{2} \\ &\iff (\gamma - v)(1 - \mu) + \mu \cdot v \geq \mu(1 - \gamma) \\ &\iff \mu \leq \frac{\gamma - v}{1 - \gamma - v} \end{aligned}$$

and also

$$\begin{aligned}
H(1 - v - \gamma) &= \frac{\gamma}{2(1 - v - \gamma)}(1 - \mu) + \frac{\mu}{2} \cdot \frac{v + \gamma}{1 - v - \gamma} \geq \frac{1}{2} \\
&\iff \gamma + \mu \cdot v \geq 1 - v - \gamma \\
&\iff \mu \geq \frac{1 - v - 2\gamma}{v}
\end{aligned}$$

which is exactly the condition in the lemma. If $1 - \gamma > 1 - v \iff v > \gamma$ then the L type cannot be used to match to the H type. So we have

$$H(1 - \gamma) = \frac{\mu}{2} \cdot \frac{\gamma}{1 - \gamma} \geq \frac{\mu}{2}$$

which holds for $\mu > 0$ if and only if $\gamma \geq \frac{1}{2}$ but the condition in μ violates the condition $1 - v - \gamma > 0$. Hence, when $\gamma < v$, AIC holds only if and only if $\mu = 0$ such that $H(1 - \gamma)$ is automatically fulfilled and

$$H(1 - \gamma - v) = \frac{\gamma}{2(1 - v - \gamma)}(1 - \mu) \geq \frac{1 - \mu}{2} \iff \gamma \geq \frac{1 - v}{2}$$

which is equivalent to the condition in the lemma since

$$\gamma \geq \frac{1 - v}{2} \implies \underline{\mu} = \frac{1 - v - 2\gamma}{v} \leq 0.$$

□

Lemma 3. Let Ψ^E be the set of obedient information structures which implement efficient outcomes. Then, there exists a producer-optimal design $\psi^* \in \operatorname{argmax}_{\psi \in \Psi^E} PS(\psi)$ such that

(i) ψ^* is supported on the prices $\{0, \gamma, L, H\}$ i.e.,

$$\operatorname{supp} \operatorname{marg}_{M^2} \psi = \{0, \gamma, L, H\}^2;$$

(ii) ψ^* is symmetric: for any m, m' ,

$$\psi_{H1}(m, m') = \psi_{H2}(m', m) \quad \text{and} \quad \psi_{L1}(m, m') = \psi_{L2}(m', m).$$

Proof of Lemma 3. We show each part in turn.

Part (i): Supported on prices $\{0, \gamma, L, H\}$. We start by observing that we must have

$$\int_{\substack{\Theta \times M^2: \\ \theta_1 - m_1 = \theta_2 - m_2 \\ \theta_1 - m_1 > 0 \\ \min\{m_1, m_2\} > 0}} d\psi = 0. \quad (\text{No ties})$$

To see this, suppose **No ties** did not hold and it is of measure $c > 0$. Now consider the uniform ϵ -downward deviation by any firm which we define as follows: for any recommendation p it receives, charge $p - \epsilon$. Note that this increases demand by at least c , and since the original demand was bounded, the losses from such a deviation is at

most ϵ . Hence we can find a small enough ϵ so that this deviation is strictly profitable, a contradiction.

Now we proceed via a contagion argument typical to the analysis of games of incomplete information with non-degenerate higher-order beliefs. This argument is adapted to the setting of Bertrand competition.

First suppose that $0 < m_1 < \gamma$ occurs with positive probability. Then, on the event that firm 1 makes the sale, by efficiency it sells to a consumer of type in E_1 and it can also make the sale by uniformly deviating upwards on the recommendation $\{0 < m_1 < \gamma\}$ to the price γ and this is a strictly profitable and likewise for firm 2. If firm 1 never makes the sale upon receipt of m_1 , we can normalize $m_1 = 0$ without loss. Hence, we have

$$\inf \left(\left(\text{supp marg}_{\mathcal{G}_{M1}} \psi \right) \cup \left(\text{supp marg}_{\mathcal{G}_{M2}} \psi \right) \setminus \{0\} \right) \geq \gamma. \quad (1)$$

Note that this argument does not rule out the possibility of atoms on γ since we have not ruled out the possibility of atoms on 0 (which is not ruled out by **No ties**). Now, suppose that $\gamma \leq m_1 < 2\gamma$ with positive probability. Since ψ is efficient, we have that on the event it makes the sale, $\theta \in E_1$. On this event, from equation 1 firm 1 thus has a strictly profitable uniform deviation on the set $\{\gamma \leq m_1 < 2\gamma\}$ up to the price $\min\{2\gamma, 1 - v\}$. Suppose without loss that $2\gamma < 1 - v$. Then,

$$\inf \left(\left(\text{supp marg}_{\mathcal{G}_{M1}} \psi \right) \cup \left(\text{supp marg}_{\mathcal{G}_{M2}} \psi \right) \setminus \{0, \gamma\} \right) \geq 2\gamma$$

noting that **No ties** rules out atoms on price 2γ . Now suppose that for some positive integer K , where $(K + 1)\gamma \leq 1 - v =: L$,

$$\inf \left(\left(\text{supp marg}_{\mathcal{G}_{M1}} \psi \right) \cup \left(\text{supp marg}_{\mathcal{G}_{M2}} \psi \right) \setminus \{0, \gamma\} \right) \geq K\gamma. \quad (2)$$

Then suppose that $K\gamma \leq m_1 < 1 - v$ with positive probability. From equation 2, since ψ is efficient, on the event firm 1 makes a sale, it also does so at the price $(K + 1)\gamma$ since $(K + 1)\gamma \leq 1 - v$ and this satisfies the consumer's IC constraint. Now define K_L which fulfils

$$K_L\gamma < 1 - v \leq (K_L + 1)\gamma.$$

We can iterate the previous argument to arrive at

$$\inf \left(\left(\text{supp marg}_{\mathcal{G}_{M1}} \psi \right) \cup \left(\text{supp marg}_{\mathcal{G}_{M2}} \psi \right) \setminus \{0, \gamma\} \right) \geq K_L\gamma.$$

Now suppose that $K_L\gamma < m_1 < (K_L + 1)\gamma$ with positive probability. By the same argument as before, the upward deviation to the price $1 - v$ must be strictly profitable. At this price, the consumer's IC is tight so an atom on $1 - v =: L$ can potentially be sustained. Hence, we have

$$\inf \left(\left(\text{supp marg}_{\mathcal{G}_{M1}} \psi \right) \cup \left(\text{supp marg}_{\mathcal{G}_{M2}} \psi \right) \setminus \{0, \gamma, L\} \right) \geq 1 - v =: L.$$

Now suppose $L < m_1 \leq L + \gamma$ with positive probability. Since ψ is efficient, if firm

1 makes a sale, $\theta = H1$. But this implies a uniform upward deviation from the event $\{L < m_1 < L + \gamma\}$ to $L + \gamma$. Hence,

$$\inf \left(\left(\text{supp marg}_{M1} \psi \right) \cup \left(\text{supp marg}_{M2} \psi \right) \setminus \{0, \gamma, L\} \right) \geq L + \gamma.$$

Moreover, observe that by **No ties** there cannot be an atom at $L + \gamma$. Now once again suppose that K_H is such that

$$L + K_H \gamma \leq 1 \leq L + (K_H + 1) \gamma$$

and iterating this argument forward, we have

$$\inf \left(\left(\text{supp marg}_{M1} \psi \right) \cup \left(\text{supp marg}_{M2} \psi \right) \setminus \{0, \gamma, L\} \right) \geq L + K_H \gamma.$$

Then suppose $L + K_H \gamma < m_1 < L + (K_H + 1) \gamma$ with positive probability. Once again, since ψ is efficient, conditional on making the sale, $\theta = H1$ so the uniform deviation from $\{L + K_H \gamma < m_1 < L + (K_H + 1) \gamma\}$ to $1 =: H$ is strictly profitable. This implies

$$\text{supp marg}_{M1} \psi \subseteq \{0, \gamma, L, H\} \quad \text{and} \quad \text{supp marg}_{M2} \psi \subseteq \{0, \gamma, L, H\}$$

as required.

Part (ii): Symmetry. Start with an arbitrary obedient structure ψ and suppose it maximizes producer surplus. Let ψ' be a symmetrized version of ψ defined naturally as:

$$\psi'_{H1}(m, m') = \psi_{H2}(m', m) \quad \psi'_{L1}(m, m') = \psi_{L2}(m', m) \quad \text{for all } m, m' \in [0, 1].$$

Observe ψ' continues to be obedient and implements the same PS as ψ since the distribution is symmetric. Now consider the information structure $\psi^* := \frac{1}{2}\psi + \frac{1}{2}\psi'$. Clearly consistency is fulfilled; it remains to check each firm's IC.

From part (i), the ties only occur when one firm charges a price of γ , and the other charges a price of 0 , and the consumer prefers the former firm. On these events, we break ties in favor of efficiency. Further, from part (i) ψ is a probability mass function. Now define

$$D_\psi(m_1, m'_1) := \int_{\Theta} \int_{\substack{\theta_1 - \theta_2 \geq m'_1 - m_2 \\ \theta_1 \geq m_1}} \psi_\theta(m_1, m_2) dm_2 d\theta$$

which is the expected demand given that firm 1 receives the message m_1 but prices at m'_1 . Firm IC requires

$$m_1 D_\psi(m_1, m_1) \geq m'_1 D_\psi(m_1, m'_1) \quad \text{for all } m_1 \in \text{supp marg}_{M1} \psi \text{ and all } m'_1.$$

and symmetrically for firm 2's recommendations. The rest of the argument follows from the linearity of the expectation operator: under the new information structure,

$$D_{\psi^*}(m_1, m'_1) := \frac{1}{2} D_\psi(m_1, p'_1) + \frac{1}{2} D_{\psi'}(m_1, p'_1)$$

hence

$$\begin{aligned}
m_1 D_{\psi^*}(m_1, m_1) &= m_1 \left(\frac{1}{2} D_{\psi}(m_1, m_1) + \frac{1}{2} D_{\psi'}(m_1, m_1) \right) \\
&\geq m'_1 \left(\frac{1}{2} D_{\psi}(m_1, m'_1) + \frac{1}{2} D_{\psi'}(m_1, m'_1) \right) \\
&= m'_1 D_{\psi^*}(m_1, m'_1)
\end{aligned}$$

as required. Furthermore, payoffs are preserved but ψ^* is now symmetric. \square

B.2 Construction of relaxed problem. With Lemma 3 in hand, the information structure is pinned down by

$$\left(\psi_{\theta}(m_1, m_2) \right)_{\theta \in \{H1, L1\}, m_1, m_2 \in P} \quad \text{where } P := \{0, \gamma, L, H\}.$$

where we leave the condition that ψ_{θ} is efficient implicit. Let $\Psi^{S,E}$ be the set of efficient and efficient structures. Write

$$\psi_{\theta}(m) := \sum_{m_2 \in P} \psi_{\theta}(m, m_2) \quad \text{and} \quad \psi^{\gamma} := \sum_{\theta \in \{L1, H1\}} \psi_{\theta}(\gamma)$$

as the mass of consumers of type $\theta \in E_1$ assigned to recommendation m for firm 1, and the mass of consumers assigned to recommendation γ for firm 1 respectively.

To characterize the producer-optimal design, we will construct and solve a relaxed problem, and show that we can attain the bound.

Construct a relaxed problem. We will develop generalized versions of AIC:

$$\gamma \psi_{H1}(1) + \mathbb{1}_{v < \gamma} \cdot (\gamma - v) \cdot \left(\frac{1}{2} - \psi^{\gamma} - \psi_{H1}(1) \right) \geq (1 - \gamma) \psi_{H1}(1) \quad (\text{G-AIC-H})$$

Claim: G-AIC-H is necessary. To see this, observe that for each price $\{m \in P : m > 1 - \gamma\}$, firm 1 should have no incentive to lower prices to $1 - \gamma$ upon receipt of the recommendation m . Hence, we require

$$\begin{aligned}
(m - (1 - \gamma)) \left(\psi_{H1}(m) + \psi_{L1}(m) \right) \\
\geq (1 - \gamma) \psi_{H2}(m, 1) = (1 - \gamma) \psi_{H1}(1, m)
\end{aligned}$$

Note since ψ is efficient, types in E_1 must accept firm 1's price and purchase product 1. The last equality is by symmetry. Then, summing over $m \in P$ where $m > 1 - \gamma$,

$$\begin{aligned}
\gamma \psi_{H1}(1) + \mathbb{1}_{v < \gamma} \cdot (\gamma - v) \left(\psi_{H1}(1 - v) + \psi_{L1}(1 - v) \right) \\
\geq (1 - \gamma) \psi_{H1}(1)
\end{aligned}$$

where the right hand side of the inequality follows from summing over $\psi_{H1}(1, m)$, and efficiency requires $\psi_{H1}(1, m) = 0$ for $m \leq 1 - \gamma$. Then observe that by doing some

accounting we have

$$\psi_{H1}(1-v) + \psi_{L1}(1-v) = \frac{1}{2} - \psi^\gamma - \psi_{H1}(1)$$

which can be rearranged to obtain **G-AIC-H**.

Next, we have

$$\begin{aligned} (v + \gamma)\psi_{H1}(1) + \gamma\left(\frac{1}{2} - \psi^\gamma - \psi_{H1}(1)\right) + \mathbb{1}_{\gamma > \frac{1-v}{2}}(2\gamma - (1-v))\psi^\gamma \\ \geq (1-v-\gamma)\left(\frac{1}{2} - \psi^\gamma\right) \end{aligned} \quad (\text{G-AIC-L})$$

Claim: G-AIC-L is necessary. To see this, observe that for each price $m \in P$ with $m > 1-v-\gamma$, firm 1 should have no incentives to deviate downwards to price $1-v-\gamma$:

$$\begin{aligned} (m - (1-v-\gamma))(\psi_{H1}(m) + \psi_{L1}(m)) \\ \geq (1-v-\gamma)\left(\sum_{m' \in P: m' \geq 1-v} \psi_{H2}(m, m') + \sum_{m' \in P: m' \geq 1-v} \psi_{L2}(m, m')\right) \\ = (1-v-\gamma)\left(\sum_{m' \in P: m' \geq 1-v} \psi_{H1}(m', m) + \sum_{m' \in P: m' \geq 1-v} \psi_{L1}(m', m)\right) \end{aligned}$$

where last equality is again by symmetry. Note that if $m > 1-v$ then $\psi_{L1}(m) = 0$.

Once again summing across all $m \in P$ with $m > 1-v-\gamma$:

$$\begin{aligned} (v + \gamma)\psi_{H1}(1) + \gamma(\psi_{H1}(1-v) + \psi_{L1}(1-v)) + \mathbb{1}_{\gamma > \frac{1-v}{2}}(2\gamma - (1-v))(\psi_{H1}(\gamma) + \psi_{L1}(\gamma)) \\ \geq (1-v-\gamma) \sum_{m \in P: m > 1-v-\gamma} \left(\sum_{m' \in P: m' \geq 1-v} \psi_{H1}(m', m) + \sum_{m' \in P: m' \geq 1-v} \psi_{L1}(m', m) \right) \\ = (1-v-\gamma) \sum_{m' \in P: m' \geq 1-v} \sum_{m \in P: m > 1-v-\gamma} (\psi_{H1}(m', m) + \psi_{L1}(m', m)) \\ = (1-v-\gamma)(\psi_{H1}(1) + \psi_{H1}(1-v) + \psi_{L1}(1-v)) \\ = (1-v-\gamma)\left(\frac{1}{2} - \psi^\gamma\right) \end{aligned}$$

The second last equality is because efficiency requires $\psi_{H1}(1, m) = 0$, $\psi_{H1}(1-v, m) = 0$ and $\psi_{L1}(1-v, m) = 0$ if $m \leq 1-v-\gamma$. Thus G-AIC-L is necessary.

We also write down several individual IC constraints which do not involve downward deviations to either the price 1 (handled by **G-AIC-H**) or price $1-v$ (handled by **G-AIC-L**). These will also be necessary for ψ to be obedient.

We first have:

$$(2\gamma - (1-v)) \cdot \psi^\gamma \geq (1-v-\gamma)\psi_{H1}(1-v, \gamma) \quad (\gamma \downarrow 1-v-\gamma)$$

if $\gamma > \frac{1-v}{2}$. This ensures that firm 1 does not wish to deviate down from γ to $1-v-\gamma$ to steal types $H2$ which are charged $1-v$ by firm 2.

We also have:

$$(1-v) \underbrace{\left(\frac{1}{2} - \psi^\gamma - \psi_{H1}(1) \right)}_{=\psi_{H1}(1-v)+\psi_{L1}(1-v)} \geq 2\gamma\psi_{H1}(1-v) \quad (1-v \uparrow 2\gamma)$$

if $\gamma > \frac{1-v}{2}$. This ensures that firm 1 does not wish to deviate up from the price $1-v$ to 2γ to sell to the high types being charged price $1-v$.

We also have:

$$(1-v) \left(\frac{1}{2} - \psi^\gamma - \psi_{H1}(1) \right) \geq (1-v+\gamma) \left(\psi_{H1}(1-v) - \psi_{H1}(1-v, \gamma) \right) \quad (1-v \uparrow 1-v+\gamma)$$

if $\gamma < v$. This ensures that firm 1 does not wish to deviate up to the price $1-v+\gamma$ to sell to the high types being charge a price of $1-v$ (with the exception of those being charged price γ by firm 2). In particular, this steals the high types being charged a price $1-v$ by firm 2 (since $1-v+\gamma < 1$).

Finally we have upward deviations from the price $1-v$ to price 1.

$$(1-v) \underbrace{\left(\frac{1}{2} - \psi^\gamma - \psi_{H1}(1) \right)}_{=\psi_{H1}(1-v)+\psi_{L1}(1-v)} \geq \psi_{H1}(1-v, 1) + \mathbb{1}_{\gamma \geq v} \cdot \psi_{H1}(1-v, 1-v) \quad (1-v \uparrow 1)$$

Our relaxed program is thus

$$\begin{aligned} \max_{\psi \in \Psi^{S,E}} & 2 \cdot \sum_{\theta \in E_1} \sum_{m, m'} m \cdot \psi_\theta(m, m') \\ \text{s.t.} & \text{G-AIC-H, G-AIC-L} \\ & \text{individual ICs} \\ & \text{consistency.} \end{aligned} \quad (\text{R})$$

where $\Psi^{S,E}$ is the set of efficient and symmetric information structures. Note that a solution to this problem does not imply obedience, but every obedient $\psi \in \Psi^{S,E}$ fulfils the constraints in **R**. To show producer-optimality, we will solve **R** and construct an information design which attains it on the original problem.

B.3 Proof of Proposition 3. We are now ready to show Proposition 3 in which we focus on the parameter space $\gamma \geq v$. Recall we split our discussion of Proposition 3 into three cases. Case A is such that $\gamma \geq \frac{1-v}{2}$ so that downward deviations to $1-v-\gamma$ are always unprofitable. Case B is such that $\underline{\gamma}(v) \leq \gamma < \frac{1-v}{2}$ so that downward deviations to $1-v-\gamma$ are potentially profitable and must be deterred by the information design. Case C is such that $\gamma < \underline{\gamma}(v)$.

Case A: $\gamma \geq \frac{1-v}{2}$. From Lemma 2, $PS^* = TS$ if and only if $\mu \leq \bar{\mu}(\gamma, v)$. Now consider the case $\mu > \bar{\mu}(\gamma, v)$.

We focus on case $\gamma > \frac{1-v}{2}$; the edge case of $\gamma = \frac{1-v}{2}$ will be discussed later. **G-AIC-H** is equivalent to

$$\psi_{H1}(1) \leq \left(\frac{1}{2} - \psi^\gamma \right) \cdot \bar{\mu}(\gamma, v).$$

Since $\gamma > v$, we have the individual constraints $1 - v \uparrow 2\gamma$, $\gamma \downarrow 1 - v - \gamma$ and $1 - v \uparrow 1$. It is without loss to set $\psi_{L1}(\gamma) = 0$. Therefore, $\psi_{H1}(1 - v) = \frac{\mu}{2} - \psi^\gamma - \psi_{H1}(1)$. We can then rewrite the objective in **R** as

$$TS^* - 2v \left(\frac{\mu}{2} - \psi^\gamma - \psi_{H1}(1) \right) - 2 \cdot (1 - \gamma) \cdot \psi^\gamma$$

which is clearly maximized by binding **G-AIC-H**. This attains value:

$$TS^* - v\mu + v \cdot \bar{\mu} - 2(1 - v - \gamma + v \cdot \bar{\mu}) \cdot \psi^\gamma$$

which is decreasing in ψ^γ .

Combining **G-AIC-H**, $\gamma \downarrow 1 - v - \gamma$ and $1 - v \uparrow 1$ we have the lower bound

$$\psi^\gamma \geq \psi_{\uparrow 1}^\gamma := \frac{(\mu - v\bar{\mu} - (1 - v))(1 - v - \gamma)}{2v(1 - \bar{\mu})(1 - v - \gamma) + 4\gamma - 2(1 - v)}$$

Combining **G-AIC-H**, $\gamma \downarrow 1 - v - \gamma$ and $1 - v \uparrow 2\gamma$ we have the lower bound

$$\psi^\gamma \geq \psi_{\uparrow 2\gamma}^\gamma := \frac{(\mu - \bar{\mu})\gamma + (\bar{\mu} - 1)(1 - v)/2}{[2\gamma - (1 - v)](1 - \bar{\mu})}$$

where we note that here $\psi_{\uparrow 1}^\gamma$ ($\psi_{\uparrow 2\gamma}^\gamma$) corresponds to a constraint imposed by the upward deviation to 1 (2γ). Further observe

$$\psi_{\uparrow 1}^\gamma \geq \psi_{\uparrow 2\gamma}^\gamma \iff \mu \leq v\bar{\mu} + (1 - v) + \tilde{\mu}.$$

where one can verify that $\tilde{\mu}$ has the expression

$$\tilde{\mu}(\gamma, v) := \frac{(1 - v)(1 - \bar{\mu})(1/2 - \gamma)[2v(1 - \bar{\mu})(1 - v - \gamma) + 4\gamma - 2(1 - v)]}{\gamma 2v(1 - \bar{\mu})(1 - v - \gamma) + (2\gamma - (1 - v))((3 - \bar{\mu})\gamma - (1 - \bar{\mu})(1 - v))}.$$

For the edge case with $\gamma = \frac{1-v}{2}$, there is no constraint $1 - v \uparrow 2\gamma$. Hence μ only has once cutoff— $\psi^\gamma \geq \psi_{\uparrow 1}^\gamma$ solves the relaxed problem whenever $v\bar{\mu} + (1 - v) \leq \mu$.

We now show the original program attains **R**. The individual constraints $1 - v \uparrow 1$ and $1 - v \uparrow 2\gamma$ take care of upward deviations from price $1 - v$. There are no profitable upward deviations from the price γ because for $\psi_{H1}(\gamma)$ we set $\psi_{H1}(\gamma, 0) = \psi_{H1}(\gamma)$. Next, the individual constraint $\gamma \downarrow 1 - v - \gamma$ pins down $\psi_{H1}(1 - v, \gamma)$.

It remains to match $\psi_{H1}(1) + \psi_{H1}(1 - v) - \psi_{H1}(1 - v, \gamma)$ mass of type $H1$ and $\psi_{L1}(1 - v)$ mass of type $L1$ to prices 1 and $1 - v$ in such a way as to deter downward deviations: upon receipt of the messages 1 and $1 - v$, firm 2 does not find it profitable to deviate to price $1 - \gamma$ or $1 - v - \gamma$. Observe $\psi_{H1}(1)$ mass of type $H1$ has reservation price $1 - \gamma$ for firm 2's product. Thus, in order to deter deviations to $1 - \gamma$ we require (dividing through by $1 - \gamma$)

$$\underbrace{\frac{\gamma}{1 - \gamma} \cdot \psi_{H1}(1) + \frac{\gamma - v}{1 - \gamma} \cdot \left(\frac{1}{2} - \psi^\gamma - \psi_{H1}(1) \right)}_{\text{Matching capacity of } \psi_{H1}(1)} \geq \psi_{H1}(1)$$

which yields exactly **G-AIC-H** which was a constraint on the relaxed problem.

Similarly, observe that $\psi_{H1}(1-v) - \psi_{H1}(1-v, \gamma)$ mass of type $H1$ and $\psi_{L1}(1-v)$ mass of type $L1$ has reservation price $1-v-\gamma$ for firm 2's product. Thus, in order to deter deviations to $1-\gamma-v$ we require

$$\underbrace{\frac{v+\gamma}{1-v-\gamma}\psi_{H1}(1) + \frac{\gamma}{1-v-\gamma}\left(\frac{1}{2} - \psi^\gamma - \psi_{H1}(1)\right)}_{\text{Matching capacity of } \psi_{H1}(1-v) - \psi_{H1}(1-v, \gamma) + \psi_{L1}(1-v)} \geq \frac{1}{2} - \psi^\gamma - \psi_{H1}(1-v, \gamma)$$

which holds because $\gamma \geq \frac{1-v}{2}$ and

$$(v+\gamma)\psi_{H1}(1) + \gamma\left(\frac{1}{2} - \psi^\gamma - \psi_{H1}(1)\right) \geq (1-v-\gamma)\left(\frac{1}{2} - \psi^\gamma\right).$$

Since total matching capacity is larger than total matching demand, by the same argument as in proof of Theorem (2), the matching can be done so that individual downward deviations are not profitable for prices 1 and $1-v$. Hence, $PS^* = R$.

Case B: $\underline{\gamma}(v) < \gamma < \frac{1-v}{2}$. We have from Lemma 2 that if $\underline{\mu} \leq \mu \leq \bar{\mu}$ then AIC holds so $\underline{PS}^* = \underline{TS}^*$.

Now suppose $\mu < \underline{\mu}$. **G-AIC-L** implies

$$\psi_{L1}(1-v) \leq \frac{\psi_{H1}(1)}{\underline{\mu}} - \psi_{H1}(1) - \psi_{H1}(1-v) \leq \frac{1-\underline{\mu}\mu}{\underline{\mu}} \frac{\mu}{2} < \frac{1-\mu}{2}$$

where recall

$$\underline{\mu}(\gamma, v) := \frac{1-v-2\gamma}{v}.$$

Hence, at least $\frac{1-\mu}{2} - \frac{1-\mu}{\underline{\mu}} \cdot \frac{\mu}{2}$ mass of type $L1$ must to purchase firm 1's product at price γ . Hence, an upper bound on R is

$$TS^* - 2(1-v-\gamma)\left(\frac{1-\mu}{2} - \frac{1-\underline{\mu}\mu}{\underline{\mu}} \frac{\mu}{2}\right)$$

which is achieved by an information structure which removes

$$\left(\frac{1-\mu}{2} - \frac{1-\underline{\mu}\mu}{\underline{\mu}} \frac{\mu}{2}\right)$$

mass of types $L1$ and $L2$ and fully reveals them to obtain a producer-surplus of γ per unit. It then applies the producer perfect information structure to remaining distribution to achieve the bound so $PS^* = R$.

Next suppose that $\mu > \bar{\mu}$. **G-AIC-H** implies

$$\psi_{H1}(1) \leq \left(\frac{1}{2} - \psi^\gamma\right) \cdot \bar{\mu}$$

and, ignoring **G-AIC-L** for the moment, we have the extra constraint $1 - v \uparrow 1$:

$$(1 - v) \left(\frac{1}{2} - \psi_{H1}(1) - \psi^\gamma \right) \geq \psi_{H1}(1 - v, 1) + \psi_{H1}(1 - v, 1 - v) = \psi_{H1}(1 - v)$$

where the last inequality is because $\psi_{H1}(1 - v, \gamma) = 0$. Once again it is without loss to set $\psi_{L1}(\gamma) = 0$. Hence, we can write the total measure of type $H1$ being charged a price $1 - v$ as

$$\psi_{H1}(1 - v) = \frac{\mu}{2} - \psi_{H1}(1) - \psi^\gamma.$$

Combined with **G-AIC-H** and $1 - v \uparrow 1$ this delivers a lower bound on ψ^γ :

$$\psi^\gamma \geq \bar{\psi}_{\uparrow 1}^\gamma := \frac{\frac{\mu}{2} - \bar{\mu}/2 - \frac{(1-v)(1-\mu)}{2v}}{1 - \bar{\mu}}$$

The objective function in the relaxed problem can is then

$$TS^* - 2v \left(\frac{\mu}{2} - \psi_{H1}(1) - \psi^\gamma \right) - 2(1 - \gamma)\psi^\gamma$$

which is maximized by binding **G-AIC-H** and attains value:

$$TS^* - v\mu + v\bar{\mu} - 2(1 - v - \gamma + v\bar{\mu})\psi^\gamma$$

which is decreasing in ψ^γ . Hence the relaxed problem is solved by choosing

$$\psi^\gamma = \min \left\{ \bar{\psi}_{\uparrow 1}^\gamma, 0 \right\}.$$

We now show this value is obtainable by constructing an information structure ψ which achieves this upper bound. There are no upward deviations for price 1. The individual IC constraints in the relaxed problem handles upward deviations from $1 - v$. For price γ , only type $H1$ will purchase product 1 at price γ since $\psi_{L1}(\gamma) = 0$. For $\psi_{H1}(\gamma)$, we can set $\psi_{H1}((\gamma, 0)) = \psi_{H1}(\gamma)$ so upward deviations for from price γ are not profitable.

It remains to match $\psi_{H1}(1) + \psi_{H1}(1 - v) - \psi_{H1}(1 - v, \gamma)$ mass of type $H1$ and $\psi_{L1}(1 - v)$ mass of type $L1$ to prices 1 and $1 - v$ in such a way as to deter downward deviations: upon receipt of the messages 1 and $1 - v$, firm 2 does not find it profitable to deviate to price $1 - \gamma$ or $1 - v - \gamma$. Observe $\psi_{H1}(1)$ mass of type $H1$ has reservation price $1 - \gamma$ for firm 2's product. Thus, in order to deter deviations to $1 - \gamma$ we require

$$\gamma\psi_{H1}(1) + (\gamma - v) \left(\frac{1}{2} - \psi_{H1}(1) - \psi^\gamma \right) \geq (1 - \gamma)\psi_{H1}(1)$$

which is exactly **G-AIC-H**.

Similarly, to deter deviations to $1 - v - \gamma$, we require

$$(v + \gamma)\psi_{H1}(1) + \underbrace{\gamma \left(\frac{1}{2} - \psi_{H1}(1) - \psi^\gamma \right)}_{\psi_{L1}(1-v) + \psi_{H1}(1-v)} \geq (1 - v - \gamma) \left(\frac{1}{2} - \psi^\gamma \right)$$

which is exactly **G-AIC-L**. By the same procedure in the proof of Theorem 2 we can construct ψ such that the individual downward deviations to prices $1 - \gamma$ and $1 - \gamma - v$ are deterred. Hence, $PS^* = R$.

Case C: $\gamma < \underline{\gamma}(v)$. In this region, $\gamma \leq \frac{1-v}{2}$. **G-AIC-H** and **G-AIC-L** simplify to

$$\gamma\psi_{H1}(1) + \mathbb{1}_{v < \gamma}(\gamma - v) \left(\frac{1}{2} - \psi^\gamma - \psi_{H1}(1) \right) \geq (1 - \gamma)\psi_{H1}(1);$$

and

$$(v + \gamma)\psi_{H1}(1) + \gamma \left(\frac{1}{2} - \psi^\gamma - \psi_{H1}(1) \right) \geq (1 - v - \gamma) \left(\frac{1}{2} - \psi^\gamma \right).$$

Now observe that setting $\mu = \frac{\psi_{H1}(1)}{\frac{1}{2} - \psi^\gamma}$ we recover **AIC-High** and **AIC-Low** in the main text. But from Lemma 2, when $\gamma < \underline{\gamma}(v)$ AIC is never fulfilled for $\mu > 0$. This implies that for **G-AIC-H** and **G-AIC-L** to be fulfilled we must have $\psi_{H1}(1) = 0$. But if so, we must have

$$\psi_{H1}(1) = \frac{1}{2} - \psi^\gamma - \psi_{H1}(1) = 0$$

which implies the relaxed program has value γ , which is achieved by giving full information so $PS^* = R$.

B.4 Producer-optimal design when $\gamma < v$. We now state and prove results about producer-optimal designs in the case $\gamma < v$ which was omitted in the main text.

Proposition 4. The following characterizes the producer surplus obtained at the producer-optimal efficient outcome, $PS^*(v, \gamma, \mu)$ for low product differentiation ($\gamma < v$):

(i) If $\frac{1-v}{2} \leq \gamma < \frac{1}{2}$:

$$PS^*(\mu, \gamma, v) = \begin{cases} TS^*(v, \mu) - v \cdot \mu & \text{if } \mu \leq \bar{\mu}^L(\gamma, v) \\ g^L(\mu, \gamma, v) & \text{if } \bar{\mu}^L(\gamma, v) \leq \mu, \end{cases}$$

where

$$\bar{\mu}^L(\gamma, v) := \frac{1 - v}{1 - v + \gamma}$$

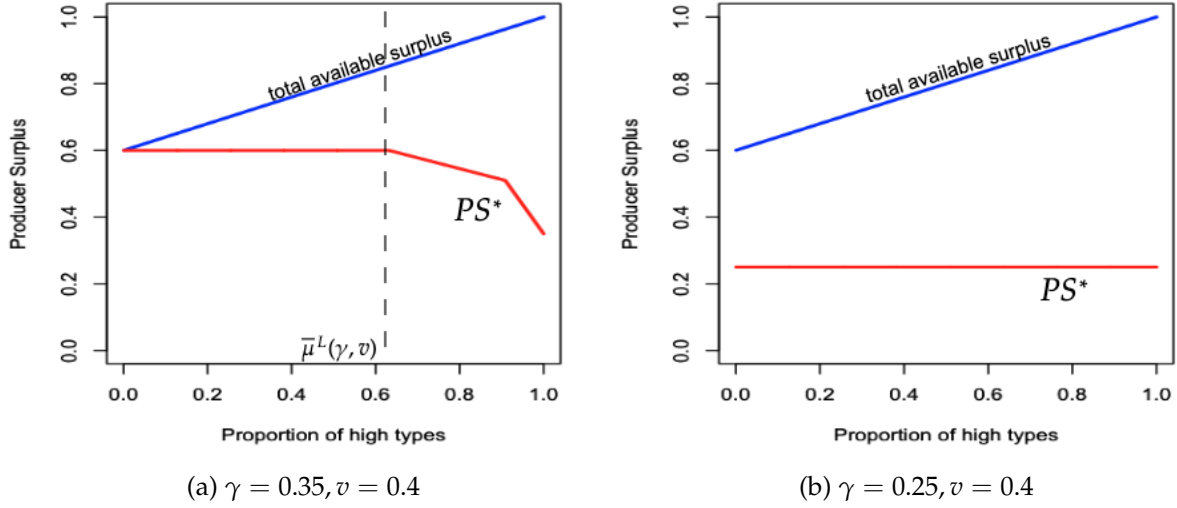
and g^L is continuous, strictly decreasing in μ , and $g^L(1, \gamma, v) = \gamma$.

(ii) If $\gamma < \frac{1-v}{2}$:

$$PS^*(\mu, \gamma, v) = \gamma.$$

When $\gamma < v$, the ability for information to soften competition is severely diminished as is reflected by Proposition 4 which is depicted in Figure For instance, consider the price recommendation of $1 - v$ for firm 1. If we allocate any high-type consumers who prefer firm 2 to this message, such a consumer would obtain $1 - \gamma - (1 - v) = v - \gamma > 0$ buying from firm 1. Thus, the designer must ensure that such consumers are not charged a price of 1 by firm 2 which restricts the power of information to segment the market.

Figure 6: Illustration of optimal producer surplus as distribution varies ($\gamma < v$)



Proof of Proposition 4. We prove each part in turn.

Part (i): $\frac{1-v}{2} \leq \gamma < v$. **G-AIC-H** simplifies to $\gamma\psi_{H1}(1) \geq (1-\gamma)\psi_{H1}(1)$ which implies $\psi_{H1}(1) = 0$ since $\gamma < \frac{1}{2}$. This also implies $\psi_{H1}(1-v, 1) = \psi_{H2}(1, 1-v) \leq \psi_{H2}(1) = 0$.

G-AIC-L simplifies to

$$\gamma\left(\frac{1}{2} - \psi^\gamma\right) + (2\gamma - (1-v)) \cdot \psi^\gamma \geq (1-v-\gamma)\left(\frac{1}{2} - \psi^\gamma\right),$$

which is always satisfied since $\gamma \geq \frac{1-v}{2}$. For the remaining constraints, we focus on the case $\gamma > \frac{1-v}{2}$ and discuss the edge case with $\gamma = \frac{1-v}{2}$ later on.

Constraint $\gamma \downarrow 1-v-\gamma$ simplifies to:

$$(2\gamma - (1-v))\psi^\gamma \geq (1-v-\gamma)\psi_{H1}(1-v, \gamma).$$

Constraint $1-v \uparrow 2\gamma$ simplifies to :

$$(1-v)\left(\frac{1}{2} - \psi^\gamma\right) \geq 2\gamma\psi_{H1}(1-v).$$

Constraint $1-v \uparrow 1-v+\gamma$ simplifies to:

$$(1-v)\left(\frac{1}{2} - \psi^\gamma\right) \geq (1-v+\gamma)\left(\psi_{H1}(1-v) - \psi_{H1}(1-v, \gamma)\right).$$

and $1-v \uparrow 1$ always holds.

We can then write the objective in **R** as $1-v-2\psi^\gamma(1-v-\gamma)$. Note it is once again without loss to set $\psi_{L1}(\gamma) = 0$. We then have

$$\psi^\gamma = \frac{\mu}{2} = \psi_{H1}(1-v) \implies \psi_{H1}(1-v) = \frac{\mu}{2} - \psi^\gamma.$$

Constraints $\gamma \downarrow 1 - v - \gamma$ and $1 - v \uparrow 1 - v + \gamma$ require

$$\psi^\gamma \geq \psi_{\uparrow 1+v-\gamma}^{\gamma,L} := (1 - v - \gamma) \frac{\mu - \mu_1}{2\gamma - 2(1 - v - \gamma)\bar{\mu}^L}.$$

Constraints $\gamma \downarrow 1 - v - \gamma$ and $1 - v \uparrow 2\gamma$ require

$$\psi^\gamma \geq \psi_{\uparrow 2\gamma}^{\gamma,L} := \frac{\mu\gamma - (1 - v)/2}{2\gamma - (1 - v)}.$$

And we have the inequality

$$\psi_{\uparrow 1+v-\gamma}^{\gamma,L} \geq \psi_{\uparrow 2\gamma}^{\gamma,L} \iff \mu \leq \bar{\mu}^L + \tilde{\mu}^L$$

where $\tilde{\mu}^L$ has the rather complicated expression

$$\tilde{\mu}^L := \frac{[(1 - v)/2 - \bar{\mu}^L \cdot \gamma](2\gamma - 2(1 - v - \gamma)\bar{\mu}^L)}{(1 - v)(1 - v - \gamma) + 2\gamma[(2 + \bar{\mu}^L)\gamma - (1 + \bar{\mu}^L)(1 - v)]}.$$

Therefore, at the solution to the relaxed problem we have

$$\psi^\gamma = \begin{cases} 0 & \text{if } \mu \leq \bar{\mu}^L \\ (1 - v - \gamma) \frac{\mu - \bar{\mu}^L}{2\gamma - 2(1 - v - \gamma)\bar{\mu}^L} & \text{if } \bar{\mu}^L < \mu \leq \bar{\mu}^L + \tilde{\mu}^L \\ \frac{\mu\gamma - (1 - v)/2}{2\gamma - (1 - v)} & \text{if } \mu > \bar{\mu}^L + \tilde{\mu}^L \end{cases}$$

and $1 - v - 2\psi^\gamma(1 - v - \gamma)$ gives an upper bound for PS^* . We now show this is tight which will conclude the proof since this upper bound agrees with Proposition 4 (i). For $\psi_{L1}(1 - v) = \frac{1-\mu}{2}$ mass of type $L1$, we let firm 2 charge them $1 - v$:

$$\psi_{L1}(1 - v, 1 - v) = \psi_{L1}(1 - v).$$

The downward deviation from price γ to price $1 - v - \gamma$ is handled in the relaxed problem by $\gamma \downarrow 1 - v - \gamma$. For the $\psi_{H1}(1 - v) = \frac{\mu}{2} - \psi^\gamma$ mass of type $H1$, $\gamma \downarrow 1 - v - \gamma$ pins down the mass which is required to be assigned to the price γ by firm 2, with the remainder assigned the price $1 - v$. **G-AIC-L** then guarantees that downward deviation to price $1 - v - \gamma$ from price $1 - v$ is deterred. Hence $PS^* = R$.

Part (ii): $\gamma < \frac{1-v}{2}$. Same as Case C in the proof of Proposition 3.

□

C: PROOFS FOR SECTION 6 (MATCHING AND INFORMATION DESIGN)

C.1 Proof of Theorem 3. *Part (i).* Let all firms access only to those consumers who value their product the most (i.e., the consideration set of consumers in E_i comprises only firm i), and give firms full information about these consumers' valuations. There is then an equilibrium where each firm i sells to all consumers in E_i at a price equals

to their respective valuations for product i . The outcome of this equilibrium is the producer optimal point.

Part (ii). We characterize the consumer-optimal outcome for an arbitrary matching design in two steps. In the first step, for a given matching design, we define the *new valuations* of each consumer type θ by setting to zero the valuation for product i whenever firm i is not in the consideration set for consumer type θ . This defines a modified economy where consumers have the *new valuations* for products and consumers' consideration sets are unrestricted. In the second step we apply Theorem 1. This gives the consumer-optimal outcome of this modified economy. Lemma 4 below shows that this corresponds to the consumer-optimal outcome for the initial economy with restricted consideration sets. We now develop these arguments formally.

Step 1: Modify the distribution of valuations. A matching scheme $\phi \in \Delta(\Theta \times 2^{\mathcal{N}})$ maps consumer types into a probability distribution over consideration sets. For an initial consumer type θ let $S \in 2^{\mathcal{N}}$ be her realized consideration set. We map this consumer type and consideration set pair, (θ, S) , into a new consumer type and the unrestricted consideration set pair (θ^S, \mathcal{N}) , where $\theta^S = (\theta_i^S)_{i \in \mathcal{N}}$ is such that

$$\theta_i^S := \begin{cases} \theta_i & \text{if } i \in S \\ 0 & \text{otherwise.} \end{cases}$$

Given the initial distribution of types, a matching scheme, and doing this mapping for all consumer type-consideration set pairs induces a new ex-ante distribution μ_ϕ over consumer types along with the matching such that all consumers have all firms in their consideration set.

Step 2: Apply the consumer-optimal structure to the modified distribution.

For a given realization of type $\theta \in \Theta$ and consideration set $S \in 2^{\mathcal{N}}$, we treat the consumer's type as if it were θ^S . We then assign messages to each firm as if the underlying type were θ^S . In particular, we choose

$$\psi \in \Delta(\Theta \times [0, 1]^n)$$

as the consumer-optimal information structure (Theorem 1) as applied to the modified distribution μ_ϕ . Finally, define λ_ϕ^* as follows:

$$\lambda_\phi^*(M \times \{S\} | \theta) := \psi(M | \theta^S) \phi(S | \theta) \quad \text{for all } M \in \mathcal{B}([0, 1]^n), S \in 2^{\mathcal{N}}$$

which specifies, for a given realization of type θ , a joint distribution over messages and consideration sets. Consistency follows immediately from construction. Furthermore, $\lambda_\phi^* \in \Lambda_\phi$. The next lemma tells us that λ^* is indeed consumer-optimal among this class of designs.

Lemma 4. The design λ_ϕ^* implements an equilibrium which obtains the highest consumer surplus and the lowest producer surplus across all equilibria that can be implemented by some design in Λ_ϕ , i.e., λ_ϕ^* implements the consumer-optimal outcome among Λ_ϕ . Furthermore, this outcome is efficient given the matching constraints ϕ .

Proof of Lemma 4. We apply the consumer-optimal information structure to the modified measure μ_ϕ and, therefore, the equilibrium outcome is efficient given μ_ϕ : for each

pair (θ, S) , the new type θ^S purchases from her favourite firm among the set of firms S , i.e., θ^S buys from firm $j \in \arg \max \theta_i^S = \arg \max_{i \in S} \theta_i$. Hence, the design implements an equilibrium that extracts all gains from trade given ϕ . These are:

$$TS^\phi := \int_{\theta \in \Theta} \left(\max_{i=1,2,\dots,n} \theta_i \right) \mu_\phi(d\theta)$$

where we integrate against the modified measure μ_ϕ (which captures the fact that under the realization (θ, S) , the consumer only has positive valuation for firms in the set S).

Further, fixing the matching scheme ϕ , notice that the minimum profits firm i can make across any information structure is

$$\underline{\Pi}_i^\phi := \sup_{p \in [0,1]} p \cdot \int \mathbf{1}(\theta_i - p \geq \max_{j \neq i} \theta_j) \mu_\phi(d\theta)$$

which is the profit that firm i makes when all other firms charge a price of zero and firm i chooses an optimal uniform price against the residual demand curve.

But since (i) λ_ϕ^* was constructed such that firm i 's profits are held down to $\underline{\Pi}_i^\phi$; and (ii) the allocation is efficient given ϕ , this must imply that consumer surplus

$$CS^\phi = TS^\phi - \sum_{i=1}^n \underline{\Pi}_i^\phi$$

is optimal. Further note that point (i) implies the consumer-optimal outcome leads to the lowest possible producer surplus across all equilibria that can be sustained by some design in Λ_ϕ . \square

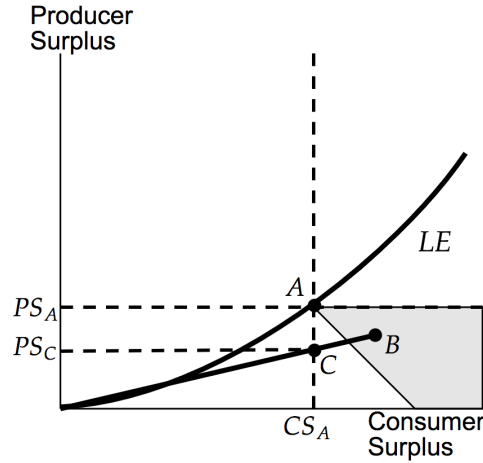
Fix a point $A := (CS_A, PS_A)$ on the lower envelope, and let the associated design implementing A be λ_A . Further let ϕ_A be the matching scheme associated with λ_A . We claim that $\lambda_{\phi_A}^*$ can also implement point A .

Suppose, towards a contradiction, that it did not. Denote by $B = (CS_B, PS_B)$ the consumer-optimal outcome that can be implemented by $\lambda_{\phi_A}^*$. From Lemma 4 we know that $CS_B \geq CS_A$ and $PS_B \leq PS_A$. In fact, it must be the case that $CS_B > CS_A$ because, if $CS_B = CS_A$ then either $PS_A = PS_B$, which contradicts that $\lambda_{\phi_A}^*$ cannot implement A or $PS_B < PS_A$, which contradicts that A is in the lower envelope of SUR . The relation between point B and point A is illustrated in Figure 7 below, where the grey area indicates the area where point B can be located.³⁸

Note next that in the producer-consumer surplus space, we can implement any convex combination of point B and the no-trade point NT by using designs which are convex combinations of $\lambda_{\phi_A}^*$ and the no-trade design (the consideration set of each consumer's type is empty). Since $CS_B > CS_A$ and $PS_B \leq PS_A$ there exists a convex combination of points B and NT , which we denote by $C = (CS_C, PS_C)$, such that $CS_C = CS_A$ and $PS_C < PS_A$ (see Figure 7 for graphical illustration). But this contradicts our assumption that point A belongs to the lower envelope of SUR .

³⁸Note that the diagonal line in the picture are all the points that produce the same total surplus; and since point B is efficient given ϕ it produces weakly higher total surplus.

Figure 7: Illustration of the proof of Theorem 3 (ii)



Part (iii). Define the map $\lambda \mapsto CS(\lambda) \in \mathbb{R}_{\geq 0}$ as the highest consumer surplus achieved in an equilibrium implemented by the design λ . Recall that, by definition, there exists a point in the lower envelope which delivers the maximum amount of consumer surplus across any design. In part (ii) we showed that every outcome in the lower envelope can be implemented by the design λ_ϕ^* for some ϕ which implies

$$\max_{\lambda \in \Lambda} CS(\lambda) = \max_{\psi \in \{\lambda_\phi^*\}_{\phi \in \Phi}} CS(\psi)$$

It remains to show that the consumer-optimal design associated with the full matching scheme implements maximum consumer surplus across $\{\psi_\phi^*\}_{\phi \in \Phi}$. We say that the matching scheme ϕ is efficient if for all $i \in \mathcal{N}$ and all $\theta \in E_i$,

$$\sum_{\substack{S \in 2^{\mathcal{N}}: \\ i \in S}} \phi(S|\theta) = 1$$

i.e., the consideration set of each consumer's type includes her favourite firm with probability one. Note that this is a necessary but not sufficient condition for the design to implement an efficient equilibrium. Denote the set of efficient matching schemes with $\Phi^E \subset \Phi$. The following lemma shows that the solution to the consumer surplus maximization problem lies within Φ^E .

Lemma 5. For each $\phi \in \Phi \setminus \Phi^E$, there exists $\phi' \in \Phi^E$ such that $\lambda_{\phi'}^*$ implements an equilibrium outcome with higher consumer surplus than λ_ϕ^* . That is:

$$\max_{\lambda \in \{\lambda_\phi^*\}_{\phi \in \Phi^E}} CS(\lambda) \geq \max_{\lambda \in \{\lambda_\phi^*\}_{\phi \in \Phi \setminus \Phi^E}} CS(\lambda).$$

Proof of Lemma 5. Since ϕ is inefficient, there exists some firm i and some positive measure of types in E_i who, with strictly positive probability, do not have firm i in their consideration set under ϕ . If ϕ delivers zero consumer surplus, then the same can be achieved with an efficient matching scheme and we are done. If ϕ delivers a positive amount of consumer surplus, then since the number of firms is finite, this implies that there exists some firm $j \neq i$ such that there is a positive measure of types within E_i

which with strictly positive probability, (i) do not have i in their consideration set; (ii) have j in their consideration set; and (iii) prefer j to all other firms in their consideration set.

Denote the type-consideration set pairs which fulfil this condition with

$$T_{ij} := \left\{ (\boldsymbol{\theta}, S) \in \Theta \times 2^{\mathcal{N}} : \boldsymbol{\theta} \in E_i, j \in S, i \notin S, \theta_j > \max_{k \in S \setminus \{i, j\}} \theta_k \right\},$$

observing that, by the argument above,

$$\int_{\boldsymbol{\theta} \in \Theta} \sum_{S: (\boldsymbol{\theta}, S) \in T_{ij}} \phi(S|\boldsymbol{\theta}) d\mu > 0.$$

We will proceed by showing that suitably modifying the access scheme to give firm i access to types $\boldsymbol{\theta} : (\boldsymbol{\theta}, S) \in T_{ij}$ strictly improves consumer welfare.

Denote

$$F_i := \left\{ (\boldsymbol{\theta}, S) \in \Theta \times 2^{\mathcal{N}} : i \in S, \theta_i > \max_{k \in S \setminus \{i\}} \theta_k \right\}$$

as the type-consideration pairs where the consumer type strictly prefers product i to other products in her consideration set. Note that $T_{ij} \cap F_i = \emptyset$.

In the consumer-optimal outcome implemented by λ_ϕ^* firm i 's profit is the same as its no-information profit under the distribution μ_ϕ which is

$$\pi_i(\lambda_\phi^*) := \max_{p \in [0,1]} p \int \sum_{S: (\boldsymbol{\theta}, S) \in F_i} \mathbf{1}(\theta_i - p \geq \max_{k \in S \setminus \{i\}} \theta_k) \phi(S|\boldsymbol{\theta}) d\mu.$$

Let us now modify the inefficient access scheme ϕ as follows: whenever $(\boldsymbol{\theta}, S) \in T_{ij}$ realizes, implement instead the consideration set $S \cup \{i\}$ and denote the resultant access scheme with ϕ' , i.e. on each $\boldsymbol{\theta} : (\boldsymbol{\theta}, S) \in T_{ij}$, we have

$$\phi'(S \cup \{i\}|\boldsymbol{\theta}) = \phi(S|\boldsymbol{\theta}).$$

We will compare profits under λ_ϕ^* and under $\lambda_{\phi'}^*$. A few observations follow. First, notice that on the realizations of $(\boldsymbol{\theta}, S) \in T_{ij}$ under the design λ_ϕ^* those consumers bought from j . However, under the design $\lambda_{\phi'}^*$ they now buy from i . This implies j 's profits must weakly decrease i.e., $\pi_j(\lambda_{\phi'}^*) \leq \pi_j(\lambda_\phi^*)$. Second, observe that under the design $\lambda_{\phi'}^*$, i 's profits are now

$$\begin{aligned} \pi_i(\lambda_{\phi'}^*) &:= \max_{p \in [0,1]} p \int_{\boldsymbol{\theta} \in \Theta} \sum_{S: (\boldsymbol{\theta}, S) \in \{F_i \cup T_{ij}\}} \mathbf{1}(\theta_i - p \geq \max_{k \in S \setminus \{i\}} \theta_k) \phi(S|\boldsymbol{\theta}) d\mu \\ &\leq \pi_i(\lambda_\phi^*) + \max_{p \in [0,1]} p \int_{\boldsymbol{\theta} \in \Theta} \sum_{S: (\boldsymbol{\theta}, S) \in T_{ij}} \mathbf{1}(\theta_i - p \geq \theta_j) \phi(S|\boldsymbol{\theta}) d\mu, \end{aligned}$$

where the inequality comes from applying the max operator to each term in the summand separately. We can then bind the improvement to i 's profits by the change in

total surplus as follows:

$$\begin{aligned}
\pi_i(\lambda_{\phi'}^*) - \pi_i(\lambda_{\phi}^*) &\leq \max_{p \in [0,1]} p \int_{\boldsymbol{\theta} \in \Theta} \sum_{S: (\boldsymbol{\theta}, S) \in T_{ij}} \mathbf{1}(\theta_i - p \geq \max_{k \in S \setminus \{i\}} \theta_k) \phi(S|\boldsymbol{\theta}) d\mu \\
&= \int_{\boldsymbol{\theta} \in \Theta} \sum_{S: (\boldsymbol{\theta}, S) \in T_{ij}} p^* \mathbf{1}(\theta_i - \theta_j \geq p^*) \phi(S|\boldsymbol{\theta}) d\mu \\
&\leq \int_{\boldsymbol{\theta} \in \Theta} \sum_{S: (\boldsymbol{\theta}, S) \in T_{ij}} (\theta_i - \theta_j) \phi(S|\boldsymbol{\theta}) d\mu \\
&= TS(\lambda_{\phi'}^*) - TS(\lambda_{\phi}^*),
\end{aligned}$$

where we use p^* to denote the solution of the maximization problem in the first equality and the second inequality follows from noting that $p^* \mathbf{1}(\theta_i - \theta_j \geq p^*) \leq \theta_i - \theta_j$. The last equality follows by observing that the increase in gains from trade obtained with the new design corresponds to the increase in consumption value obtained by the consumers who now purchase from i instead of j .

Denote $PS(\lambda_{\phi}^*)$ as the producer surplus implemented by λ_{ϕ}^* . We have:

$$\begin{aligned}
PS(\lambda_{\phi'}^*) - PS(\lambda_{\phi}^*) &= \sum_{k \in \mathcal{N}} \left(\pi_k(\lambda_{\phi'}^*) - \pi_k(\lambda_{\phi}^*) \right) \\
&= \left(\pi_i(\lambda_{\phi'}^*) - \pi_i(\lambda_{\phi}^*) \right) + \left(\pi_j(\lambda_{\phi'}^*) - \pi_j(\lambda_{\phi}^*) \right) \\
&\leq TS(\lambda_{\phi'}^*) - TS(\lambda_{\phi}^*),
\end{aligned}$$

where the inequality follows from the argument above that j 's profits decrease and i 's profits increase by less than the change in total surplus.

Now denoting $CS(\lambda_{\phi}^*)$ as total consumer surplus under λ_{ϕ}^* , we similarly have

$$\begin{aligned}
CS(\lambda_{\phi'}^*) - CS(\lambda_{\phi}^*) &= \left(TS(\lambda_{\phi'}^*) - PS(\lambda_{\phi'}^*) \right) - \left(TS(\lambda_{\phi}^*) - PS(\lambda_{\phi}^*) \right) \\
&= \left(TS(\lambda_{\phi'}^*) - TS(\lambda_{\phi}^*) \right) - \left(PS(\lambda_{\phi'}^*) - PS(\lambda_{\phi}^*) \right) \geq 0.
\end{aligned}$$

Since there are a finite number of firms, we can repeat this procedure a finite number of times until the final matching scheme is efficient. \square

We now show that among efficient matching schemes, the full matching scheme implements the maximum consumer surplus. To see this, observe that total surplus is the same across all efficient matching schemes since for all $\phi \in \Phi^E$,

$$TS^{\phi} = \int_{\boldsymbol{\theta} \in \Theta} \sum_{S \in 2^{\mathcal{N}}} \max_{j \in S} \theta_j \phi(S|\boldsymbol{\theta}) d\mu = \int_{\boldsymbol{\theta} \in \Theta} \max_{j \in \mathcal{N}} \theta_j d\mu.$$

On the other hand, from Theorem 1 and Lemma 4, the expected profits of firm i under the design λ_{ϕ}^* is

$$\underline{\Pi}^{\phi} = \max_{p \in [0,1]} p \cdot \int_{\boldsymbol{\theta} \in E_i} \sum_{\substack{S \in 2^{\mathcal{N}}: \\ i \in S}} \mathbf{1}(\theta_i - p \geq \max_{j \in S \setminus \{i\}} \theta_j) \phi(S|\boldsymbol{\theta}) d\mu.$$

Observe that pointwise (fixing $\theta \in E_i$), we have that

$$\sum_{\substack{S \in 2^{\mathcal{N}}: \\ i \in S}} \mathbf{1}(\theta_i - p \geq \max_{j \in S \setminus \{i\}} \theta_j) \phi(S | \theta) \geq \mathbf{1}(\theta_i - p \geq \max_{j \in \mathcal{N} \setminus \{i\}} \theta_j),$$

since restricting the consumer's consideration set decreases competition, i.e., the max operator is increasing in the set order, and efficient matching scheme λ has firm i in type θ 's consideration set with full probability. Hence, the design λ_ϕ^* corresponding to the full matching scheme minimizes firm i 's profits and thus total producer surplus. Putting everything together, we have

$$CS^* := \max_{\lambda \in \Lambda} CS(\lambda) = \max_{\lambda \in \{\lambda_\phi^*\}_{\phi \in \Phi}} CS(\lambda) = \max_{\lambda \in \{\lambda_\phi^*\}_{\phi \in \Phi^E}} CS(\lambda) = CS(\lambda_{\phi^F}^*)$$

where we use $\phi^F \in \Phi^E$ to denote the full matching scheme.

Part (iv). Part (i) of Theorem 3 shows that the producer optimal point $(0, TS)$ belongs to the set SUR , Part (iii) shows that the consumer optimal point is in the efficient frontier and part of the lower envelope of SUR , Part (ii) characterizes the lower-envelope. Note that SUR is convex: A $(\alpha, 1 - \alpha)$ convex combination of two implementable welfare outcomes can be implemented by using the corresponding designs with respective probabilities α and $1 - \alpha$. Hence, SUR is the convex hull generated by the producer optimal point and the lower envelope (LE).