# Market Segmentation through Information*

Matthew Elliott
Cambridge

Andrea Galeotti
London Business School

Andrew Koh
MIT

Wenhao Li
European University Institute

This version: February 23, 2025

**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. Our analysis demonstrates the power that users' data can endow intermediaries with, and speaks directly to current regulatory debates of digital marketplaces.

*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 $(\theta_A, \theta_B)$. 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 $(1, 1/2)$

[1] While 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.

[2] For 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 $p_A(\theta_A, \theta_B) = \max\{\theta_B - \theta_A, 0\}$. For this distribution of consumer values 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.[3]

**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 de-

[3]A recent paper by Bergemann, Brooks, and Morris (2023) builds on our analysis of the consumer-optimal design by extending our construction to incorporate the case where there is uncertainty about costs.

signer 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 inframarginal 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 possibly 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 that for small violations of aggregate incentive compatibility, the producer-optimal efficient outcome is obtained by applying the producer-perfect design to a suitably transformed distribution of consumer values rather than the actual distribution. In the transformed distribution some consumers have their valuations for all products uniformly reduced. This leaves these consumers with consumer surplus equal to the amount by which their values have been depressed, while the remaining surplus is extracted as producer surplus (Proposition 3).

A feature of the producer-perfect information design when it is implementable on the actual distribution is that firms receive a payoff of zero whenever they set a higher price than was recommended. However, when this design is applied to the transformed distribution, upward deviations can matter once enough consumers have had their values sufficiently reduced. This is the point at which producer optimal-design no longer coincides with implementing a producer-perfect design on the transformed distribution. We fully characterize what such designs look like within 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 4). 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 when there does not exist a producer perfect information design granting the intermediary the additional power to restrict which consumers can access which products[4] 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 in the way in which information is used to soften competition compared to intensifying it. 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 like Google's proposed "Privacy Sand Box." Hence, enhancing users' privacy need be no impediment to extracting consumer surplus—to the contrary, it may facilitate it.[5]

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

[5] It 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).

**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.[6]

Two papers that also consider how information shapes market competition are Bergemann, Brooks, and Morris (2017) and Bergemann, Brooks, and Morris (2021a). 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 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. (2021a) 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 (2021a) is a special case of our model in which product differentiation is derived via search frictions.[7]

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.)[8] Similarly, in Bergemann, Brooks, and Morris (2021a) 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

[6]A 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).

[7]The paper Bergemann et al. (2021b) studies whether market outcomes like market power and price volatility can be robustly inferred without knowing the information structure. They find it is not possible to predict market power without knowing the information structure, but one can robustly predict price volatility.

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

achieve by sharing the data with firms. Complementary to this, Ali, Lewis, and Vasserman (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) Montes et al. (2019), Jones and Tonetti (2020), Bounie, Dubus, and Waelbroeck (2020), Bergemann et al. (2021b)[9]; 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 $\mathcal{N} = \{1, \ldots, n\}$ each of which produces a single product at constant marginal cost, which we normalize to zero for each firm.[10] There is a continuum of consumers with unit mass each of whom demands a single unit inelastically.[11] Consumers have different valuations for different firms: type $\boldsymbol{\theta} \in \Theta = [0,1]^n$ has value $\theta_1$ for firm 1, $\theta_2$ for firm 2 and so on. The distribution of consumers over $\Theta$ is given by $\mu \in \Delta(\Theta)$.[12] Note that $\mu$ could have atoms and, therefore, this formulation nests both discrete and continuous types.

[9]This paper studies whether market outcomes like market power and price volatility can be robustly inferred without knowing the information structure. They find it is not possible to predict market power without knowing the information structure, but one can robustly predict price volatility.

[10]In Online Appendix III 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.

[11]All results translate into an alternate setting with a single consumer of uncertain type.

[12]We make the mild assumption that $\mu$ can be decomposed into a discrete measure and an absolutely continuous measure (which thus admits a density).

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

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

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

$$\psi \in \Delta\Big(\Theta \times \boldsymbol{M}\Big) \quad \text{such that} \quad \text{marg}_{\boldsymbol{M}} \psi = \mu$$

where $\boldsymbol{M} := \prod_i M_i := [0,1]^n$ is an $n-$dimensional message space. Let $\boldsymbol{\Psi}$ be the set of all information structures. Write $\psi(\cdot|m_i) \in \Delta(\Theta \times \boldsymbol{M}_{-i})$ 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 $\theta$.

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 $i \in \{1, \ldots, n\}$ is $p_i : M_i \to [0,1]$.[13] A mixed strategy for firm $i$ is $\sigma_i : M_i \to \Delta([0,1])$. Each consumer 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.[14]

**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 $\gamma_H$ 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 $\gamma_L$ if they buy product $j \neq i$.

[13]Given 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)).

[14]Bergemann 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 be helpful to reframe certain information design problems as matching problems (see also Dworczak 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.

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 $\gamma_H$ and $\gamma_L$ capture the level of horizontal differentiation in the market and can be interpreted as switching costs. The parameters $v$ and $\mu$ 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 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 $TS^* = \sum_{i=1}^n \int_{\boldsymbol{\theta} \in E_i} \theta_i d\mu$ 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 $i$ ($\theta_i$), 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 $\theta_i$.

We now construct the consumer-optimal information structure $\psi^*$ for the oligopoly case. Define the projection $\Lambda_i(\theta) = \theta_i - \max_{j \neq i} \theta_j$ which sends each type $\boldsymbol{\theta} \in E_i$ 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 $i$.[15] For each firm $i$ and each $\boldsymbol{\theta} \in E_i$,

$$\psi^*\Big((\cdot, \boldsymbol{m}_{-i} = \mathbf{0})|\boldsymbol{\theta}\Big) = \sigma^*_{\mu_i^*}\Big(\cdot\,|\Lambda_i(\boldsymbol{\theta})\Big).$$

In words, for each $\boldsymbol{\theta} \in E_i$, 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 ($\mu_i^*$); and (ii) all firms other than $i$ receive the message 0. We note that the design $\psi^*$ 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 $\psi^*$ is consumer-optimal.[16] 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 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 $\gamma_L$ and demand $\mu/2$ at price $\gamma_H$. 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 $\psi^*$. 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.

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

[16] We observe that, by construction, the producer-surplus implemented under $\psi^*$ is the lowest producer surplus that can be implemented in any equilibrium for any information structure.

Note that the mass of high-value consumers $x$ has been chosen so that firm 1 is indifferent between charging $\gamma_H$ and $\gamma_L$ upon receiving the price recommendation $\gamma_L$.[17] Under $\psi^*$, 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.[18]

## 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 $\theta$ pays $\max_i \theta_i$.

Condition P is 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 $\Gamma(\boldsymbol{\psi})$ denote the Bayesian pricing game induced by the information structure $\boldsymbol{\psi}$. Let $\boldsymbol{\Gamma}^*$ denote the set of induced games in which there exists an equilibrium satisfying condition P, and let $\boldsymbol{\Psi}^* := \{\boldsymbol{\psi} : \Gamma(\boldsymbol{\psi}) \in \boldsymbol{\Gamma}^*\}$ be the set of information structures that can be used to fulfil condition P. We refer to $\boldsymbol{\psi} \in \boldsymbol{\Psi}^*$ 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 $\boldsymbol{\Psi}^* \neq \emptyset$.

Suppose an information structure induces a producer-perfect outcome. Then consumers of type $\boldsymbol{\theta} \in E_i$ 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 a set a price above $\hat{p}_i$ to all consumers of types $\boldsymbol{\theta} \in E_i$ 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 $\boldsymbol{\theta'} \notin E_i$. 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 $\boldsymbol{\theta'} \notin E_i$ who value $i$'s product weakly

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

[18] The 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.

above $\hat{p}_i$. Thus a sufficient condition for no firm to ever want to deviate downwards like this is

$$\int_{\boldsymbol{\theta}\in E_i:\theta_i>\hat{p}_i}(\theta_i-\hat{p}_i)d\mu \geq \hat{p}_i\int_{\boldsymbol{\theta}\in\Theta\setminus E_i:\theta_i\geq\hat{p}_i}d\mu$$
$$\text{for all } \hat{p}_i<\sup\{\theta_i:\boldsymbol{\theta}\in E_i\} \text{ and all firms } i. \tag{AIC}$$

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 $\theta_i$ 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.[19]

**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 $\psi$ must agree with the actual distribution of consumers:

$$\text{marg}_{\Theta}\psi=\mu \tag{Consistency}$$

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 $\boldsymbol{\theta},\boldsymbol{\theta}'\in E_i$ with different values $\theta_i\neq\theta_i'$ for product $i$:

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

We have argued that each firm $i$ charges types $\boldsymbol{\theta}\in E_i$ a price of $\theta_i$. 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**).

[19] In Bergemann, Brooks, and Morris (2017), the analogous outcome to our producer-perfect outcome is an outcome in which the winning bidder pays 0 for the item. There is no information structure which implements this outcome.

We start with **Consumer IC**. Consider a consumer of type $\boldsymbol{\theta} \in E_j$ and suppose all firms follow the price recommendations. By **Separation**, firm $j$ charges $\theta_j$ 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 $\theta_j$, we need that $\theta_i$ is lower than $m_i$. This is what **Consumer IC** states:

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

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 $E_i$:

$$\underbrace{(m_i - \hat{p}_i) \int_{E_i \times \boldsymbol{M}_{-i}: \theta_i = m_i} d\psi(\cdot|m_i)}_{\text{Inframarginal losses}} \geq \underbrace{\hat{p}_i \int_{\Theta\setminus E_i \times \boldsymbol{M}_{-i}: \theta_i \geq \hat{p}_i} d\psi(\cdot|m_i)}_{\text{Business-stealing gains}} \qquad \textbf{(Firm IC)}$$

for all firms $i \in \mathcal{N}$, all $m_i \in M_i$[20] and all $\hat{p}_i < m_i$. Note that this implies that for almost all consumers, firm $i$ receives a message $m_i \in \{\theta_i : \boldsymbol{\theta} \in E_i\}$. Thus, our information design problem can be recast as a problem of matching types $\boldsymbol{\theta} \notin E_i$ to messages $\{\theta_i : \boldsymbol{\theta} \in E_i\}$ for each firm $i$.

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 $\psi$ which, for all firms $i \in \mathcal{N}$, satisfies **Separation**, **Consistency**, **Consumer IC** and **Firm IC**.

Let $\psi$ 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**:

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

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

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

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:

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

To find the overall capacity for matching consumer types $\boldsymbol{\theta} \notin E_i$ 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 $\psi$ restricted to types $E_i$ concentrates on the messages $[\hat{p}_i, 1]$ for firm $i$ (by **Separation**) so we can replace $\psi$ with $\mu$.

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.[21] The extreme case is one where firms offer homogeneous

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

products; that is, for each consumer type $\boldsymbol{\theta}$, $\theta_i = \theta_j$ for all products $i$ and $j$, but some consumer types may have higher valuation than others.[22]

**Proposition 1** (Product differentiation is necessary for producer profits)**.** Suppose firms offer homogeneous products. Then for any $\psi \in \boldsymbol{\Psi}$ and any equilibrium induced by $\psi$, 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 $\mu$, 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 $\mu$ whenever, for each firm $i$, (i) the distribution of valuations for firm $i$ among consumers in $E_i$ under $\hat{\mu}$ dominates that under $\mu$ such that for all $x$,

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

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

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

**Proposition 2** (Polarization aids segmentation)**.** If a producer-perfect information structure exists under $\mu$, and $\hat{\mu}$ is more polarized than $\mu$, 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.[23] 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.

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

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

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.[24]

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:

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

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

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

AIC-High and AIC-Low simplify to

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

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

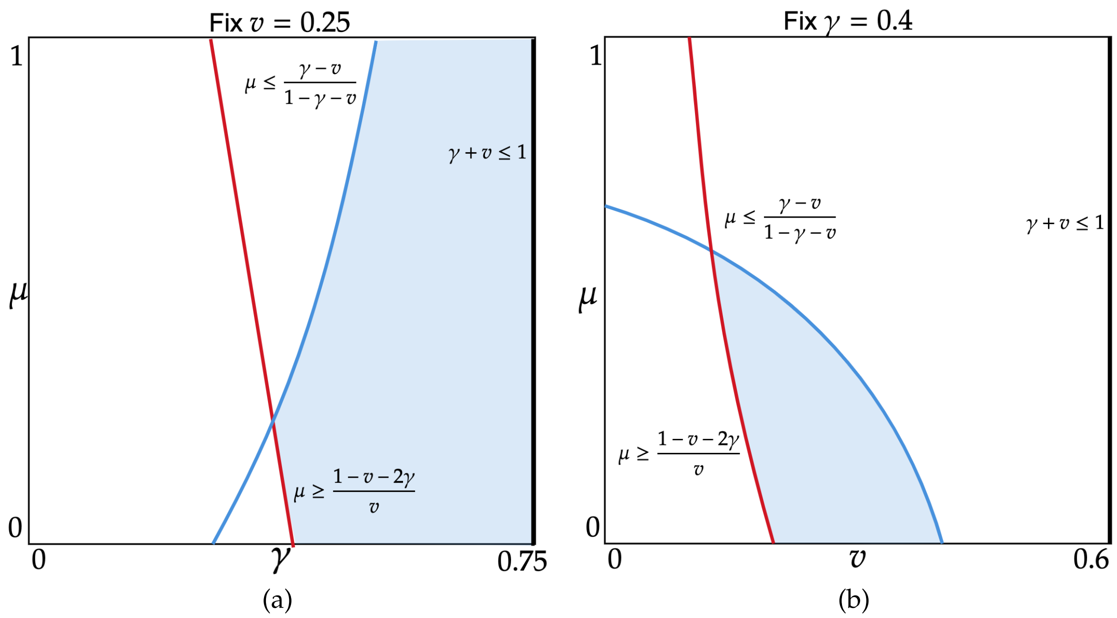
[24]This is similar to the argument that, in the presence of switching costs, firms with a smaller customer base can be a stronger competitive constraint on market behaviour than more established firms (Klemperer, 1995). Based in part on this logic the UK antitrust authorities prohibited the acquisition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

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 $\alpha$ of high types, and a fraction $\beta$ 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 $\alpha$ and $\beta$ so that it is profitable for firm 1 to follow these price recommendations. This requires that the fraction of high types $\mu$ takes an intermediate value (condition (ii)). For intermediate values of $\mu$, we can find $\alpha$ and $\beta$ 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: Parameters under which a producer-perfect outcome exists




(a) (b)

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 $\gamma$ makes consumers' preferences more polarized (as in Definition 1) and so it facilitates AIC (Proposition 2).[25] Further note that the effect of $v$ on AIC is ambiguous (Panel (b)): increasing $v$ decreases the inframarginal losses from low-types to deviate

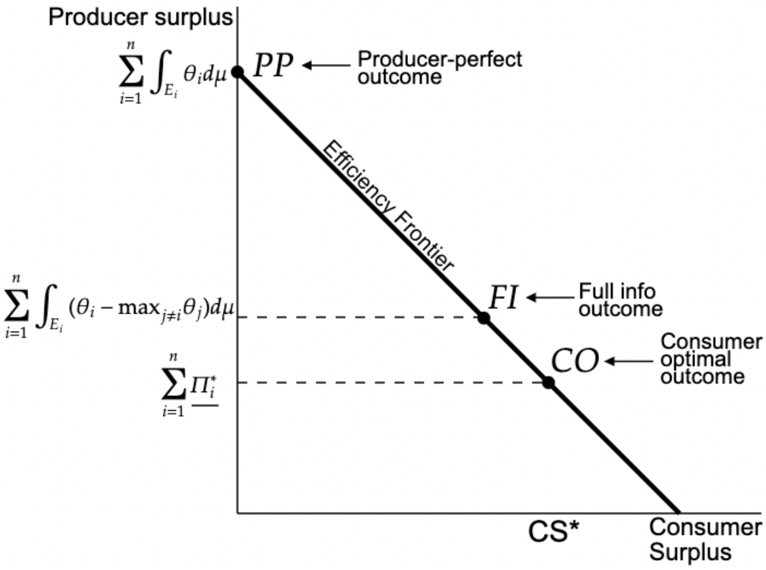
[25]For each firm $i$, an increase in $\gamma$ keeps the distribution of valuations for $i$ in $E_i$ unchanged, but decreases the valuations for $i$ in $E_j$

to $1-\gamma$ and so it tightens AIC-High, however it increases the inframarginal losses 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 V 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 $(\theta_1, \theta_2)$ are anti-correlated, i.e., $\theta_2 = 1-\theta_1$ and $\theta_1$ is uniformly distributed over the unit interval. Also in this case the AIC holds. When, instead, $\theta_1$ is drawn from a truncated (at 0 and 1) normal distribution with mean $1/2$ and variance $\sigma^2$, 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.

Figure 2: Efficient Information Structures under AIC




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 $\mu$ into $\mu_{PP} = \lambda \cdot \mu$ and $\mu_{CO} = (1-\lambda)\cdot\mu$ and then apply the producer-perfect information design to $\mu_{PP}$ and the consumer-optimal design to $\mu_{CO}$. Since AIC condition holds for $\mu$ it also holds for $\mu_{PP}$ because this is simply a renormalization of measure.

## 5 Efficient frontier when Aggregate Incentive Compatibility fails

We now consider implementable efficient outcomes when AIC is violated. We first derive a general upper-bound on the maximum amount of producer surplus which can be extracted across any implementable efficient outcome (subsection 5.1); we then show that for small deviations to AIC, this upper bound is tight and, therefore, is producer-optimal. Our derivation applies the logic of AIC to a transported distribution of valuations.

We then identify a general tension between simultaneously deterring business stealing (downward deviations) and deterring firms from exploiting their ensuing local market power (upward deviations). It is only the need to deter upward deviations that drives a wedge between our upper-bound and the maximum extractable producer surplus. To understand this better, in subsection 5.2 we characterize the producer-optimal design within our switching cost model.

**5.1 Producer-optimal information design for small violations of AIC.** We will suppose, for simplicity, that there are $n = 2$ firms. Our starting point is that $\mu$ violates AIC so that, per Theorem 2, the producer-perfect outcome is not implementable. Let $PS^*$ denote the highest possible producer surplus which can be implemented in an efficient outcome. Our key observation is that for small violations of AIC the outcome $PS^*$ is obtained by replicating the information design described for the producer-perfect outcome but for a modified distribution of types. In this modified distribution, some consumers are treated as if they have a lower value for all products than they do. As a result they are now charged a lower price but continue to buy the same product. This means that such consumers are, in effect, pooled together with other consumers who have a lower willingness to pay.

We now introduce the relevant definitions to formalize and state this result.

**Definition 2.** For distributions $\mu, \mu' \in \Delta(\Theta)$, say that $\mu'$ is *diagonally dominated* by $\mu'$ (denoted $\mu' \preceq_D \mu$) if:

$$\mathsf{DIAG}^i_D \mu' \leq_{FOSD} \mathsf{DIAG}^i_D \mu \quad \text{for } D \in [0,1], i \in \{1,2\} \qquad \text{(Diagonal dominance)}$$

and

$$\int_{\substack{\theta \in \Theta: \\ \theta_i - \theta_j \subseteq \mathcal{D}}} (d\mu - d\mu') = 0 \quad \text{for any } \mathcal{D} \in \mathcal{B}([0,1]), i \in \{0,1\} \qquad \text{(Measure preservation)}$$

where $\mathsf{DIAG}^i_D : \Delta(\Theta) \to \Delta([0,1])$ is the diagonal conditional operator

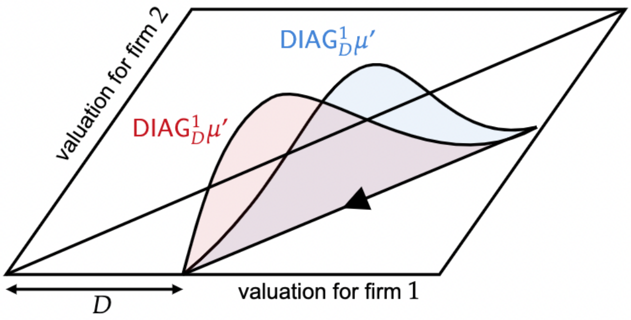
$$\mathsf{DIAG}^i_D \mu := \mathbb{P}_\mu\Big( \cdot \,\Big|\theta_i - \theta_{-i} = D \Big) \in \Delta([0,1])$$

which gives the marginal distribution of valuations for firm $i$'s product along the slice of the type space where the difference of valuation is $D$, and $\leq_{FOSD}$ is the first-order stochastic dominance relation.

Figure 3 illustrates diagonal dominance. Here we pick firm 1 and $D \in [0,1]$ and compare the conditional distribution along the diagonal. The red curve denotes the distribution $\mathsf{DIAG}^1_D\mu'$ which is first-order stochastically dominated by that of $\mathsf{DIAG}^1_D\mu$ (blue) and furthermore, the measure of agents over this diagonal is the same across $\mu$ and $\mu'$.[26]

Diagonal dominance requires that this relation holds for all $D \in [0,1]$ and all $i \in \{1,2\}$.

Figure 3: Diagonal dominance




For simplicity, we will suppose that types are supported on the $K-$grid:

$$\text{supp}\mu = \Theta^K := \left\{\frac{1}{K}, \frac{2}{K}, \dots \frac{K}{K}\right\}^n,$$

where we emphasize that $K$ can be taken large so the grid is arbitrarily fine. Now consider the following program:

$$\sup_{\mu' \in \Delta(\Theta): \mu' \preceq_D \mu} \int_{\theta \in \Theta} \max_i \theta_i d\mu' \tag{D}$$
$$\text{s.t. } \mu' \text{ satisfies AIC.}$$
$$\text{supp}\mu' \subseteq \Theta^K.$$

The problem (D) returns the fictional distribution $\mu'$ with the highest total surplus subject to the constraint that (i) $\mu'$ is diagonally dominated by $\mu$; (ii) $\mu'$ satisfies AIC; and (iii) $\mu'$ is supported on $\Theta^K$. Observe that if the original distribution $\mu$ fulfills AIC, then it is also the solution to (D). Thus, the interpretation of the solution to (D) is that it specifies, for each firm $i$ and diagonal $D$, the marginal distribution over willingness to pay (specified by $\mathsf{DIAG}^i_D\mu$) and equilibrium prices charged (specified by $\mathsf{DIAG}^i_D\mu'$) so that information can be designed to deter business stealing via downward deviations.[27]

**Proposition 3.** Let $PS^*$ be the highest producer surplus achievable via an information design which induces an efficient equilibrium and let $\mu^*$ be a solution to (D).

[26] This has bite if the distribution along the diagonal under $\mu$ is positive measure. The definition of diagonal dominance also handles cases where $\mu$ is atomless.

[27] The additional restriction that $\text{supp}\mu' \subseteq \Theta^K$ ensures that sale prices are from the grid $\{1/K, 2/K, \dots 1\}$ which, as we will show, is a necessary condition for efficiency across any information design.

(i) The problem (D) is an upper bound on the highest producer which can be extracted conditional on efficiency:

$$(D) \geq PS^*.$$

(ii) The bound is tight for small violations of AIC: there exists $\epsilon > 0$ such that

$$\|\mu^* - \mu\|_\infty < \epsilon \implies (D) = PS^*.$$

Moreover, whenever the upper-bound is achievable, an information design which achieves it applies a producer-perfect information design to a distribution which solves (D).

The basic idea underlying Proposition 3 is to take the types that cause AIC to be violated and to implement a producer perfect information design *as if* their valuations for all products had been reduced equally. This transformation turns the original distribution $\mu$ into a fictional distribution $\mu' \preceq_D \mu$ so under the producer-perfect design applied to $\mu'$, some consumers are charged a price below their maximum willingness to pay. This, in turn, makes it less tempting for other firms to try and steal them away since the requisite price cut is steeper—and hence less profitable. Indeed, for large enough transformations, consumers' valuations for their less preferred products is zero—this corresponds to implementing the full information design for consumers' actual values—which guarantees the existence of a distribution $\mu' \preceq_D \mu$ under which AIC holds.

The reason that (D) is an upper bound follows from a simple sequence of observations. First, it is without loss of generality to focus on obedient information designs in which firms are recommended prices and find it optimal to follow these recommendations. Moreover, we show in the appendix that efficiency requires that such designs recommend prices in the set $\{1/K, 2/K, \ldots 1\}$. Finally, we have seen that AIC is a necessary condition for obedience since it simply sums over individual downward deviation constraints. Since these are all necessary conditions for obedience, the program must be an upper bound.

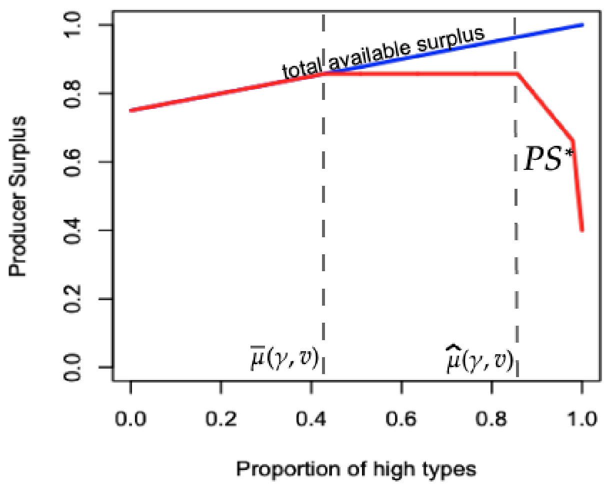
The reason why the upper bound (D) is not always a tight is that while in a producer perfect information design applied to the true distrbution of types a firm is never remotely tempted to set a price higher than recommended (as doing so will violate the consumer's IC and yield zero sales), this is no longer true under the transformed distribution $\mu' \preceq_D \mu$. As firms are now charging prices to consumers below their actual values, upward pricing deviations come into play. It is when (and only when) these upward deviation incentive compatibility constraints would be violated that (D) is not achievable.

For small violations of AIC, however, the upper-bound (D) is always achievable—we show this by constructing an information design in which consumer types who prefer firm $i$ are pooled with one another as follows: the distribution over prices along each diagonal ($\mathsf{DIAG}^i_D \mu^*$) can be matched to the distribution over types ($\mathsf{DIAG}^i_D \mu$) in an assortative fashion so that upward deviations remain slack. Furthermore, we match prices recommended to firm $j$ to prices recommended to firm $i$ in the same way as we did under the producer-perfect design, but now applied to the transformed distribution. Thus, the way through which market segmentation softens price competition continues to be—as in the case of producer-perfection—to deter business stealing downward deviations.

For large violations of AIC, it can be impossible to guarantee that all upward deviations remain slack—the producer-optimal information design must then perform a delicate balancing act of simultaneously fulfilling both upward- and downward-deviations. To understand what such designs look like, we now develop a full characterization of the efficient frontier within the switching-cost environment.

**5.2 Producer-optimal information design in switching cost environment.** A complete analysis of the efficient frontier within the switching-cost environment is developed in Online Appendix I. Here, we focus on Figure 4 that illustrates the maximum amount of producer surplus that can be obtained across any efficient outcome as a function of the proportion of high-types for $v = 0.25$ and $\gamma_H = \gamma_L = \gamma = 0.4$. The discussion makes a methodological point by illustrating the essential role that upward deviations play in constraining the amount of producer surplus which can be extracted.

Figure 4: Optimal producer surplus as distribution varies, $\gamma_H = \gamma_L = 0.4, v = 0.25$




When the proportion of high-types is smaller than threshold $\overline{\mu}(\gamma, v)$ we can implement the producer-perfect outcome (AIC holds.) At the boundary case $\mu = \overline{\mu}(\gamma, v)$ AIC just holds and, in particular, AIC-High holds with equality:

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

If we now move to an economy with a few more high-value types (i.e., we increase $\mu$ above $\overline{\mu}(\gamma, v)$), condition AIC-High fails and the producer-perfect outcome is not implementable. Proposition 3 tells us that if $\mu$ is not much larger than $\overline{\mu}(\gamma, v)$ the producer optimal outcome is obtained by taking the producer-perfect information design and assigning some high-value consumers to the low-value message: the measure $(\mu - \overline{\mu}(\gamma, v))/2$ of high-value consumers for firm $i$ are assigned to the message $1 - v$ for firm $i$. Hence, the producer optimal outcome creates a producer surplus of $1 - v(1 - \overline{\mu}(\gamma, v))$, and the high-type consumers who are treated as low-types enjoy information rent $v(\mu - \overline{\mu}(\gamma, v))$.

The construction implies that there are no incentives for firms to deviate downwards. However, whereas checking downward deviations was both necessary and sufficient to implement a perfect-optimal outcome, this is no longer the case when the measure of high types is sufficiently large. Under $\mu > \overline{\mu}(\gamma, v)$ 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 $(\mu - \overline{\mu}(\gamma, v))/2$. 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.[28] Thus, the segmentation must ensure that firms do not want to deviate upwards to exploit their newfound local market power—we must deter deviations to the price 1 when recommended to charge the price $1 - v$.[29] This requires that

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

which binds when the mass of high types increases to $\mu = \hat{\mu}(\gamma, v)$.

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 $\mu$ increases as is reflected in the decreasing portions of the red line in Figure 4. 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.

The characterization in the Online Appendix I formalizes these insights. Our results characterizing producer-optimal information structures when AIC is violated relate to the elegant paper of Bergemann, Brooks, and Morris (2021a) who study producer-optimal information in search markets in which there is uncertainty about how many firms the consumer has access to.[30] 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

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

[29] Note that if both firms charge price 1, this is not an equilibrium since one firm would have incentive to deviate downwards. Indeed, the key tension is that to prevent downward deviations we must induce firms to offer information rents; but doing so introduces the incentives to deviate upwards.

[30] 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. e.g., Armstrong and Vickers (2019).

vertical heterogeneity $v$—and products might be more or less differentiated—as captured by the degree of horizontal differentiation $\gamma$. This yields different and complementary insights relative to the search environment of Bergemann, Brooks, and Morris (2021a). 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.[31] 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

$$\lambda \in \Delta\left(\Theta \times \boldsymbol{M} \times 2^{\mathcal{N}}\right)$$

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

For a given type $\boldsymbol{\theta} \in [0,1]^n$, the design $\lambda$ induces a measure $\lambda(\cdot|\boldsymbol{\theta}) \in \Delta([0,1]^n \times 2^{\mathcal{N}})$ over both vector-valued messages $\boldsymbol{m} \in [0,1]^n$, 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 $\boldsymbol{\theta}$. Consumers observe offers only from firms in their consideration set.

The design $\lambda$ 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 $\theta$ 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 $\boldsymbol{\theta}$. It will sometimes be helpful to start with a fixed

[31] See Bergemann and Bonatti (2019) for a discussion on the distinction between information and access design. Bergemann, Brooks, and Morris (2021a) (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.

matching scheme $\phi$, in which case we are restricting the space of designs to

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

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

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 3.** The feasible surplus set $SUR \subset \mathbb{R}^2_{\geq 0}$ comprises the pairs of producer surplus (PS) and consumer surplus (CS) that can be implemented as an equilibrium outcome of some design $\lambda \in \mathbf{\Lambda}$. 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 4.** 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 $\phi$-consumer optimal design is a design which implements the highest consumer suplus among all designs in $\mathbf{\Lambda}_\phi$.

**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 $\phi$-consumer-optimal design for some $\phi$.

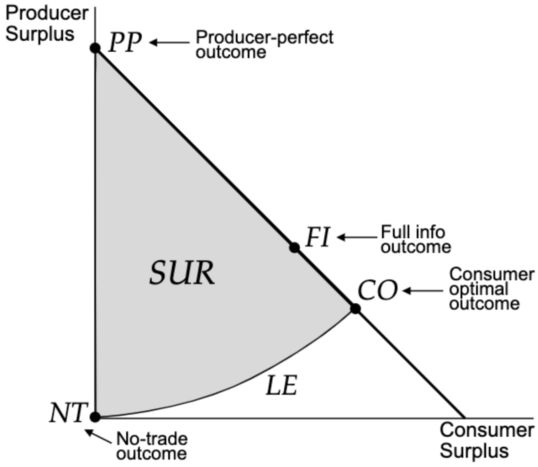
(iii) The consumer-optimal point (CO) is obtained through the unrestricted matching denoted $\phi^*(\cdot|\theta) = \delta_{\mathcal{N}}$ for all $\boldsymbol{\theta} \in \Theta$, paired with the $\phi^*$-consumer optimal design.

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

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.

Figure 5: Illustration of $SUR$




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 $\theta$

(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.[32]

[32]Online Appendix II formalizes these arguments.

## 7 Concluding Remarks

**7.1 Takeaways for regulators.** We have explored how platforms can use information about market participants to shape price competition. By packaging information about consumers' preferences in different ways, the platform can relax or intensify market competition to obtain different ratios 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. The outright prevention of the use of such information will typically sacrifice efficiency and prevent a platform that wishes to increase consumer surplus from intensifying price competition well beyond the complete information benchmark case (with a corresponding increase in 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 get 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 regulate the use of consumer information. Our analysis shows that the disclosure of information creates qualitatively different market segments when price competition is being intensified as opposed to being relaxed. Regulations that require consumers to be grouped into segments 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 would prevent the worst outcomes from consumers being realized without preventing well-intentioned intermediaries from increasing efficiency and enhancing consumer surplus.[33] We derived this implication by normalizing firms' marginal costs to zero. When firms have not too different positive marginal costs the prescription is still valid. When firms have very different marginal costs, the authority also needs to aggregate information on these costs.

Finally, information designs that relax price competition often rely on private signals, whereas the consumer-optimal outcome can always be implemented with public signals. A more direct intervention is a non-discrimination requirement at the level of the information provided to firms by an intermediary. That is, regardless of the way in which the intermediary packages information about consumer preferences, the aggregated information should be public among 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

[33]For instance, the regulators can formulate rules of conduct prescribing that machine learning techniques used to aggregate consumers into groups do so with either the objective of grouping similar consumers together, as in the consumer-optimal design, or constraints serving the same purpose.

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.[34] 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.[35] 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 is being used to facilitate discriminatory pricing has drawn regulatory interest. China's new anti-monopoly guidelines—tailored exclusively to reigning in tech firms—explicitly outlines the phenomena of data being used to "achieve coordinated behaviour" (State Administration for Market Regulation, 2021).[36] 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.[37]

## 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


[34] See 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.

[35] A 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).

[36] There 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.

[37] See European Comission (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.

——— (2021a): "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., T. HEUMANN, AND S. MORRIS (2021b): "Information, market power, and price volatility," *The RAND Journal of Economics*, 52, 125–150.

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.

CARLSSON, H. AND E. VAN DAMME (1993): "Global games and equilibrium selection," *Econometrica: Journal of the Econometric Society*, 989–1018.

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 COMISSION (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. SONNENSCHEIN (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.

RUBINSTEIN, A. (1989): "The Electronic Mail Game: Strategic Behavior Under" Almost Common Knowledge"," *The American Economic Review*, 385–391.

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 (prodcer-perfect) designs, as well as comparative statics on AIC (Propositions 1 and 2). Appendix I 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 an lower bound on firm $i$'s profit. Given information structure $\psi$ and opponent firms follows strategy profile $\sigma$, firm $i$ can always ignore any information from information structure $\psi$. In particular, fix any joint distribution over types and other firms' prices $\psi_{-i} \in \Delta(\Theta \times \boldsymbol{M}_{-i})$. The firm can always obtain a profit

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

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

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

We are ready to prove Theorem 1. When the consumer is in set $E_i$, $\psi^*$ will recommend firms other than $i$ to price at zero. Given that opponent firms follow the price recommendation, type $\boldsymbol{\theta} \in E_i$ will have a residual value of $\Lambda_i(\boldsymbol{\theta})$ for product $i$. In words, type $\boldsymbol{\theta}$ purchase product $i$ at price $p_i$ if $\Lambda_i(\boldsymbol{\theta}) \geq p_i$. 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 $\mu_i^*$. Firm $i$ will exactly obtains its lower bound profit $\underline{\Pi_i^*}$:

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

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

$\psi^*$ then applies the uniform profit preserving extremal information structure (Bergemann, Brooks, and Morris, 2015) to $\mu_i^*$. 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.[38] 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 $\text{CS}^*$. 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

[38]Since 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 $\psi$ is an producer-perfect information structure. To see that **Firm IC** is necessary, recall that **Firm IC** is:

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

For type $\boldsymbol{\theta} \notin E_i$, firms other than $i$ will charge type $\theta$ at least her value for their product (by **Separation** and **Consumer-IC**). Therefore, each type $\boldsymbol{\theta} \notin E_i$ 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 $\psi$ is a valid information structure. **Separation**, **Consistency** ensure that each type $\boldsymbol{\theta} \in E_i$ 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. □

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

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

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

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

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

On the other hand, we have

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

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

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

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

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

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

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

where

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

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:

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

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

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

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

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 acommondate the total mass of types not in $E_i$ with value for pruduct $i$ above $c$. $\gamma$ 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 $\theta_i$ in $E_i$. Further note that $\gamma$ is weakly increasing and $\gamma(c) \geq c$.

For all $\boldsymbol{\theta} \in E_i$, probability mass $\psi_i(m_i = \theta_i|\boldsymbol{\theta}) = 1$ fulfilling **separation**. Recall that we assumed $\mu$ 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:

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

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

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

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

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

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

Note $H_i$ is always differentiable even if $\mu$ has atoms.

When $m_i$ is an atom ($m_i \in V_i$), $\psi_i(m_i|\boldsymbol{\theta})$ 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 $\boldsymbol{\theta}$ is such that $\theta_i \notin V_i^-$ then

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

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

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

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 $\theta_i$ 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_{\boldsymbol{\theta} \notin E_i, \theta_i \geq \hat{p}_i} \psi_i(m_i|\boldsymbol{\theta}) d\mu &\leq \int_{\boldsymbol{\theta} \notin E_i, \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i|\boldsymbol{\theta}) d\mu && \text{(by } \gamma(\theta_i) \geq \theta_i\text{)} \\
&= \int_{\boldsymbol{\theta} \notin E_i, m_i > \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i|\boldsymbol{\theta}) d\mu \quad \text{(by } \psi_i(m_i|\boldsymbol{\theta}) = 0 \text{ if } m_i \leq \gamma(\theta_i)\text{)} \\
&= \sum_{\boldsymbol{\theta} \notin E_i, \boldsymbol{\theta} \in V, m_i > \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i|\boldsymbol{\theta}) \mu(\boldsymbol{\theta}) \\
&\quad + \int_{\boldsymbol{\theta} \notin E_i, m_i > \gamma(\theta_i) \geq \hat{p}_i} \psi_i(m_i|\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= \sum_{k : m_i > \gamma(v_k) \geq \hat{p}_i} \psi_i(m_i|v_k) \mu(\boldsymbol{\theta} \notin E_i, \boldsymbol{\theta} \in V, \theta_i = v_k) + \\
&\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 \\
&&& \text{(by } \psi_i(m_i|\boldsymbol{\theta}) \text{ only depending on } \theta_i\text{)} \\
&= \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^-\text{)} \\
&= \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) \\
&&& \text{(by } H_i(\gamma(c)) = Q_i(c)\text{)} \\
&= \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} \left. \frac{\partial G_i(c, m_i)}{\partial c} \right|_{c = \gamma(\theta_i)} (-1) d\gamma(\theta_i) \\
&&& \text{(by the construction of } \psi_i(m_i|\boldsymbol{\theta})\text{)} \\
&= \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 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 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., $\bar{\theta} = \inf\{\theta \in [0,1] | \mu([0,\theta]^n) = 1\}$. If $\bar{\theta} = 0$, Proposition 1 directly follows. For the rest of the proof we suppose $\bar{\theta} > 0$. Fix an arbitrary information structure $\psi$ and first observe that in any equilibrium induced by $\psi$, 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

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

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

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

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 $F^{\psi}$, choose $\epsilon > 0$ sufficiently small such that

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

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 $i$:[39]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

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

[39] Feldman, 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 $\psi$, 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 $\epsilon > 0$. □

*Proof of Proposition 2.* Since producer-perfect information structure exists under $\mu$, by Theorem 2, AIC holds:

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

Define

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

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

We have

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

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 $\theta_i$ and $\tilde{F}_{E_i}$ first order stochastically dominates $F_{E_i}$; the third step is by noticing that the distribution of $\theta_i$ induced by measure $\mu$ conditional on $E_i$ is the same as the distribution of $\theta_i$ 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 defintion 1. Therefore, AIC also holds under $\hat{\mu}$. By Theorem 2, producer-perfect information structure exists under $\hat{\mu}$. $\square$

## B: Proof of results in Section 5

Here we prove Proposition 3; the remaining results in Section 5 are in the Online Appendix.

*Proof of Proposition 3.* We begin with the following lemma which we prove below.

**Lemma 2.** If $\text{supp}\mu \subseteq \Theta^K$, then for any information design which implements an efficient outcome, realized sales prices are supported on $\subseteq \{1/K, 2/K, \ldots 1\}$.

*Proof of Lemma 2.* It is without loss to focus on direct recommendations (Bergemann and Morris, 2016) and let the information structure $\psi \in \Delta(\Theta \times \prod P_i)$ induce an efficient outcome. Write $S_i$ to denote the support of $i$'s sale price. Notice that since $\mu$ is supported on $\Theta^K$, the difference in valuations are at least $1/K > 0$. We will show that

$$\text{for all firms } i \in \mathcal{N},\, S_i \subseteq \Big\{1/K, 2/K, \ldots 1\Big\}.$$

We proceed by inducting on $L = 1, \ldots K$

<u>Base step:</u> $L = 1$. Suppose

$$S_i \cap [0, 1/K) \neq \emptyset.$$

Then, on the event that they make the sale, by efficiency it must be to types $\theta \in E_i$. But then firm $i$ would also have done so by setting a price of $1/K$. Then consider the following upward deviation: upon receipt of any recommendation in $S_i \cap [0, 1/K)$, set the price $1/K$ instead. This is a strictly profitable upward deviation which violates obedience.

<u>Inductive step</u>. For $L \geq 1$, we want to establish

$$\Bigg\{\text{for all } i \in \mathcal{N},\, S_i \bigcap \Big\{ \bigcup_{1 \leq J \leq L} \Big\{\Big(0, \frac{J}{K}\Big)\Big\}\Big\} = \emptyset\Bigg\}$$
$$\implies \Bigg\{\text{for all } i \in \mathcal{N},\, S_i \bigcap \Big\{ \bigcup_{1 \leq J \leq L+1} \Big\{\Big(0, \frac{J}{K}\Big)\Big\}\Big\} = \emptyset\Bigg\}$$

To see this, notice that conditional on making the sale on recommendation

$$p \in \Big(\frac{L}{K}, \frac{L+1}{K}\Big),$$

the firm could have made the sale charging a price of $(L+1)/K$ because from our induction hypothesis, on the event that $j$ makes the sale, firm $-i$'s price must be either above $p$ or in the set $\{1/K, 2/K, \ldots L/K\}$. Hence, we can once again consider the deviation up to the price $(L+1)/K$ whenever the firm receives a recommendation in

$(L/K, (L+1)/K)$. This is strictly profitable which violates obedience, a contradiction. This delivers the implication.

Now inducting, we have

$$\text{for all firms } i \in \mathcal{N}, S_i \bigcap \left\{ \left\{0\right\} \cup \bigcup_{1 \le J \le K} \left\{ \left(0, \frac{J}{K}\right) \right\} \right\} = \emptyset$$

as required. $\square$

Proof of Proposition 3 (i)

Let $\psi \in \Delta(\Theta \times \prod P_i)$ denote some obedient information design. This induces a joint distribution

$$\text{marg}_{E_i \times P_i} \psi =: \psi_i \in \Delta(E_i \times P_i)$$

which specifies the joint distribution between types and firm $i$'s *realized sales* since, from the condition of efficiency, types in $E_i$ must buy from firm $i$. We will construct the distribution $\mu' \in \Delta(\Theta)$ as follows. Define the operator

$$\tau_i : \Theta \times P_i \to \Theta \quad \text{s.t.} \quad \tau(\theta, p) = \Big(\theta_i = p, \theta_j = p - |\theta_i - \theta_j|\Big)$$

where $\tau_i(\theta, p)$ takes the type $\theta \in E_i$ and price charged $p \in P_i$ and returns the effective type. Now write

$$\mu_i' := \psi_i \circ \tau_i^{-1} \in \Delta(E_i)$$

which is simply the pushforward of $\psi_i$ under $\tau$. Summing up,

$$\mu' := \mu\Big(E_1\Big) \cdot \mu_1' + \mu\Big(E_2\Big) \cdot \mu_2'.$$

The interpretation of $\mu'$ is that under the information design $\psi$, *effective type* $\theta = (x, y) \in E_1$ is charged a price of $x$ by firm 1. Thus, firm 2 must charge a price of $y$ to make the sale. Now observe that, from efficiency and consumer IC, each type in $E_i$ must buy from firm $i$ at a price weakly less than $\theta_i$ which in turn implies $\mu' \preceq_D \mu$. We next show that it satisfies AIC.

**Lemma 3.** $\psi$ is obedient only if $\mu'$ satisfies AIC.

*Proof of Lemma 3.* Note that by construction, the expected profits upon receipt of the recommendation $p$ is proportional to

$$p \int_{\theta \in E_i : \theta_i = p} d\mu'.$$

Hence, a necessary condition for obedience is that for each recommendation $p$,

$$(p - p') \int_{\theta \in E_i : \theta_i = p} d\mu' =: p' G(p, p').$$

where $G(p, p')$ is the highest mass which firm $i$ can steal conditional on deviating from

$p$ to $p'$. Then, define the aggregating matching capacity to the price $p'$ as:

$$H_i(p') := \int_{p>p'} G(p,p')dp = \int_{p>p'} \frac{\Big((p-p')\int_{\theta\in E_i:\theta_i=p} d\mu'\Big)}{p'} dp$$

and so to ensure that uniform downward deviations are unprofitable (a necessary condition for obedience) we require:

$$H_i(p') \geq \int_{\substack{\theta\in\Theta\setminus E_i:\\ \theta_i-p'\geq 0}} d\mu'.$$

which is just AIC for the distribution $\mu'$. Note that the RHS of the expression is the profits from a uniform downward deviation to $p'$ since, from our construction of $\mu'$, *effective type* $\theta = (x,y) \in E_1$ is charged a price $x$ by firm 1 and, if charged a price $y$ by firm 2, would find it optimal to buy from 2. □

We have shown that for any obedient recommendation $\psi$, we have that the induced $\mu'$ (i) is diagonally dominated by $\pi$; (ii) satisfies AIC (Lemma 3); and (iii) has support contained within $\Theta^K$ (Lemma 2). Since these are all necessary conditions, (D) $\geq PS^*$.

Proof of Proposition 3 (ii) We will construct an information structure $\psi$ which implements (D). For each diagonal $D \in [0,1]$ and each firm $i$, we can define a comonotone matching from prices charged to consumers who prefer firm $i$, and consumer types.[40] Then, this specifies the joint distribution

$$\begin{aligned}&\pi_i \in \Delta(E_i \times P_i)\\ &\quad \text{s.t. } \mathrm{marg}_{E_i}\pi_i = \mu|_{E_i},\ \mathrm{marg}_{P_i}\pi_i = \mathrm{marg}_{\Theta_i}\mu^*|_{E_i}\end{aligned}$$

where throughout we write $\mu|_E$ to denote the (normalized) measure over $\Delta(\Theta)$ restricted to $E \subseteq \Theta$ i.e., $\mu|_E(A) := \frac{1}{\mu(E)}\mu(A\cap E)$ for any measurable sets $A, E$.

Now take any $p \in \mathrm{supp}\,\mathrm{marg}_{P_i}\pi$ i.e., a price in the support of the marginal of the transported distribution. First note that for each firm $i$ the set of diagonals where there is any mass is finite because $\mu$ is finite; let this set be denoted by $\Delta_i \subset [0,1]$. For $D \in \Delta_i$, write $\mathsf{proj}(D)$ to denote the corresponding 1−dimensional slice of the type space

$$\mathsf{proj}(D) := \Big\{\theta \in E_i : \theta_i - \theta_j = D\Big\}.$$

Notice that profits from obedience can be lower-bounded as follows:

$$\begin{aligned} p \cdot \sum_{\substack{\theta\in E_i:\\ \theta_i \geq p}} \pi_i(\theta,p) &\geq p\cdot \mu^*|_{E_i}(p) \\ &\geq p\cdot\Big(M - \Big\|\mu^*|_{E_i} - \mu|_{E_i}\Big\|_\infty\Big) \\ &> p\cdot\Big(M - \frac{\epsilon}{\mu(E_i)}\Big) \end{aligned}$$

[40] Since we are working with discrete distributions, this will not in general be unique but we will take an arbitrary selection from the set of comonotone transport plans.

where $M := \min_\theta \mu(\theta) > 0$ since $\mu$ is supported on a finite set, and the second inequality is because $\mu|_{E_i}$ and $\mu^*|_{E_i}$ are comonotonically matched.

Now consider the profit from deviating upwards to any price $p' > p$.

$$\begin{aligned}
p' \cdot \sum_{\substack{\theta \in E_i: \\ \theta_i \geq p'}} \pi_i(\theta, p) &= p' \sum_{D \in \Delta_i} \sum_{\theta \in \mathsf{proj}(D): \theta_i \geq p'} \pi_i(\theta, p') \\
&\leq p' \sum_{D \in \Delta_i} \frac{\mu\big(\mathsf{proj}(D)\big)}{\sum_{D' \in \Delta_i} \mu\big(\mathsf{proj}(D')\big)} \cdot \|\mathsf{DIAG}^i_D \mu^* - \mathsf{DIAG}^i_D \mu\|_\infty \\
&\leq |\Delta_i| \cdot p' \cdot \|\mu_i^* - \mu_i\|_\infty \\
&< |\Delta_i| \cdot p' \cdot \epsilon
\end{aligned}$$

where the first inequality is from the fact that types along each diagonal are matched comonotonically. Then noting that $p' \leq 1$ while $p \geq 1/K$, we can choose

$$0 < \epsilon < M \cdot \min_i \frac{\mu(E_i)}{K \cdot |\Delta_i|}$$

so that all upward deviations from any recommendation $p$ to any candidate deviation $p'$ are unprofitable, where we emphasize that the bound on $\epsilon$ depends only on $\mu$.

We have shown that we can construct $\pi_i \in \Delta(E_i \times P_i)$ such that upward deviation constraints are fulfilled. We now lift this construction to the information structure $\psi \in \Delta(\Theta \times \prod P_i)$ as follows. We will first construct

$$\psi|_{E_i} \in \Delta\Big(E_i \times \prod P_i\Big).$$

Observe that the marginal over $E_i$ is pinned down by the original distribution $\mu$. Moreover, the joint over $E_i \times P_i$ has already be constructed as $\pi_i \in \Delta(E_i \times P_i)$. It remains to specify the joint over $P_i \times P_j$. To do so, notice that since $\mu^*$ fulfils AIC, we can apply the sufficiency direction of Theorem 2 to $\mu^*$ to obtain a producer-perfect information structure for $\mu^*$. Now for this information structure, we can back out the joint distribution over the price that firm $i$ charges to types in $E_i$, and the price that firm $j$ charges. We set this as the joint over $P_i \times P_j$. Then choose

$$\psi = \sum_i \mu(E_i) \cdot \psi|_{E_i} \in \Delta\Big(\Theta \times \prod P_i\Big).$$

Obedience on the part of each firm yields profits of

$$\sum_i \frac{1}{\mu(E_i)} \int \theta_i d\mu^*_{E_i} = (\text{D})$$

since $\mu^*$ is by definition the solution to (D). □

## 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 $\boldsymbol{\theta}$ by setting to zero the valuation for product $i$ whenever firm $i$ is not in the consideration set for consumer type $\boldsymbol{\theta}$. 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 $\boldsymbol{\theta}$ let $S \in 2^{\mathcal{N}}$ be her realized consideration set. We map this consumer type and consideration set pair, $(\boldsymbol{\theta}, S)$, into a new consumer type and the unrestricted consideration set pair $(\boldsymbol{\theta}^S, \mathcal{N})$, where $\boldsymbol{\theta}^S = (\theta_i^S)_{i \in \mathcal{N}}$ 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 $\mu_\phi$ 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 $\boldsymbol{\theta} \in \Theta$ and consideration set $S \in 2^{\mathcal{N}}$, we treat the consumer's type as if it were $\boldsymbol{\theta}^S$. We then assign messages to each firm as if the underlying type were $\boldsymbol{\theta}^S$. In particular, we choose

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

as the consumer-optimal information structure (Theorem 1) as applied to the modified distribution $\mu_\phi$. Finally, define $\lambda_\phi^\star$ as follows:

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

which specifies, for a given realization of type $\boldsymbol{\theta}$, a joint distribution over messages and consideration sets. Consistency follows immediately from construction. Furthermore, $\lambda_\phi^\star \in \boldsymbol{\Lambda}_\phi$. The next lemma tells us that $\lambda^\star$ is indeed consumer-optimal among this class of designs.

**Lemma 4.** The design $\lambda_\phi^\star$ 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 $\boldsymbol{\Lambda}_\phi$, i.e., $\lambda_\phi^\star$ implements the consumer-optimal outcome among $\boldsymbol{\Lambda}_\phi$. Furthermore, this outcome is efficient given the matching constraints $\phi$.

*Proof of Lemma 4.* We apply the consumer-optimal information structure to the modified measure $\mu_\phi$ and, therefore, the equilibrium outcome is efficient given $\mu_\phi$: for each pair $(\boldsymbol{\theta}, S)$, the new type $\boldsymbol{\theta}^S$ purchases from her favourite firm among the set of firms $S$, i.e., $\boldsymbol{\theta}^S$ 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 $\phi$. These are:

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

where we integrate against the modified measure $\mu_\phi$ (which captures the fact that under the realization $(\boldsymbol{\theta}, S)$, the consumer only has positive valuation for firms in the set $S$).

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

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

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) $\lambda_\phi^\star$ was constructed such that firm $i$'s profits are held down to $\underline{\Pi}_i^\phi$; and (ii) the allocation is efficient given $\phi$, 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 $\boldsymbol{\Lambda}_\phi$. □

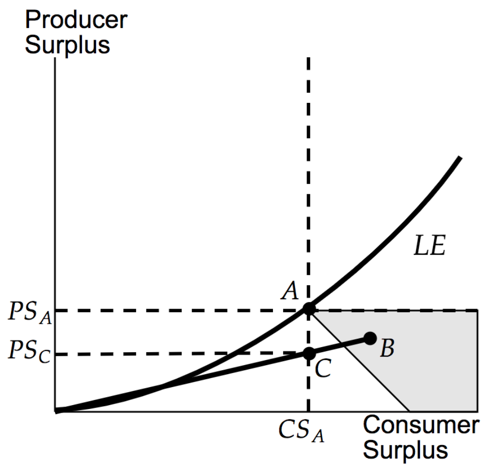
Fix a point $A := (CS_A, PS_A)$ on the lower envelope, and let the associated design implementing $A$ be $\lambda_A$. Further let $\phi_A$ be the matching scheme associated with $\lambda_A$. We claim that $\lambda_{\phi_A}^\star$ can also implement point $A$.

Suppose, towards a contradiction, that it did not. Denote by $B = (CS_B, PS_B)$ the consumer-optimal outcome that can be implemented by $\lambda_{\phi_A}^\star$. 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 $\lambda_{\phi_A}^\star$ 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 6 below, where the grey area indicates the area where point $B$ can be located.[41]

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 $\lambda_{\phi_A}^\star$ 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

[41]Note that the diagonal line in the picture are all the points that produce the same total surplus; and since point B is efficient given $\phi$ it produces weakly higher total surplus.

Figure 6: Illustration of the proof of Theorem 3 (ii)




$PS_C < PS_A$ (see Figure 6 for graphical illustration). But this contradicts our assumption that point $A$ belongs to the lower envelope of $SUR$.

*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 $\lambda$. 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 $\lambda^\star_\phi$ for some $\phi$ which implies

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

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

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

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 $\mathbf{\Phi}^E \subset \mathbf{\Phi}$. The following lemma shows that the solution to the consumer surplus maximization problem lies within $\mathbf{\Phi}^E$.

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

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

*Proof of Lemma 5.* Since $\phi$ 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 $\phi$. If $\phi$ delivers zero consumer surplus, then the same can be achieved with an efficient matching scheme and we are done. If $\phi$ 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 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

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

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

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

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 $\lambda^\star_\phi$ firm $i$'s profit is the same as its no-information profit under the distribution $\mu_\phi$ which is

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

Let us now modify the inefficient access scheme $\phi$ 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 $\phi'$, 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 $\lambda^\star_\phi$ and under $\lambda^\star_{\phi'}$. A few observations follow. First, notice that on the realizations of $(\boldsymbol{\theta}, S) \in T_{i,j}$ under the design $\lambda^\star_\phi$ those consumers bought from $j$. However, under the design $\lambda^\star_{\phi'}$ they now buy from $i$. This implies $j$'s profits must weakly decrease i.e., $\pi_j(\lambda^\star_{\phi'}) \leq \pi_j(\lambda^\star_\phi)$. Second, observe that under the design $\lambda^\star_{\phi'}$, $i$'s profits are now

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

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:

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

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 $PS(\lambda^\star_\phi)$ as the producer surplus implemented by $\lambda^\star_\phi$. We have:

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

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 $CS(\lambda^\star_\phi)$ as total consumer surplus under $\lambda^\star_\phi$, we similarly have

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

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. □

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 $\phi \in \mathbf{\Phi}^E$,

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

On the other hand, from Theorem 1 and Lemma 4, the expected profits of firm $i$ under the design $\lambda^\star_\phi$ 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 $\boldsymbol{\theta} \in E_i$), we have that

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

since restricting the consumer's consideration set decreases competition, i.e., the max operator is increasing in the set order, and efficient matching scheme $\lambda$ has firm $i$ in type $\boldsymbol{\theta}$'s consideration set with full probability. Hence, the design $\lambda^\star_\phi$ corresponding to the full matching scheme minimizes firm $i$'s profits and thus total producer surplus. Putting everything together, we have

$$CS^\star := \max_{\lambda \in \boldsymbol{\Lambda}} CS(\lambda) = \max_{\lambda \in \{\lambda^\star_\phi\}_{\phi \in \boldsymbol{\Phi}}} CS(\lambda) = \max_{\lambda \in \{\lambda^\star_\phi\}_{\phi \in \boldsymbol{\Phi}^E}} CS(\lambda) = CS(\lambda^\star_{\phi^F})$$

where we use $\phi^F \in \boldsymbol{\Phi}^E$ 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 $\alpha$ and $1-\alpha$. Hence, $SUR$ is the convex hull generated by the producer optimal point and the lower envelope ($LE$).

# Online Appendix to
# 'Market Segmentation through Information'

**Andrea Galeotti, Matthew Elliott, Andrew Koh, Wenhao Li**

FOR ONLINE PUBLICATION ONLY

**Outline.** This online appendix collects further extensions, robustness, and examples related to 'Market Segmentation through Information' henceforth referred to as the main text. Online Appendix I develops a formal analysis of the switching cost model. Online Appendix II shows that in large markets, the surplus set $SUR$ obtained via matching and information design continue to approximately hold even when the monopolist is constrained to match each consumer to at least $K$ offerings. Online Appendix III formalizes the claim that normalizing marginal costs to zero in the main text is without loss. Online Appendix IV shows that within the switching cost model, the producer-optimal design conditional on efficiency coincides with producer-optimal design unconditioned on efficiency when the information structure is finite and firms play pure strategies. Online Appendix V works through several parametric examples for when AIC holds or fails.

## I: Formal Analysis of producer optimal in the switching cost model

**I.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 valuations.

Recall that an information structure is a joint distribution $\psi \in (\Theta \times M^2)$, and furthermore when 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 $\theta$. 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 6.** AIC holds in the switching cost model if and only if

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

*Proof of Lemma 6.*

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

and

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

so

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

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 $\mu$ 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

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

□

We will tiebreak throughout in favor of efficiency. Each type $\theta \in E_i$ will break ties in favor of product $i$. Under this tiebreaking rule, for an obedient information structure $\psi$, let $PS(\psi)$ denote the (expected) producer surplus when all players obey the recommendation.

Further define $\Psi^E$ as the set of obedient information structures which implement efficient outcomes i.e., $\psi \in \Psi^E$ implies that in the Bayes Correlated Equilbiria in which all players follow price recommendations, each type $\theta \in E_i$ purchases from firm $i$ with probability 1.

**Lemma 7.** There exists an efficient producer-optimal design $\psi^*$ which solves $\sup_{\psi \in \Psi^E} PS(\psi)$ such that

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

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

(ii) $\psi^*$ is symmetric: for any $m, m' \in \{0, \gamma, L, H\}$,

$$\psi^*_{H1}(m, m') = \psi^*_{H2}(m', m) \quad \text{and} \quad \psi^*_{L1}(m, m') = \psi^*_{L2}(m', m).$$

We emphasize that on recommendations for which the firm never makes the sale, there is some freedom to recommend prices outside the set $\{0, \gamma, 1 - v, 1\}$. Nonetheless, before proving Lemma 7, we will establish that we can modify such information structures so that these recommendations which lead to no sales are replaced with the zero recommendation while preserving both obedience and equilibrium producer surplus.

**Lemma 8** (Replacing no-sale messages with the zero recommendation)**.** Let $\hat{\Psi} \subseteq \Psi^E$ denote the set of efficient and obedient information structures that never recommend positive prices resulting in a firm having no sales. That is, for each $\hat{\psi} \in \hat{\Psi}$, each firm $i$, and for any Borel measurable set $A_i \subseteq (0, 1]$,

$$\hat{\psi}(E_i, A_i, M_{-i}) = 0 \implies \hat{\psi}(\Theta, A_i, M_{-i}) = 0.\text{[1]}$$

Then, for any obedient and efficient information structure $\psi \in \Psi^E$, there exists another obedient efficient information structure $\hat{\psi} \in \hat{\Psi}$ with the same producer surplus: $PS(\psi) = PS(\hat{\psi})$.

*Proof of Lemma 8.* Under $\psi \in \Psi^E$, suppose there exists a set of positive price recommendations $A_i \subseteq (0, 1]$ such that firm $i$ never makes the sale i.e., $\psi(E_i, A_i, M_{-i}) = 0$ but firm $i$ receives such recommendations with positive probability i.e., $\psi(\Theta, A_i, M_{-i}) > 0$. Now, consider the information structure $\psi'$ which modifies $\psi$ as follows: whenever the designer sends recommendations in $A_i$ to firm $i$, it sends the recommendation 0 instead.

We first verify that $\psi'$ is still obedient. For firm $i$, since $\psi$ is obedient by conjecture, upon receipt of any message $m_i \in A_i$, it is optimal for firm $i$ to obey the recommendation with revenue almost surely zero. Furthermore, upon receipt of message $m_i = 0$, it was also optimal for firm $i$ to obey under $\psi$. Hence, from the law of total probability, it must be optimal for firm $i$ to obey the recommendation $m_i = 0$ under $\hat{\psi}$. Moreover, for the message $m_i \notin A_i \cup \{0\}$, firm $i$ must still find it optimal to obey under $\hat{\psi}$ since the conditional distribution over $\Theta \times M_2$ is the same as under $\psi$.

We now verify the obedience of firm $j$. First observe that for any $m_i \in A_i$, such that under $\psi$, firm $i$ cannot obtain any positive demand by deviating to the price 0, recalling that we break ties in favor of efficiency, and that conditional on $m_i \in M_i$, the consumers concentrate on $E_j$ (since $\psi$ is efficient). If this were not true, there exists some price

[1] Note under efficient information structures, if some messages result in firm $i$ having no sales, those messages are never sent by types in $E_i$.

$m_i' > 0$ that firm $i$ can profitably deviate to, contradicting the obedience of $\psi$. Thus, when firm $i$ deviates to price 0 following message $m_i$ under $\psi$, no consumer changes her purchase decision.

Now, observe that for any recommendation $m_j \in M_j$, firm $j$ obtains the same expected profits from obedience under $\hat{\psi}$ as under $\psi$ since we have just shown that by replacing messages $A_i$ with 0, no consumer changes her purchase decision. But by deviating to any alternate price, firm $j$ obtains weakly lower profit under $\hat{\psi}$ than under $\psi$ because firm $i$ now sets a price of 0 under $\hat{\psi}$. We have shown that $\hat{\psi}$ is obedient for firms $i$ and $j$. Finally, since no consumer changes her purchase decision, $\hat{\psi}$ leads to an efficient allocation with the same producer surplus, as required. □

By Lemma 8, it is without loss to restrict our attention to efficient and obedient information structures in the set $\hat{\Psi}$. We are now ready to prove Lemma 7.

We will proceed via a "contagion/infection argument" by constructing regions of price recommendations from which firm 1 deviating upwards is strictly profitable. Then, this implies firm 2 must also deviate upwards on those regions; but if both firms deviate upwards, this, in turn, implies that the region of price recommendations which are (interim) strictly dominated expands, and we can run this process of iteratively eliminating interim dominated strategies forward. We use this to show that, conditional on efficiency, the only prices which survive this process are $\{0, \gamma, 1-v, 1\}$. This argument is closely related to the argument employed in Lemma 2, but the switching model requires a little more finesse. The logic of the argument is closely connected to those employed in incomplete information games (Rubinstein, 1989) and, in particular, global games (Carlsson and Van Damme, 1993) where dominance regions leads to "contagion" via the interim elimination of strictly dominated strategies.

*Proof of Lemma 7.* We show each part in turn.

Part (i): Supported on $\{0, \gamma, 1-v, 1\}$. By Lemma 8, it is without loss to restrict our attention to efficient and obedient information structures in the set $\hat{\Psi}$. Suppose that $\psi \in \hat{\Psi}$ is obedient, efficient.

**Step 1:** $\psi$ only recommends prices in the set

$$\widehat{P} := \underbrace{\Big\{k\gamma\Big\}_{k=0}^{k=\overline{K}}}_{\text{(A)}} \cup \underbrace{\Big\{1-v+k\gamma\Big\}_{0\leq k\leq \bar{K}-\underline{K}-1} \cup \Big\{1-v+(\bar{K}-\underline{K})\gamma\Big\}}_{\text{(B)}} \cup \underbrace{\Big\{1\Big\}}_{\text{(C)}},$$

where

$$\bar{K} := \max\Big\{k \in \mathbb{N} : k\gamma \leq 1\Big\} \quad \text{and} \quad \underline{K} := \max\Big\{k \in \mathbb{N} : k\gamma \leq 1-v\Big\}.$$

We will consider the case $1-v+\gamma \leq 1$ and $\underline{K} \geq 2$; the other cases are simpler and proceed similarly. Note that (A) are the multiples of $\gamma$, (B) are multiples of $\gamma$ with remainder $1-v$ up to 1, and (C) is the highest possible valuation. Hence, $\{0, \gamma, 1-v, 1\} \subseteq \text{(A)} \cup \text{(B)} \cup \text{(C)}$.

Step 1A: rule out $[0, 1-v) \setminus \hat{P}$. We will start by eliminating $(0, \gamma)$, $(\gamma, 2\gamma)$, and so on.

Base step: observe that we cannot have $\psi(\Theta, (0,\gamma), M_{-i}) > 0$. Suppose not. Since $\psi \in \hat{\Psi}$, we must have $\psi(E_i, (0,\gamma), M_{-i}) > 0$. Since $\psi$ is efficient and obedient, upon receiving recommendations in set $(0, \gamma)$, firm $i$ sells to all types in $E_i$ when firm $i$ obeys recommendations. But consider the uniform upward deviation by firm $i$ to the price $\gamma$: since we tiebreak in favor of efficiency, whenever firm $i$ would have made the sale obeying the recommendation $m_i \in (0, \gamma)$, it would also have made the sale at price $\gamma$. But this is a strictly profitable deviation, contradicting the obedience of $\psi$.

Inductive step: for integer $1 \leq K < \underline{K}$, suppose that $\psi$ never recommends prices in the set

$$\bigcup_{0 \leq k \leq K-1} \Big(k\gamma, (k+1)\gamma\Big).$$

We will show that the prices $(K\gamma, (K+1)\gamma)$ must also be ruled out. Suppose not so that $\psi(\Theta, (K\gamma, (K+1)\gamma), M_{-i}) > 0$. Since $\psi \in \hat{\Psi}$, we must have $\psi(E_i, (K\gamma, (K+1)\gamma), M_{-i}) > 0$. Since $\psi$ is efficient and obedient, upon receiving recommendations in set $(K\gamma, (K+1)\gamma)$, firm $i$ sells to all types in $E_i$ when firm $i$ obeys recommendations. In addition, for types in $E_i$ to purchase product $i$ at prices in set $(K\gamma, (K+1)\gamma)$, firm $j$'s price must be strictly above $(K-1)\gamma$. By our inductive hypothesis, firm $j$ is never recommended prices between $((K-1)\gamma, K\gamma)$. Therefore, on the event that firm $i$ makes the sale, firm $j$'s price is at least $K\gamma$. But since ties are broken in favor of efficiency, the uniform upward deviation to the price $(K+1)\gamma$ is strictly profitable.

This inductive argument implies that prices in the set

$$\bigcup_{0 \leq k \leq \underline{K}-1} \Big(k\gamma, (k+1)\gamma\Big)$$

must be ruled out. We can also rule out prices in $(\underline{K}\gamma, 1-v)$: if with positive probability firm $i$ is recommend to set prices in this interval, then on the event that firm $i$ makes the sale by obeying recommendations, the type must come from set $E_i$ and firm $j$'s price must be strictly above $(\underline{K}-1)\gamma$. Since firm $j$ never sets prices in $((\underline{K}-1)\gamma, \underline{K}\gamma)$, firm $j$'s price is at least $\underline{K}\gamma$. Firm $i$ once again has a strictly profitable upward deviation to the price $1-v$.

To take stock, we have ruled out all prices $[0, 1-v) \setminus \hat{P}$. We will now rule out prices above $1-v$:

<u>Step 1B: rule out $[1-v, 1] \setminus \hat{P}$.</u>

Base step: rule out $(1-v, (\underline{K}+1)\gamma) \cup ((\underline{K}+1)\gamma, 1-v+\gamma)$. First observe that we canrule out prices $(1-v, (\underline{K}+1)\gamma)$: if with positive probability firm $i$ is recommend to set prices in this interval, then on the event that firm $i$ makes the sale, the consumer type must be in set $E_i$ and have valuation 1 for $i$'s product, and firm $j$'s price must be strictly above $1-v-\gamma$. From the previous step, firm $j$ never sets prices in $(1-v-\gamma, \underline{K}\gamma)$. Hence firm $j$'s price is at least $\underline{K}\gamma$. Therefore, firm $i$ once again has a strictly profitable upward deviation to the price $(\underline{K}+1)\gamma$.

Moreover, we can rule out prices $((\underline{K}+1)\gamma, 1-v+\gamma)$ by again observing that since $\psi \in \hat{\Psi}$, conditioned on firm $i$ making the sale, we have already ruled out firm $j$ pricing at $(\underline{K}\gamma, 1-v)$. Hence, if firm $i$ makes the sale, firm $j$ must be pricing at least at price $1-v$. But then firm $i$ has a strictly profitable uniform upward deviation to the price $1-v+\gamma$.

Inductive step: we suppose that $1-v+(\bar{K}-\underline{K})\gamma < 1$; the case where it is above 1 is easily handled. Now suppose that for integer $0 \leq K < \bar{K}-\underline{K}-1$, suppose $\psi$ never recommends prices in the set

$$\Big(1-v+K\gamma, (\underline{K}+K+1)\gamma\Big) \cup \Big((\underline{K}+K+1)\gamma, 1-v+(K+1)\gamma\Big)$$

then we will show that prices

$$\underbrace{\Big(1-v+(K+1)\gamma, (\underline{K}+K+2)\gamma\Big)}_{\text{(I)}} \cup \underbrace{\Big((\underline{K}+K+2)\gamma, 1-v+(K+2)\gamma\Big)}_{\text{(II)}}$$

must also be ruled out.[2] We first rule out (I): if with positive probability firm $i$ is recommended to set prices in this set, then from the inductive hypothesis, firm $j$ never sets prices in the set $(1-v+K\gamma, (\underline{K}+K+1)\gamma)$. Hence, on the event that firm $i$ makes the sale, firm $j$ must be pricing at least at $(\underline{K}+K+1)\gamma$ but then firm $i$ has a strictly profitable upward deviation to $(\underline{K}+K+2)\gamma$.

We now rule out (II) similarly: if with positive probability firm $i$ is recommended to set prices in this set, then from the inductive hypothesis, firm $j$ never sets prices in the set $((\underline{K}+K+1)\gamma, 1-v+(K+1)\gamma)$. Hence, on the event that firm $i$ makes the sale, firm $j$ must be pricing at least at $1-v+(K+1)\gamma$ but then firm $i$ has a strictly profitable upward deviation to $1-v+(K+2)\gamma$.

Then, inductiving on $K$ up to $\bar{K}-\underline{K}-1$, it remains to rule out the prices $(1-v+(\bar{K}-\underline{K})\gamma, 1)$: suppose that firm $i$ is recommended those prices with positive probability. Then, by the same argument as before, on the event that firm $i$ makes the sale, we know that firm $j$ must be charging at least $\bar{K}\gamma$ which again implies firm $i$ has a strictly profitable uniform upward deviation to price 1. This concludes Step 1.

**Step 2:** $\psi$ only recommends prices in the set $\{0, \gamma, 1-v, 1\}$.

We begin with the following observation:

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

Equation (No ties) says that $\psi$, being obedient, cannot with positive probability make a recommendation such that the consumer (i) is indifferent between both firms ($\theta_1 - m_1 = \theta_2 - m_2$); and (ii) obtains strictly positive surplus ($\theta_1 - m_1 > 0$). To see this, suppose (No ties) did not hold and it is of measure $c > 0$, and $c_i > 0$ mass of types comes from set $E_i$. Now consider the uniform $\epsilon-$downward deviation by firm $j$ which we define as follows: for any recommendation $m$ it receives, charge $m - \epsilon$. Note that this increases demand by at least $c_i$, and since the original demand was bounded by 1, the losses from such a deviation is at most $\epsilon$. Hence we can find a small enough $\epsilon$ so that this deviation is strictly profitable, a contradiction.

Now suppose that

$$\psi\Big(\Theta, \hat{P}\setminus\{0,\gamma,1-v,1\}, M_{-i}\Big) > 0.$$

which, since $\psi \in \hat{\Psi}$, implies $\psi(E_i, \hat{P}\setminus\{0,\gamma,1-v,1\}, M_{-i}) > 0$. Then, for each recommendation $m_i$, on the event that firm $i$ makes the sale, we have $\theta_i - m_i > 0$ (since $1-v$

[2]Note that when $K=0$, this recovers our base step.

and 1 are ruled out). But since firm $i$ makes the sale, by Equation (No ties) this also implies $m_j > m_i - \gamma$. Further note that $m_j \in \widehat{P}$ from Step 1. But this implies firm $i$ can profitably deviate upwards by $\min_{p,p' \in \hat{P}} |p - p'| > 0$ and still make the sale whenever it would previously made it, a contradiction.

Part (ii): Symmetry. Start with an arbitrary obedient structure $\psi$ and suppose it maximizes producer surplus. Let $\psi'$ be a symmetrized version of $\psi$ 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 $\psi'$ continues to be obedient and implements the same PS as $\psi$ 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 $\gamma$, 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) $\psi$ 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

$$m_1 D_\psi(m_1, m_1) \geq m'_1 D_\psi(m_1, m'_1) \quad \text{for all } m_1 \in \operatorname{supp} \operatorname{marg}_{M_1} \psi \text{ and all } m'_1.$$

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

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

as required. Furthermore, payoffs are preserved but $\psi^*$ is now symmetric. □

**I.2 Construction of relaxed problem.** With Lemma 7 in hand, the information structure is pinned down by

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

where we leave the condition that $\psi_\theta$ is efficient implicit. Let $\Psi^{S,E}$ be the set of efficient and symmetric 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 $\gamma$ 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:

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

**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

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

Note since $\psi$ 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 $m > 1-\gamma$,

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

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

$$\begin{aligned}(v+\gamma)\psi_{H1}(1) + \gamma\Big(\frac{1}{2} - \psi^{\gamma} - \psi_{H1}(1)\Big) + \mathbb{1}_{\gamma>\frac{1-v}{2}}\Big(2\gamma-(1-v)\Big)\psi^{\gamma}&\\ \geq \Big(1-v-\gamma\Big)\Big(\frac{1}{2}-\psi^{\gamma}\Big)&\end{aligned} \tag{G-AIC-L}$$

**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 $1-v-\gamma$:

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

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 $m > 1-v-\gamma$:

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

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 $\psi$ to be obedient.

We first have:

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

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

We also have:

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

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

We also have:

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

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 $\gamma$ 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.

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

Our relaxed program is thus

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

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.

**I.3 Proof of Proposition 4.** Here we focus on the case where $\gamma \geq v$; we develop the case of $\gamma < v$ in Appendix I.4 as the main ideas are similar. Let

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

**Proposition 4.** Suppose $v \leq \gamma < 1/2$. The following characterizes the producer surplus obtained at the producer-optimal efficient outcome:

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

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

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

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

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

We are now ready to show Proposition 4 in which we focus on the parameter space $\gamma \geq v$. Recall we split our discussion of Proposition 4 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 $\gamma < \underline{\gamma}(v)$.

<u>Case A: $\gamma \geq \frac{1-v}{2}$.</u> From Lemma 6, $PS^* = TS$ if and only if $\mu \leq \overline{\mu}(\gamma, v)$. Now consider the case $\mu > \overline{\mu}(\gamma, v)$.

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

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

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:

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

which is decreasing in $\psi^{\gamma}$.

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

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

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

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

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

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

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

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

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

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 $\gamma$ 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 $1-\gamma$)

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

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

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

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

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

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 6 that if $\underline{\mu}\leq\mu\leq\overline{\mu}$ then AIC holds so $PS^*=TS^*$.

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

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

where recall

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

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

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

which is achieved by an information structure which removes

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

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

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

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

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 $\psi^{\gamma}$:

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

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:

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

which is decreasing in $\psi^{\gamma}$. Hence the relaxed problem is solved by choosing

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

We now show this value is obtainable by constructing an information structure $\psi$ 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 $\gamma$, only type $H1$ will purchase product 1 at price $\gamma$ 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 $\gamma$ 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

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

which is exactly G-AIC-L. By the same procedure in the proof of Theorem 2 we can construct $\psi$ 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 6, 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 $\gamma$, which is achieved by giving full information so $PS^* = R$.

**I.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 5.** 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 $\frac{1-v}{2} \leq \gamma < \frac{1}{2}$:

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

where

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

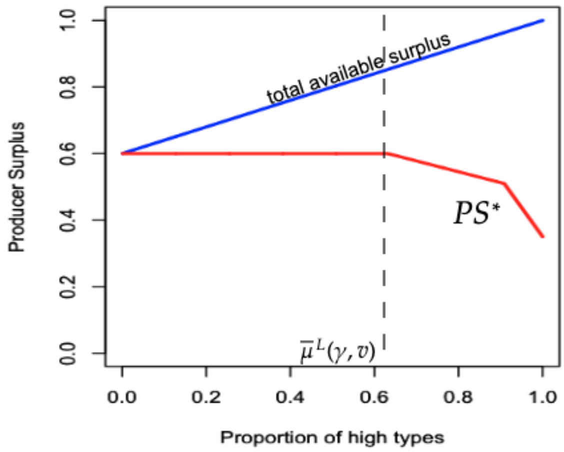
and $g^L$ is continuous, strictly decreasing in $\mu$, and $g^L(1,\gamma,v) = \gamma$.

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

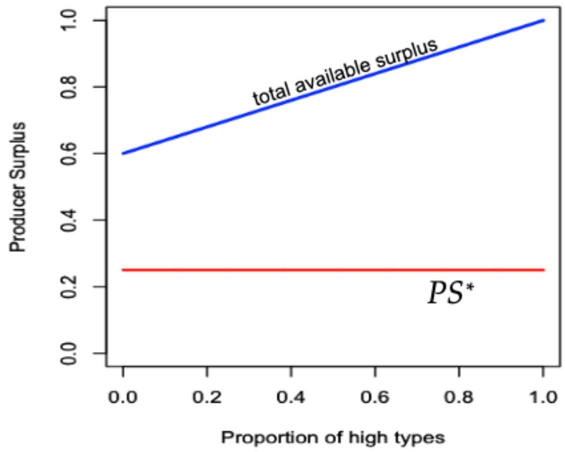
$$PS^*(\mu,\gamma,v) = \gamma.$$

When $\gamma < v$, the ability for information to soften competition is severely diminished as is reflected by Proposition 5 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

Figure 7: Illustration of optimal producer surplus as distribution varies ($\gamma < v$)




(a) $\gamma = 0.35, v = 0.4$




(b) $\gamma = 0.25, v = 0.4$

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.

*Proof of Proposition 5.* 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

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

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:

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

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

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

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

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

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

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

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

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

And we have the inequality

$$\psi^{\gamma,L}_{\uparrow 1+v-\gamma}\geq \psi^{\gamma,L}_{\uparrow 2\gamma} \iff \mu\leq\overline{\mu}^L+\tilde{\mu}^L$$

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

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

Therefore, at the solution to the relaxed problem we have

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

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 5 (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 $\gamma$ 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 $\gamma$ 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 4.

□

In the main text (Section 6) we considered an environment in which the platform had unrestricted power to match consumers to firms. For instance, the producer-optimal outcome was attained by choosing a matching which matches each consumer to her favorite firm and none other. We now consider what surplus points within $SUR$ are eliminated by the constraint that all consumers must have access to at least $K \leq N$ firms. This constraint may reflect a regulation imposed on an intermediary that influences matching between firms and consumers. Alternatively, it may represent consumers search choices. For example, $K$ could be the number of products shown on the first page of a search query in an online marketplace.

We model such settings by defining

$$\mathbf{\Psi}_{\geq K} := \Big\{\lambda \in \mathbf{\Psi} : S \in \text{supp}(\text{marg}_{2^{\mathcal{N}}}\lambda) \implies |S| \geq K\Big\}$$

as the set of all designs in which every consumer type is shown at least $K$ offers. We study equilibrium welfare outcomes under this restriction, which we define as follows.

**Definition 5.** The $K$-feasible surplus set $SUR_K \subset \mathbb{R}^2_{\geq 0}$ are the pairs of producer surplus (PS) and consumer surplus (CS) that can be implemented as an equilibrium outcome of some design $\lambda \in \mathbf{\Lambda}_{\geq K}$.

We will henceforth assume that consumers' valuation for each firms' product is drawn independently from a full-support continuous and density $g : [0,1] \to [\underline{g}, \overline{g}]$ for some $\underline{g}, \overline{g} > 0$. That is, for any $B := \prod_{i \in \mathcal{N}} B_i \in \mathcal{B}(\Theta)$

$$\mu(B) = \prod_{i \in \mathcal{N}} \int_{\theta_i \in B_i} g(\theta_i) d\theta_i$$

This imposes (i) symmetry on the relative attractiveness of each firms' product; and (ii) independence in valuations across different firms' products. As will be clear from the analysis, we could weaken this slightly to having some correlation or having firm-specific distributions, i.e., value for firm $i$ drawn from $g_i$ independently. Let $l : \mathcal{B}(\mathbb{R}^2) \to [0, +\infty)$ denote the Lebesgue measure. The following is our main result for this environment.

**Proposition 6.** For any fixed $K$, when the number of firms is sufficiently large almost every outcome in $SUR$ is achievable, i.e., For any $\epsilon > 0$ and $K$, there exists $\bar{n}_{\epsilon,K}$ such that for all $n \geq \bar{n}_{\epsilon,K}$,

$$l\Big(SUR \setminus SUR_K\Big) \leq \epsilon.$$

The matching and information designs which implement each point in $SUR_K$ is detailed in Lemma 9 (Producer perfect), Lemma 10 (No trade), and Lemma 11 (Consumer optimal).

*Proof of Proposition 6.* For a given realization of consumer type $\boldsymbol{\theta} \in \Theta$, order $\theta_{(1)} > \theta_{(2)} > \ldots > \theta_{(n)}$ where we let $\boldsymbol{\theta}_{(j)}$ be the $j-$th highest element. Further define $(j)_{\boldsymbol{\theta}}$ as the $j$-th favourite firm of type $\boldsymbol{\theta}$. Note that because the distribution is atomless, there are no ties almost surely.

Lemma 9 shows that for sufficiently large $n$ the producer perfect outcome is exactly $PO = (0, TS)$

**Lemma 9.** For every given $K$, there exists $\bar{n} > 0$ such that if the number of firms is larger than $\bar{n}$, the producer-optimal point $PO = (0, TS)$ is implementable. The producer-optimal point can be implemented with the following design:

(i) Matching: match the consumer $\boldsymbol{\theta}$ 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.

***Proof of Lemma 9.*** We wish to show that facing the modified distribution of valuations under the matching scheme specified in part (i) of Lemma 9, each firm $i$, upon receipt of the public message $m \in [0,1]$, prefers to obey and charge $p = m$, given that all other firms charge $m$.

We begin with several observations. First, if the firm $i$ charges $m$, its expected revenue is $m/n$ (by symmetry, each firm is equally likely to be the consumer's favourite). Second, if the firm deviates to some price $p \in (0, m)$, notice that by the matching scheme we specified, the only event on which the firm could potentially business-steal is when it is among the $K-1$ least favourite firms for the consumer. We can write firm $i$'s payoff from deviating to the price $p \in (0, m)$ as follows:

$$
\begin{aligned}
REV(p) &= \frac{p}{n} + p \cdot \sum_{j=1}^{K-1} \mathbb{P}(\theta_i \geq p \text{ and } \theta_i \text{ is } j\text{-lowest} \mid \theta_{(n)} = m) \\
&= \frac{p}{n} + p \cdot \frac{\sum_{j=1}^{K-1} \int_p^m \overbrace{\binom{n-2}{j-1} \Big(G(m) - G(q)\Big)^{n-j-1} G(q)^{j-1} g(q)(n-1) g(m)}^{\text{PDF of } \theta_i = q \text{ and } \theta_i \text{ is } j\text{-th lowest draw and max is } m} \, dq}{\underbrace{n g(m) G(m)^{n-1}}_{\text{prob. highest draw is } m}} \\
&= \frac{p}{n} + p \cdot \frac{\sum_{j=1}^{K-1} \int_p^m \binom{n-2}{j-1} \Big(G(m) - G(q)\Big)^{n-j-1} G(q)^{j-1} g(q)(n-1) dq}{n G(m)^{n-1}} \\
&= \frac{p}{n} + p \cdot \frac{n-1}{n} \cdot \frac{\sum_{j=1}^{K-1} \int_p^m \binom{n-2}{j-1} \Big(1 - G(q)/G(m)\Big)^{n-j-1} [G(q)/G(m)]^{j-1} g(q) dq}{G(m)}
\end{aligned}
$$

For firm $i$ to obey recommendation $m$, we need $\dfrac{m}{n} \geq REV(p)$ for all $p < m$; Equivalently:

$$
(m-p)G(m) \geq p \cdot (n-1) \cdot \int_p^m I(q, m, n, K) g(q) dq, \qquad \text{(Obedience constraint)}
$$

where

$$I(q,m,n,K) := \sum_{j=1}^{K-1} \binom{n-2}{j-1} \Big(1 - G(q)/G(m)\Big)^{n-j-1} [G(q)/G(m)]^{j-1}.$$

Our goal will be to show the existence of a positive constant $n^{PPO}$ such that for all $n \geq n^{PPO}$ a deviation to any price $p < m$ is unprofitable, , for any realization $m$. Our approach is to split potential deviations into two cases. The first case considers deviations to prices below $m/(c\sqrt{n})$ where $c$ is a constant independent of $m$ and $n$ that we specify appropriately later. The second case considers deviations to prices between $m/(c\sqrt{n})$ and $m$.

Case 1: Deviating to $p \leq \frac{m}{c\sqrt{n}}$ is unprofitable. Split the integral in the obedience constraint into two intervals: $\Big(p, \frac{m}{c\sqrt{n}}\Big]$ and $\Big(\frac{m}{c\sqrt{n}}, m\Big]$. The integral over the first interval is at most $\bar{g}\dfrac{m}{c\sqrt{n}}$, since $I(q,m,n,K) \leq I(q,m,n,n) = 1$ from the Binomial theorem. For the second part, notice that:

$$\binom{n-2}{j-1} \leq \binom{K-2}{j-1} \cdot \binom{n-2}{K-2}$$

and $I(q,m,K,K) = 1$ which implies $I(q,m,n,K) \leq \binom{n-2}{K-2}\Big(1 - G(q)/G(m)\Big)^{n-K}$. Hence, we can bind the second integral as follows:

$$\begin{aligned}
\int_{m/(c\sqrt{n})}^{m} I(q,m,n,K)g(q)dq &\leq \int_{m/(c\sqrt{n})}^{m} \binom{n-2}{K-2}\Big(1 - G(q)/G(m)\Big)^{n-K} g(q)dq \\
&\leq \binom{n-2}{K-2}\Big(1 - G(m/(c\sqrt{n}))/G(m)\Big)^{n-K} \int_{m/(c\sqrt{n})}^{m} g(q)dq \\
&\leq \binom{n-2}{K-2}\Big(1 - G(m/(c\sqrt{n}))/G(m)\Big)^{n-K} \bar{g} \cdot \Big(1 - \frac{1}{c\sqrt{n}}\Big)m \\
&\leq \binom{n-2}{K-2}\left(1 - \frac{\int^{m/(c\sqrt{n})} g(q)dq}{\int^{m} g(q)dq}\right)^{n-K} \bar{g} \cdot \Big(1 - \frac{1}{c\sqrt{n}}\Big)m \\
&\leq n^{K-2}\Big(1 - \frac{\underline{g}}{c\bar{g}\sqrt{n}}\Big)^{n-K} \bar{g}m.
\end{aligned}$$

Hence, for $p < \dfrac{m}{c\sqrt{n}}$, our obedience equation is fulfilled from the following series of inequalities:

$$\begin{aligned}
(m-p)G(m) &\geq m\Big(1 - \frac{1}{c\sqrt{n}}\Big)\underline{g}m && (G(m) \geq \underline{g}m) \\
&\geq \Big[\frac{1}{c^2} + \frac{1}{c}n^{K-1}\Big(1 - \frac{\underline{g}}{c\bar{g}\sqrt{n}}\Big)^{n-K}\Big]\bar{g}m^2
\end{aligned}$$

(By choosing $c$ appropriately and taking $n$ large; see equation 1 below)

$$
\begin{aligned}
&\geq \frac{m}{c\sqrt{n}}(n-1)\cdot\left[\bar{g}\frac{m}{c\sqrt{n}}+n^{K-2}\Big(1-\frac{\underline{g}}{c\bar{g}\sqrt{n}}\Big)^{n-K}\bar{g}m\right] \\
&\geq p\cdot(n-1)\cdot\left(\int_p^{m/(c\sqrt{n})} I(q,m,n,K)g(q)dq+\int_{m/(c\sqrt{n})}^{m} I(q,m,n,K)g(q)dq\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(Bounds for each integral developed above)} \\
&\geq p\cdot(n-1)\int_p^m I(q,m,n,K)g(q)dq,
\end{aligned}
$$

where the second inequality is implied by

$$
1 \geq \left[\frac{1}{c^2}+\frac{1}{c}n^{K-1}\Big(1-\frac{\underline{g}}{c\bar{g}\sqrt{n}}\Big)^{n-K}\right]\cdot\frac{\bar{g}}{\underline{g}}+\frac{1}{c\sqrt{n}}. \tag{1}
$$

which we will show for an appropriate choice of $c$ and large $n$. The following fact is useful

**Claim 1.**

$$
\lim_{n\to\infty} n^{K-1}\Big(1-\frac{\underline{g}}{c\bar{g}\sqrt{n}}\Big)^{n-K}=0.
$$

***Proof of Claim 1.*** Taking logarithms and exponentiating,

$$
\Big(1-\frac{\underline{g}}{c\bar{g}\sqrt{n}}\Big)^{n-K}=\exp\left((n-K)\cdot\log\left(1-\frac{\underline{g}}{c\bar{g}\sqrt{n}}\right)\right)
$$

now observe that we can rewrite

$$
\log\left(1-\frac{\underline{g}}{c\bar{g}\sqrt{n}}\right)=\frac{1}{1+x^*}\cdot\left(\frac{-\underline{g}}{c\bar{g}\sqrt{n}}\right)\leq\frac{-\underline{g}}{c\bar{g}\sqrt{n}}\quad\text{for some } x^*\in\left(\frac{-\underline{g}}{c\bar{g}\sqrt{n}},0\right).
$$

Putting things together,

$$
n^{K-1}\Big(1-\frac{\underline{g}}{c\bar{g}\sqrt{n}}\Big)^{n-K}\leq n^{K-1}\cdot\exp\left(\frac{-\underline{g}}{c\bar{g}\sqrt{n}}\cdot(n-K)\right)
$$

which converges to zero as required. $\square$

Choose $c=(3\bar{g}/\underline{g})^{1/2}$ and note that we can find $\underline{n}_1$ such that for all $n\geq\underline{n}_1$,

$$
\max\left\{\left[\frac{1}{c^2}+\frac{1}{c}n^{K-1}\Big(1-\frac{\underline{g}}{c\bar{g}\sqrt{n}}\Big)^{n-K}\right]\cdot\frac{\bar{g}}{\underline{g}},\frac{1}{c\sqrt{n}}\right\}\leq\frac{1}{2}
$$

Hence, deviations to $p<m/(c\sqrt{n})$ are unprofitable for all $n\geq\underline{n}_1$.

Case 2: Deviating to $p>\frac{m}{c\sqrt{n}}$ is unprofitable. The obedience constraint can be equivalently expressed as:

$$
\frac{(m-p)}{p}-\frac{n-1}{G(m)}\int_p^m I(q,m,n,K)g(q)dq\geq 0.
$$

Notice that this condition is satisfied for $p = m$. Hence, it is sufficient to show that

$$\frac{\partial}{\partial p}\left[\frac{(m-p)}{p} - \frac{n-1}{G(m)}\int_p^m I(q,m,n,K)g(q)dq\right] \leq 0$$

because this would guarantee that the obedience constraint is satisfied for any $p \in (m/(c/\sqrt{n}), 1]$. We show this next:

$$\begin{aligned}
&\frac{\partial}{\partial p}\left[\frac{(m-p)}{p} - \frac{n-1}{G(m)}\int_p^m I(q,m,n,K)g(q)dq\right] \\
&\qquad = -\frac{m}{p^2} + \frac{n-1}{G(m)} I(p,m,n,K)g(p) = \frac{1}{G(m)}\left[-\frac{mG(m)}{p^2} + (n-1)I(p,m,n,K)g(p)\right] \\
&\qquad \leq \frac{1}{G(m)}\left[-\frac{m\underline{g}m}{m^2} + (n-1)\binom{n-2}{K-2}\Big(1 - G(p)/G(m)\Big)^{n-K} g(p)\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad \left(\text{from } I(q,m,n,K) \leq \tbinom{n-2}{K-2}\Big(1 - G(q)/G(m)\Big)^{n-K}\right) \\
&\qquad \leq \frac{1}{G(m)}\left[-\underline{g} + n^{K-1}\Big(1 - \frac{\underline{g}}{c\bar{g}\sqrt{n}}\Big)^{n-K}\bar{g}\right]
\end{aligned}$$

From Claim 1, there exists some $\underline{n}_2$ such that for all $n \geq \underline{n}_2$

$$\frac{1}{G(m)}\left[-\underline{g} + \underbrace{n^{K-1}\Big(1 - \frac{\underline{g}}{c\bar{g}\sqrt{n}}\Big)^{n-K}\bar{g}}_{\to 0 \text{ by Claim 1}}\right] \leq 0,$$

which implies that for all $n \geq \underline{n}_2$

$$\frac{\partial}{\partial p}\left[\frac{(m-p)}{p} - \frac{n-1}{G(m)}\int_p^m I(q,m,n,K)g(q)dq\right] \leq 0$$

Combining the analyses for case 1 and case 2 we obtain that all price deviations are deterred for $n \geq n^{PPO} := \max\{\underline{n}_1, \underline{n}_2\}$ which completes the proof. □

The next Lemma 10 shows that we can obtain an arbitrarily good approximation of the no trade surplus point $NT = (0,0)$. This is done by picking the $K$ least favourite firms for each type's consideration set so that the total gains from trade under this matching scheme converges to zero.

**Lemma 10.** For all $\epsilon > 0$, there exists $\bar{n}_{\epsilon,K} > 0$ such that for all $n > \bar{n}_\epsilon$, there exists $\psi^{NT} \in \mathbf{\Psi}_{\geq K}$ in which for all equilibria, both consumer and producer surplus is upper-bounded by $\epsilon$ hence this approximates the $NT$ point $(0,0)$. $\lambda^{NT}$ takes the following form: For type $\boldsymbol{\theta}$,

(i) Matching: match consumer $\boldsymbol{\theta}$ to their $K$ least favourite firms i.e.,

$$\begin{aligned}
\phi^{NT}(\cdot|\boldsymbol{\theta}) &:= \text{marg}_S \lambda^{NT}(\cdot, S|\boldsymbol{\theta}) \\
&\quad \text{puts full probability on } \Big\{(n-K+1)_{\boldsymbol{\theta}}, (n-K+2)_{\boldsymbol{\theta}}, \ldots (n)_{\boldsymbol{\theta}}\Big\}
\end{aligned}$$

(ii) Information: send an arbitrary public message.

***Proof of Lemma 10.*** First observe that under the matching scheme specified in part (i) of Proposition 10, the expected gains from trade are given by the $n-K+1$th highest realization, $\theta_{n-K+1}$. As such,

$$\max\{CS, PS\} \leq \mathbb{E}[\theta_{n-K+1}] \leq \epsilon/2 + \mathbb{P}(\theta_{n-K+1} > \epsilon/2) \cdot 1 \leq \epsilon/2 + \underbrace{\binom{n}{K}(1-G(\epsilon/2))^{n-K}}_{\to 0 \text{ as } n \to +\infty.}$$

so pick $n_\epsilon^{NT}$ such that for all $n \geq n_\epsilon^{NT}$ the second term is $\leq \epsilon/2$ which completes the proof. □

Finally, Lemma 11 shows that the consumer optimal point can be implemented approximately through some design $\psi \in \mathbf{\Psi}_{\geq K}$.

**Lemma 11.** For all $\epsilon > 0$, there exists $\bar{n}_{\epsilon,K} > 0$ such that for all $n > \bar{n}_\epsilon$, there exists $\psi^{CO} \in \mathbf{\Psi}_{\geq K}$ which implements the equilibria with welfare outcome $(CS, PS)$ such that $CS \geq 1-\epsilon$, and $PS \leq \epsilon$. As such, this approximates the consumer optimal point $CO$. $\psi^{CO}$ takes the following form. For type $\boldsymbol{\theta}$,

(i) Matching: match consumer $\boldsymbol{\theta}$ to its $K$ most favourite firms i.e.,

$$\phi^{NT}(S|\boldsymbol{\theta}) := \text{marg}_S \lambda^{NT}(\cdot, S|\boldsymbol{\theta}) \text{ puts full probability on } \Big\{(1)_{\boldsymbol{\theta}}, (2)_{\boldsymbol{\theta}}, \dots (K)_{\boldsymbol{\theta}}\Big\}.$$

(ii) Information: for consumer $\boldsymbol{\theta}$ and consideration set $S$, send the public message $\boldsymbol{\theta}$ to all firms i.e., give firms full information.

***Proof of Lemma 11.*** Observe that there is a Bayes Correlated Equilibrium induced by the information structure where all firms but the consumer's favourite firm $(1)_{\boldsymbol{\theta}}$ charges a price of zero, and firm $(1)_{\boldsymbol{\theta}}$ charges the price $\theta_{(1)} - \theta_{(2)}$ and the consumer breaks ties in favour of her favourite firm.

First observe that conditioned on the highest draw $\theta_{(1)} =: m \geq \epsilon/6$, the probability that the second highest draw is greater than $\epsilon/6$ away is

$$\begin{aligned}
\mathbb{P}(\theta_{(1)} - \theta_{(2)} \geq \epsilon/6 \mid \theta_{(1)} = m) &= \mathbb{P}(n-1 \text{ independent draws are} \leq m - \epsilon/6 \mid \theta_{(1)} = m) \\
&= \left(\frac{G(m-\epsilon/6)}{G(m)}\right)^{n-1} \leq \left(\frac{G(m) - \underline{g}\epsilon/6}{G(m)}\right)^{n-1} \\
&\leq \left(1 - \underline{g}\epsilon/6\right)^{n-1}
\end{aligned}$$

which tends to zero as $n \to \infty$. Since this is true for all $m \in [\epsilon/6, 1]$, there exists $\bar{n}_{1,\epsilon}$ (which does not depend on $m$) such that for all $n \geq \bar{n}_{1,\epsilon}$ and all $m \in [\epsilon/6, 1]$,

$$\mathbb{P}(\theta_{(1)} - \theta_{(2)} \geq \epsilon/6 \mid \theta_{(1)} = m) \leq \epsilon/6.$$

Now applying the law of total expectation twice,

$$\mathbb{E}[\theta_{(1)} - \theta_{(2)}] = \mathbb{E}[\theta_{(1)} - \theta_{(2)} \mid \theta_{(1)} < \epsilon/6] \cdot \mathbb{P}(\theta_{(1)} < \epsilon/6)$$

$$
\begin{aligned}
&\quad + \mathbb{E}[\theta_{(1)} - \theta_{(2)} \mid \theta_{(1)} \geq \epsilon/6] \cdot \mathbb{P}(\theta_{(1)} \geq \epsilon/6) \\
&\leq \epsilon/6 + \mathbb{E}[\theta_{(1)} - \theta_{(2)} \mid \theta_{(1)} \geq \epsilon/6] \\
&\leq \epsilon/6 + \mathbb{P}(\theta_{(1)} - \theta_{(2)} < \epsilon/6 \mid \theta_{(1)} \geq \epsilon/6) \cdot \epsilon/6 \\
&\quad + \mathbb{P}(\theta_{(1)} - \theta_{(2)} \geq \epsilon/6 \mid \theta_{(1)} \geq \epsilon/6) \cdot 1 \\
&\leq \epsilon/6 + \epsilon/6 + \epsilon/6
\end{aligned}
$$

which implies that expected producer surplus is upper bounded by $\epsilon/2$ for all $n \geq \bar{n}_1$. It remains to obtain a lower bound on the total gains from trade, TS:

$$
\begin{aligned}
TS = \mathbb{E}[\theta_{(1)}] &\geq \mathbb{P}(\theta_{(1)} \geq 1-\delta) \cdot (1-\delta) \\
&= \underbrace{\Big(1 - (G(1-\delta))^n\Big)}_{\to 1 \text{ as } n \to +\infty} \cdot (1-\delta)
\end{aligned}
$$

and so pick $\bar{n}_{2,\epsilon}$ and so that for all $n \geq \bar{n}_{2,\epsilon}$, $TS \geq \epsilon/2$.

Finally, under the matching scheme, each type has her $K$ favourite firms in her consideration set, and we give full information to all firms, this also brings about the efficient outcome which implies that consumer surplus is just total surplus less producer surplus:

$$
CS = TS - PS \geq (1-\epsilon/2) - \epsilon/2 = 1 - \epsilon
$$

for all $n \geq n_\epsilon^{CO} := \max(\bar{n}_{1,\epsilon}, \bar{n}_{2,\epsilon})$ as required. □

We now conclude the proof of Proposition 6. From Lemma 9, there exists $n^{PO}$ so that for all $n \geq n^{PO}$, the producer-optimal point is implemented exactly. From Lemma 10, for any $\epsilon$, there exists $n_\epsilon^{NT}$ so that for all $n \geq n_\epsilon^{NT}$, $CS \leq \epsilon$, $PS \leq \epsilon$. From Lemma 11. there exists $n_\epsilon^{CO}$ so that for all $n \geq n_\epsilon^{CO}$, $CS \geq 1-\epsilon$ and $PS \leq \epsilon$. Observe that since CS is lower-bounded by zero, this implies that the lower envelope must be approximately linear. Finally, $SUR_K$ is convex. This implies that for all $n \geq \bar{n} := \max\{n^{PO}, n_\delta^{NT}, n_\delta^{CO}\}$ for an appropriately chosen $\delta$ (which depends on $\epsilon$),

$$
l\Big(SUR \setminus SUR_K\Big) \leq l\Big(PO - NT - CO \setminus SUR_K\Big) \leq \epsilon
$$

as required, where $PO - NT - CO$ is the triangle connecting the points $PO = (1,0)$, $NT = (0,0)$, and $CO = (0,1)$.[3] □

### III: Non-zero marginal costs

In the main text we normalized marginal costs to zero. Here we show that our results continue to obtain where marginal costs are positive and heterogeneous across firms.

Suppose that firms face constant costs $c := \{c_1, c_2, \ldots c_n\} \in [0,1]^n$ which is common-knowledge. Define the translation $\sigma_c : \mathbb{R}^n \to \mathbb{R}^n$ where

$$
\sigma_{-c}(x) = x - c := (x_1 - c_1, x_2 - c_2, \ldots x_n - c_n).
$$

[3] Since the gap between $SUR_K$ and $PPO - NT - PCO$ is upper-bounded by distance $\delta$ on all three edges of the triangle, picking $\delta = \epsilon/4$ would suffice.

Analogously, define the translation $\eta_{-c} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ where

$$\eta_{-c}((x, y)) = (\sigma_{-c}(x), \sigma_{-c}(y))$$

i.e., $\eta_{-c}$ is just a product-wise application of $\sigma_{-c}$. We will, as in the main text, take as a primitive the distribution $\mu \in \Delta(\Theta)$ over consumer types which represent their valuations for different firms. Let $\mu_{-c} := \mu \circ \sigma_{-c}^{-1}$ be the distribution over valuations of each type net of marginal costs (which can be different across products). Note that valuations net of marginal costs might now be negative but this is without loss because firms never price below marginal costs in equilibrium. Nonetheless we extend our type space to $\overline{\Theta} := \mathbb{R}^n$ and message space to $\overline{\boldsymbol{M}} := \mathbb{R}^n$.

For the joint distribution $\psi \in \Delta(\Theta \times \boldsymbol{M})$ over types and prices, define $\psi_{-c} \in \Delta(\overline{\Theta} \times \overline{\boldsymbol{M}})$ as

$$\psi_{-c} = \psi \circ \eta_{-c}^{-1}.$$

Observe $\text{marg}_{\overline{\Theta}} \psi_{-c} = \mu \circ \sigma_{-c}^{-1} = \mu_{-c}$.

**Proposition 7.** The joint distribution $\psi \in \Delta(\Theta \times [0, 1]^n)$ is implementable for distribution $\mu$ under marginal costs $c$ if and only if the joint distribution joint distribution $\psi_{-c} \in \Delta(\overline{\Theta} \times \overline{\boldsymbol{M}})$ is implementable for distribution $\mu_{-c} := \mu \circ \sigma_{-c}^{-1}$ under zero marginal costs.

*Proof.* Recall we defined $\psi(\cdot|m_i)$ as the conditional distribution over $\Delta(\Theta \times \boldsymbol{M}_{-i})$ given message $m_i$. As usual we will associate messages with price recommendations and without loss restrict our attention to obedient information structures (Bergemann and Morris, 2016). We will assume ties are zero-measure without loss, so write the sale condition as weak inequalities. Observe

$$\begin{aligned}(m_i' - c_i) &\int_{\theta_i - m_i' \geq \max\{0, \max_{j \neq i}(\theta_j - m_j)\}} d\psi(\cdot|m_i) \\ &= (m_i' - c_i) \int_{\theta_i - c_i - (m_i' - c_i) \geq \max\{0, \max_{j \neq i}(\theta_j - c_j - (m_j - c_j))\}} d\psi_{-c}(\cdot|m_i - c_i)\end{aligned}$$

so the expected payoff from charging $m_i'$ upon receipt of message $m_i$ under $\psi$ (with positive costs) and the expected payoff from charging $m_i' - c_i$ upon receipt of message $m_i - c_i$ under $\psi_{-c}$ (with zero costs) is identical.

Therefore if $\psi_{-c}$ is implementable under zero marginal costs then $\psi$ is implementable under marginal costs $c$. □

**Proposition 8.** Producer and consumer surplus under $\psi$ with marginal costs $c$ are identical to those under $\psi_{-c}$ with zero marginal costs.

*Proof.* Writing out definitions,

$$\begin{aligned}PS(\psi, c) &= \sum_{i \in \mathcal{N}} \int_{\theta_i - m_i \geq \max\{0, \max_{j \neq i}(\theta_j - m_j)\}} (\underbrace{m_i}_{\substack{\text{price} \\ \text{rec.}}} - c_i) d\psi \\ &= \sum_{i \in \mathcal{N}} \int_{\theta_i - c_i - (m_i - c_i) \geq \max\{0, \max_{j \neq i}(\theta_j - c_j - (m_j - c_j))\}} \underbrace{(m_i - c_i)}_{\substack{\text{price} \\ \text{rec.}}} d\psi_{-c} = PS(\psi_{-c}, 0)\end{aligned}$$

where the second equality follows from noting that the integral condition on either side are of equal measure by construction. A similar argument yields

$$TS(\psi, c) = \mathbb{E}(\max_i(\theta_i - c_i)) = TS(\psi_{-c}, 0)$$

so $CS(\psi, c) = CS(\psi_{-c}, 0)$ as required. $\square$

Together, Propositions 7 and 8 implies that the environment with consumer types distributed $\mu$ and positive marginal costs $c$ is isomorphic to the environment with consumer types distributed $\mu_{-c} := \mu \circ \sigma_{-c}^{-1}$ and zero marginal costs.

## IV: Producer optimality without efficiency

Recall we defined $PS^*$ in the main text as the maximum producer surplus among all information structures which induce efficient outcomes in the switching model. In this appendix we show that this coincides with the maximum amount of producer surplus across all finite information structures. In particular, Lemma 12 shows that finiteness and no mixing implies that prices are supported on $\{0, \gamma, L, H\}$. Lemma 13 then shows that producer surplus maximizing structures supported on $\{0, \gamma, L, H\}$ implies efficiency. This then implies the producer optimal design across all finite structures coincides with that across those constrained to implement efficient outcomes.

**Definition 6.** Information structure $\psi \in \Delta(\Theta \times M^2)$ is finite if

$$|\text{supp}\psi| < +\infty.$$

Finite information structures are supported on a finite set of price recommendations. With a finite message space, however, it is no longer true that studying obedient recommendations is without loss for characterizing the set of Bayes Correlated Equilibria. We will nonetheless focus on messages as price recommendations which implicitly implies that firms do not mix upon receipt of recommendations.

**Lemma 12.** For all finite and obedient information structures, $\int_{\Theta \times \{0,\gamma,L,H\}^2} d\psi = 1$.

*Proof of Lemma 12.* We once again have

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

which was shown in the proof of Lemma 7. This rules out positive-measure of ties away from the boundary where prices are zero. Now suppose that

$$\int_{\Theta \times M_1 \setminus \{H, L, \gamma, 0\} \times M_2} d\psi > 0$$

so firm 1 is recommended prices $m_1 \notin \{H, L, \gamma, 0\}$ on a positive measure set. Since $\Theta$ is finite and firm 1 must be making the sale sometimes otherwise it has a strictly positive downward deviation from such prices, or firm 2 is charging 0 on this set and it has a strictly positive upward deviation. We then have

$$\int_{\substack{\Theta \times M_1 \setminus \{H,L,\gamma,0\} \times M_2: \\ \gamma + m_2 > m_1 \\ \text{or } \min\{m_1,m_2\}=0}} d\psi > 0$$

where the inequality within the integral is strict by equation No ties which means that upon receipt of $m_1 \notin \{H, L, \gamma, 0\}$, the change in demand in response to a small increase in prices is zero. Then consider the $\epsilon$-upward deviation by firm 1 such that, upon receipt of the recommendation $m_1 \notin \{H, L, \gamma, 0\}$ firm 1 chooses $m_1 + \epsilon$. For small enough $\epsilon > 0$ this deviation is strictly profitable.

A symmetric argument follows for firm 2. We have shown that if the information structure is finite, only prices $\subseteq \{0, \gamma, L, H\}$ can be sustained. $\square$

From Lemma 12 we restrict our attention to symmetric information structures supported on the prices $\{H, L, \gamma, 0\} =: P$.

**Lemma 13.** If $\psi$ maximizes producer-surplus among all information structures supported on the prices $\{H, L, \gamma, 0\}^2$ then it induces an efficient outcome.

*Proof of Lemma 13.* Start with an information structure $\psi$ supported on $\{H, L, \gamma, 0\}^2$ which maximizes producer surplus. By the same argument in Lemma 7 (ii) it is without loss to suppose it is symmetric. Now suppose it is inefficient.

Suppose there is some type $\theta \in E_1$ which buys from firm 2 with positive probability: that is,

$$\psi_\theta(p_1, p_2) > 0 \quad \text{where} \quad \theta - p_1 < \theta - \gamma - p_2$$

for some $p_1, p_2 \in \{0, \gamma, L, H\}$. Letting $\theta'$ denote the symmetric counterpart of $\theta$ (e.g., if $\theta = L1$ then $\theta' = L2$), by the symmetry of $\psi$ we have $\psi_{\theta'}(p_2, p_1) = \psi_\theta(p_1, p_2) > 0$. We proceed by considering cases.

Case 1: $p_2 = \gamma$. Then consider the modification from $\psi$ to $\psi'$ such that

$$\psi'_\theta(\gamma, 0) = \psi_\theta(p_1, p_2) \quad \psi'_{\theta'}(0, \gamma) = \psi_{\theta'}(p_2, p_1).$$

and $\psi' = \psi$ for all other recommendations. In words, this reverses the misallocation taking the probability that type $\theta$ buys from firm 2 at price $\gamma$ and allocating the sale to firm 1.

It is easy to see that firm ICs remain fulfilled. To see this, first consider the price $p_1 \notin \{0, \gamma\}$. On the information structure $\psi$, upon receipt of $p_1$, firm 1 did not make the sale when the consumer's type is $\theta$. Hence, if $\psi$ fulfilled obedience at $p_1$, it continues to do so under $\psi'$. Next consider the price $p_1 = \gamma$. Observe that compared to the structure $\psi$, firm 1 now receives higher demand from obedience while the additional business stealing from downward deviations are identical. Furthermore, upward deviations from $\gamma$ were unprofitable under $\psi$ so must remain so under $\psi'$ since on the additional demand $\psi'_\theta(\gamma, 0)$, the consumer's IC is tight. Finally consider the price $p_1 = 0$ and observe that if firm 1 did not find it profitable to deviate upwards under $\psi$, it continues to find it unprofitable under $\psi'$.

Case 2: $p_2 = L$. Then the only possibility is $p_1 = H$ for the consumer of type $\theta \in E_1$ to buy from firm 2. We have two subcases. First suppose $\theta = L1$. Then we have $\psi_{L1}(H, L) > 0$ and the consumer buys from firm 1. Now peform the following modification from $\psi$ to $\psi'$ where

$$\psi'_{L1}(L, H) = \psi_{L1}(H, L) \quad \psi'_{L2}(H, L) = \psi_{L1}(L, H).$$

and $\psi' = \psi$ for all other recommendations. It is clear that deviations from prices $p_1 \neq L$ remain fulfilled since firm 1's obedience payoff from charging the price $L$ is unchanged, and the downward deviations weakly slacken. Upward deviations remain fulfilled since the consumer IC is tight. Next suppose $\theta = H1$. Then we have $\psi_{H1}(H, L) > 0$ and the consumer buys from firm 1 in the case $1 - \gamma - (1 - v) = v - \gamma > 0$. But in this case, the price $H$ cannot be sustained since for any symmetric structure $\psi$,

$$H \cdot \psi_{H1}(H, H) < (H - \gamma) \cdot 2\psi_{H1}(H, H)$$

since $\gamma < v \implies \gamma < 1/2$. Hence, $\psi_{H1}(H, L) = 0$. $\square$

## V: Parametrized environments when AIC holds

We will focus on duopoly case with symmetric distribution of valuations.

**V.1 Horizontal Differentiation: Uniform.** We start by considering the case of perfectly anti-correlated distributions over $[0, 1]$, where a consumer of type $\theta \in [0, 1]$ has preference $\theta$ for firm 1's product, and $1 - \theta$ for firm 2's product. It is straightforward to check that the uniform distribution over $[0, 1]$ fulfils the condition in Theorem 1:

$$\begin{aligned} H(\hat{\theta}) &= \int_{1/2}^{1} \underbrace{\frac{m - \hat{\theta}}{\hat{\theta}} f(m)}_{G(\hat{\theta}, m)} \, dm \\ &= \left[ \frac{1}{2\hat{\theta}} m^2 - m \right]_{1/2}^{1} \\ &= \frac{3}{8\hat{\theta}} - 1/2 > F(\hat{\theta}) := \int_{\hat{\theta}}^{1/2} dm = 1/2 - \hat{\theta} \quad \text{for all } \hat{\theta} \in [0, 1/2]. \end{aligned}$$

**V.2 Horizontal Differentiation: Normal.** We now consider the truncated normal distribution over $[0, 1]$ with mean $1/2$ and variance $\sigma^2$. In particular, letting $X \sim N(1/2, \sigma^2)$. The truncated variable $\bar{X}$ is distributed as the conditional distribution of $X$ conditional on $X \in [0, 1]$. Hence $\bar{X}$ has density:

$$\hat{f}(m) = \begin{cases} \dfrac{\frac{1}{\sigma}\phi(\frac{m - 1/2}{\sigma})}{S(\sigma)} & m \in [0, 1] \\ 0 & m \notin [0, 1] \end{cases}$$

where $S(\sigma) := \mathbb{P}[X \in [0, 1]] = \Phi(\frac{1}{2\sigma}) - \Phi(\frac{-1}{2\sigma})$, and $\Phi$ and $\phi$ are the CDF and PDF of a standard normal random variable respectively. We will be interested in computing the ranges of $\sigma^2$ under which the condition of Theorem 2 is fulfilled.

As a first observation, observe that as $\sigma \to \infty$, $\hat{f}(m) \to 1$ for all $m \in [0,1]$ which, we know from the calculations above, fulfils the condition in Theorem 2. We now turn to characterizing the values of $\sigma$ under which the condition of Theorem 2 remain fulfilled.

**Remark 1.** When valuations are perfectly anti-correlated and distributed as the truncated normal distribution with parameter $\sigma$, larger $\sigma$ corresponds to more polarised distributions of valuations.

*Proof.* By symmetry, it suffices to show that $\int_{\theta}^{1/2} \hat{f}(m)dm$ decreases in $\sigma$ for all $\theta \in [0,1/2]$ i.e., for all $\theta \in [1/2,1]$, the mass of consumers over $E_2$ who have preferences greater than $\theta$ for 1's product decreases. We show this through a direct calculation. Noting

$$\int_{\theta}^{1/2} \hat{f}(m)dm = \frac{\Phi(\frac{1/2-\theta}{\sigma}) - 1/2}{2[\Phi(\frac{1}{2\sigma}) - 1/2]}$$

and so

$$\frac{\partial}{\partial \sigma}\left(\int_{\theta}^{1/2} \hat{f}(m)dm\right) = \frac{A}{4\sigma^2[\Phi(\frac{1}{2\sigma}) - 1/2]^2}$$

where

$$A = (2\theta - 1)\phi\left(\frac{1/2-\theta}{\sigma}\right)\left[\Phi\left(\frac{1}{2\sigma}\right) - 1/2\right] + \phi\left(\frac{1}{2\sigma}\right)\left[\Phi\left(\frac{1/2-\theta}{\sigma}\right) - 1/2\right]$$

It remains to show $A \leq 0$ for all $\theta \in [0,1/2]$.

$$\frac{\partial A}{\partial \theta} = \phi\left(\frac{1/2-\theta}{\sigma}\right)B$$

$$B = \left(2\Phi\left(\frac{1}{2\sigma}\right) - 1\right)\left(1 - \frac{(1/2-\theta)^2}{\sigma^2}\right) - \frac{1}{\sigma}\phi\left(\frac{1}{2\sigma}\right)$$

where we used that $\phi'(x) = -x\phi(x)$ for all $x \in \mathbb{R}$. Now noticing that (i) $B$ is strictly increasing in $\theta$ for $\theta \in [0,1/2]$; (ii) $\phi > 0$, and (iii) $A|_{\theta=0} = A|_{\theta=1/2} = 0$, we have the result. $\square$

Now checking the condition in Theorem 2,

$$\begin{aligned}
H(\hat{\theta}) &= \int_{1/2}^{1}\left(\frac{m-\hat{\theta}}{\hat{\theta}}\hat{f}(m)\right)dm \\
&= \frac{1}{\hat{\theta}}\int_{1/2}^{1} m\hat{f}(m)dm - \int_{1/2}^{1}\hat{f}(m)dm \\
&= \frac{1}{\hat{\theta}}\int_{1/2}^{1} m\hat{f}(m)dm - 1/2 \\
&= \frac{1}{\hat{\theta}S(\sigma)}\int_{1/2}^{1}\frac{m}{\sigma}\phi\left(\frac{m-1/2}{\sigma}\right)dm - 1/2
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\hat{\theta}S(\sigma)}\int_0^{1/2\sigma}\left(y+\frac{1}{2\sigma}\right)\phi(y)\sigma dy - 1/2 && (y=(m-1/2)/\sigma)\\
&= \frac{\sigma}{\hat{\theta}S(\sigma)}\int_0^{1/2\sigma} y\phi(y)dy + \frac{1}{2\hat{\theta}S(\sigma)}\int_0^{1/2\sigma}\phi(y)dy - 1/2\\
&= -\frac{\sigma}{\hat{\theta}S(\sigma)}\int_0^{1/2\sigma}\phi'(y)dy + \frac{\Phi(\frac{1}{2\sigma})-\Phi(0)}{2\hat{\theta}S(\sigma)} - 1/2 && (x\phi(x) = -\phi'(x))\\
&= \frac{\sigma}{\hat{\theta}S(\sigma)}\left(\phi(0)-\phi(\frac{1}{2\sigma})\right) + \frac{1}{4\hat{\theta}} - 1/2 && \text{(Fundamental Theorem of Calculus)}
\end{aligned}
$$

which we compare against

$$
F(\hat{\theta}) = \int_{\hat{\theta}}^{1/2}\hat{f}(m)dm = S(\sigma)^{-1}\left(\Phi\left(\frac{1/2-\hat{\theta}}{\sigma}\right)-1/2\right).
$$

Then the condition in Theorem 2 is fulfilled if and only if $H(\hat{\theta}) \geq F(\hat{\theta})$ for all $\hat{\theta} \in [0,1/2]$ or, equivalently, if

$$
\frac{\sigma}{\hat{\theta}}\left(\phi(0)-\phi(\frac{1}{2\sigma})\right) + (\frac{1}{4\hat{\theta}}-1/2)S(\sigma) \geq \Phi\left(\frac{1/2-\hat{\theta}}{\sigma}\right) - 1/2 \quad \text{for all } \hat{\theta}\in[0,1/2]. \tag{2}
$$

Observe that

$$
\lim_{\sigma\to 0}\frac{\sigma}{\hat{\theta}}\left(\phi(0)-\phi(\frac{1}{2\sigma})\right) + (\frac{1}{4\hat{\theta}}-1/2)S(\sigma) = \frac{1}{4\hat{\theta}} - 1/2 < \lim_{\sigma\to 0}\Phi\left(\frac{1/2-\hat{\theta}}{\sigma}\right) - 1/2 = 1/2
$$

whenever $\hat{\theta} > 1/4$, so the producer-optimal outcome is never attainable since there is virtually no variation in preferences. Conversely, as $\sigma \to \infty$, both sides are zero in the limit and the condition is trivially fulfilled.

Now notice that the above remark implies that there is some $\bar{\sigma} > 0$ such that for all $\sigma \geq \bar{\sigma}$, the condition in Theorem 2 is fulfilled; conversely, for $\sigma < \bar{\sigma}$, the condition is not fulfilled. We now turn to solving for $\bar{\sigma}$.

Rewriting Equation 2,

$$
\sigma\left[\phi(0)-\phi\left(\frac{1}{2\sigma}\right)\right] + \frac{\Phi(\frac{1}{2\sigma})-1/2}{2} \geq C(\hat{\theta},\sigma)
$$

where

$$
C(\hat{\theta},\sigma) = \hat{\boldsymbol{\theta}}\left[\Phi\left(\frac{1/2-\hat{\theta}}{\sigma}\right)+\Phi\left(\frac{1}{2\sigma}\right)-1\right]
$$

$$
\frac{\partial C(\hat{\theta},\sigma)}{\partial\hat{\theta}} = \Phi\left(\frac{1/2-\hat{\theta}}{\sigma}\right)+\Phi\left(\frac{1}{2\sigma}\right)-1-\frac{\hat{\theta}}{\sigma}\phi\left(\frac{1/2-\hat{\theta}}{\sigma}\right)
$$

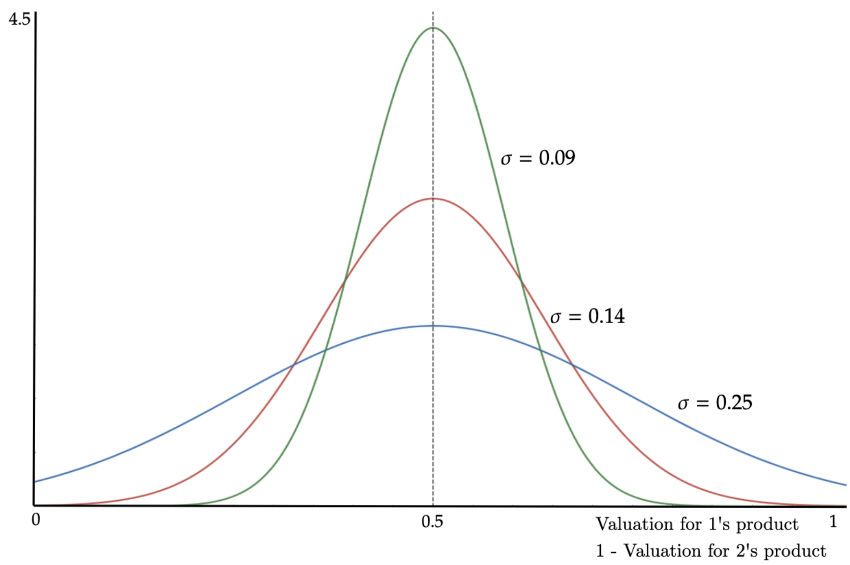
which is strictly decreasing in $\hat{\theta}$ for $\hat{\boldsymbol{\theta}} \in [0,1/2]$. Further noting $\frac{\partial C(\hat{\theta},\sigma)}{\partial\hat{\theta}}|_{\hat{\theta}=0} = 2\Phi(\frac{1}{2\sigma}) - 1 > 0$ and $\frac{\partial C(\hat{\theta},\sigma)}{\partial\hat{\theta}}|_{\hat{\theta}=1/2} = \Phi(\frac{1}{2\sigma}) - \Phi(0) - \frac{1}{2\sigma}\phi(0) < 0$, $C(\hat{\theta},\sigma)$ is maximized at the unique root of $\frac{\partial C(\hat{\theta},\sigma)}{\partial\hat{\theta}} = 0$.

Hence, $\bar{\sigma}$ is characterized by the equations:

$$\bar{\sigma}[\phi(0) - \phi(\frac{1}{2\bar{\sigma}})] + \frac{\Phi(\frac{1}{2\bar{\sigma}}) - 1/2}{2} = \theta^*[\Phi\left(\frac{1/2 - \theta^*}{\bar{\sigma}}\right) + \Phi(\frac{1}{2\bar{\sigma}}) - 1]$$
$$\Phi\left(\frac{1/2 - \theta^*}{\bar{\sigma}}\right) + \Phi(\frac{1}{2\bar{\sigma}}) - 1 - \frac{\theta^*}{\bar{\sigma}}\phi\left(\frac{1/2 - \theta^*}{\bar{\sigma}}\right) = 0.$$

Solving for the smallest $\bar{\sigma}$, we have $\bar{\sigma} \simeq 0.14$. The corresponding PDF of this distribution is shown in Figure 8 below (red line). For all $\sigma \geq \bar{\sigma}$, Theorem 2 is fulfilled (e.g., the blue line); for all $\sigma < \bar{\sigma}$, Theorem 2 is not (e.g., the green line).

Figure 8: Illustration of the truncated normal distribution




**V.3 Bivariate Uniform.** We now depart from the perfectly anti-correlated case and study how the condition in Theorem 2 interacts with the degree of variation in consumer preferences. To do so, we study the uniform distribution over $[a, b]^2$ with $a = 1/2 - \delta$, $b = 1/2 + \delta$, for $\delta \in (0, 1/2]$. In particular, for any $(\theta_1, \theta_2) \in [a, b]^2$, $f(\theta_1, \theta_2) = 1/4\delta^2$.

**Remark 2.** Since the model is invariant to preference rescaling, this is equivalent to a distribution with uniform preferences over

$$\left[\frac{1/2 - \delta}{1/2 + \delta}, 1\right]^2$$

where $\delta$ controls the degree of variation in preferences: when $\delta \to 0$, we approach the dirac delta on the point $(1, 1)$; when $\delta = 1/2$, we obtain the uniform distribution will full support over $[0, 1]^2$.

It turns out that in this setting, the condition in Theorem 2 is fulfilled if and only if $\delta = 1/2$. By direct calculation in Mathematica,

$$H(\hat{\theta}) - F(\hat{\theta}) = \frac{(10\delta - 4\hat{\theta} - 1)(2\delta - 2\hat{\theta} + 1)^2}{129\delta^2\hat{\theta}} \geq 0 \quad \text{for all } \hat{\theta} \in [a, b]$$

$$
\begin{aligned}
&\iff \min_{\hat{\theta}\in[a,b]} 10\delta - 4\hat{\theta} - 1 \geq 0 \\
&\iff \delta \geq 1/2.
\end{aligned}
$$

This implies that in the duopoly case with uniform and uncorrelated preferences, the intermediary is just able to structure information to achieve the producer-optimal outcome.