UNIVERSITY OF CAMBRIDGE
Faculty of Economics

CAMBRIDGE WORKING PAPERS IN ECONOMICS

Market segmentation through information

Matthew Elliott
Andrea Galeotti
Andrew Koh

Abstract

Prodigious amounts of data are being collected by internet companies about their users' preferences. We consider the information design problem of how to share this information with traditional companies which, in turn, compete on price by offering personalised discounts to customers. We provide a necessary and sufficient condition under which the internet company is able to perfectly segment and monopolise all such markets. This condition is surprisingly mild, and suggests room for regulatory oversight.

Reference Details

CWPE 2105
Published 14 January 2021

Key Words Information design, market segmentation, price discrimination
JEL Codes D43, D83, L13

Website www.econ.cam.ac.uk/cwpe

Market segmentation through information

Matthew Elliott, Andrea Galeotti and Andrew Koh*

November 19, 2020

Abstract

Prodigious amounts of data are being collected by internet companies about their users' preferences. We consider the information design problem of how to share this information with traditional companies which, in turn, compete on price by offering personalised discounts to customers. We provide a necessary and sufficient condition under which the internet company is able to perfectly segment and monopolise all such markets. This condition is surprisingly mild, and suggests room for regulatory oversight.

1 Introduction

The last two decades have witnessed the emergence of a new internet based business model where the value of the business is derived from the information it collects about its users.[1] The information collected by these internet companies is central to its offering—it permits advertising to be precisely targeted by enticing users to click through with special, personalized offers.

The extraordinary profitability of this business model has raised new antitrust concerns. Such concerns have thus far centred largely on the dominance of these internet companies *qua* internet companies e.g. the dominance of Google in the market for digital advertising.[2] By contrast, considerably less attention has been paid to the attendant effects on competition in traditional markets. Can the owner of information provide it in a way which creates rents by reducing competition to the detriment of consumers? Is this something that could realistically happen—and might be happening—across very many markets, or are the requisite conditions sufficiently demanding such as to remain beyond practical concern? Might the ability of such internet companies to create such rents explain their (colossal) valuations?[3]

We make progress on this problem under the assumption that the internet company's profits are increasing in the firm surplus obtained in traditional markets. This abstracts

*Elliott: Cambridge University, email mle30@cam.ac.uk; Galeotti: London Business School; Koh: Oxford University. First version: August 2019. We wish to thank Nageeb Ali, Ben Golub, Konstantin Guryev, Navin Kartik and Ludovic Renou. Jörg Kalbfuss provided excellent research assistance. Elliott acknowledges funding from the European Research Council under the grant EMBED #757229. Galeotti acknowledges funding from the European Research Council under the grant #724356.

[1] The revenues of these businesses are primarily sourced from traditional companies who pay for access to users (via advertisements) rather than users themselves, who typically enjoy free access to the internet company's services e.g. a platform to connect with friends.

[2] For instance, the US Justice Department is preparing to bring an antitrust lawsuit against Google over alleged anticompetitive practices in its advertising business.

[3] Britain's Competition and Markets Authority found that the cost of digital advertising for firms was worth £500 ($640) per household per year. While some of these costs are no doubt driven by the market concentration among internet companies, we explore a distinct but complementary force.

from strategic interactions between the internet company and traditional firms, and centres our analysis on the following information design problem: an information designer has perfect information about consumer preferences for differentiated products, and chooses what information to pass along to firms. After firms receive information, they compete by setting prices.

Theorem 1 provides a necessary and sufficient condition under which the information designer can induce perfect segmentation and monopolization of traditional markets. This condition is relatively mild, and when it is not fulfilled, we show that an intuitive alternate design can still extract significantly more surplus than the setting in which firms have no additional information about consumer preferences. Furthermore, the proof is constructive and it illustrates how such an information structure can be designed.

1.1 Example. We illustrate the main insights with a simple example. Consider the canonical model of horizontal differentiation in Hotelling (1929). Two competing firms, 1 and 2, are located, respectively, at the extreme points of the interval $[0, 1]$. A mass $1 - 2\eta$ of consumers is uniformly distributed on the interval. These are non-captive consumers who have unit demand and consumption value for either product normalised to 1. Horizontal differentiation is captured by linear transportation cost t so a consumer located at x pays transportation costs tx to purchase from firm 1 and $t(1-x)$ to purchase from firm 2. In addition, a mass of consumers $\eta \in [0, 1/2]$ are captive for each firm. These consumers have inelastic unit demand and a consumption value of 1 for the product of the firm they are captive to, and no consumption value for the product of the other firm.

Ex-ante, firms do not know the location of any particular consumer. But there exists an information designer who possesses this information. Can the information designer reveal information to the firms about the customers' preferences so that, in the ensuing pricing game (i) the equilibrium is efficient; and (ii) firms extract all the available surplus?

Simple Information Design. Consider the following information structure where firm 1 (2 resp.) receives precise information about the preferences of all non-captive consumers in $[0, 1/2]$ ($[1/2, 1]$ resp.), but receives no information for all other consumers (including their captive customers).

After receiving this information, both firms play a price setting game.[4] Given this information structure, consider the strategy where firm 1 (2 resp.) offers a personalised discount to customers located in $[0, 1/2]$ ($[1/2, 1]$ resp.) which just covers their transportation costs, and no discounts to all other customers (thus charging them a price of 1). Under this strategy profile, the market is perfectly segmented and monopolised: firms sell to their captive consumers at list price 1 and, at the same time, extract all the available surplus from the non-captive consumers located closer to them than the other firm.

When is this strategy profile an equilibrium? Consider, without loss of generality, firm 1. A crucial observation is that from firm 1's point of view, it is unable to distinguish consumers in the interval $[1/2, 1]$—those which are closer to firm 2—and its own captive consumers. This profile is an equilibrium when it prefers to charge price 1 to this bundle

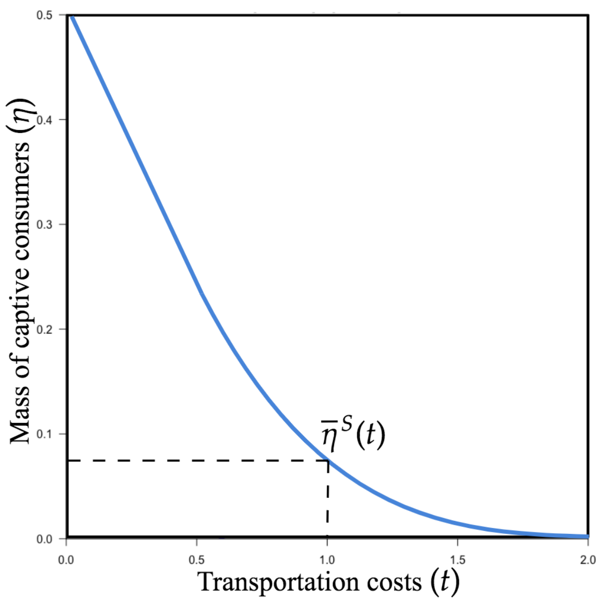
[4]We can think of firms as setting a uniform list price of 1 equal to valuations, and offering targeted discounts. Since it is dominated for firms to set any price greater than 1, this is without loss.

rather than offering a discount which would steal some consumers in the interval $[1/2, 1]$ at the cost of charging each consumer a lower price. As firm 2 extracts all surplus from consumers in the interval $[1/2, 1]$ by charing a price equal to their respective valuations, if firm 1 deviated and charged these consumers a price p, a consumer located at $(1-p)/t$ would be indifferent between buying from firm 1 and firm 2. Hence firm 1 does not have a profitable deviation to charge a price p instead of the price 1 if and only if

$$\eta \geq p\left(\eta + (1-2\eta)\left[\frac{1-p}{t} - \frac{1}{2}\right]\right).$$

When this obtains for all $p \in [0, 1/2]$, firm 1 will charge consumers in the interval $[1/2, 1]$ a price of 1, hence allowing firm 2 to extract the full surplus from such consumers. For every $t > 0$, there exists a threshold $\bar{\eta}^S(t) < 1/2$ such that if the fraction of captive consumers is above this threshold, the above strategy profile is an equilibrium; Figure 1 plots this function, and the formal analysis is provided in Supplementary Appendix II. A natural benchmark to consider is $t = 1$ i.e. both firms have gains from trade with every consumer in the interval $[0, 1]$. For this case, it is sufficient for each firm to have just 7.5% of the total mass of consumers captive.

Figure 1: The minimum mass of captive customers required for the simple information design described to induce perfect segmentation and monpolization of the market.



The information structure that we have proposed exhibits the following properties. First, it gives detailed and exclusive information to each firm about the subset of consumers the firm should sell to in an efficient allocation. Second, it matches a firms' captive customers to all consumers the firm is tempted, but for whom it is inefficient for it to sell to. Third, firms extract all consumer surplus from all its sales. As such, the information designer in effect recreates the standard differentiated product market monopoly problem by bundling consumers together: firms choose between extracting more surplus from high value consumers by excluding low value consumers ('intensive margin'), or extracting less surplus from more consumers by setting a lower price ('extensive margin'). The information designer sets up this problem so that the optimal choice for both firms is to exclude the low value customers, leading the market to be segmented.

1.2 The role of captive consumers. Our simple information design works in some, but not the widest possible range of settings e.g. for lowest possible mass of captive consumers, or lowest possible transportation costs. To see why, observe that the information design we considered only made use of captive consumers to hold down each firm's incentives to price more aggressively to the segment of consumers captured by its competitor. Our general characterization in Theorem 1 shows that information structures that are most generally able to attain efficiency and full extraction also utilize non-captive consumers to the same effect, though they share the key features documented above.[5]

1.3 Related Literature. While there is a rich literature on price discrimination from Pigou (1920), our paper contributes to a more recent strand studying how the informational environment interacts with consumer and firm surplus. Bergemann et al. (2015) studies price discrimination when a monopolist obtains additional information about consumer valuations and show that every combination of consumer and producer surplus in the 'surplus triangle' can be obtained by some segmentation. By contrast, we work in a n-firm setting in which firms obtain (generically different) additional information about valuations and compete for customers.[6] Introducing competition richens the setting,[7] but also makes the problem more demanding than the monopoly case by requiring consideration of an induced pricing game among firms.[8]

In a similar vein, Ali et al. (2020) consider a disclosure game in which a consumer chooses some verifiable information about her preferences to convey to competing firms. They show access to 'rich messages'—the ability to reveal only partial information—can in effect play firms against each other and intensify competition, thereby increasing consumer surplus. By contrast, we study the conditions under which suitably designed information structures can *weaken* competition, thereby increasing firm surplus. Further,

[5]Nevertheless, captive consumers will serve as a convenient benchmark in Section 4 to study how the conditions for perfect segmentation and monopolization vary with transportation costs, as well as compare the relative performance of different information structures.

[6]There is large literature on information sharing among oligopolists asking to what extent competing firms with partial information might share information via an intermediary (see, e.g. Novshek and Sonnenschein (1982), Vives (1988), Raith (1996)). Our setting differs considerably because (i) we explicitly model an information designer who chooses how information is partitioned among firms—by contrast, the intermediary in their setting simply facilitates the dissemination of information firms themselves choose to share; and (ii) we consider more granular information at the level of individual valuations rather than, say, about a stochastic demand function.

[7]The very same markets in which we might expect feasible personalized pricing e.g. online retailers are also those with low consumer search costs—and hence high degrees of competition.

[8]Armstrong and Vickers (2019) consider the effect on consumer welfare when sellers are able to discriminate against their respective captive consumers. They assume all firms observe a common signal about whether, and to which seller a consumer is captive to. By contrast, we emphasise the possibility—and limits—of information design in more general type spaces by allowing the designer to send firms generically different signals. Although we use captive consumers in the example to illustrate key forces, they are in general non-essential to our main result. See also Albrecht (2019) who considers a Bertrand duopoly setting.

we take consumer valuations as known to the information designer rather than disclosed strategically.[9]

While we focus on a setting in which firms are uncertain about consumer valuations, Roesler and Szentes (2017) study the converse problem in which consumers rather than firms have uncertain valuation. For the monopoly case, they characterise the signal structure which is best for consumers. Armstrong and Zhou (2019) extend this setting to the duopoly case, and characterise both firm-optimal and consumer-optimal signal structures.

More generally, our paper relates to a wider literature in information design. Bergemann and Morris (2013, 2016) consider general many-player settings and examine how the informational environment maps to resultant equilibria.[10] In the special case with a single receiver, Kamenica and Gentzkow (2011) show that 'concavification' of the designer's payoff as a function of receiver's posteriors generally binds the designer's maximum attainable utility and characterises the optimal signal structure. However, there are well-known difficulties applying such arguments to infinite-dimensional settings characteristic of differentiated products. We make progress by showing that it can be helpful to reframe a seemingly intractable information design problem as a matching problem, and believe this approach might be fruitfully applied in other contexts.

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), Lizzeri (1999), Taylor (2004) Calzolari and Pavan (2006), Bergemann and Bonatti (2015), Bergemann et al. (2018), Bergemann et al. (2019), Fainmesser and Galeotti (2019), Acemoglu et al. (2019), Jones and Tonetti (2020); also see Bergemann and Bonatti (2019) for a summary). These papers explicitly model the economic forces arising between strategic agents e.g. consumers' concern for privacy, firms' incentives to share data, or data intermediaries' ability to extract rent. By contrast, we abstract from strategic interactions and focus on establishing tight conditions under which rich information structures allows the information designer to segment the market in a way that firms can extract all efficient trading gains.[11]

2 Model

2.1 Setup. There is finite set of firms, indexed $\mathcal{N} = \{1, \dots, n\}$ which each produces a single product. There is a unit mass of consumers distributed on $\Theta = [0, 1]^n$. Each consumer demands a single unit inelastically and type $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n) \in \Theta$ obtains value $\theta_i \in [0, 1]$ from purchasing from firm i. To keep track of which consumers value which product most highly, we partition the type space Θ into n regions, $E_1, \dots, E_n$, where $E_i := \{\boldsymbol{\theta} \in \Theta : \max(\theta_1, \theta_2, \dots, \theta_n) = \theta_i\}$ are the types of consumers who value firm

[9]While we view this as a realistic assumption, it will be evident that whenever a firm-optimal equilibrium is possible, all firms find it optimal to set a list price equal to the highest possible valuation so personalized pricing is conducted via personalized discounts. As such, insofar as disclosure is verifiable, consumers have incentive to report their valuations to enjoy lower prices e.g. by enabling their cookies.

[10]In the language of Bergemann and Morris (2013, 2016), our problem can be restated as finding necessary and sufficient conditions for a Bayes Correlated Equilibria which fulfils efficiency and full surplus extraction.

[11]While this is somewhat more narrow, our results can be naturally nested within a broader framework which captures strategic interactions.

i's good most. Although we focus on the interpretation with a continuum of consumers throughout this paper, all results can be translated into an alternate setting with a single consumer of uncertain type.

Firms choose prices from a discrete price grid with k increments between the minimal product valuation of 0 and maximum valuation of 1. The discrete price grid serves two purposes. It simplifies the exposition by eliminating the need to compare zero measure sets, and removes concerns that results we would obtain in a continuous version of the model do not correspond to the limit of a discrete model. For completeness, we solve the continuous version of the model in Appendix B and show that an exact analog of our main result continues to hold.

Let $V_k = \{0, \delta, \ldots, (k-1)\delta, k\delta = 1\}$ denote the available prices. Assuming it is not possible to make consumers commit to paying a random price, we restrict our attention to the coarsening of consumer valuations for a given product i to $v_i(\theta_i) := \max\{v_i \in V_k : v_i \leq \theta_i\}$ which is the highest price a consumer of type θ is willing to pay for firm i's product. Denote $\boldsymbol{v}(\boldsymbol{\theta}) := (v_1(\theta_1), \ldots v_n(\theta_n))$ which we call the consumer's *effective type*. Finally, let $f(\boldsymbol{v}(\boldsymbol{\theta})) : V_k^n \to [0,1]$ denote the mass of consumers with valuations in the space $[v_1, v_1 + \delta) \times [v_2, v_2 + \delta) \times \ldots \times [v_n, v_n + \delta) \subseteq \Theta$.[12] When it is unlikely to cause confusion, we will often abuse notation and use $\boldsymbol{v}(\boldsymbol{\theta}) \in \Theta'$ to denote $\{\boldsymbol{v}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta'\}$ where $\Theta' \subseteq \Theta$. Finally, to simplify the exposition, we assume consumers are distributed atomlessly with full support over Θ which also implies f has full support over V_k^n.[13] We have thus coarsened the type space Θ to the effective type space V_k^n where there is a positive mass of consumers for each effective type.

An *information designer*—an internet company, say—chooses how to distribute information about consumer preferences across firms. We make the following assumption on the designer's payoff.

Assumption 1. The information designer's payoff is increasing in total firm surplus.

For each firm, the information designer commits in advance to a signal structure it will provide about each type of consumer. For each firm $i \in \{1, \ldots, n\}$, the information designer chooses a message function

$$\psi_i : \Theta \to \Delta(M)$$

where $M \supseteq V_k$ is the message space and $\Delta(M)$ is the set of feasible probability distributions over M. Call $m \in M$ a message realisation. Denote the set of all sets of message functions with $\boldsymbol{\Psi}$ where $\boldsymbol{\psi} := (\psi_i)_{i=1}^n$ is a typical member. The information designer learns the exact valuation of each consumer for each product,[14] and releases messages to each firm $i \in \{1, \ldots, n\}$ in accordance with the message function. Given the messages received for each consumer, firms then play a pricing game: a pricing function for firm $i \in \{1, \ldots, n\}$ is $p_i : M \to V_k$. Let P_i be the space of possible pricing functions for firm i. A strategy for firm i is $\sigma_i : M \to \Delta(P_i)$.

[12] With the exception of $v_i = 1$, in which case we simply consider the point $\{1\}$ on the ith product.

[13] This is in general not necessary.

[14] Note that it does not matter whether the information designer learns about each consumer's type before or after committing the message functions.

Our goal is to characterize conditions under which the information designer is able to choose message functions such that, in the resultant subgame, each consumers of type $\boldsymbol{\theta} \in \Theta$:

C1 **(Efficient allocation)** buys from firm i if and only if $\boldsymbol{\theta} \in E_i$; and

C2 **(Full Surplus Extraction)** pays $\max \boldsymbol{v}(\boldsymbol{\theta})$.

Note that in practice, internet firms can either provide information to supplement firms' existing access to consumers, or provide both information and access to consumers.[15] We focus on the former case in which firms can make personalized offers to consumers without depending on the information designer e.g. through discount codes via customer mailing lists. This can be interpreted as the worst case for the information designer. In Proposition 2, we consider the possibility that the information designer obtains the ability to segment access in addition to designing information and show that this weakens the conditions under which an equilibrium fulfilling conditions C1 and C2 can be induced.

Let $\boldsymbol{\Gamma}^*$ denote the set of induced games in which there exists an equilibrium satisfying conditions C1 and C2, and let

$$\boldsymbol{\Psi}^* := \{\boldsymbol{\psi} \,:\, \Gamma(\boldsymbol{\psi}) \in \boldsymbol{\Gamma}^*\}$$

be the set of message functions that the information designer can use to fulfil both conditions. We now turn our attention towards characterising exact conditions under which an information designer is able to achieve conditions C1 and C2 i.e. $\boldsymbol{\Psi}^* \neq \emptyset$. When $\boldsymbol{\Psi}^* \neq \emptyset$ we say that the information designer can perfectly segment the market.

3 Characterisation

In order to satisfy conditions C1 and C2 we have to sell all consumers their most preferred product and charge each consumer the highest possible price they would be willing to pay for this product. Thus each firm i must learn the effective valuation $\bar{v}_i$ for each consumer of type E_i. Hence we can, without loss of generality, fix the message received for such consumers with $\bar{v}_i$. This yields the natural interpretation of messages as price recommendations.

If firm i simply receives information about the valuations of consumers in E_i, it can do better by additionally capturing consumers it is not meant to sell to under conditions C1 (those of types $\Theta \setminus E_i$). This is because conditions C1 and C2 imply such consumers are being charged prices equal to their maximum effective valuation, the maximum surplus for such a consumer is δ. As such, firm i can generically offer such consumers discounts to undercut their competitor, thereby violating condition C2. In order to prevent firms from making such sales, we need to assign consumers of types $\Theta \setminus E_i$ (those firm i should *not* sell to) messages also shared by consumers of types E_i (those firm i should sell to). We do so with the following matching function:[16]

$$g_i(\boldsymbol{v}'(\boldsymbol{\theta}')|\bar{v}_i) := \psi_i(m = \bar{v}_i|\boldsymbol{v}'(\boldsymbol{\theta}'))f(\boldsymbol{v}'(\boldsymbol{\theta}')),$$

[15] See Bergemann and Bonatti (2019) for a discussion of this distinction.

[16] We have $g_i : \{v_i(\theta_i) : \boldsymbol{\theta} \in E_i\} \to \Delta\{\boldsymbol{v}(\boldsymbol{\theta}) \in \Theta \setminus E_i\}$ so this is unique only up to types.

which denotes the mass of consumers of effective type $\boldsymbol{v}'(\boldsymbol{\theta}') \in \Theta \setminus E_i$ which are matched to types $\{v_i(\theta_i) = \bar{v}_i : \boldsymbol{\theta} \in E_i\}$—and hence for whom firm i receives message $\bar{v}_i$.

A first observation is that if $v_i' > \bar{v}_i$, then $g(\boldsymbol{v}'|\bar{v}_i) = 0$ because after receiving message $\bar{v}_i$, condition C2 requires that firm i sets price $\bar{v}_i$. But given that the pricing of other firms satisfies condition C2, consumers of type $\boldsymbol{v}'(\boldsymbol{\theta}) \in \Theta \setminus E_i$ are only obtaining surplus between $[0, \delta)$, and would prefer to buy from firm i to obtain surplus greater than δ thereby violating condition C1. We thus restrict our attention to matching functions which fulfil this condition.

Next, we need to ensure that the matching functions do not incentivize any firm i to deviate and charge a price strictly lower that v_i to capture some of the extra consumers after receiving a given message realisation. Conditions C1 and C2 require that for all firms i, this needs to hold for all possible messages $\bar{v}_i \in V_k$ and all possible deviations given a message $\bar{v}_i$ to some price $\hat{v}_i < \bar{v}_i$.

However, as an artefact of our discretization, the case where $v_i' = \bar{v}_i$ requires some care. Consider a consumer of effective type in $\{\boldsymbol{v}'(\boldsymbol{\theta}') \in E_j : v_i' = \bar{v}_i\}$. By condition C2, they must be receiving surplus $s_j \in [0, \delta)$ from buying from firm $j \neq i$. However, since $v_i' = \bar{v}_i$, they also receive positive surplus $s_i \in [0, \delta)$ from buying from firm i. Whether $s_i > s_j$ or $s_j \geq s_i$ will depend on the consumer's exact valuation given by their underlying type on the continuous space Θ. We will approach the problem by first supposing all such consumers have valuations $s_i > s_j$, and hence buy from firm i. This maximally restricts the matching which will be possible without violating incentive compatibility. We then suppose all such consumers have valuations $s_i < s_j$, and hence buy from firm $j \neq i$. This will introduce some slack into our incentive compatibility constraints.

First we write down the tighter incentive compatibility constraints which can be viewed as a sufficient condition for incentive compatibility. We require for all $\bar{v}_i \in V_k$,

$$\underbrace{\bar{v}_i \sum_{\substack{\boldsymbol{v}' \in E_i: \\ v_i' = \bar{v}_i}} f(\boldsymbol{v}')}_{\text{Profits from charging } \bar{v}_i} \geq \underbrace{\hat{v}_i \left(\sum_{\substack{\boldsymbol{v}' \in E_i: \\ v_i' = \bar{v}_i}} f(\boldsymbol{v}') + \overline{G}_i(\hat{v}_i|\bar{v}_i) \right)}_{\text{Profits from charging } \hat{v}_i < \bar{v}_i} \tag{$\overline{\boldsymbol{IC}}$}$$

where

$$\overline{G}_i(\hat{v}_i|\bar{v}_i) := \sum_{\substack{\boldsymbol{v}' \in \Theta \setminus E_i: \\ v_i' \geq \hat{v}_i}} g(\boldsymbol{v}'|\bar{v}_i)$$

is an upper bound on the additional mass of consumers obtained by deviating to setting price $\hat{v}_i \leq \bar{v}_i$.

Next we write down the looser incentive compatibility constraint which can be viewed as a necessary condition for incentive compatibility. Formally, we require for all $\bar{v}_i \in V_k$,

$$\underbrace{\bar{v}_i \sum_{\substack{\boldsymbol{v}' \in E_i: \\ v_i' = \bar{v}_i}} f(\boldsymbol{v}')}_{\text{Profits from charging } \bar{v}_i} \geq \underbrace{\hat{v}_i \left(\sum_{\substack{\boldsymbol{v}' \in E_i: \\ v_i' = \bar{v}_i}} f(\boldsymbol{v}') + \underline{G}_i(\hat{v}_i|\bar{v}_i) \right)}_{\text{Profits from charging } \hat{v}_i < \bar{v}_i} \tag{$\underline{\boldsymbol{IC}}$}$$

where

$$\underline{G}_i(\hat{v}_i|\bar{v}_i) := \sum_{\substack{\boldsymbol{v}' \in \Theta \setminus E_i: \\ v'_i > \hat{v}_i}} g(\boldsymbol{v}'|\bar{v}_i)$$

is a lower bound on the additional mass of consumers obtained by deviating to setting price $\hat{v}_i \leq \bar{v}_i$.

Finally, the total mass of consumers matched to some message v_i must be consistent with the actual distribution of consumers. Formally, for all firms i and all effective types $\boldsymbol{v}' \in \{V_k^n : v'_i > 0\}$,

$$\sum_{\bar{v}_i \in V_k} g(\boldsymbol{v}'|\bar{v}_i) = f(\boldsymbol{v}'). \qquad \textbf{(Consistency)}$$

To see this, note that if $\sum_{v_i} g(\boldsymbol{v}'|v_i) < f(\boldsymbol{v}')$, firm i can in effect infer the effective types of some consumers of type $\boldsymbol{v}' \in \Theta \setminus E_i$ because these consumers have not been garbled together with consumers of effective types E_i. There then exists some price deviation firm i can offer to capture these consumers, thereby violating conditions C1 and C2. These observations on what message functions must look like to induce an equilibrium fulfilling C1 and C2 are formalised in the following Lemma.

Lemma 1.

(i) The information designer can perfectly segment the market if there exist messages $\boldsymbol{\psi} \in \boldsymbol{\Psi}$ such that for all firms $i \in \{1, \dots, n\}$, $\overline{\boldsymbol{IC}}$ and **Consistency** are fulfilled.

(ii) The information designer cannot perfectly segment the market if there does not exist messages $\boldsymbol{\psi} \in \boldsymbol{\Psi}$ such that for all firms $i \in \{1, \dots, n\}$, $\underline{\boldsymbol{IC}}$ and **Consistency** are fulfilled.

The proof of Lemma 1 is deferred to Appendix A. It proceeds by arguing $\overline{\boldsymbol{IC}}$ and **Consistency** are by construction sufficient to induce an equilibrium in the subsequent subgame fulfilling conditions C1 and C2. By $\overline{\boldsymbol{IC}}$, every firm finds it optimal to obey the pricing recommendation given by each of its message realisations. Further, **Consistency** ensures that in the process of setting these prices to extract the full surplus from consumers in E_i, firm i has, in the process, also set prices $\bar{v}_i > v'_i$ for nearly all types $\boldsymbol{v}' \in \Theta \setminus E_i$.[17] But for such prices, consumers of type $\boldsymbol{v}' \in E_j$, $j \neq i$ obtain strictly negative profits from purchasing from firm i, and so can do no better than purchasing from firm j at price v_j. In this regard, the matching function g_i has successfully garbled all consumers of types $\Theta \setminus E_i$ together with those in E_i, but in an incentive compatible manner. Further, both $\underline{\boldsymbol{IC}}$ and **Consistency** are necessary. $\underline{\boldsymbol{IC}}$ imposes a lower bound on firm i's additional surplus from deviating to a lower price. If this is violated, then firm i has a profitable deviation for some positive mass of consumers. **Consistency** requires that the matching function g_i successfully garbles types $\Theta \setminus E_i$ together with those in E_i. If this is not fulfilled, then this implies some positive mass of consumers of types $\Theta \setminus E_i$ are assigned message $M \setminus V_k$. But if so, firm i can infer the identities of these consumers, and will find it profitable to undercut its competitors.

[17] Except for types which have zero valuation for firm i's product.

Lemma 1 pins down what, if exists, first-best message functions should look like. We now turn to the question of whether such message functions can be constructed. Define the following class of functions which binds both $\overline{\boldsymbol{IC}}$ and $\underline{\boldsymbol{IC}}$ for all possible price deviations.

$$G_i^*(\hat{v}_i|\bar{v}_i) := \begin{cases} \left(\dfrac{\bar{v}_i - \hat{v}_i}{\hat{v}_i}\right) \sum\limits_{\substack{\boldsymbol{v}' \in E_i: \\ v_i' = \bar{v}_i}} f(\boldsymbol{v}') & \text{if } \bar{v}_i > \hat{v}_i \\ 0 & \text{otherwise.} \end{cases}$$

It will also be helpful to define

$$H_i^*(\hat{v}_i) := \sum_{v_i \in V_k} G_i^*(\hat{v}_i|\bar{v}_i)$$

which is the total mass of consumers with valuations greater than or equal to $\hat{v}_i$ which have also been matched to consumers of types E_i. We are now ready to state our main result.

Theorem 1.

(i) The information designer can perfectly segment the market if for all firms $i \in \{1, \ldots, n\}$ and all consumer valuations $\hat{v}_i \in V_k$

$$H_i^*(\hat{v}_i + \delta) \geq \sum_{\substack{\boldsymbol{v}' \in \Theta \setminus E_i: \\ v_i' \geq \hat{v}_i}} f(\boldsymbol{v}')$$

(ii) The information designer cannot perfectly segment the market if there exists a firm $i \in \{1, \ldots, n\}$ and consumer valuation $\hat{v}_i \in V_k$ such that

$$H_i^*(\hat{v}_i) < \sum_{\substack{\boldsymbol{v}' \in \Theta \setminus E_i: \\ v_i' \geq \hat{v}_i}} f(\boldsymbol{v}')$$

(iii) In the limit as the price grid becomes fine, the sufficient condition from part (i) and necessary condition from part (ii) for perfect market segmentation coincide

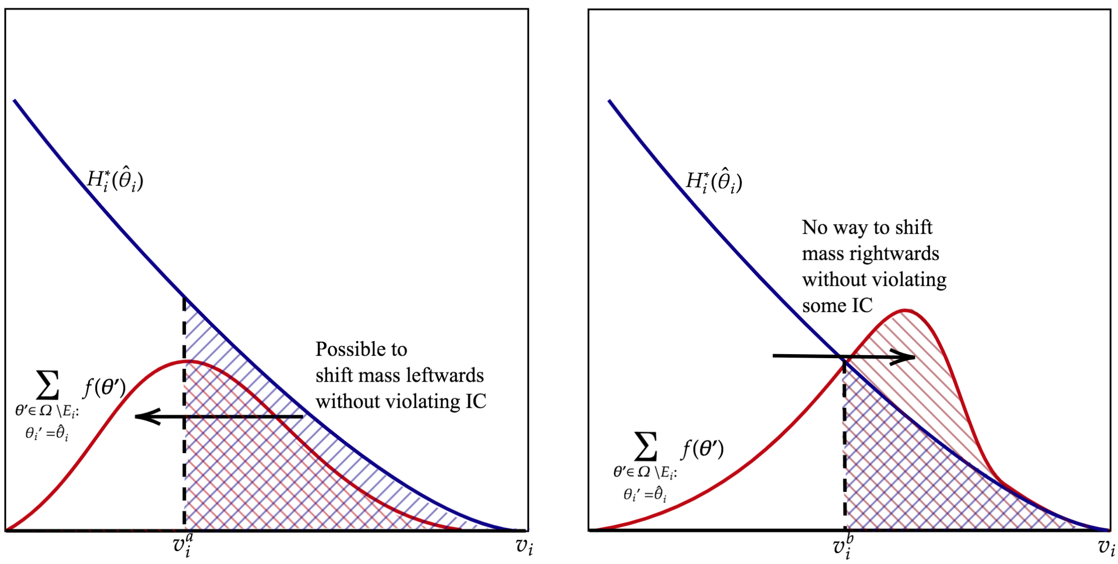
$$\lim_{k \to \infty} H_i^*(\hat{v}_i + \delta; k) = \lim_{\delta \to 0} H_i^*(\hat{v}_i + \delta) = H_i^*(\hat{v}_i).$$

Theorem 1 offers sufficient and necessary conditions for there to exist message functions which can induce an equilibrium fulfilling C1 and C2 in the resultant subgame. Part (i) states that if the translated function $H_i^*(\hat{v}_i + \delta)$ is first-order stochastically dominated by the actual distribution of consumers f, this is sufficient for there to exist such message functions. The translation by δ corresponds to the 'slack' obtained from assuming that all indifference in effective valuations between i and j for a type in E_j is broken in favour of firm i. Part (ii) offers a partial converse, and states that if $H_i^*(\hat{v}_i)$ is not first-order stochastically dominated by the actual distribution of consumers f, there does not exist any such message function. Condition (iii) states that these two conditions are identical as we make the price grid arbitrarily fine i.e. $\delta \to 0$.

Before we prove this, Figure 2 below illustrates the key underlying ideas. We show the limit case as the price grid becomes fine for simplicity of exposition. In both panels, the blue line labelled $h_i^*(\hat{v}_i)$ represents the total mass of consumers of effective types $\boldsymbol{v} \in \Theta \backslash E_i$ assigned message $\hat{v}_i \in V_k$ (viewed from the ith axis).[18] The red lines represents the actual mass of consumers of effective types $\Theta \setminus E_i$ with valuations $\hat{v}_i$ for firm i's product.

Panel (a) illustrates how a distribution fulfilling the conditions in Lemma 1 can be constructed when the condition in Theorem 1 is fulfilled. Note that for every $v_i^a \in V_k$, the mass of consumers in $\Omega \setminus E_i$ assigned some message greater than v_i^a exceeds the total mass actual consumers with valuations of at least v_i^a for firm i's product. As such, the designer can perform 'leftward' shifts of the mass of consumers matched to E_i to satisfy the condition that the mass of matched consumers must be consistent with the actual distribution. Doing so simultaneously ensures all incentive compatibility constraints continue to be satisfied. Conversely, panel (b) shows that when the condition in Theorem 1 is unfulfilled, it is impossible to construct a distribution fulfilling both **Consistency** and ***IC***. At v_i^b for instance, the mass actual consumers with valuations at least v_i^b for firm i's product (shaded red area) exceeds the total mass of consumers in $\Omega \setminus E_i$ assigned some message of at least v_i^b (shaded blue area). As such, since H_i^* has been constructed such that all incentive compatibility constraints are tight, any 'rightward' shift of the distribution to satisfy consistency must entail violating at least one incentive compatibility constraint.

Figure 2: Illustration of Theorem 1



(a) (b)

Proof of Theorem 1. We first show that if the condition in part (i) of Theorem 1 is fulfilled, $\boldsymbol{\Psi^*} \neq \emptyset$. In view of Lemma 1, it is sufficient to show there exists a set of

[18] We can write this explicitly as $h_i^*(\hat{v}_i) = (H_i^*(\hat{v}_i + \delta) - H_i^*(\hat{v}_i))/\delta$.

matching functions fulfilling incentive compatibility and consistency. We show this can be constructed by first considering the set of matching functions $(\overline{G_i^*})_{i=1}^n$ where $\overline{G_i^*} = G_i^*$ for each i. By construction, this binds the incentive compatibility constraint everywhere when indifferences in effective valuations are settled in favour of firm i. However, this matching might violate consistency. But since the condition in Theorem 1 obtains, $H^*(\hat{v}_i + \delta) \geq \sum_{\Theta \setminus E_i : v_i \geq \hat{v}_i} f(\boldsymbol{v})$ and we can find an alternate set of matching functions $(\overline{G_i'})_{i=1}^n$ through a series of 'leftward' shifts such that

(i) $\overline{G_i'}(\hat{v}_i|\bar{v}_i) \leq \overline{G_i^*}(\hat{v}_i|\bar{v}_i)$ for all $\bar{v}_i \in V_k$ and all $\hat{v}_i \in \{v \in V_k \,:\, v \leq \bar{v}_i\}$; and

(ii) $\overline{H_i}(\hat{v}_i + \delta) = \sum_{\Theta \setminus E_i : v_i \geq \hat{v}_i} f(\boldsymbol{v})$ for all $\hat{v}_i \in V_k$.

Since $\overline{G_i^*}(\hat{x}_i|\bar{x}_i)$ binds the incentive compatibility constraint everywhere, (i) ensures they remain satisfied for all effective types $\bar{v}_i \in V_k$ since each potential deviation can only be less profitable under $\overline{G_i'}$; (ii) implies for all $\hat{v}_i \in V_k$,

$$\overline{H_i'}(\hat{v}_i + \delta) := \sum_{\bar{v}_i \in V_k} \overline{G_i'}(\hat{v}_i + \delta|\bar{v}_i) := \sum_{\bar{v}_i \in V_k} \sum_{\substack{\boldsymbol{v'} \in \Theta \setminus E_i : \\ v_i' + \delta > \hat{v}_i}} g(\boldsymbol{v'}|\bar{v}_i) = \sum_{\Theta \setminus E_i : v_i \geq \hat{v}_i} f(\boldsymbol{v})$$

which is sufficient to fulfil consistency since (ii) implies we can find a corresponding matching such that $\sum_{\bar{v}_i \in V_k} g_i(\boldsymbol{v'}|\bar{v}_i) = f(\boldsymbol{v'})$. Intuitively, we can 'spread' the mass of matched consumers over the remaining $n-1$ dimensions until consistency is fulfilled. Since we can always perform this procedure for all firms $i \in \{1, \ldots, n\}$, the condition in Theorem 1 (i) is sufficient.

We now show that if the condition in part (ii) of Theorem 1 is fulfilled, $\boldsymbol{\Psi^*} = \emptyset$. Now consider the matching $(\underline{G_i^*})_{i=1}^n$ where $\underline{G_i^*} = G_i^*$ for each i. By construction, this binds the incentive compatibility constraint everywhere when indifferences in effective valuations are broken in favour of firm $j \neq i$. By hypothesis, there exists some firm $i \in \mathcal{N}$ and some $\hat{v}_i \in V_k$ such that

$$H_i^*(\hat{v}_i) < \sum_{\substack{\boldsymbol{v'} \in \Theta \setminus E_i : \\ v_i' \geq \hat{v}_i}} f(\boldsymbol{v'})$$

which implies that under the matching corresponding to $\underline{G_i^*}$,

$$H_i^*(\hat{v}_i) := \sum_{\bar{v}_i \in V_k} \underline{G_i^*}(\hat{v}_i|\bar{v}_i) := \sum_{\bar{v}_i \in V_k} \sum_{\substack{\boldsymbol{v'} \in \Theta \setminus E_i : \\ v_i' \geq \hat{v}_i}} g(\boldsymbol{v'}|\bar{v}_i) < \sum_{\substack{\boldsymbol{v'} \in \Theta \setminus E_i : \\ v_i' \geq \hat{v}_i}} f(\boldsymbol{v'})$$

hence

$$\sum_{\bar{v}_i \in V_k} g(\boldsymbol{v'}|\bar{v}_i) < f(\boldsymbol{v'})$$

thereby violating consistency. We therefore require an alternate function $\underline{G_i}'$ such that $\underline{G_i}'(\hat{v}_i|\bar{v}_i) \geq \underline{G_i}^*(\hat{v}_i|\bar{v}_i)$ with strict inequality for some $\bar{v}_i$, $\hat{v}_i$. But recall $\underline{G_i}^* = G_i^*$ hence the incentive compatibility constraint is tight everywhere. Under $\underline{G_i}'$, there will then exists some $\bar{v}_i$, $\hat{v}_i$ for which it is broken, thereby violating condition $\underline{IC}$. It is thus impossible to simultaneously satisfy consistency and $\underline{IC}$ so by Lemma 1 (ii), we conclude $\boldsymbol{\Psi^*} = \emptyset$.

Finally, part (iii) of Theorem 1 follows because the distribution of types over Θ is atomless. □

3.1 Comparative Statics. To avoid introducing addition notation, we state our comparative statistics in the limit case where $\delta \to 0$ though these results apply more generally.

Proposition 1. Consider two distributions $\tilde{f}$ and f such that

(i) For all $i \in \mathcal{N}$,

$$\sum_{\substack{\boldsymbol{v}' \in E_i: \\ v_i' \geq \hat{v}_i}} f(\boldsymbol{v}') \leq \sum_{\substack{\boldsymbol{v}' \in E_i: \\ v_i' \geq \hat{v}_i}} \tilde{f}(\boldsymbol{v}') \quad \text{for all } \hat{v}_i \in V_k,$$

(ii) For all $i, j \in \mathcal{N}$,

$$\sum_{\substack{\boldsymbol{v}' \in E_j: \\ v_i' \geq \hat{v}_i}} f(\boldsymbol{v}') \geq \sum_{\substack{\boldsymbol{v}' \in E_j: \\ v_i' \geq \hat{v}_i}} \tilde{f}(\boldsymbol{v}') \quad \text{for all } \hat{v}_i \in V_k,$$

then as $\delta \to 0$, $\boldsymbol{\Psi^*} \neq \emptyset$ under f implies $\boldsymbol{\Psi^*} \neq \emptyset$ under $\tilde{f}$.

The proof of Proposition 1 is deferred to Appendix A. Proposition 1 states that as preferences become more polarised, i.e. consumers who originally preferred firm i's now prefer it more, while consumers who originally preferred j's product now prefer i's product less, it becomes easier for the information designer to access information structures which extract all potential surplus from consumers. A straightforward means to polarise preferences might be through product differentiation or advertising e.g. Johnson and Myatt (2006). Proposition 1 suggests advertising and related mechanisms to influence preferences might be complementary with information design.[19]

Proposition 2. Suppose the information designer can also restrict firm i's access to some consumers. Denote $r_i : V_k^n \to [0,1]$ where $r_i(\boldsymbol{v}') \in [0,1]$ is the mass of consumers of type $\boldsymbol{v}'$ which firm i is prevented from accessing. If all $i \in \mathcal{N}$,

$$\sum_{\substack{\boldsymbol{v}' \in \Theta \setminus E_i: \\ v_i' \geq \hat{v}_i}} r_i(\boldsymbol{v}') \leq \sum_{\substack{\boldsymbol{v}' \in \Theta \setminus E_i: \\ v_i' \geq \hat{v}_i}} \tilde{r}_i(\boldsymbol{v}')$$

then as $\delta \to 0$, $\boldsymbol{\Psi^*} \neq \emptyset$ under $(r_i)_{i=1}^n$ implies $\boldsymbol{\Psi^*} \neq \emptyset$ under $(\tilde{r}_i)_{i=1}^n$.

Proof of Proposition 2. If firm i cannot access consumers distributed according to r_i,the RHS of the condition in Theorem 1 becomeds $\sum_{\substack{\boldsymbol{v}' \in \Theta \setminus E_i: \\ v_i' \geq \hat{v}_i}} [f_i(\boldsymbol{v}' - r_i(\boldsymbol{v}')]$. □

Proposition 2 is intuitive, and states that if the information designer can, in addition to designing information, additionally restrict each firm's access to some consumers, this also weakens the conditions under which perfect segmentation and monopolization for other non-restricted consumers is possible. For instance, if internet companies are able to restrict the product offerings viewed by certain consumers, this not only directly restricts price competition among firms, but also has the indirect effect of making segmentation through information easier for other consumers to whom the internet company is unable to restrict access.

[19]More generally, when both firms and consumers have incomplete information about valuations, the information designer might find it optimal to design information for consumers in a way which polarises preferences in expectation e.g. via the design of product rating systems. This might suggest that internet companies' ability to design information for both firms and consumers might be complementary, and presents a fruitful avenue for future work.

4 Continuous Version of General Design in Hotelling Setting

The general model stated in Section 2 was discretized to aid exposition. However, we develop a continuous counterpart of the discrete in Appendix B. This continuous model works directly with mass $1-\sum_i \eta_i$ of consumers distributed over the underlying type space $\Theta = [0,1]^n$, and mass $\eta_i \geq 0$ of captive consumers of type $(v_1 = 0, \ldots, v_i = 1, \ldots, v_n = 0)$ for each $i \in \mathcal{N}$. The continuous version of Theorem 1 can thus be stated as

Theorem 1$^{\text{c}}$ (Continuous analogue of Theorem 1). $\boldsymbol{\Psi^*} \neq \emptyset$ if and only if for all $i \in \{1, \ldots, n\}$ and all $\hat{x}_i \in [0,1]$,

$$H_i^{c*}(\hat{v}_i) \geq \int_{\Theta \setminus E_i : x_i \leq \hat{x}_i} f^c(\boldsymbol{v}')d^n\boldsymbol{v}'.$$

We develop the model formally and prove this in Appendix B but point out the similarities to Theorem 1 here. $H_i^{c*}(\hat{v}_i)$ is the continuous analogue of $H_i^*(\hat{v}_i)$ and denotes the maximum mass of consumers over $\Theta \setminus E_i$ which can be matched to consumers in E_i—and hence garbled together—while, at the same time, ensuring almost all incentive compatibility constraints remains fulfilled. As before, this is done by constructing a function analogous to G_i^* which binds almost all incentive comparability constraints associated with the message $\bar{v}_i \in [0,1]$ and then integrating this function over $\bar{v}_i$. The continuous model generates the same key insights as the discrete case, but will additionally allow us to map these conditions directly onto canonical models of product differentiation which are typically set on continuous type spaces.

4.1 Mapping to Hotelling. We now map our general model into the Hotelling setting introduced in the introduction with 2 firms and perfectly anti-correlated transportation costs. The type space is now $[0,1]$ and a consumer of type $x \in [0,1]$ obtains utility $1 - tx$ $(1 - t(1-x)$ resp.) purchasing from firm 1 (2 resp.). Each firm has a mass $\eta \in [0, 1/2]$ of captive consumers, and the remaining mass of $1 - 2\eta$ consumers are distributed uniformly over $[0,1]$. As such, $E_1 = [0, 1/2]$ and $E_2 = [1/2, 1]$. In addition, define $N_1 := [1/2, \min\{1, 1/t\}]$, the consumer types firm 1 cannot sell to in an efficient allocation, but for which there are nonetheless gains from trade. Analogously, $N_2 := [\max\{0, 1-1/t\}, 1/2]$.[20] Finally, given t, let $\bar{\eta}^G(t)$ be the threshold function denoting the minimum mass of captive consumers for each firm such that the condition in Theorem 1$^{\text{c}}$ is fulfilled.

The condition in Theorem 1$^{\text{c}}$ then reads that for almost all $\hat{x}_1 \in N_1$,

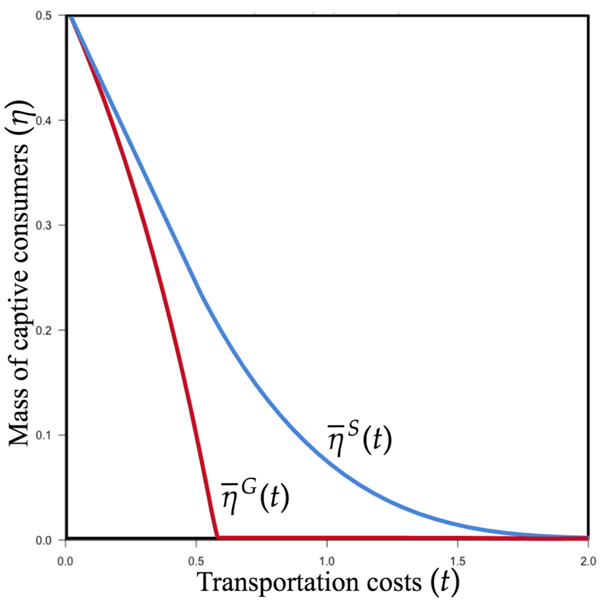
$$(1-2\eta)\frac{4t\hat{x}_1 - t}{8(1-t\hat{x}_1)} + \eta\frac{t\hat{x}_1}{1-t\hat{x}_1} \geq \int_{1/2}^{\hat{x}_1}(1-2\eta)d\hat{x}_1.$$

Figure 3 below plots the threshold functions under simple $(\bar{\eta}^S)$ and general $(\bar{\eta}^G)$ designs. The full analysis is deferred to Supplementary Appendix II.

Note the marked decrease in the requisite mass of captive consumers given transport costs t. In particular, for the canonical benchmark $t = 1$, while the simple design detailed in

[20] In the general model, we associated the type space directly with valuations, so $N_i = \Theta \setminus E_i$.

Figure 3: Threshold functions for simple vs optimal information design



the introduction required 7.5% of captive consumers, the general design does not require any captive consumers. This is because the general design fully utilises all—and not just captive—consumers of types E_i to hold down firm 1's incentive to steal consumers of types N_1.

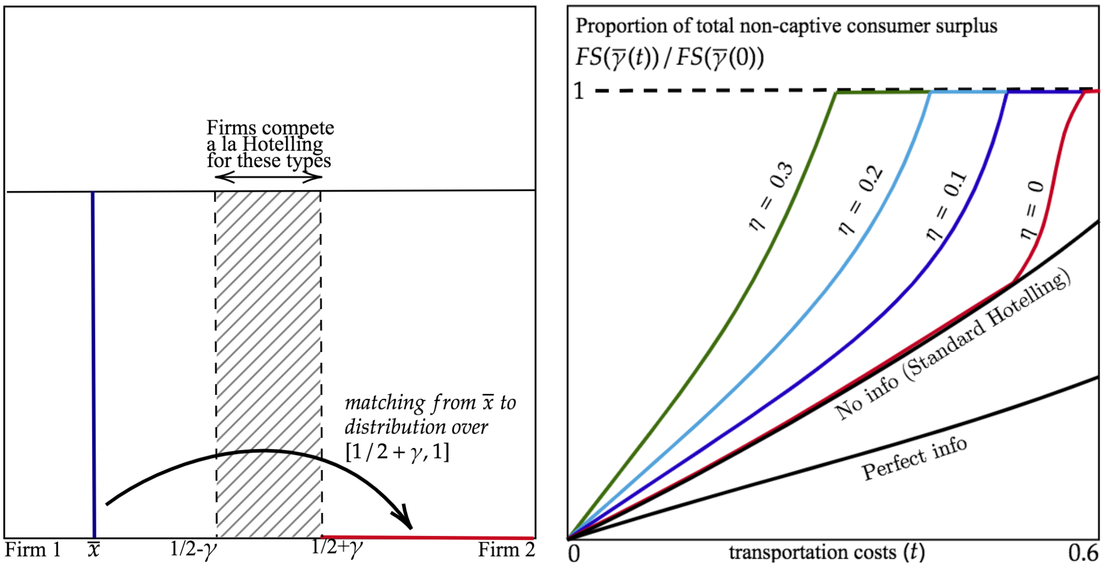
4.2 Lower bound on extractable surplus. We continue to work in the Hotelling setting and suppose the condition in Theorem 1^c is not fulfilled i.e. for a given (t, η) pair, $\eta < \bar{\eta}(t)$. Consider the alternate design under which consumers of types $[1/2 - \gamma, 1/2 + \gamma]$ are bundled together. Both firms then compete a la Hotelling by setting a uniform price for these types. We then perform our matching as before, but exclude types $[1/2 - \gamma, 1/2 + \gamma]$. An illustration of this alternate scheme is shown in Figure 4 (a) which shows how each type $\bar{x} \in E_1 \cap [0, \frac{1}{2} - \gamma] = [0, 1/2 - \gamma]$ is matched to a distribution over $N_1 \cap [\frac{1}{2} + \gamma, 1] = [1/2 + \gamma, \min\{1, 1/2\}]$. For a given pair (t, η), we can then find some threshold $\bar{\gamma}(t)$ such that for all $\gamma > \bar{\gamma}(t)$, there exists a set of message functions which induces a subgame fulfilling efficiency and full surplus extraction for all types $[0, \frac{1}{2} - \gamma] \cup [\frac{1}{2} + \gamma, 1]$. For this threshold $\bar{\gamma}(t)$, we can then find the total extractable surplus from non-captive consumers which we denote $FS(\bar{\gamma}(t))$. The full analysis deferred to Supplementary Appendix III.

Figure 4 (b) compares the proportion of total surplus which can be extracted from non-captive consumers under the alternate scheme against no-information (Hotelling), and perfect-information benchmarks. For low masses of captive consumers, the alternate design utilizes captive consumers to great effect and extracts substantially more surplus.

5 Conclusion

We have characterised a necessary and sufficient condition under which perfect segmentation and monopolization can be achieved. These conditions are relatively weak, and

Figure 4: Alternate design and lower bound on total firm surplus



(a) Alternate Design

(b) Total Firm Surplus

weaken further as consumer preferences become more polarised, or the designer can additionally restrict firms' access to consumers. When this condition is not fulfilled, we showed that an intuitive alternative can still allow firms to extract significantly more surplus than in a canonical Hotelling setting. These results suggest that absent regulatory oversight, segmentation and monopolization is something which could realistically happen—and might be happening—across very many markets.

Although the general information design which works in the widest possible range of settings can be complex relative to the simple design, implementation in both cases is equally straightforward and simply requires all firms to obey pricing recommendations. Further, conditional on all firms' participation in such a scheme, in the resultant equilibrium, no (possibly singleton) subset of firms, taking the obedience of other firms as given, can deviate to the benefit of all members. In this regard, firms can sidestep characteristic impediments to collusion e.g. issues of monitoring and enforcement. This suggests an additional object of regulatory concern.

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 Economic Research.

Admati, A. R. and P. Pfleiderer (1986): "A monopolistic market for information," *Journal of Economic Theory*, 39, 400–438.

Albrecht, B. C. (2019): "Price Competition and the Use of Consumer Data," .

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.

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.

CALZOLARI, G. AND A. PAVAN (2006): "On the optimality of privacy in sequential contracting," *Journal of Economic Theory*, 130, 168–204.

FAINMESSER, I. P. AND A. GALEOTTI (2019): "Pricing network effects: Competition," *Johns Hopkins Carey Business School Research Paper*.

HOTELLING, H. (1929): "Stability in Competition," *The Economic Journal*, 39, 41–57.

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. AND M. GENTZKOW (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.

LIZZERI, A. (1999): "Information revelation and certification intermediaries," *The RAND Journal of Economics*, 214–231.

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.

ROESLER, A.-K. AND B. SZENTES (2017): "Buyer-optimal learning and monopoly pricing," *American Economic Review*, 107, 2072–80.

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.

Appendices

A Omitted Proofs from Discrete Model

Proof of Lemma 1. **Part (i):** We first show that consistency and $\overline{IC}$ are sufficient to induce an equilibrium fulfilling C1 and C2. By construction, every firm $i \in \mathcal{N}$ receives precise information about the identities of consumers which value their product most. If g_i fulfils incentive compatibility, then conditional on receiving this information, firm i prefers to follow the pricing recommendation rather than deviating to a lower price to capture a larger slice of consumers (which includes those firm i should not sell to in equilibrium). But by our consistency requirement, we have $\sum_{\bar{v}_i \in V_k} g(\boldsymbol{v}'|\bar{v}_i) = f(\boldsymbol{v}')$ which implies that after obeying pricing recommendations for each message realisation $\bar{v}_i \in V_k$, all consumers of type $\boldsymbol{v}' \in \Theta \setminus E_i$ have been successfully garbled together with those in E_i. As such, firm i cannot undercut its competitors to capture these consumers without also hurting its profit margins on those of types E_i—which, by incentive compatibility, is unprofitable. Finally, since this obtains for all firms, we have $\boldsymbol{\Psi}^* \neq \emptyset$.

Part (ii): Recall we assumed $V_k \subseteq M$, and every firm learns about the effective valuation $\bar{v}_i$ for each consumer in E_i. This is without loss of generality since C1 and C2 require firm i to sell to all consumers of type E_i and extract all possible surplus. This implies for any $\boldsymbol{v}, \boldsymbol{v}' \in E_i$ such that $\bar{v}_i \neq \bar{v}_i'$, firm i must set different prices. As such, we need $\psi(m = \phi(\bar{v}_i)|\boldsymbol{v}(\boldsymbol{\theta})) = 1$ where ϕ performs a unique transformation of the preimage. It is then without loss to choose $\phi(\bar{v}_i) = \bar{v}_i$ with the natural interpretation of messages as price recommendations. It will thus suffice to show that both $\overline{IC}$ and consistency are necessary.

First, $\overline{IC}$. Recall this is the incentive compatibility condition under the assumption that all indifferences in effective valuations are broken in favour of firm $j \neq i$. In particular, $\underline{G}_i(\hat{v}_i|\bar{v}_i)$ represents a lower bound on the minimum extra mass of consumers firm i captures from deviating to any price $\hat{v}_i < \bar{v}_i$, no matter the underlying distribution of consumers within each price grid. As such, if $\overline{IC}$ is not fulfilled for all firms, all prices recommendations, and all price deviations, then there must exist some firm $i \in \mathcal{N}$, some recommendation $\bar{v}_i$, and some deviation $\hat{v}_i$ such that firm i can do better by charging price $\hat{v}_i < \bar{v}_i$ which successfully captures some positive mass of consumers of types $\Theta \setminus E_i$, which violates both C1 and C2.

Finally, consistency. If $M = V_k$, consistency is trivially fulfilled since every consumer, by definition, must be assigned a message realisation. We thus direct our attention to richer message spaces with $M \supset V_k$. Since every consumer must be assigned some message, we have $\sum_{\bar{v}_i \in V_k} g(\boldsymbol{v}'|\bar{v}_i) \leq f(\boldsymbol{v}')$. Now suppose the inequality is strict for some firm $i \in \mathcal{N}$ and some $\boldsymbol{v}' \in \{V_k^n : v_i' > 0\}$ and denote $S := \{\boldsymbol{v}' \in \Theta \setminus E_i : \sum_{\bar{v}_i \in V_k} g(\boldsymbol{v}'|\bar{v}_i) < f(\boldsymbol{v}')\}$ which represents the types which firm i should not sell to, but which have not been matched to any consumer in E_i via g_i (and hence assigned message realisation in $M \setminus V_k$). Now consider the following deviation by firm i to charge all such consumers with message realisations $M \setminus V_k$ price $\min\{v_i' : \boldsymbol{v}' \in S\}$. Under this deviation, since consumers are distributed with full support over Θ, some positive measure set of consumers find it profitable to buy from firm i instead at a strictly positive price. As such, this represents a strictly profitable deviation which violates both C1 and C2. □

Proof of Proposition 1. Suppose $\mathbf{\Psi^*} \neq \emptyset$ under $f_{\mathcal{N}}$ and fix $\hat{v}_i$. By Theorem 1 (ii), we have

$$\begin{aligned} H_i{}^*(\hat{v}_i; f) := \sum_{\bar{v}_i \in V_k} G_i^*(\hat{v}_i | \bar{v}_i; f) &:= \sum_{\bar{v}_i \in V_k} \mathbb{1}_{\bar{v}_i > \hat{v}_i} \left(\frac{\bar{v}_i - \hat{v}_i}{\hat{v}_i} \right) \sum_{\substack{\boldsymbol{v'} \in E_i: \\ v_i' = \bar{v}_i}} f(\boldsymbol{v'}) \\ &= \sum_{\substack{\boldsymbol{v'} \in E_i: \\ v_i' > \hat{v}_i}} \left(\frac{v_i' - \hat{v}_i}{\hat{v}_i} \right) f(\boldsymbol{v'}) \\ &\leq \sum_{\substack{\boldsymbol{v'} \in E_i: \\ v_i' > \hat{v}_i}} \left(\frac{v_i' - \hat{v}_i}{\hat{v}_i} \right) \tilde{f}(\boldsymbol{v'}) = H_i^*(\hat{v}_i; \tilde{f}) \end{aligned}$$

and

$$\sum_{\substack{\boldsymbol{v'} \in \Theta \setminus E_i: \\ v_i' \geq \hat{v}_i}} f(\boldsymbol{v'}) = \sum_{j \neq i} \sum_{\substack{\boldsymbol{v'} \in E_j: \\ v_i' \geq \hat{v}_i}} f(\boldsymbol{v'}) \geq \sum_{j \neq i} \sum_{\substack{\boldsymbol{v'} \in E_j: \\ v_i' \geq \hat{v}_i}} \tilde{f}(\boldsymbol{v'}) = \sum_{\substack{\boldsymbol{v'} \in \Theta \setminus E_i: \\ v_i' \geq \hat{v}_i}} \tilde{f}(\boldsymbol{v'})$$

so by Theorem 1 (i), parts (i) and (ii) imply $\mathbf{\Psi^*} \neq \emptyset$ under $\tilde{f}$. □

B Continuous Model

In this Appendix, we develop a continuous counterpart to the discrete model introduced in Section. 2

B.1 Setup. There are $n < \infty$ firms indexed with the set $\mathcal{N} = \{1, \ldots, n\}$, and a continuum of consumers with unit mass distributed over the type space where $\boldsymbol{\theta} = (v_1, v_2, \ldots, v_n)$ is a typical member of Θ. Each consumer demands a single unit inelastically, so a consumer of type $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n) \in \Theta$ has value $\theta_i \in [0,1]$ from purchasing good i. Each firm i has mass $\eta_i \geq 0$ of captive consumers of type $C_i := (\theta_1 = 0, \ldots, \theta_i = 1, \ldots, \theta_n = 0)$. These consumers obtain the maximum value of 1 from buying from firm i, and no utility from buying from firms $j \neq i$. The remaining mass $1 - \sum \eta_i$ of non-captive consumers are distributed with bounded density $f^c(\boldsymbol{\theta})$ with full support over Θ. We will use the Dirac delta function $\delta(m-t)$ where $\int_{t-\varepsilon}^{t+\varepsilon} \delta(m-t) = 1$ for all $\varepsilon > 0$ to represent atoms with positive mass at t.

Partition the hypercube $\Theta = [0,1]^n$ into n regions where $E_i := \{\boldsymbol{\theta} \in \Theta : \max(\theta_1, \theta_2, \ldots, \theta_n) = \theta_i\}$ are the types of consumers (including captive consumers) who value firm i's product most, and so should purchase from firm i in an efficient allocation. Note $\Theta = \bigcup_{i=1}^n E_i$, and $E_i \cap E_j$ is of measure zero for all $i, j \in \{1, \ldots, n\}$.

There is also an *information designer* and as before, we assume

Assumption 1. The information designer's payoff is increasing in total firm surplus.

For each firm, the information designer commits in advance to a signal structure it will provide about each type of consumer. For each firm $i \in \{1, \ldots, n\}$, the information designer chooses a message function

$$\psi_i^c : \Theta \to \Delta(M^c)$$

where $M^c \supseteq [0,1]$ is the message space and $\Delta(M)$ is the set of feasible probability distributions over M. Call $m \in M$ a message realisation. Denote the set of all sets of message functions with $\boldsymbol{\Psi}$. The information designer then learns the exact valuation of each consumer for each product,[21] and releases messages to each firm $i \in \{1, \ldots, n\}$ accordance to the signal structure. Given the messages for each consumer, firms then play a pricing game: a pricing function for firm $i \in \{1, \ldots, n\}$ is $p_i^c : M^c \to \mathbb{R}_+$. Let P_i^c be the space of possible pricing functions for firm i. A strategy for firm i is $\sigma_i^c : M^c \to \Delta(P_i^c)$. Denote the subgame induced by the message functions $\boldsymbol{\psi^c} := (\psi_i^c)_{i=1}^n$ with $\Gamma^c(\boldsymbol{\psi}^c)$.

Our goal is, as before, to characterize conditions under there exists $\boldsymbol{\psi}^c$ such that in $\Gamma^c(\boldsymbol{\psi}^c)$, each consumer of type $\boldsymbol{\theta} \in \Theta$,

- C1 **(Efficient allocation)** buys from firm i only if $\boldsymbol{\theta} \in E_i$; and
- C2 **(Full Surplus Extraction)** pays $\max_{v_i} \boldsymbol{v}(\boldsymbol{\theta})$.

[21] Note that it does not matter whether the information designer learns about each consumer's type before or after committing the message functions.

B.2 Characterisation. Let $\boldsymbol{\Gamma^{c*}}$ denote the set of induced games in which there exists an equilibrium satisfying conditions C1 and C2, and let

$$\boldsymbol{\Psi^{c*}} := \{\boldsymbol{\psi^c} \,:\, \Gamma^c(\boldsymbol{\psi^c}) \in \boldsymbol{\Gamma^{c*}}\}$$

be the set of message functions that the information designer can use to fulfil both conditions. We now turn our attention towards characterising exact conditions under which an information designer is able to achieve conditions C1 and C2 i.e. $\boldsymbol{\Psi^{c*}} \neq \emptyset$.

Define the matching $g_i^c : \{\theta_i : \boldsymbol{\theta} \in E_i\} \to \Delta(\Theta \setminus E_i)$ which matches all consumers with types in E_i of with valuation θ_i for firm i's product to a distribution over the types $\Theta \setminus E_i$. Matched consumers are assigned identical messages. Now assume types $\boldsymbol{\theta} \in E_i$ are allocated message $m = \theta_i$ with probability 1. We subsequently show this is without loss of generality. Observe that the density of consumers at $\boldsymbol{\theta'} \in \Theta \setminus E_i$ matched to $\bar{\theta}_i \in \{\theta_i : \boldsymbol{\theta} \in E_i\}$ (and hence assigned message $\bar{\theta}_i$) must be equivalent to the density of consumers which are assigned message $\bar{\theta}_i$ given the underlying type $\boldsymbol{\theta'} \in \Theta \setminus E_i$ scaled by the actual density of consumers at $\boldsymbol{\theta'}$. This can be succinctly written as

$$g_i^c(\boldsymbol{\theta'}|\bar{\theta}_i) = \psi_i(m = \bar{\theta}_i|\boldsymbol{\theta'})f^c(\boldsymbol{\theta'}) \quad \text{for all } \bar{\theta}_i \in \{\theta_i \,:\, \boldsymbol{\theta} \in E_i\},\, \boldsymbol{\theta'} \in \Theta \setminus E_i.$$

For non-captive consumers of types $E_i \setminus C_i$, we can now represent firm i's demand given the message realisation m, matching functions $\boldsymbol{g^c}$, and price p_i with

$$\mathcal{D}_i(p_i = \hat{\theta}_i|\boldsymbol{g^c}, m = \bar{\theta}_i) = \frac{\overbrace{\int_{E_i \setminus C_i} f^c(\theta'_i = \bar{\theta}_i, \boldsymbol{\theta'}_{-i})d^{n-1}\boldsymbol{\theta'_{-i}}}^{\text{Demand from the set } E_i \setminus C_i.} + \overbrace{\int_{\Theta \setminus E_i : \theta'_i \geq \hat{\theta}_i} g_i^c(\boldsymbol{\theta'}|\bar{\theta}_i)d^n\boldsymbol{\theta'}}^{\text{Demand from the set } \Theta \setminus E_i}}{\psi_i^c(m = \bar{\theta}_i)},$$

where $1 > \bar{\theta}_i \geq \hat{\theta}_i$. For captive consumers of type C_i, we can analogously represent firm i's conditional demand with

$$\mathcal{D}_i(p_i = \hat{\theta}_i|\boldsymbol{g^c}, m = \bar{\theta}_i = 1) = \frac{\eta_i + \int_{\Theta \setminus E_i : \theta'_i \geq \hat{\theta}_i} g_i^c(\boldsymbol{\theta'}|1)d^n\boldsymbol{\theta'}}{\psi_i^c(m = 1)},$$

where $1 = \bar{\theta}_i \geq \hat{\theta}_i$, noting that $\psi_i(m = 1)$ is of positive mass if and only if it has positive mass of captive consumers, i.e. $\psi_i^c(m = 1) = \eta_i \delta(m - 1)$. We now state an analogue of Lemma 1 which shows that it is without loss of generality to look for message functions of a particular kind.

Lemma $\mathbf{1^c}$. $\boldsymbol{\Psi^*} \neq \emptyset$ if and only if there exists $\boldsymbol{\psi} \in \boldsymbol{\Psi}$ such that

(i) **(Price recommendation)** For all $i \in \{1, \ldots, n\}$ and all $\bar{\boldsymbol{\theta}} \in E_i$, $\psi_i^c(m = \bar{\theta}_i|\bar{\boldsymbol{\theta}}) = \delta(m - (1 - \bar{\theta}_i))$ i.e. firm i receives message $m = \bar{\theta}_i$ with probability 1.

(ii) There exists a set of matching functions $\boldsymbol{g^c}$ such that for all $i \in \{1, \ldots, n\}$,

(a) **(Incentive Compatibility)** for almost all $\bar{\theta}_i \in [0, 1]$

$$\underbrace{\bar{\theta}_i \mathcal{D}_i(p_i = \bar{\theta}_i|\boldsymbol{g^c}, m = \bar{\theta}_i)}_{\text{Surplus from setting price } \bar{\theta}_i.} \geq \underbrace{\hat{\theta}_i \mathcal{D}_i(p_i = \hat{\theta}_i|\boldsymbol{g^c}, m = \bar{\theta}_i)}_{\text{Surplus from setting price } \hat{\theta}_i.}$$

for all $\hat{\theta}_i \leq \bar{\theta}_i$;

(b) **(Consistency)** for almost all $\boldsymbol{\theta'} \in \Theta \setminus E_i$,

$$\int_{(0,1]} g_i^c(\boldsymbol{\theta'}|\bar{\theta}_i)d\bar{\theta}_i + g_i^c(\boldsymbol{\theta'}|0) = f^c(\boldsymbol{x'}).$$

We prove Lemma 1[c] at the end of this Appendix. By Lemma 1[c], is without loss of generality to look for messages satisfying conditions (i) and (ii)(a)-(b). Condition (i) states that each firm i should receive precise and accurate information about preferences for i's product for all consumers it should serve in an efficient allocation. This makes full surplus extraction feasible. Condition (ii)(a) requires that conditional on receiving message $m = \bar{\theta}_i$, firm i prefers to extract all surplus from types $\{\boldsymbol{\theta} \in E_i : \theta_i = \bar{x}_i\}$ by following the pricing recommendation rather than lowering its price to $\hat{\theta}_i$ and obtaining additional demand from consumers outside the set E_i. Condition (ii)(b) requires the set of matching functions to be consistent with the distribution of consumer types.

We now turn to the question of whether such message functions can be constructed. Consider firm $i \in \{1, \ldots, n\}$ and choose a set of message functions—and hence matching functions—such that incentive compatibility constraint binds everywhere.[22] For non-captive consumers, define

$$\begin{aligned} G_i^*(\hat{\theta}_i|\bar{\theta}_i) &:= \int_{\Theta\setminus E_i:\theta_i'\geq\hat{\theta}_i} g_i^*(\boldsymbol{\theta'}|\bar{\theta}_i)d^n\boldsymbol{\theta'} \\ &= \begin{cases} \left(\dfrac{\bar{\theta}_i - \hat{\theta}_i}{\hat{\theta}_i}\right) \displaystyle\int_{E_i} f^c(\theta_i' = \bar{\theta}_i, \boldsymbol{\theta'}_{-i})d^{n-1}\boldsymbol{\theta'_{-i}} & \text{if } \hat{\theta}_i < \bar{\theta}_i < 1 \ , \\ 0 & \text{if } \hat{\theta}_i \geq \bar{\theta}_i, \end{cases} \end{aligned}$$

noting that this object is always of zero mass.[23] For captive consumers, define

$$G_i^{c*}(\hat{\theta}_i|\bar{\theta}_i = 1) = \eta_i\left(\frac{\bar{\theta}_i - \hat{\theta}_i}{\hat{\theta}_i}\right)$$

noting that this can be of positive mass since there can be an atom of captive consumers at C_i. It will also be helpful to develop notation for the total mass of consumers in the set $\Theta \setminus E_i$ which (i) have value greater than $\hat{\theta}_i$ from purchasing from firm i; and (ii) are matched to consumers in E_i. To this end, define

$$H_i^{c*}(\hat{\theta}_i) := \underbrace{\int_{[0,1)} G_i^{c*}(\hat{\theta}_i|\bar{\theta}_i)d\bar{\theta}_i}_{\substack{\text{non-captive} \\ \text{consumers}}} + \underbrace{G_i^{c*}(\hat{\theta}_i|\bar{\theta}_i = 1)}_{\substack{\text{captive} \\ \text{consumers}}}$$

noting that $H_i^{c*}(\hat{x}_i)$ is decreasing and of positive mass. We are now ready to state the continuous counterpart to Theorem 1.

Theorem 1[c]. $\boldsymbol{\Psi^*} \neq \emptyset$ if and only if for all $i \in \{1, \ldots, n\}$ and all $\hat{x}_i \in [0, 1]$,

$$H_i^{c*}(\hat{\theta}_i) \geq \int_{\Theta\setminus E_i:\theta_i'\geq\hat{\theta}_i} f^c(\boldsymbol{\theta'})d^n\boldsymbol{\theta'}.$$

[22] Note that we can do this because we are not yet imposing consistency.

[23] To see this, note that $\int_{E_i} f^c(\theta_i' = \bar{\theta}_i, \boldsymbol{\theta'}_{-i})d^{n-1}\boldsymbol{\theta'_{-i}}$ is the mass of a $n-1$ dimensional hyperplane over E_i.

The key ideas underlying Theorem 1^c are similar to that of Theorem 1.

Proof of Theorem 1^c. **<u>If:</u>** Suppose the condition in Theorem 1^c is fulfilled. In view of Lemma 1^c, it is sufficient to show that there exists a set of matching functions fulfilling conditions (i) and (ii)(a)-(c). We show that these matching functions can always be constructed by first considering $G^{c*}(\hat{\theta}_i|\bar{\theta}_i)$. By construction, this just satisfies the incentive compatibility constraints of condition (ii)(a). However, it might violate the consistency condition (ii)(b). But since the condition in Theorem 1^c obtains, we can find an alternate function $\tilde{G}^c(\hat{x}_i|\bar{x}_i)$ such that

(i) $\tilde{G}^c(\hat{\theta}_i|\bar{\theta}_i) \leq G^{c*}(\hat{\theta}_i|\bar{\theta}_i)$ for all $\bar{\theta}_i \in [0,1]$; and

(ii) $\int_{[0,1)} \tilde{G}^c(\hat{\theta}_i|\bar{\theta}_i)dx_i + \tilde{G}^c(\hat{\theta}_i|\bar{\theta}_i = 1) = \int_{\Theta\setminus E_i:\theta'_i\geq\hat{\theta}_i} f^c(\boldsymbol{\theta'})d^n\boldsymbol{\theta'}$ for all $\hat{\theta}_i \in [0,1]$.

Since $G^{c*}(\hat{\theta}_i|\bar{\theta}_i)$ binds the incentive compatibility constraint everywhere, (i) ensures they remain satisfied for almost all types in the interval $[0,1]$. (ii) implies we can always find a corresponding matching distribution $\tilde{h}^c_i(\boldsymbol{\theta'}) := \int_{[0,1)} \tilde{g}^c_i(\boldsymbol{\theta'}|\bar{\theta}_i)d\bar{\theta}_i + \tilde{g}^c_i(\boldsymbol{\theta'}|\bar{\theta}_i = 1)$ such that $\tilde{h}_i(\boldsymbol{\theta'}) = f^c(\boldsymbol{\theta'})$ for almost all $\boldsymbol{\theta'} \in \Theta \setminus E_i$, hence fulfilling the consistency condition of Lemma 1^c. Since we can always perform this procedure for all firms $i \in \{1,\ldots,n\}$, the condition in Theorem 1^c is sufficient.

<u>Only if:</u> Again, consider $G^{c*}(\hat{\theta}_i|\bar{\theta}_i)$ and corresponding function $H^{c*}_i(\hat{\theta}_i)$. Noting that H^*_i is continuous and monotonically decreasing, set $\alpha_i = H^{c*-1}_i(\int_{\boldsymbol{\theta'}\in\Theta\setminus E_i} f^c(\boldsymbol{\theta'})d^n\boldsymbol{\theta'})$. This is the value at which our cumulative function H^{c*}_i is equal to the total mass of consumers distributed over $\Theta \setminus E_i$. The condition in Theorem 1^c trivially fulfilled for $\hat{\theta}_i \in [0,\alpha_i]$ so it is without loss to consider $\hat{\theta}_i \in [\alpha_i, 1]$. Now suppose, towards a contradiction, that $\boldsymbol{\Psi^{c*}} \neq \emptyset$ but there exists some $i \in \{1,\ldots,n\}$ and some $\hat{\theta}_i \in [\alpha_i, 1]$ such that

$$H^{c*}_i(\hat{\theta}_i) < \int_{\Theta\setminus E_i:\theta'_i\geq\hat{\theta}_i} f^c(\boldsymbol{\theta'})d^n\boldsymbol{\theta'}.$$

There then exists a positive mass set of consumers of types $S \subset \{\boldsymbol{\theta} \in \Theta \setminus E_i : \theta \geq \hat{\theta}_i\}$ which will not be assigned some message $m \in [0,1]$ hence violating condition (ii)(b) of Lemma 1^c. Intuitively, if this obtains, firm i can profitably deviate by stealing some positive mass of consumers in this set from firms $j \neq i$ thereby violating C1 and C2. As such, we require an alternate function $\tilde{\tilde{G}}^c_i(\hat{\theta}_i|\bar{\theta}_i)$ such that $\tilde{\tilde{G}}^c_i(\hat{\theta}_i|\bar{\theta}_i) \geq G^{c*}_i(\hat{\theta}_i|\bar{\theta}_i)$ with strict inequality over some positive measure set. This corresponds to 'leftward' shifts of the matching distribution. But then recall $G^{c*}_i(\hat{\theta}_i|\bar{\theta}_i)$ was constructed such that firm i is just indifferent between obeying the pricing recommendation and deviating by setting price $\hat{\theta}_i \leq \bar{\theta}_i$. As such, fulfilling consistency requires violating incentive constraints for some positive measure set. It is thus impossible to simultaneously satisfy conditions (ii)(a) and (ii)(c) of Lemma 1^c, a contradiction which then implies $\boldsymbol{\Psi^{c*}} = \emptyset$. □

B.3 Proof of Lemma 1^c. The proof of Lemma 1^c proceeds similarly to that of Lemma 1 though we state it here for completeness.

Proof. **If:** By construction, conditions (i), (ii)(a)-(b) are sufficient to induce an equilibrium in the resultant subgame fulfilling conditions C1 and C2. We now show it is without loss of generality to consider messages of this form. **Only if:** We wish to show if there is no set of message functions fulfilling conditions (i), (ii)(a)-(b), then $\boldsymbol{\Psi^*} = \emptyset$. To see part (i) is necessary, recall $\boldsymbol{\Psi^*}$ is the set of all sets of message functions which can induce an equilibrium fulfilling conditions C1 and C2 in the induced subgame. C1 and C2 require firm i to sell to all consumers of type E_i and extract all possible surplus. As such, messages for such types must be sufficiently precise such that for any $\bar{\boldsymbol{\theta}} \in E_i$, $\psi(m = \phi(\bar{\theta}_i)|\bar{\boldsymbol{\theta}}) = \delta(m - \phi_i^c(\bar{\theta}_i))$ where ϕ performs a unique transformation of the pre-image. This also implies for any $\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\theta}}' \in E_i$ such that $\bar{\theta}_i \neq \bar{\theta}_i'$, firm i must set different prices. It is therefore without loss of generality to choose $\phi(\bar{\theta}_i) = \bar{\theta}_i$.

We now consider condition (ii)(b). If $M = \{\theta_i : \boldsymbol{\theta} \in E_i\} = [0, 1]$, it is trivially fulfilled. We thus direct our attention to richer message spaces with $M \supset [0, 1]$. Since every consumer must be assigned some message, consumers of type $\boldsymbol{\theta}$ assigned messages in the interval $[0, 1]$ must be weakly less than the actual density of actual consumers, i.e. $\int_{\bar{\theta}_i : E_i \setminus C_i} g_i(\boldsymbol{\theta}|\bar{\theta}_i) d\bar{\theta}_i + g_i(\boldsymbol{x}|\bar{\theta}_i = 1) \leq f^c(\boldsymbol{\theta})$. As such, to show that condition (ii)(b) is necessary, it will suffice to show that if $\boldsymbol{\Psi^{c*}} \neq \emptyset$ and condition (i) obtains, then the set

$$S := \left\{ \boldsymbol{\theta} \in \Theta \setminus E_i \; : \; \int_{[0,1)} g_i(\boldsymbol{\theta}|\bar{\theta}_i) d\bar{x}_i + g_i(\boldsymbol{\theta}|\bar{\theta}_i = 1) < f^c(\boldsymbol{\theta}) \right\}$$

is zero measure for all firms $i \in \{1, \ldots, n\}$.

To see this, suppose, towards a contradiction, that the set S is not zero measure for some $i \in \{1, \ldots, n\}$ i.e. some consumers with types in the set S are not matched to any consumers in E_i (and so assigned some message $M \setminus [0, 1]$). Then consider the following deviation by firm i to charge all consumers in S some price $p \in (0, \sup\{\theta_i : \boldsymbol{\theta} \in S\})$. This is strictly profitable because it successfully obtains a positive mass of additional consumers. To see this, recall $\boldsymbol{\psi}^c \in \boldsymbol{\Psi^{c*}}$ implies conditions C1 and C2 are fulfilled, so all firms $j \neq i$ extract the full surplus from all consumers in the set E_j. This in turn implies all consumers in the set $\Theta \setminus E_i$ obtain zero surplus in equilibrium whereas under the proposed deviation, consumers in some positive mass subset of consumers of types $S \subseteq \Theta \setminus E_i$ obtain strictly positive surplus. As such, if condition (ii)(b) is violated, then $\boldsymbol{\Psi^{c*}} = \emptyset$.

Finally, suppose condition (ii)(a) is violated. There will then exist some firm i and some positive measure set $T \subseteq [0, 1]$ such that for all message realisations $t \in T$, firm i finds it strictly profitable to deviate from the pricing recommnedation and sell to consumers in the set $\Theta \setminus E_i$ at a price strictly lower than their valuation. This then violates conditions C1 and C2. As such, if condition (ii)(a) is violated, then $\boldsymbol{\Psi^{c*}} = \emptyset$. Hence, it is without loss of generality to consider messages satisfying conditions (i) and (ii)(a)-(ii)(b). □

Supplementary Appendix for 'Market segmentation through information': For Online Publication Only

In this supplementary appendix we provide calculations supporting Figures 1, 3, and 4 in the main text.

I Threshold for Simple Information Design (Figure 1)

Recall

- **Simple Information Design.** Consider the following information structure where firm 1 (2 resp.) receives precise information about the preferences of all non-captive consumers in $[0, 1/2]$ ($[1/2, 1]$ resp.), but receives no information for all other consumers (including their captive customers).

Given this information structure, we postulate the following strategy profile: firm 1 (2 resp.) charges (i) a price of $1 - tx$ $(1 - (t(1-x)$ resp.) if the firm has information about their types; (ii) a price $p = 1$ to all other consumers.

We now derive the conditions under which this strategy profile is an equilibrium. Under the proposed strategy profile, the profit of each of the two firms is

$$\Pi = \eta + \frac{1-2\eta}{2}\int_0^{1/2}(1-tx)dx.$$

Without loss of generality, consider the best possible deviation of firm 1. Note that firm 1 cannot gain by changing the price offered to the non-captive consumers located at $[0, \frac{1}{2}]$. As such, the only possible deviation is to charge a different uniform price $\hat{p} < 1$ to the other consumers. Offering price $\hat{p} \in [1 - \frac{t}{2}, 1)$ is clearly strictly dominated since, at this price, the demand of firm 0 is perfectly inelastic with value η. Hence it can charge $\hat{p} = 1$ while maintaining the same demand. Further, offering price $\hat{p} < 1 - t$ is also strictly dominated since firm 1 will sell to all non-captive consumers from $[1/2, 1]$ in addition to its captive consumers η. But if so, then firm 0 can set $\hat{p} = 1 - t$ while maintaining the same demand.

As such, we can restrict our attention to deviations $\hat{p} \in [v - t, v - \frac{t}{2})$. Fix a price in this range. Then the marginal consumer of type $x^*(\hat{p})$ is indifferent between buying from firms 1 and 2 so that

$$1 - tx^*(\hat{p}) - \hat{p} = 0 \iff x^*(\hat{p}) = \frac{1-\hat{p}}{t},$$

The demand for firm 1 at price $\hat{p}$ is thus

$$D(\hat{p}) = \eta + \left(x^*(\hat{p}) - \frac{1}{2}\right)(1-2\eta) = 2\eta - \frac{1}{2} + (1-2\eta)\frac{1-\hat{p}}{t}.$$

So firm 1 selects $\hat{p}$ to maximize

$$\hat{p}D(\hat{p}) = \hat{p}\left[2\eta - \frac{1}{2} + (1-2\eta)\frac{1-\hat{p}}{t}\right].$$

Taking the derivatives with respect to p and assuming that p is interior, we have that

$$\hat{p}^I = \frac{1}{2} - \frac{t}{4}\left(\frac{1-4\eta}{1-2\eta}\right).$$

We now check that $\hat{p}^I$ is, indeed, interior by considering the following three cases:

Case A. (Upper price boundary) $\hat{p}^I \geq 1 - \frac{t}{2}$ if and only if $\frac{1}{t} \leq \frac{1}{2(1-2\eta)}$ if and only if $\eta \geq \frac{2-t}{4}$. Then, the best deviation is to set $\hat{p}^* = 1 - \frac{t}{2}$. Clearly this deviation cannot be profitable because, as we have already argued, at that price firm 1 does sell to any customers between $[\frac{1}{2}, 1]$, and, at the same time, is losing margins from its captive customers. We conclude that if $\eta \geq \frac{2-t}{4}$ then the postulated the strategy profile is an equilibrium.

Case B. (Lower price boundary) $\hat{p} < 1 - t$ if and only if $\frac{1}{t} > \frac{3-4\eta}{2(1-2\eta)}$ if and only if $\eta < \frac{2-3t}{4(1-t)}$. Then the best deviation is to set $\hat{p}^* = 1 - t$. That is, firm 1's best deviation is to charge a price so that will be able maximise its margin conditional on selling to all customers in $[1/2, 1]$. Firm 1's profits from doing so is

$$\hat{\Pi} = \frac{1-2\eta}{2}\int_0^{1/2}(1-tx)dx + \left[\eta + \frac{1-2\eta}{2}\right][1-t].$$

Hence, firm 0 does not wish to deviate if and only if $\Pi \geq \hat{\Pi}$ or, equivalently,

$$\eta \geq \left[\eta + \frac{1-2\eta}{2}\right][1-t] \iff \eta \geq \frac{1}{2} - \frac{t}{2}$$

As such, we conclude if $\frac{1}{2} - \frac{t}{2} \leq \eta \leq \frac{2-3t}{4(1-t)}$ the postulated strategy profile is an equilibrium.

Case C. (Interior) $\hat{p}^I \in [1-t, v-t/2)$ if and only if $\eta \in \left(\frac{2-3t}{4(1-t)}, \frac{2-t}{4}\right]$. Then the best deviation is to set $\hat{p}^* = \hat{p}^I$. The profit that the firm obtains by such a deviation is

$$\hat{\Pi} = \frac{1-2\eta}{2}\int_0^{1/2}(1-tx)dx + (\hat{p}^I)^2\frac{1-2\eta}{t}$$

The deviation is not profitable whenever $\Pi_A \geq \hat{\Pi}_A$ or, equivalently,

$$\eta \geq \frac{1-2\eta}{t}\left[\frac{1}{2} - \frac{t}{4}\left(\frac{1-4\eta}{1-2\eta}\right)\right]^2$$

We conclude if $\eta \in \left(\frac{2-3t}{4(1-t)}, \frac{2-t}{4}\right]$ and $\eta \geq \frac{1-2\eta}{t}\left[\frac{1}{2} - \frac{t}{4}\left(\frac{1-4\eta}{1-2\eta}\right)\right]^2$ then the postulated strategy profile is an equilibrium.

From Cases A-C, we use these conditions to construct a (continuous) piecewise function $\bar{\eta}^S(t)$ as show in Figure 3.

II Threshold for Optimal Information Design (Figure 3)

We now map our general model into the Hotelling setting introduced in the introduction with 2 firms and perfectly anti-correlated transportation costs. The type space is now $[0,1]$ and a consumer of type $x \in [0,1]$ obtains utility $1-tx$ ($1-t(1-x)$ resp.) purchasing from firm 1 (2 resp.). We hence define $\theta_i = 1 - tx$ and work with $x \in [0,1]$ directly. Each firm has a mass $\eta \in [0, 1/2]$ of captive consumers, and the remaining mass of $1-2\eta$ consumers are distributed uniformly over $[0,1]$. As such, $E_1 = [0, 1/2]$ and $E_2 = [1/2, 1]$. In addition, define $N_1 := [1/2, \min\{1, 1/t\}]$, the consumer types firm 1 cannot sell to in an efficient allocation, but for which there are nonetheless gains from trade. Analogously, $N_2 := [\max\{0, 1-1/t\}, 1/2]$.[24] Finally, given t, let $\bar{\eta}^G(t)$ be the threshold function denoting the minimum mass of captive consumers for each firm such that the condition in Theorem 1^c is fulfilled.

We construct G_1 as before by tightening almost all incentive compatibility constraints

$$G_1(\hat{x}_1|\bar{x}_1) := \begin{cases} (1-2\eta)\dfrac{t(\hat{x}_1 - \bar{x}_1)}{1-t\hat{x}_1} & \text{if } \hat{x}_1 \geq \bar{x}_1 > 0; \\ 0 & \text{if } \hat{x}_1 < \bar{x}_1; \\ \eta\dfrac{t\hat{x}_1}{1-t\hat{x}_1} & \text{if } \bar{x}_1 = 0, \end{cases}$$

where $\hat{x}_1 \in N_1$, $\bar{x}_1 \in [0, 1/2]$. and once again define $H_1(\hat{x}_i) = \int_{(0,\frac{1}{2}]} G_1(\hat{x}_1|\bar{x}_1) d\bar{x}_i + G_1(\hat{x}_1|0)$. The condition in Theorem 1^c then reads that for almost all $\hat{x}_1 \in N_1$,

$$(1-2\eta)\frac{4t\hat{x}_1 - t}{8(1-t\hat{x}_1)} + \eta\frac{t\hat{x}_1}{1-t\hat{x}_1} \geq \int_{1/2}^{\hat{x}_1} (1-2\eta) d\hat{x}_1 = (1-2\eta)(\hat{x}_1 - 1/2).$$

Now divide through by $(1-2\eta)$ and define

$$\begin{aligned} f(\hat{x}_1, \eta, t) &:= \frac{4t\hat{x}_1 - t}{8(1-t\hat{x}_1)} + \frac{\eta}{1-2\eta}\frac{t\hat{x}_1}{1-t\hat{x}_1} - (\hat{x}_1 - 1/2) \\ &= \frac{8t\hat{x}_1^2 - 8\hat{x}_1 - t + 4}{8(1-t\hat{x}_1)} + \frac{\eta}{1-2\eta}\frac{t\hat{x}_1}{1-t\hat{x}_1} \geq 0 \\ &\iff \tilde{f}(\hat{x}_1, \eta, t) := \frac{8t\hat{x}_1^2 - 8\hat{x}_1 - t + 4}{8} + \frac{\eta}{1-2\eta}(t\hat{x}_1) \\ &= t\hat{x}_1^2 + \left[\frac{t\eta}{1-2\eta} - 1\right]\hat{x}_1 + \frac{4-t}{8} \geq 0 \end{aligned}$$

First assume solution is interior, which gives

$$\frac{\delta \tilde{f}(\hat{x}_1, \eta, t)}{\delta \hat{x}_1} = 2t\hat{x}_1 + \left[\frac{t\eta}{1-2\eta} - 1\right] = 0 \quad \text{so} \quad \hat{x}_1^I = \frac{1 - \dfrac{t\eta}{1-2\eta}}{2t}.$$

[24] In the general model, we associated the type space directly with transportation costs, so $N_i = \Theta \backslash E_i$.

Recall we require $\hat{x} \in N_1 := [1/2, \min\{1, 1/t\}]$ and first suppose $t \leq 1$ so $\min\{1, 1/t\} = 1$ i.e. there are gains from trade with every consumer.

Case A. (Lower boundary) $\hat{x}_1^I < \frac{1}{2}$ if and only if $\eta > \frac{t-1}{t-2}$ in which case $\tilde{f}(\hat{x}^* = \frac{1}{2}, \eta, t) = t\left(\frac{1}{4} + \frac{\eta}{1-2\eta}\right) \geq 0$ for all t, η.

Case B. (Upper boundary) $\hat{x}_1^I > 1$ if and only if $\eta < \frac{1-2t}{2-3t}$ in which case $\tilde{f}(\hat{x}_1^* = 1, \eta, t) \geq 0$ if and only if $\hat{\eta}(t) \geq \frac{7t-4}{6t-8}$.

Case C. (Interior) $\hat{x}_1^I \in [\frac{1}{2}, 1]$ if and only if $\eta \in \left[\frac{1-2t}{2-3t}, \frac{t-1}{t-2}\right]$. For such η, $\hat{x}_1^I = \hat{x}_1^*$ so $\tilde{f}(\hat{x}_1^* = \hat{x}_1^I, \eta, t) \geq 0$ if and only if

$$\bar{\eta}(t) \geq \left(\frac{t}{1 - 2t\sqrt{\frac{4-t}{8t}}} + 2\right)^{-1}$$

From Cases A-C, we can then construct a (continuous) piecewise function as shown in Figure 3. Finally, it should then be evident that for any $t > 1$, $\eta(t) = 0$.

III Calculation for the Lower Bound on Total Firm Surplus (Figure 4)

We construct G_1 identically as in Supplementary Appendix II by tightening all incentive compatibility constraints

$$G_1(\hat{x}_1|\bar{x}_1) := \begin{cases} (1-2\eta)\dfrac{t(\hat{x}_1 - \bar{x}_1)}{1 - t\hat{x}_1} & \text{if } \hat{x}_i > \bar{x}_i > 0; \\ 0 & \text{if } \hat{x}_i \leq \hat{x}_1; \\ \eta\dfrac{t\hat{x}_1}{1 - t\hat{x}_1} & \text{if } \bar{x}_1 = 0, \end{cases}$$

where $\hat{x}_1 \in [\frac{1}{2} + \gamma, \min\{1, 1/t\}]$, $\bar{x}_1 \in (0, \frac{1}{2} - \gamma]$. So

$$\int_{(0,\frac{1}{2}-\gamma]} G(\hat{x}_1|\bar{x}_1) d\bar{x}_i = (1-2\eta)\int_{(0,\frac{1}{2}-\gamma]} \frac{t(\hat{x}_1 - \bar{x}_1)}{1 - t\hat{x}_1} = (1-2\eta)\frac{2t\hat{x}_1(\frac{1}{2}-\gamma) - t(\frac{1}{2}-\gamma)^2}{2(1-t\hat{x}_1)}$$

The condition in Theorem 1^c then reads for all $\hat{x} \in [1/2 + \gamma, \min\{1, 1/t\}]$,

$$(1-2\eta)\frac{2t\hat{x}_1(\frac{1}{2}-\gamma) - t(\frac{1}{2}-\gamma)^2}{2(1-t\hat{x}_1)} + \eta\frac{t\hat{x}_1}{1-t\hat{x}_1} \geq \int_{1/2+\gamma}^{\hat{x}_1} (1-2\eta) d\hat{x}_1 = (1-2\eta)(\hat{x}_1 - 1/2 - \gamma).$$

Now divide through by $(1-2\eta)$ and define

$$\begin{aligned} f(\hat{x}_1, t, \gamma, \eta) &:= \frac{2t\hat{x}_1(\frac{1}{2}-\gamma) - t(\frac{1}{2}-\gamma)^2}{2(1-t\hat{x}_1)} + \frac{\eta}{1-2\eta}\frac{t\hat{x}_1}{1-t\hat{x}_1} - (\hat{x}_1 - 1/2 - \gamma) \geq 0 \\ \iff \tilde{f}(\hat{x}_1, t, \gamma, \eta) &:= t\hat{x}_1(\frac{1}{2}-\gamma) - \frac{t}{2}(\frac{1}{2}-\gamma)^2 + \frac{\eta t\hat{x}_1}{1-2\eta} - (1-t\hat{x})(\hat{x}_1 - 1/2 - \gamma) \\ &= \hat{x}_1^2 t + \hat{x}_1\left[\frac{\eta t}{1-2\eta} - 1 - 2t\gamma\right] - \frac{t}{2}(\frac{1}{2}-\gamma)^2 + \frac{1}{2} + \gamma \geq 0 \end{aligned}$$

and checking, when $\gamma = 0$, we get back the same equation as before. Now we would like to solve for the requisite function $\bar{\gamma} : [0, 2] \to [0, 1/2]$ which gives the minimum γ required given t and fixing η. First assume $\tilde{f}$ is minimised when $\hat{x}_1$ is interior and differentiating,

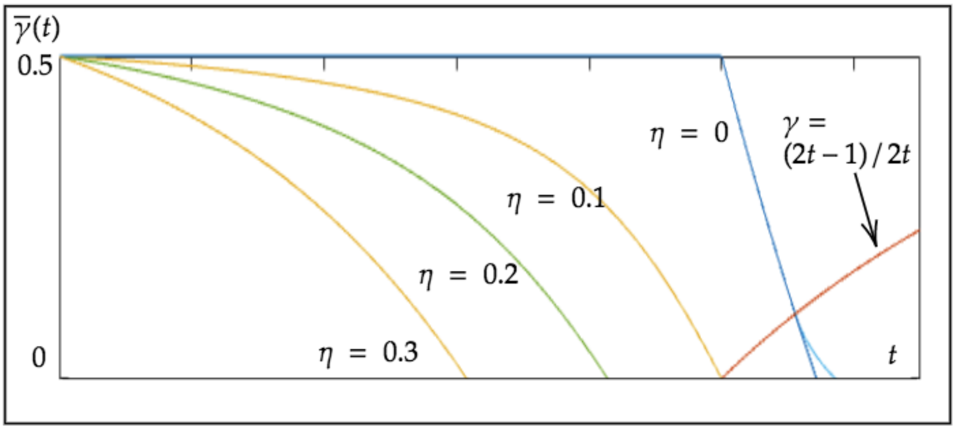
$$\frac{\delta \tilde{f}(\hat{x}_1, \eta, t)}{\delta \hat{x}_1} = 2t\hat{x}_1 + \left[\frac{\eta t}{1-2\eta} - 1 - 2t\gamma\right] = 0 \quad \text{so} \quad \hat{x}_1^I = \frac{1 + 2t\gamma - \frac{t\eta}{1-2\eta}}{2t}$$

We now check whether $\hat{x}_1^I$ does, indeed, lie within $[1/2 + \gamma, \min\{1, 1/t\}]$. From Supplementary Appendix II, $\eta^G(0) < 1$. hence the condition in Theorem 1[c] only fails for $t \leq 1$. It is thus without loss to consider the interval $[1/2 + \gamma, 1]$. Now $\hat{x}_1^I \in [1/2 + \gamma, 1]$ if and only if $\gamma \in [0, \frac{2t-1}{2t}]$. We then find the threshold $\bar{\gamma}(t)$, and check that $\bar{\gamma}$ does in fact lie in that interval. Otherwise, we will work with the boundary condition $\hat{x}_1 = 1$ instead. Substituting $\hat{x}_1^I$ back into $\tilde{f}$:

$$\tilde{f}(\hat{x}^I, t, \gamma, \eta) = \left(\frac{1 + 2t\gamma - \frac{t\eta}{1-2\eta}}{2t}\right)^2 t - 2t\left(\frac{1 + 2t\gamma - \frac{t\eta}{1-2\eta}}{2t}\right)^2 - \frac{t}{2}(\frac{1}{2} - \gamma)^2 + \frac{1}{2} + \gamma = 0.$$

$\bar{\gamma}(t)$ can then be solved and is plotted below in Figure III for various values of η. Note that for $\eta < 0.1$, we have to consider piecewise functions where $\hat{x}_1^*$ which minimises $\tilde{f}$ is in the interior ($\hat{x}_1^* = \hat{x}_1^I$, light blue line) for certain values of t, and on the boundary ($\hat{x}_1^* = 1$, dark blue line) for others.

Figure 5: $\gamma(t)$ threshold function



We now map γ to total surplus extracted by defining

$$TS(\gamma) = 2\int_0^{\frac{1}{2}-\gamma} (1 - tx)dx + 4\gamma^2 t$$
$$= 1 - 2\gamma - t(\frac{1}{2} - \gamma)^2 + 4\gamma^2 t$$

noting that $TS(0) = 1 - \frac{t}{4}$ and $TS(1/2) = t$. The last term $4\gamma^2 t$ solves the Hotelling game played over types $[0, 1/2 - \gamma, 1/2 + \gamma]$. More precisely, let $x^*(p_1, p_2)$ be the indifferent consumer.

$$1 - t\bar{x} - p_1 = 1 - t(1 - \bar{x}) - p_2 \quad \text{so} \quad x^*(p_1, p_2) = \frac{p_2 - p_1 + t}{2t}.$$

So firm 1 makes profits of

$$\Pi_1(p_1, p_2) = p_1 D_1(p_1, p_2) = p_1 \left(\frac{p_2 - p_1 + t}{2t} - 1/2 + \gamma \right)$$

from these types. Solving for firm 1's best response,

$$\frac{\delta \Pi_1(p_1, p_2)}{\delta p_1} = \frac{p_2}{2t} - \frac{p_1}{t} + \gamma = 0 \quad \text{so} \quad p_1^*(p_2) = p_2/2 + \gamma t$$

and by symmetry, $p^* = 2\gamma t$ so equilibrium profits are $2\gamma^2 t$ and total profits are $4\gamma^2 t$. Finally, we plot $TS(\hat{\gamma}(t))/TS(0)$ for various masses of captive consumers as given by Figure 4 (b) in the main text.

Finally, we solve for the perfect information benchmark. If both firms 1 and 2 know the identities of consumers of type $x \in E_1 = [0, \frac{1}{2}]$, then in equilibrium, firm 1 charges price $1 - tx - (1 - t(1 - x)) = t(1 - 2x)$ which is equivalent to the difference in the consumer's net valuation for firm 1 and 2's product. As such, the total surplus extractable from $[0, 1]$ is

$$2t \int_0^1 1 - 2x dx = t/2$$

which is then scaled accordingly by $TS(0)$.