Intrafirm comparison shopping

Saara Hämäläinen*

February 22, 2019

Abstract

We present a price competition model with intrafirm comparison shopping. In the simplest model version, firms offer two products each. Consumers have limited time budgets to search for a product that offers them the highest net utility, i.e., the quality of the product minus the price, which they observe after some random time. We find that the observation lag leads to intrafirm price dispersion: firms benefit from offering a discount on a random subset of products. This *discount set pricing* (i) engages consumers in comparing products *inside stores*, decreasing the time left for comparison *across stores*, and (ii) enables more refined price discrimination between consumers who observe different product information. The strategy is generally more beneficial with a larger number of products.

**Keywords:** Multiproduct price competition; Comparison shopping; Substitute products; Search frictions in-store; Price variation in-store; Quality variation in-store. **JEL-codes:** D43, D83.

---

*University of Helsinki and Helsinki GSE, P.O. Box 17 (Arkadiankatu 7), FI-00014 University of Helsinki, FINLAND, Email: saara.hamalainen@helsinki.fi. I thank Tuomas Takalo, Hannu Vartiainen, Juuso Välimäki, Vaiva Petrikaité, Pauli Murto, Diego Moreno, Klaus Kultti, Mats Godenhielm, Michele Crescenzi, and various conference and seminar audiences for comments. An earlier radically different version of this paper used to circulate under the title "Obfuscation by substitutes". I am grateful to Finnish Academy of Science and Letters, OP-Pohjola Group Research Foundation, Yrjö Jahnsson Foundation, Emil Aaltonen Foundation, and Finnish Cultural Foundation for financial aid. Any shortcomings are my own.
1 Introduction

*Today the world of fashion is scattered across thousands of websites around the globe. The opportunity is to create a curated customer experience for shoppers where they can find all the things they love in one place.* (Chris Morton, Forbes Magazine, Inamedinova (2016))

Price, quality and variety are the key criteria affecting shopping decisions. Majority of consumers say they try to find the lowest prices before a purchase.\(^1\) This is particularly easy on mobile internet, which greatly facilitates multichannel search. However, a somewhat paradox finding from recent customer surveys is that, while there are also specialized price comparison sites, many consumers reveal they primarily engage in comparison shopping *within a store*: preferring to apply the tools provided by their current store (e.g., by using the sorting function on product listings and considering ”sponsored products related to this item”). Fewer than half of the customers even claim they always compare prices *across stores*.

This article contributes to the previous literature by investigating how search incentives and price competition in multiproduct oligopolies are affected by possible comparison shopping both within a store and across stores. To develop a framework for analysis, we extend the fixed sample price search model of Burdett and Judd (1983) in two dimensions: First, each firm has multiple substitute products available that a consumer may consider for buying. Second, a consumer samples products sequentially and can determine which products to study the first. We can thus analyze how much a consumer invests in comparing products within a store as opposed to searching across different stores. As in Burdett and Judd (1983), the eventual number of sampled products is random but, here, the composition of the sample depends on search priority.

Because consumers generally observe more than one product but do not manage to sample all products in the market, competition intensity depends crucially on whether consumers mostly sample offers within a store or across stores. The main finding of this article is that firms can relax competition intensity simply by carrying multiple related products and offering a discount on a random (minor) subset of products. This generates intrafirm dispersion in the consumer surplus provided by a firm’s different products, which (i) engages consumers in comparison shopping *within stores*, restricting the possibility for comparison *across stores*, and (ii) enables more sophisticated price discrimination across consumers who observe different product information. Search technology ultimately determines how many products firms discount in equilibrium.

Such a pricing strategy is generally more beneficial with a larger number of products. The paper thus shows that firms benefit from offering numerous quite similar products. The practice is widely-spread:

\(^1\)See, e.g., Jayasinghe (2016) at Strax Insights and Pilon (2016) at AskYourTargetMarket.
• The Online Drugstore, for instance, provides more than two hundred search results for "ibuprofen". Still, the medical effect of a similar dosage of ibuprofen should remain the same. So, as the use instructions are standardized for medical products, mainly the design of the package, brand and price seem to vary.

• Walmart Grocery also provides a huge selection of products for home delivery. A search for "milk 2 %" provides 73 results in the category Dairy Milk alone. More than a dozen brands of milk are available in different sized containers. It is not immediate to identify the lowest prices per a fluid ounce of milk.²

• In the fashion industry, when an item is in vogue, like oversized cardigans and skinny jeans in the early 2010s or printed clothes and yoga pants in the mid 2010s, stores like Asos, Macy’s, Zara, Zalando, and H&M offer many not very different versions of these demanded items.³ Products can be filtered by size, color and price, but quality varies. Sometimes stores highlight the lowest prices but, as a rule, the cheapest products are of the lowest quality. Sales are often a means for firms to dispose off returned or unsold items. So, lowest prices or showy promotional items may not always come with the highest consumption utilities. Consumers recognize this and many devote substantial amounts of time to finding good value for money.⁴

Research confirms (Baye et al., 2006a,b; Kaplan et al., 2018) that prices indeed vary both within and across stores, however, little is known about how a firm chooses prices for multiple substitute products under price competition.⁵ There is anecdotal evidence, though, that many firms resort to simple rules of thumb like "apply markup pricing on most products and discount prices on some products". Doing so might work quite well, actually. Our model suggests a rationale for firms to offer a discount on a subset of a larger set of related substitute products. This provides a precise prediction to test that prices should vary not only inside stores but also in each product category.⁶

In the model, we consider two stores selling a number of product qualities each. Consumers are searching for a product that offers them the highest net utility, i.e., the quality of the product minus the price, which takes some random time to ascertain. Consumers have limited time budgets but, at each moment before running out of time, they can freely choose in which store to observe new products – and which one to sample next.

⁴A survey by Freedman (2015) at Insurance.com studied shopping times and savings per minute in comparison shopping. The subjects would be willing put, for example, on average, 41 min into shopping for a better price on clothing items.
⁵A monopoly selling two substitutes sets higher prices than a monopoly with just one of the products (e.g., Motta (2004)) to account for the negative selling externality between substitutes arising because a consumer only purchases one substitute.
⁶To our knowledge, there is only indirect evidence on within-a-store substitution between alternative products. Seiler and Pinna (2017), however, find that an additional minute in a supermarket reduces consumer expenditures by $2.1.
We concentrate on almost frictionfree internet environments where the main search cost is the time cost of search. Thus, it is assumed that consumers can switch between stores essentially costlessly, by a click of a mouse, so that the only information processing costs lie inside and not between stores.\(^7\) The focus of the paper is on describing the most profitable pricing strategies with comparison shopping. In the simplest model version, this entails that stores provide one product with a regular price and another product with a discount price; the discount is a randomized percentage of a quality adjusted maximum price and sometimes no discount is given.

This pricing strategy benefits the firms by modifying consumers’ switching incentives. Intuitively, knowing that firms have a mix of regular priced products and discount priced products, consumers have incentives to stay in a store until they find a discount priced product. By offering different consumer surplus on different products, a firm can thus induce a kind of hold up problem on consumers. The hold up is only temporary because consumers would prefer to switch, eventually, when they believe they have sampled the best product the firm provides. However, because consumers have limited time budgets, the time spend on comparison shopping within a store reduces the time spend on comparison shopping across stores, which alleviates the intensity of price competition.

Generally, observing a regular price is both good news and bad news to a consumer. The good news is that there remain fewer products to sample so that finding the discount price is closer (a consumer becomes more optimistic). The bad news is that it is starting to look more likely that the store has only regular prices available (a consumer becomes more pessimistic). The paper shows that the optimistic effect dominates because consumers start off being relatively pessimistic and expect to find mostly regular prices in every store. This requires that a sufficiently small number of (remaining) products carry a discount price.\(^8\) Additionally, we find that variable markups also enable more detailed consumer screening because differently informed consumers pay different prices on average.

A crucial equilibrium property is that, because consumers decide themselves in which order they consider the products, firms optimally randomize the set of products on which they offer a discount. Otherwise, consumers sample products in the order of expected consumer surplus and switch between stores immediately after sampling the first product. Early switching hurts a firm. Specifically, the only reason why a firm might wish to guide consumers first to the highest consumer values is to target faster a consumer who is comparing products across stores after observing a competing offer. But, if most con-

\(^7\)It is easy to see that our framework produces heterogeneity in consumer search outcomes of the same kind as Burdett and Judd (1983). The model variant where the time lag is exponentially distributed also resembles bandit learning models (Bergemann and Välimäki, 2006) because the firms’ pricing strategies basically define a Bandit problem to consumers.

\(^8\)Higher prices sustain consumer search also in Garcia and Janssen (2018) where manufacturers sell to different stores at different prices. Consumers observing a high price infer that the chance they next observe a low price is considerably high. In this paper, we consider the effects of higher prices on stimulating search incentives within a store.
sumers are "captured" in intrafirm comparisons, these "shoppers" are not a very valuable consumer segment. Thus, a firm may not want to induce ordered search in its store.

**Literature**

Evidence is accumulating documenting that consumers frequently fail to find the lowest prices or their preferred products (Ellison and Ellison, 2009; Reutskaja et al., 2011; Kalaycı and Potters, 2011). This can have serious effects on the annual fees of privatized social insurance Hastings et al. (2013) or on the prices of generic pharmaceutical products Bronnenberg et al. (2015), to name just a few applications; for a review of theory, data, and policy issues, see Grubb (2015). Our paper discusses a new way of undermining search efficiency by generating incentives for extensive intrafirm comparison shopping.

Previous literature considers a firm’s incentive to raise the costs of information acquisition (Ellison, 2005; Gabaix and Laibson, 2006; Carlin, 2009; Hagiu and Jullien, 2011) (i) to induce a consumer holdup problem (Wilson, 2010; Ellison and Woltzky, 2012), (ii) to price discriminate efficiently (Petrikaitė, 2018; Taylor, 2017), or (iii) to exploit consumers’ limited abilities to compare products (Piccione and Spiegler, 2012; Chioveanu and Zhou, 2013). Our main departure from the literature is that a firm can only manipulate search costs indirectly – through its joint price strategy for a set of related substitute products.

The idea that proliferation of alternatives can reduce search efficiency is not completely new. Ireland (2007) allows firms to post many prices for the same good on a search engine. Consumers observe a random sample of prices but cannot see if they come from a single seller or from different ones. Thus, to benefit from possible consumer confusion, firms quote two perfectly correlated prices. In our paper, the prices in a given store are not correlated but dispersed to encourage intrafirm shopping and consumers can affect the sample of prices they observe. Carlin and Ederer (2018) develop a dynamic model of search fatigue. Search costs depend on how many products were sampled last period. The equilibrium is characterized by cycles where, in odd periods, firms tire consumers with numerous cheap products. In even periods, firms can charge monopoly prices because, after the frenzy, exhausted consumers choose a firm in random. In our model, firms have no need to change their prices synchronously in time.

A closely related literature concentrates on equilibrium price dispersion (Salop and Stiglitz, 1977; Burdett and Judd, 1983; Stahl, 1989). The new features here are that (i) our model has price and quality variance inside individual stores, and that (ii) consumer information and competition intensity depend on pricing; usually, they are derived from an exogenous information partition following, e.g., Varian (1980). Additionally, the literature has concentrated on price comparisons between stores, abstracting from what occurs within a firm, regarding a firm basically as a black box.9 A notable exception...

---

9But see Salop (1977) who discusses consumer screening by permitting slight differences (in prices or
is Petrikaitė (2018). She finds that a firm has an incentive to differentiate its horizontally differentiated products also by their prices and search costs. This enables sequential screening; consumers never switch early. Another example is Menzio and Trachter (2018) who consider a setup where intrafirm price dispersion arises as a way of discriminating between different shopping habits of consumers.\footnote{Similar to McAfee (1995), Shelegia (2012), and Rhodes (2014), the consumers in Menzio and Trachter (2018) are searching for a bundle of products.}

In the rest of this paper, we first set up the model in Section 2 and define equilibria and develop a familiar benchmark. We then move to our main part of the analysis in Section 3, which describes equilibria with discount set pricing and intrafirm comparison shopping. Section 4 discusses various immediate generalizations of our setting and Section 5 concludes the analysis. Most proofs appear in the Appendix, which also considers three more extensions ("Discount set pricing with numerous products", "Different numbers of available products" and "Scale economies in search technology").

## 2 Model

There are two firms indexed by $i$ selling equal numbers $K$ of related products with different qualities ($q^i_1, \ldots, q^i_K$) and a unit mass of consumers looking for one product each. The set of available qualities is exogenous but can differ across the firms. Buying a product $k$ from firm $i$ provides the consumer with utility $v^i_k = q^i_k - p^i_k$, which depends on the quality of the product $q^i_k$ and the price $p^i_k$. For tractability, the total trade surplus equals unity, $q^i_k - c^i_k \equiv 1$, for all products such that, if $p^i_k \in [c^i_k, q^i_k]$ is the price, $m^i_k = p^i_k - c^i_k \in [0, 1]$ is the markup. Payoffs are linear in prices, qualities and markups. If a mass $B^i_k$ of consumers buys the product $k$ of firm $i$, the firm’s payoff is the weighted sum $\Pi^i = \sum_{k=1}^K B^i_k m^i_k$, whereas the payoff of a consumer is $v^i_k = 1 - m^i_k \in [0, 1]$.

Firms set prices for each product quality individually. A typical consumer wants to purchase the product with the highest net utility but cannot detect this product immediately because observing the quality of the product (a potentially multidimensional characteristic) and the check-out price (a sum of possibly several different charges) has a positive time cost.\footnote{Actually, we only need to assume that utilities are initially unobserved by consumers so, if the price is observed, the quality is not or, if the quality is observed, the price is not – but we do not need both assumptions. At a technical level, our model is just a price competition model but, as long as qualities are uncertain, the model works even if prices are easy to find. See the extension section for further discussion.}

To identify the product with the best value, consumers sample products sequentially. They can decide the firm $i$ and which of the firm’s products $k$ to study next given the history of prices and qualities $h_t$ they have observed up to the present time $t$. The order of search matters because, consumers have limited time budgets and they may run out quality) among seemingly identical products.
of time before they manage to sample all relevant products in the market. Observing product information is a random process, so, some consumers may sample more products than others. The fraction of consumers who observe \( N \) products in total is denoted by \( B_N \geq 0 \). Our model is thus like a random sample size model, e.g., Burdett and Judd (1983), but product observations are collected sequentially.

This is a versatile modeling approach, which allows for various particular search technologies \((B_n)^{\infty}_{n=0} \), putting additional structure on \( B_n \)'s. Especially, the results in this paper are often illustrated using Poisson sampling as an example:

**Example 1 (Poisson sampling)** Search follows a Poisson sampling technology if product information arrives at a Poisson rate \( \theta \) when a consumer is searching for products in a store. Hence, if a consumer decides to consider product \((i, k)\) next, it takes in expectation time \( t \sim \text{Exp}(\theta) \) for her to observe \((q_{ik}, p_{ik})\) and estimate the markup \( m_k^i \) and utility \( v_k^i \). The parameter \( \theta \) captures the quality of search technology as larger \( \theta \) entails higher probability of observing more products: \( \frac{\partial B_N(\theta)}{\partial \theta} > 0 \) for all \( N > 0 \). Search might either stop at a fixed deadline or at a deadline shock, which hits at a point when a consumer becomes tired of searching. Specifically, the fraction of consumers who observe \( N \) products is given by\(^{12}\)

\[
B_N(\theta) = \frac{\theta^N}{N!} e^{-\theta},
\]

in case a consumer has a fixed deadline at time \( t = 1 \), and

\[
B_N(\theta) = \eta \int_0^\infty (\theta t)^N \frac{N!}{N!} e^{-\theta t} e^{-t} dt,
\]

in case a consumer is hit by a deadline shock at \( t \sim \text{Exp}(1) \).

We consider an online search setup where switching between different stores is relatively inexpensive so that consumers have a meaningful choice regarding whether to compare products within a store or across stores. To focus on this consumer decision, it is assumed that switching between different stores bears only a small cost \( \varepsilon > 0 \), which is taken to the limit \( \varepsilon \rightarrow 0 \), so that a consumer can indeed freely choose the order of studying different products. The switching cost can be thought as the cost of a mouse-click to open a new firm’s online store and the cost of switching attention from one firm’s website to another.

In the paper, the switching cost basically acts as a tie-breaking rule: because of the switching cost, if a consumer is indifferent between searching in different stores, she refrains from switching; or, she can also stop searching altogether is she sees no gain in switching back to a store where she has previously been. Additionally, the switching cost also introduces a tiny recall cost: if a consumer has observed two equally desirable

\(^{12}\)The probability that information \((q_{ik}, p_{ik})\) is revealed to a consumer who searches in firm \( i \) for product \( k \) during some short time \( dt > 0 \) is fixed and given by \( \theta dt \). The probability of observing information on more than one product at a time is an event of order \((dt)^2\), i.e., negligible.
products in different stores, she prefers buying the product from her last store to avoid
switching again. This simplifying assumption helps us in calculating different consumer
segments.

Similar to any Stahl (1989) like model, it is assumed that the cost of starting searching
in the first store is zero.\textsuperscript{13} However, we do make the bookkeeping assumption that
collectors may stop searching: so, if consumers do not expect to gain from search, they
will just stop.

The following timeline summarizes the game:

1. Each firm sets prices \((p^i_k, \ldots, p^i_K)\) for \(K \geq 2\) different (exogenous) qualities \((q^i_1, \ldots, q^i_K)\).
   Product information \((p^i_k, q^i_k)\) remains unobservable to consumers until they find it
   but the number of different products \(K\) is common knowledge.

2. Consumers search optimally until their deadlines \(d\). At any instant before the dead-
   line, \(t < d\), they choose which product \((i, k)\) they attend to at that moment \(t\);
   observing product information \((p^i_k, q^i_k)\) takes some (random) time.

3. At the deadline, \(t = d\), purchase decisions are made and payoffs realize.

This is an extensive form game of imperfect information where the consumer has to
to solve a dynamic optimization program. Note that, because the best value for money
wins the consumer, our model is essentially a price competition model: following the
standard logic (e.g., Varian (1980)), firms may hence try to discriminate between hetero-
genously informed consumers by employing mixed markup strategies. Below, we restrict
the analysis to particular symmetric strategies but generally the strategies can be defined
as follows:

\begin{itemize}
    \item A strategy of a consumer is a mapping \(\sigma\) from histories to products \(\sigma : h_t \mapsto (i, k)\).
    \item A history \(h_t = (q^i_k, p^i_k, t(i, k), s(i, k) \leq t)\) records sampled product information \((i, k)\) and the
    sampling time \(t(i, k)\) of each product \((i, k)\).
    \item A strategy of a firm is a probability distribution \(M^i\) over \([0, 1]^K\) determining the
    markups \(m^i_k\) selected for a firm’s all different product qualities \((i, k)\). The markup defines
    the quality adjusted price: \(p^i_k = q^i_k + m^i_k - 1\).
\end{itemize}

\section*{Equilibrium}

We analyze perfect Bayesian equilibria. Strategies must be best responses to beliefs
and beliefs should be consistent with strategies on the equilibrium path.\textsuperscript{14} We consider

\textsuperscript{13}Otherwise, our model is quite different from sequential search models (e.g., Stahl (1989)) and follows
more closely the spirit of random sample size models (e.g., Burdett and Judd (1983), which do not have
per product sampling costs: for concreteness, we could think in our Poisson example that a consumer’s
search cost is zero before the deadline and infinite after the deadline.

\textsuperscript{14}A working paper version (available upon request) contains a lengthy discussion of the optimality
conditions.
symmetric equilibria. Here, symmetry introduces the following restrictions (the fourth of which we relax afterwards):

**Definition 1** (Symmetry) (i) Firms use the same pricing (markup) strategy, \( M^i = M \). (ii) Consumers employ identical search strategies \( \sigma \). (iii) When indifferent between which firm to start from, consumers randomize between the firms and, when indifferent between which product to sample, consumers randomize between the products; and (iv) Firms’ markup strategies \( M \) are symmetric with respect to their different products, i.e., invariant to a permutation (a bijective mapping) \( k'(k) : \{1, ..., K\} \to \{1, ..., K\} \) in the indexing of products \( k \).

Assuming symmetry with respect to products (iv), we can think that a firm chooses the markups of its different products by first observing a realization \( (m^i_1, ..., m^i_K) \) of \( M \) and then randomly selecting on which of its different products \( (i, k) \) to assign each markup \( m^i_k \).

We show later that the assumption is not actually an equilibrium restriction but rather an equilibrium characteristic: it induces a consumer to observe the products of a firm randomly (obviously without replacement). This random search order is also in the firm’s interest because, otherwise, consumers go straight to the lowest expected markup and switch.

To see how much a consumer invests in intrafirm comparison shopping as opposed to comparing products across stores, it turns to be important to understand what kind of markup observations prompt a consumer to switch between stores. To formalize the analysis, it is useful to define the concept of a switching set, i.e., the set of histories after which a consumer switches: with some notation, the set is given by \( \hat{H} = \{ h_t | \text{proj}_1(\sigma(h_t)) \neq \text{proj}_1(\sigma(h_{t-dt})), dt > 0, dt \to 0 \} \).

In the symmetric equilibria we consider, the switching sets are especially simple: assuming a consumer has not observed all the products of a firm (after which she would naturally switch to the other firm’s store), a consumer’s decision to switch depends on the latest sampled markup only; we denote by \( \hat{M} \) the set of markups that induce a consumer to switch.

---

15 We use the standard assumption that beliefs about markups are “passive” and thus the same on and off the equilibrium path. This is justified because (i) all markups within \([m, 1]\) are in use in equilibrium and (ii) the lowest markup \( m \) always sells to a consumer. So, no firm would apply \( m < m \) (an off path markup) because charging \( m \) (an on path markup) has a higher payoff.

15 Because a strategy of a consumer is a mapping \( \sigma : h_t \mapsto (i, k) \), the projection \( \text{proj}_1(\sigma(h_t)) \) picks its first coordinate, which shows the store where the consumer searches; the store will change when a consumer switches.
**Benchmark**

We proceed to show that a symmetric equilibrium with comparison shopping always exists. This is done by constructing a benchmark equilibrium that has some standard properties. The benchmark has comparison shopping only across stores whereas we later build an equilibrium with comparison shopping also within stores. Additionally, there always exists a continuum of Diamond (1971) like equilibria with constant markups and no comparison shopping.

**Remark 1** There exists a continuum of symmetric equilibria where the markups are fixed and the same across the market: \( m^i_k = m \in \left[ \frac{B_1}{1-B_0}, 1 \right] \) for all \((i, k)\).

(i) For \( m = 1 \), a consumer prefers not to search at all.

(ii) If \( m \in \left[ \frac{B_1}{1-B_0}, 1 \right) \), a consumer samples one product.

**Proof.** A firm has no incentive to lower the price because a consumer does not engage in comparison shopping. Also, a firm has no incentive to deviate upward from \( m < 1 \) if a consumer expects the other firm to offer \( m < 1 \) because then the consumer would continue to search. \( \square \)

The existence of a Diamond like equilibrium demonstrates that consumers’ beliefs about markups are important to maintain search incentives because we have assumed that consumers do not continue sampling products if they do not expect any gain from that. However, adding any small fraction of consumers who observe more products would eliminate the above equilibria by introducing price competition. Indeed, the standard properties of equilibrium price distribution (e.g. Varian (1980); Stahl (1989)) hold if consumers search in multiple stores:

**Lemma 1** Suppose that \( B_N > 0 \) for \( N \leq 2K \) and (i) a positive fraction of consumers starts from each store and continues searching to the other store, and (ii) observing a lower markup is more likely to trigger consumer switching than observing a higher markup. Then,

- Each firm applies a mixed pricing strategy for its lowest markup, \( m^i_k \), and randomizes it over the same interval support, \([m, \bar{m}]\).

- The lowest markup is positive \( m \in (0, 1) \) and the highest markup is maximal \( \bar{m} = 1 \).

- If \( \hat{M} = [m, \bar{m}] \), the distribution function of markup \( m^i_k \) is continuous except possibly at the upper bound of the support, \( \bar{m} = 1 \).

Note that it is natural to think that a lower markup is more likely to trigger switching than a higher markup because after seeing a low markup a consumer is more likely to conclude that she has sampled the highest net utility the store has to offer. Our later
analysis shows how this can make it quite beneficial for firms to offer higher markups if consumers search and switch optimally.

With comparison shopping across stores, a familiar-looking equilibrium arises:

**Lemma 2** (Uniform discount pricing; Varian (1980), Stahl (1989), Ireland (2007))

There exists a symmetric equilibrium, where all products in a store have the same random markup, $m^i_k = \ldots = m^i_K = m^i \sim M$ and where consumers observe one product per one store.

The equilibrium markup distribution is

\[
M : \left[ \frac{C}{C + S}, 1 \right] \to [0, 1], \quad M(m) = Pr(m^i < m) = 1 + \frac{C}{S} - \frac{\Pi}{S} \frac{1}{m},
\]

where a firm’s profit is given by $\Pi = C$.

Above, $C = B_1/2$ denotes the fraction of "captives" who observe one firm’s product and $S = 1 - B_0 - B_1$ denotes the fraction of "shoppers" who compare two different firms’ products.

On the surface, the equilibrium in Lemma 2 is reminiscent of the equilibrium in Ireland (2007), discussed in the earlier literature review, where firms carry multiple identical prices. However, unlike in Ireland (2007), where consumers obtain a bunch of different prices from different stores without knowing from which firms they are, here the strategy of uniform pricing creates no confusion, nor higher profit, because a consumer can choose which products to observe.

Technically, apart from the fact that firms carry multiple products in their store, the equilibrium in Lemma 2 is equivalent with Varian (1980) and Stahl (1989). In their classic papers, consumers are also divided into captives and shoppers. There is price competition over shoppers only. With comparison shopping only across stores, the consumer division is basically exogeneous. This is the standard case in the literature.

However, the situation becomes different when consumers also spend time comparing products within a store. This increases the fraction of captives (who compare products within a store) and reduces the fraction of shoppers (who compare discounts across stores). The extent of intrafirm comparison shopping therefore determines the equilibrium competition intensity. We turn to this novelty case in the next section.

**Example 2** (Poisson sampling) Assuming Poisson sampling under a fixed deadline $d = 1$, a firm’s profit becomes $\Pi = \frac{B_1}{2} = \frac{\theta e - \theta}{2}$. Thus, profits are largest at $\theta = 1$ which maximizes the fraction of captives who sample exactly one product.

Profits vanish, $\Pi(\theta) \to 0$, if search technology is either very inefficient, $\theta \to 0$ (most consumers observe zero products), or very efficient, $\theta \to \infty$ (most consumers find multiple products).

One perhaps notable distinguishing feature between these profit-wise similar cases is
that, in the former, lowest markups converge to one (all buyers are captives) but, in the latter, all markups are in the support of M (all buyers are shoppers):

$$\text{supp}(M) = \left[ \frac{C}{C + S}, 1 \right] = \left[ \frac{\theta e^{-\theta}}{1 - e^{-\theta}}, 1 \right] \rightarrow \begin{cases} \{1\}, & \text{as } \theta \rightarrow 0, \\ [0, 1], & \text{as } \theta \rightarrow \infty. \end{cases}$$

To derive the fractions of shoppers and captives, in the above Poisson example, recall that consumer search proceeds as follows: Consumers first approach one firm. Thereafter, they search in this firm’s store until they have analyzed one of its products, then move to the other firm’s store, and search in that firm’s store until they have observed one of its product – or until no time is left.

This entails that, when it is time to stop at the deadline $d = 1$,

$$C = \frac{B_1}{2} = \frac{\theta e^{-\theta}}{2}$$

consumers have observed only the product of firm $i = 1$,

$$C = \frac{B_1}{2} = \frac{\theta e^{-\theta}}{2}$$

consumers have observed only the product of firm $i = 2$,

$$R = B_0 = e^{-\theta}$$

frustrated consumers have observed neither product whereas the remaining shoppers

$$S = \sum_{N=2}^{\infty} B_N = 1 - R - 2C = 1 - e^{-\theta} - \theta e^{-\theta}$$

have observed both.

3 Discount set equilibria

3.1 General results on discount set pricing

All equilibria subsequently considered in this paper fall into a wider class of discount set equilibria:

**Definition 2** [Discount set pricing] In a discount set equilibrium, $\text{DSE}(K,R)$, the products that a firm carries are divided in random into two sets: the products in one set have always regular prices, which offer no positive expected utility to a consumer, $m_1^1 = ... = m_R^1 = 1$, whereas the products in another set may have discount prices,
which offer the same randomized utilities to a consumer, \( m_{R+1}^i = \ldots = m_K^i =: m_d^i \), where \( 0 < R < K \).\(^{16}\)

Prices and search are characterized more specifically as follows:

1. (Discount prices) The random markup \( m_d^i \) is distributed according to a continuous distribution \( M_d^i \) over an interval support \( [m_d, 1] \), which can have an atom of size \( a \geq 0 \) at the highest applied markup \( m_d^i = 1 \).

2. (Pricing regimes) Because the discount percentage \( 1 - m_d^i \) can be zero, with probability \( a \geq 0 \), a firm has only regular prices, i.e., \( m_d^i = 1 \), whereas, with probability \( 1 - a > 0 \), a firm has also discount prices, i.e., \( m_d^i < 1 \).

3. (Search & switching behavior) Consumers start searching from a random firm \( i \) and search in that firm’s store either until they observe a discount price \( m_d^i < 1 \) or until they have sampled all its products; thereafter consumers switch to the other firm’s store and search there either until they observe a discount price \( m_d^i < 1 \) or until they have sampled all its products.

**Proposition 1** In a DSE\((K, R)\), the equilibrium distribution of the markup \( m_d^i \) is

\[
M_d : \left[ \frac{C_d}{C_d + S_d} , 1 \right] \rightarrow [0, 1], M_d(m) = Pr(m_d^i < m|m_d^i < 1) = 1 + \frac{C_d}{S_d} - \frac{\Pi - C_r}{S_d} \frac{1}{m},
\]

possibly with a positive mass point at the highest markup: \( a = Pr(m_d^i = 1) \geq 0 \). The lowest markup is given by \( m = \frac{C_d}{C_d + S_d} \) and a firm’s profit equals \( \Pi = C_r + C_d > B_1/2 \).

Above, \( C_r, C_d \) denote captives who pay a regular price and a discount price, respectively, and \( S_d \) denote shoppers who compare the discount prices of two different stores.

All \( C_r, C_d, S_d, \) and \( a \) are expressed as functions of \( (B_N)_{N=0}^\infty \) in the Appendix (See Eqs. (8) and (12), for the example case with \( R = K - 1 \); the findings generalize to any \( R \)).

The equilibrium price distribution in Proposition 1 can be obtained with the same procedure as in the proof of Lemma 2. We only have to take into account that the lowest prices that captives have observed may be either regular or discount prices. In addition, not all consumers who observe multiple prices in different stores observe discount prices in different stores. They may thus still be captives of their first store if they have observed a discount price in there previously but not in their current store. This enlarges firms’ profits from \( C = \frac{B_1}{2} \) in Lemma 2 because a firm \( i \) is now sheltered from competition not only for consumers who observe just its one price but also for (i) those consumers who observe several prices in firm \( i \)’s store before switching (because it takes time to find the

\(^{16}\) Some might call a regular price also a monopoly price but we prefer utilizing this less dramatic terminology.
first discount price) and (ii) those consumers who observe several prices in firm \(-i\)'s store after switching (because it takes time to find a second discount price).

Generally, the demand of a firm \(i\) may now come from six different consumer segments.

- \(C_r^i\) demand from captives who start from store \(i\) and find no discount price,
- \(C_d^i\) demand from captives who start from store \(i\) and find a discount price only from store \(i\),
- \(S_d^i\) demand from shoppers who start from store \(i\) and find a discount price from both stores \(i\) and \(-i\).

These segments define the totals of captives and shoppers as follows: \(C_r = C_r^1 + C_r^2\), \(C_d = C_d^1 + C_d^2\), and \(S_d = S_d^1 + S_d^2\). The average size of each consumer segment depends on the firm’s discount set strategy, i.e., on the number of available products \(K\) and the number of regular priced products \(R\). They determine how long it takes to find a discount price in a store and, thus, how long it takes to turn a consumer from a captives \(C_r\) (with a regular price) into a captives \(C_d\) (with a discount price) and finally into a shopper.

In practice, deriving the captives and shoppers is a straightforward exercise delegated to the Appendix.\(^{17}\) For instance, we need to consider what happens to consumers whose first store \(i\) is currently providing only regular prices. These consumers are first captured by their initial store (counted as \(C_r^i\)), until they switch, and thereafter they are captured by their second store (counted as \(C_d^{-i}\)), from which they purchase either to avoid switching again (in the case of \(C_r^{-i}\)) or because it offers a higher utility (in the case of \(C_d^{-i}\)).

We want to stress that there could be multiple discount set equilibria depending on the mix of discount prices and regular prices that firms have. To show that a particular equilibrium exist we thus need to show that the following conditions hold: (i) A firm does not prefer to offer a different mix of discount prices or regular prices. (ii) A consumer does not prefer to switch earlier than before sampling a discount price. In the later analysis of the paper, we provide specific conditions to show that a certain discount set equilibrium (e.g. \(K, R = (2, 1)\) and \(K, R = (k, k - 1)\) for \(k = 2, 3, \ldots\) and \(k \to \infty\)) exist. Generally, the following Lemmata summarize the tradeoffs of consumers and firms:

**Proposition 2** A DSE\((K, R)\) exists if the following optimality conditions hold:

1. **(Optimal search)** The expected number of regular prices \(R\) is sufficiently high relative to the expected number of discount prices \(K - R\): e.g., if just one product is usually on discount, \(R = K - 1\), then consumers have no incentives to deviate.

\(^{17}\)Note that we here need to track consumers along all conceivable equilibrium search paths within and across stores. When each store has numerous products that consumers observe in random sequences, combinatorical methods are required to determine the mass of each consumer segment.
2. (Optimal prices) The benefit from poaching shoppers arriving from the other store by offering additional discount prices is sufficiently low: i.e., if Condition (3) as described below is satisfied, then firms have no incentives to deviate.

We next discuss why the depicted search (pricing) behavior is the best response to the described pricing (search) behavior.

Search incentives with discount sets

Because the discount is the same on all discount priced products in a store, the proof of Proposition 2.1 amounts to showing that, if a consumer has not yet observed a discount price in either of the stores but has sampled more regular prices in store $i$ than in store $-i$, her chances of finding one store’s discount price are higher if she continues searching in store $i$ rather than switches into store $-i$. Her chances of finding two stores’ discount prices, on the other hand, are independent of the order in which she searches. This is so because she then anyway has to find a discount price from each of the stores. This probability does not depend on whether she first observes one firm’s discount price or the other firm’s discount price as long as she switches after observing a discount price in one store. The consumer can find either zero, one or two distinct discount prices.

In principle, finding a regular price in store $i$ could make a consumer either more optimistic or more pessimistic about her chances of finding a discount price in store $i$. The optimistic effect arises because, with probability $1 - a$, the store has $R$ discount prices and $K - R$ regular prices. Thus, because products are optimally sampled without replacement, the chances of finding a discount price among the remaining products become higher every time after a consumer finds a regular price: in this pricing regime, they start from $R/K$ and they grow into $R/(K - 1)$ and $R/(K - 2)$, and so on. In isolation, this effect induces a consumer to revise her beliefs about the values of the remaining products in a store upwards over time.

The pessimistic effect arises because, with probability $a$, the store does not have discount prices. During her search, a consumer thus also updates her beliefs about whether this is the case. Because observing a regular price is more likely if the store does not have a discount price, taken alone, this effect induces a consumer to revise her beliefs about the remaining products downwards. However, we can show that the optimistic effect dominates the pessimistic effect as long as consumers do not expect discount prices to be many. Consumers therefore do not become overly pessimistic about a store when they find a regular price. This makes them continue in their first store until they indeed find a discount price or until they have observed all the products in a store.

Pricing incentives with discount sets

We derive in the Appendix (necessary and sufficient) Condition (3) under which a firm does not benefit if it reduces its prices so that its discount set contains more than
one discount price. Obviously, a store could deviate from a single discount price to any additional number of discount prices \( l \), which means that Condition (3) must hold for all \( l \in \{1, ..., K - 1\} \).\(^{18}\)

By introducing an extra discount price, a firm loses some captives who now pay a lower price and switch earlier to the other firm’s store, but also wins new shoppers who compare its price to a previously observed competing discount price. To sustain a discount set equilibrium, the profit loss from captives should thus be larger than the profit gain from shoppers.

\[
|\Delta C_i^d(l)| (1 - p) + |\Delta C_{r-i}^d(l)| (1 - p) \geq |\Delta S_{d-i}^-(l)| p, \quad \text{for all } l \in \{2, ..., K\}
\]

Above, \( |\Delta C_i^d(l)| \), \( |\Delta C_{r-i}^d(l)| \) and \( |\Delta S_{d-i}^-(l)| \) denote the changes in the fraction of captives (negative) and the fraction of shoppers (positive) that arise with introducing \( l \) additional discount prices. Notice that captives who observe a discount price \( C_i^d \) are obviously also affected by increasing the number of discount prices. They switch earlier. However, because the most favorable deviation is to two lowest prices, that always sell to shoppers, and because the demands, e.g., from starting consumers \( C_{r}^i + C_{d}^i + S_d^i \) add up to \((1 - B_0)/2\), we only need to consider changes in \( C_{r}^i, C_{r-i}^d \) and \( S_d^{-i} \).

Note also that, although randomizing the markups of products independently in a store to differentiate their consumer utilities may seem like the optimal strategy for motivating a consumer to scrutinize all products of a firm, it cannot be sustained as equilibrium but in some knife edge cases. Generally, a firm has no incentive to offer different positive discounts because profits from a product are linear in markups between the firm’s current regular price \( m_{i_k}^r = 1 \) and discount price \( m_{i_k}^d = m^i < 1 \) and strictly lower for prices inferior to \( m^i \) due to the self-competition effect that arises.\(^{19}\)

### 3.2 Discount set pricing with two products

This section considers pricing when each firm has two products. We describe a discount set equilibrium where the firms typically have a regular price and a discount price but may also have just regular prices. In the former case, we say that the firm is in the hi-lo regime and, in the latter case, we say the firm is in the hi-hi regime.

As described by Proposition 1, the markup associated with the discount price \( m_{d}^i \) is continuously distributed according to \( M_d \). Thus, using information about the captives and

\(^{18}\)Not surprisingly, a single-crossing condition for \( l \) does not apply regardless of the search technology, so that we cannot simplify (3). However, we do not need to consider the case of adding more regular prices because offering all regular prices happens already on the equilibrium path.

\(^{19}\)For \( K = 2 \), for instance, independent pricing requires that \( B_1(1 - m_1^i) = 1/2B_2((1 - M(m_2^i))(m_1^i - (1 - M(m_1^i)))m_2^i) \) for all \((m_1^i, m_2^i) \in supp(M)\) because, otherwise, the firm has a profitable deviation in its higher price \( m_1^i \) to unity or to its lower price \( m_2^i \) (details from author upon request). In other words, a firm benefits from a move towards a discount set strategy.
shoppers in this case as described in Appendix (Table 3), the existence and properties of equilibrium can be linked to the primitive search technology \( \{B_N\}_{N=0}^{\infty} \).

**Corollary 1** (of Proposition 1) *Assuming that firms have \( K = 2 \) products, a discount set equilibrium exists iff

\[
\frac{2B_1 + 2aB_3}{B_2 + B_3} > \frac{B_1 + \left( \frac{3}{2} + \frac{a}{2} \right)B_2 + \left( \frac{1}{2} + \frac{5a}{2} \right)B_3 + 4aB_{4+}}{B_2 + 3B_3 + 4B_{4+}}
\]

where the atom size is

\[ a = Pr(m_i^d = 1) = \max \left\{ \frac{B_2 - B_3}{B_2 + 3B_3 + 4B_{4+}}, 0 \right\}, \]

and the profit of a firm

\[
\Pi = \begin{cases} 
\frac{1}{2} (B_1 + B_2) + \frac{a}{2} (B_3 + B_{4+}), & \text{for } a > 0, \\
\frac{1}{2} B_1 + \frac{3}{8} B_2 + \frac{1}{8} B_3, & \text{for } a = 0,
\end{cases}
\]

where \( B_{4+} := \sum_{N=4}^{\infty} B_N \).

The fraction of captives and shoppers can be derived from the equilibrium distribution of search results depicted in Table 1, which compresses a lot of information about how consumer search proceeds within a store and across stores in a discount set equilibrium. A careful reader should take a good look at it. Each of the rows in Table 1 follows a consumer along a possible sampling sequence under a specific pricing pattern.

The table provides a search outcome (e.g., \( m_i^d \)), which a consumer may end with, alongside with its probability (e.g., \( \frac{1}{2} \frac{1-a}{2} B_1 \)). The first row, for example, traces a consumer along the sequence \((m_i^r, m_i^d, m_i^r, m_i^r)\). Initially, the consumer starts from a firm \( i \) in the \( hi-lo \) regime and samples first its regular price (with probability \( \frac{1-a}{2} \)). The consumer then observes the firm \( i \)'s discount price. Afterward, the consumer switches to a firm \(-i\) in the \( hi-lo \) regime and samples first its regular price (with probability \( \frac{1-a}{2} \)). Finally, the consumer observes the firm \(-i\)'s discount price. In this similar vein, the following two rows study alternative continuation paths for \((m_i^r, m_i^d)\), i.e., \((m_i^r, m_d^i, m_d^i)\) and \((m_i^r, m_d^i, m_i^r, m_i^r)\), where the order of sampling the products in the last firm is different, either by chance or because the second firm is in the \( lo-lo \) regime, etc.

To derive the demand of a firm, we have color-coded the results (Table 1 has on-path results and Table 2 has off-path results): blue is used for firm \( i \)'s captives and red is used for firm \(-i\)'s captives and green for shoppers. The totals as shown in Table 3 can be obtained by addition, accounting for both a firm’s demand from starting consumers (as firm \( i \)) and its demand from switching consumers (as firm \(-i\)) in the \( hi-lo \) regime.
To see the effects of intrafirm shopping at work, it can be illuminating to compare
the demand of a firm with discount set pricing in Table 1 to that of a firm with standard
uniform discount pricing:

<table>
<thead>
<tr>
<th>store $i$</th>
<th>store $-i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{2}B_1$</td>
<td>$(m^i_d)$</td>
</tr>
<tr>
<td>$\frac{1}{2}B_2$</td>
<td>$(m^i_d, m^i_{d'})$</td>
</tr>
<tr>
<td>$\frac{1}{2}B_3$</td>
<td>$(m^i_d, m^i_{d'}, m^i_{r''})$</td>
</tr>
<tr>
<td>$\frac{1}{2}B_4+ (m^i_d, m^i_{d'}, m^i_{d''}, m^i_{d''})$</td>
<td></td>
</tr>
</tbody>
</table>

With discount set pricing even some consumers who observe all products are captives whereas with standard uniform discount pricing all consumers who observe multiple products are shoppers.

<table>
<thead>
<tr>
<th>store $i$</th>
<th>store $-i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{2} \frac{1-a}{2} B_1$</td>
<td>$(m^i_d)$</td>
</tr>
<tr>
<td>$\frac{1}{2} \frac{1-a}{2} B_2$</td>
<td>$(m^i_d, m^i_{d'})$</td>
</tr>
<tr>
<td>$\frac{1}{2} \frac{1-a}{2} B_3$</td>
<td>$(m^i_d, m^i_{d'}, m^i_{d''})$</td>
</tr>
<tr>
<td>$\frac{1}{2} \frac{1-a}{2} B_4+ (m^i_d, m^i_{d'}, m^i_{d''}, m^i_{d''})$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>store $i$</th>
<th>store $-i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{2} \frac{1-a}{2} B_1$</td>
<td>$(m^i_d)$</td>
</tr>
<tr>
<td>$\frac{1}{2} \frac{1-a}{2} B_2$</td>
<td>$(m^i_d, m^i_{d'})$</td>
</tr>
<tr>
<td>$\frac{1}{2} \frac{1-a}{2} B_3$</td>
<td>$(m^i_d, m^i_{d'}, m^i_{d''})$</td>
</tr>
<tr>
<td>$\frac{1}{2} \frac{1-a}{2} B_4+ (m^i_d, m^i_{d'}, m^i_{d''}, m^i_{d''})$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>store $i$</th>
<th>store $-i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{2} B_1$</td>
<td>$(m^i_d)$</td>
</tr>
<tr>
<td>$\frac{1}{2} B_2$</td>
<td>$(m^i_d, m^i_{d'})$</td>
</tr>
<tr>
<td>$\frac{1}{2} B_3$</td>
<td>$(m^i_d, m^i_{d'}, m^i_{d''})$</td>
</tr>
<tr>
<td>$\frac{1}{2} B_4+ (m^i_d, m^i_{d'}, m^i_{d''}, m^i_{d''})$</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The unconditional equilibrium distribution of (positive) search results: firm $i$’s captives are marked in bluer (light shade for $C^i_d$, darker for $C^i_d'$), firm $-i$’s captives are marked in red (light shade for $D^i_{r''}$, darker for $D^i_{d''}$), and shoppers in green.

Next, we briefly discuss the intuition for Corollary 1:

The condition for existence is just a reformulation of (3) which gives us in this case

$$\frac{\Delta C^i_d}{\Delta S^i_d} = \frac{\frac{1}{4}B_1 + \frac{a}{4}B_3}{(1-a)\left(\frac{3}{8}B_2 + \frac{5}{8}B_3\right)} > \frac{m}{1 - m} = \frac{\frac{1}{4}B_1 + (\frac{3}{8} + \frac{a}{8})B_2 + (\frac{1}{8} + \frac{5a}{8})B_3 + aB_4+}{(1-a)\left(\frac{3}{8}B_2 + \frac{5}{8}B_3 + B_4+\right)}$$

Some notable examples of search technologies satisfying the existence condition are

Example 3 1. A Stahl (1989) model with only captives who observe one product, $B_1 = 1 - \mu$, and shoppers who observe all products, $B_4+ = \mu$, such that $B_2 = B_3 = 0$.  

18
Table 2: The distributions of search results when firm $i$ deviates to $(m^i_d, m^i_d)$ (above) and when firm $-i$ deviates to $(m^d_i, m^d_i)$ (below).

<table>
<thead>
<tr>
<th>Store $i$</th>
<th>Store $-i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{2} B_1 (m^i_d)$</td>
<td>$\frac{1}{2} - \frac{a}{2} B_2 (m^i_d, m^{-i}_d)$</td>
</tr>
<tr>
<td>$\frac{1}{2} - \frac{a}{2} B_2 (m^i_d, m^{-i}_d)$</td>
<td>$\frac{1}{2} - \frac{a}{2} B_3 (m^i_d, m^{-i}_d)$</td>
</tr>
<tr>
<td>$\frac{1}{2} a B_2 (m^i_d, m^{-i}_d)$</td>
<td>$\frac{1}{2} a B_3 (m^i_d, m^{-i}_d)$</td>
</tr>
</tbody>
</table>

2. Uniform sampling probabilities: $B_N = b$ for $N < n_0$ so that $B_1 = B_2 = B_3 = b$ and $B_{4+} = 1 - 3b$ for $b \in (0, \frac{1}{3})$.

3. Poisson sampling for all values of $\theta$ that we have numerically considered up to $\theta > 1000$ and beyond.

By looking at the existence condition, we see that, to sustain a discount set equilibrium:

First, $B_1$ has to be sufficiently large relative to $(B_2 + B_3)/2$ to make sure that the decrease in the fraction of captives (made of starting consumers who observe a single one product) is high enough compared to the increase in the fraction of shoppers (made of switching consumers who observe two or three products) if a firm deviates from discount set pricing to uniform discount pricing.

Second, $B_{4+}$ has to be sufficiently large relative to $B_1$ to make sure that the intensity of competition (determined by the relative fractions of captives $C_d$ and shoppers $S_d$) is still so tense that the loss $1 - p$ in the price charged to the captives $\Delta C_d$ is high enough compared to the gain $p - 0$ in the price covered from the shoppers $\Delta S_d^{-i}$ if a firm deviates from $(m^i_1, m^i_2) = (1, p)$ to $(m^i_1, m^i_2) = (p, p)$.

The size of the atom $a \geq 0$ at $(m^i_d, m^i_d) = (1, 1)$ is determined so that the profit in the hi-hi regime equals the profit in the hi-lo regime. Thus, the benefit of capturing the starting captives (composed of $B_2$) by charging a higher regular price $m^i_d = 1$ and

---

20Note that, even when a discount set equilibrium fails to exist because there is a profitable deviation to two discount prices, there may exist a more complex variant of a discount set equilibrium with three regimes over which a firm could mix: in the hi-hi regime, they would have two regular prices, in the hi-lo regime, one regular price and one discount price and, in the lo-lo regime, two discount prices. The details appear in the author’s thesis.
delaying the switching of consumers, is exactly offset by the benefit of capturing the switching captives (composed of $B_{2+}$) by charging a lower discount price $m_i^d = 1 - \epsilon$ and undercutting a later sampled regular price. The probability of the hi-hi regime $a$ thus adjusts to balance the benefits of inducing comparison shopping within a store as opposed to across stores. Indeed, $(1 - a)/2 < 1/2$ gives the probability that a consumer switches after sampling one product whereas $1 - (1 - a)/2 > 1/2$ gives the probability that a consumer switches after sampling two products. Under a Poisson sampling process, the expected switching time could thus be significantly delayed compared to the case where each store applies a uniform discount, as shown in Figure 1.a. This shows that, while there is no significant cost associated with switching here, pricing variance can still create an effective barrier to switching.

Figure 1: Temporary consumer holdup problem: The expected switching time of a consumer as a function of $\theta$ for $K = 1, 2$ (left), the likelihood of sampling two discount prices as a function of $t$ for $K = 1, 2$ and $\theta = 3$ (right).

Figure 2: Third degree price discrimination: The lowest markup the average consumer has found as a function of $N = 1, 2, 3, 4$ (left), the lowest markup the average consumer has found as a function of $t$ (right) for $\theta = 2$.

Moreover, once a consumer has observed one discount and switched, the time to find a
competing price discount is longer. This manifestation of a (temporary) consumer holdup problem is illustrated by Figure 1.b. It compares the Poisson probability that a consumer has observed two discount prices by time $t$, with discount set pricing,

$$(1-a)^2 \left( \frac{1}{2} \right)^2 \frac{(\theta t)^2}{2!} e^{-\theta t} + (1-a)^2 \left( \frac{1}{2} \right)^2 + 2 \left( \frac{1}{2} \right)^2 \frac{(\theta t)^3}{3!} e^{-\theta t} + (1-a)^2 \sum_{N=4}^{\infty} \frac{(\theta t)^N}{N!} e^{-\theta t},$$

and, with uniform discount pricing,

$$1 - e^{-\theta t} - (\theta t) e^{-\theta t}.$$

Additional price variation also enables the firms to price discriminate more effectively between consumer groups who observe different product information. Figure 2.a describes the expected prices that are paid by consumers who observe a total of one, two, three or four products, respectively. Consumers who manage to sample more products typically obtain better value for their money.

As seen in Figure 2.b, the lowest markup that the average consumer has discovered up to time $t$ is decreasing in $t$. So, although the prices are determined before search begins, equilibrium strategies are played out as if a consumer is assisted by a firm’s sales clerk who first tries to persuade her to buy one of the store’s highest margin products and only later demonstrates cheaper alternatives.

To sum up, discount set pricing is a simple-to-apply pricing strategy which elevates firms’ profits by alleviating price competition and facilitating price discrimination. This positive effect on profits is shown in Figure 3. The non-monotone pattern arises because firms’ profits are maximized when most consumers are captives, which occurs at $\theta = 1$ with uniform discount pricing (the average captive observes one product) and at $\theta \approx 1.25$ with discount set pricing (the average captive observes more products); $\theta$ measures the quality of the search technology.

Figure 3: A firm’s profit for $K = 1$ (with uniform discount pricing) and $K = 2$ (with discount set pricing) as a function of the rate of the Poisson sampling process.
4 Extensions

In this section we consider briefly some immediate extensions:

1. **Advertizing prices:** Making prices observable does not change the equilibrium because prices do not convey any information about consumer values. To see this, consider a discount set equilibrium where product 1 has quality 3.4 and product 2 has quality 4.5. A firm $i$ assigns randomly the following markups $m_1^i = 1$ to one of the products (say, product 1) and $m_2^i = 0.7$ to one of the products (say, product 2). Thus, the prices are 3.4 for product 1 and 4.2 for product 2. However, because the markups were randomly assigned and the consumer expects this, telling the consumer the prices, does not give her any information about consumer values; it is assumed that a consumer has uniform prior about possible qualities $q_i \geq 1$.

2. **Advertizing markups:** The working mechanism through which discount set pricing increases profits clearly requires that firms’ profit margins are unobservable. Otherwise, consumers have no incentives to engage in intrafirm comparison shopping and firms would not benefit from carrying several products. In other words, if a firm was able to commit to pointing a consumer directly to the most valuable products, the firm would lose the additional profit collectable in a discount set equilibrium, which it can obtain if the consumer believes that the firm might still have a better product. The firm has thus no incentive to invest in clearing any uncertainty that its consumers may have about which of its products offer the highest value.

3. **Identical qualities:** The case of identical qualities is intrinsically interesting because it renders our model a classic homogenous product price competition model. Homogenous products applications fit into our modeling framework if firms apply an “add-on” pricing strategy or if the total price is otherwise hard to understand. Such strategies are described, e.g., by Ellison (2005). Our paper then shows a novel example of a setting where equilibrium price dispersion arises between identical products, here both within stores and across stores.

4. **Differences in trading surplus:** Selling products with different trading surpluses would clearly require a different pricing strategy than what we have described pre-

---

21 We abstract from the special case, where one of the product qualities is almost one because then a firm might set an equilibrium price below unity, revealing a positive consumer utility.

22 Except perhaps for reasons outside the model, such as competing for consumer traffic, etc. That is, by the way, not an unnatural application and we study it more in our paper Hämäläinen (2018), where the firms can compete on their Poisson rates $\theta^i$. However, tackling both questions at once is outside the scope of the paper.

23 Booking a flight is perhaps a familiar example where many people have experienced difficulties in calculating the final price (inclusive of canceling and baggage fees, etc.) of a rather standardized and conventional product (say, an economy class flight from city A to B) despite using a comparison site as a starting point.
viously in this paper. A firm would then no longer be indifferent about which product a consumer purchases but would strictly prefer selling the product with the highest trade surplus. Thus, assigning a particular markup realization \( m^i_d \sim M_d \) to a random product would not be optimal anymore.

For example, suppose a firm has four products with the different consumer utilities and costs \((c^i_k, q^i_k)\) given by \((c^1_1, q^1_1) = (1, 2)\), \((c^2_1, q^2_1) = (3, 4)\), \((c^3_1, q^3_1) = (1, 3)\), and \((c^4_1, q^4_1) = (3, 5)\). Then, a firm clearly prefers to assign a random discount (say, 0.3) to either product 3 or 4, leaving it a profit margin of 1.7, than to one of the products 1 or 2, which would only give it a profit of 0.7. Indeed, unless the gain from delaying consumer switching is very high, the firm would actually prefer to get rid of products 1 and 2.

As a result, we believe that it is more accurate to think that firms have products with different trading surpluses because the market is divided into different consumer segments and each surplus class targets a specific consumer segment. So, in our earlier example, products 1 and 2 could be aimed at mid income families and products 3 and 4 could target more wealthy people. It would then make sense to implement discount set pricing separately to each market.

5. **Decreasing the number of products:** Observe firstly that, assuming two products, reducing the number of products by deviating from one product to two products has exactly the same effect as a deviation from one discount price to two discount prices, which by definition cannot be profitable under discount set pricing equilibrium. Thus, our earlier analysis already covers this case and shows that firms have no incentives to limit their product numbers but, rather, have incentives to make their consumers think they can discover some real bargains by searching longer inside their store.

6. **Increasing the number for products:** This case merits its own section and is, therefore, carefully analyzed in the Appendix. All in all, we find that depending on the underlying search technology there could be different discount set pricing equilibria \( DSE(K, R) \) where the number of products \( K \) and the size of the discount set \( K - R \) can be different. The multiplicity of possible outcomes should not come as a surprise to anybody familiar with price competition models in more complex settings and dynamic pricing games.

In their seminar paper, Mailath et al. (1993) argue that a reasonable way to select among various equilibria would be to focus on those that the players who need to coordinate among equilibria like the most, i.e., to *undefeated equilibria*. Here, coordination relies mostly on firms. So, to say the least, a point if favor of discount set pricing is that, according to Proposition 1, despite which discount set equilibrium
is played, it defeats the benchmark of uniform discount pricing from the point of view of the firms.

To be specific, we find that the following corollary holds:

**Corollary 2** (of Proposition 1) Suppose there exists a $DSE(K^1, R^1)$ and a $DSE(K^2, R^2)$ for some $K^1 > R^1$ and $K^2 > R^2$ under a certain fixed search technology. Then,

- If the number of products is the same in both $K^1 = K^2$ but the discount sets are different in size so that $R^1 > R^2$, then a firm has a larger profit in $DSE(K^2, R^2)$ which has a smaller discount set.

- If the discount sets have the same size in both $R^1 = R^2$ but the numbers of products is different so that $K^1 > K^2$, then a firm has a larger profit in $DSE(K^1, R^1)$ which has a larger product set.

The result arises because a smaller discount set or a larger product set introduces a longer spell of intrafirm comparison shopping.

## 5 Conclusion

Electronic retailing has become a numbers game. Firms may carry hundreds of products nowadays. As a consequence, sorting out the ones that provide the highest value for money may not be fast. In this paper, we study how this can affect the extent to which stores are exposed to the competitive pressure of comparison shopping. Some recent survey evidence suggests that consumers sometimes compare products only within one store instead of shopping across stores.

The paper proceeds to develop a formal analysis of how intrafirm comparison shopping affects the intensity of competition. It describes a profitable selling technique called discount set pricing, which amplifies consumers’ incentives for intrafirm comparison shopping. We show that offering a discount on a subset of products introduces a temporary holdup problem (this alleviates price competition) and creates dispersion in search outcomes (this facilitates price discrimination).

The fear is that the negative effects of such a harmless looking selling practice may easily fly under the radar of competition authorities and slip certain suggested remedies, such as banning vector pricing to prevent consumer obfuscation (e.g., Grubb, 2015), because adding more choice does not make the prices of individual products particularly difficult to observe but it does make it harder to find the best prices in the choice set. To capture the problem, empirical research could perhaps study (i) how comparison shopping proceeds within a store and across stores (say, by using click path data from experiments) and (ii) whether margins vary within stores between substitute products (say, by using product level data from retailers), as is suggested by our theory.
Appendix

Discount set pricing with numerous products

In this section we present a two-part sufficient condition on the search technology $B$ which guarantees that a discount set equilibrium $(DSE(K, R))$ exists for any number of product qualities $K$ larger than some cutoff $K^*$. The first part of the condition requires that typically a consumer samples only a moderate number of products so that the tail probability of observing extremely many qualities converges to zero fast enough, $\sum_{N=K+1}^{\infty} B_N = O\left(\frac{1}{K}\right)$. The assumption seems reasonable. According to Proposition 1, it entails that even the highest value offered by a firm, $v_K$, in a discount set equilibrium is almost surely a regular price, $a \rightarrow 1$ as $K \rightarrow \infty$.24 Also, when a firm has extremely many regular priced products, most consumers spend all their time in the first store. By condition 3, a deviation to offering more discount prices is therefore not beneficial since there are not many shoppers $S_{d}^{-1}$ arriving from the other store whom a firm $i$ could poach by a lower price. The second part of the condition requires additionally that the variance of the finding process $D^2[N] = E[N^2] - E[N] = \sum_{N=0}^{\infty} N^2B_N - \sum_{N=0}^{\infty} N B_N$ is not too small relative to its expected value $E[N]$; otherwise, it turns out that the firms may benefit from carrying only discount prices.

**Proposition 3** (Sufficient condition) Assume the following conditions hold:

1. $\sum_{N=K+1}^{\infty} B_N = O\left(\frac{1}{K}\right)$,
2. $E[N]/D^2[N] \leq 1 - B_0$.

Then, there exist a cutoff $K^*$ such that an $DSE(K, K-1)$ exists for all $K \geq K^*$.

When this two-part sufficient condition holds, it is of obvious interest to analyze the limiting equilibrium $DS(K, K-1)$ as $K \rightarrow \infty$. We find that firms can then extract full surplus and divide the market peacefully among themselves. Both conditions hold, for instance, when $B = \{B_N\}_{N=0}^{\infty}$ is given by an appropriately weighted Poisson process. We discuss this more in the following section.

The conditions in Proposition 3 ascertain that the existence condition (3) is satisfied for any large enough number of products $K$ because, then,

$$|\Delta C_r^i| (1 - \frac{m}{k}) \rightarrow 0 \text{ as } K \rightarrow \infty$$

slower than

$$|\Delta S_{d}^{-1}| \frac{m}{k} \rightarrow 0 \text{ as } K \rightarrow \infty$$

since then

$$a \rightarrow 1 \text{ as } K \rightarrow \infty$$

faster than $O\left(\frac{1}{K}\right)$.

---

24. The Landau big O notation $O\left(\frac{1}{K}\right)$ puts an upper bound on the convergence rate of the sum $\sum_{N=K+1}^{\infty} B_N$ by requiring that there exists a fixed number $c$ such that $|\sum_{N=K+1}^{\infty} B_N| \leq \frac{c}{K}$ when $K$ gets sufficiently large.

25. The Landau big O notation $O\left(\frac{1}{K}\right)$ puts an upper bound on the convergence rate of the sum $\sum_{N=K+1}^{\infty} B_N$ by requiring that there exists a fixed number $c$ such that $|\sum_{N=K+1}^{\infty} B_N| \leq \frac{c}{K}$ when $K$ gets sufficiently large.
In other words, we analyze whether a firm’s profits increase when it adds more discount prices. Recall that above, \( \Delta C^i_1 (1 - m) \) denotes the change in profits from consumers who start from our firm and \( \Delta S^{-i}_d/m \) denotes the change in profits from consumers who come from the other store after sampling its discount price.

Generally, as the number of products \( K \) increases, the event of finding one discount price is of order \( O(1/\pi) \) and the event of finding two discount prices is an order \( O(1/\pi^2) \) event. Thus, the profit changes that come with an additional discount price are also of two completely different orders, \( |\Delta C^i_1| \sim 1/\pi \) and \( |\Delta S^{-i}_d| \sim 1/\pi^2 \).

Proposition 3 hence obtains after we show that \( 1 - m \sim 2/K \); the details are in Appendix.26 An additional requirement for this to hold is that the variance of the sampling process must not be too small relative to its expectation for, otherwise, the firm has a profitable deviation to carrying only discount prices. This depends in subtle ways on the convergence rates of \( |\Delta C^i_1| \), \( |\Delta S^{-i}_d| \), and the price ratio \( 1 - m \). Namely, if a firm has just discount prices available,

\[
K \frac{1 - m}{m} \to D^2[N] \text{ and } K \frac{|\Delta S^{-i}_d|}{|\Delta C^i_1|} \to \frac{E[N]}{1 - B_0} \text{ as } K \to \infty.
\]

Thus, the second condition in Proposition 3 is thereby necessary to introduce enough competition to make this kind of deviation not profitable. Sufficient variation in search results around the average helps to avoid cases where the lowest markups \( m \) converge too rapidly to the regular level unity. Intuitively, even if most consumers observe products from just one store there should still remain sufficiently many who find two discount prices to put downward pressure on \( m \).

The general message of Proposition 3 is that, when the number of qualities gets very large, it becomes unattractive to lower prices to poach switching consumers because there are so few of them. Searching for the first discount price takes all the time for most consumers. Almost nobody has time to switch.

In effect, when a firm has a very high number of alternatives, it becomes almost impossible to find even one discount price. Therefore, the fraction of captives \( C_i^i \) who only find regular prices converges to \( 1 - B_0 \). As a result, each firm gets an equal share of these consumers in the symmetric limit DSE:

**Corollary 3** (Limiting equilibrium) If there exist a number \( K^* \) such that \( DSE(K, K - 1) \) exists for all \( K \geq K^* \), then firms extract full surplus in the limit where the number of qualities explodes: \( \Pi \to \frac{1 - B_0}{2} \) as \( K \to \infty \).

In other words, if the sufficient condition presented in Proposition 3 holds, firms can extract full surplus at the limit as the number of qualities becomes infinitely large. This limiting equilibrium resembles the one in Diamond (1971) because both firms have infinitely many qualities with regular prices. Hence, finding the one discount price that could be inside a given store, becomes much like looking for a needle in a haystack. Nevertheless, as a distinction from the well known case of Diamond (1971), where nobody would have an incentive to search when the gain is virtually zero, in our model where consumers tend to use all their time until the deadline, the residual price variation is sufficient to keep all of the consumers searching.

**Discount set pricing with Poisson sampling**

In this section we return to the Poisson setting to analyze the profits in a discount set equilibrium with more than two qualities. We find that the sufficient condition in Proposition 3 can be fulfilled for this case although some modifications have to be made. Particularly, to satisfy the

---

26Recall about orders of convergence that big \( O \) denotes an upper bound \( (f(K) = O(g(K)) \) if \( |f(K)| \leq c |g(K)| \) for \( K \) large), big \( \Omega \) a lower bound \( (f(K) = \Omega(g(K)) \) if \( |f(K)| \geq c |g(K)| \) for \( K \) large); \( f \sim g \) says that \( \frac{f(K)}{g(K)} \to 1 \) as \( K \to \infty \).
second part of Proposition 3 we consider a search technology derived from a weighted Poisson process:

**Definition 3** In a weighted Poisson setting $B(w\theta)$, the original intensity parameter $\theta$ is modified by a random variable $w$. It is assumed that the mean of $w$ is $E[w] = 1$ and its variance $D^2[w] > 0$. With this modification, the probability of discovering $N$ qualities becomes $B_N = \int \frac{(w\theta)^N}{N!} e^{-w\theta} dF$ where $F$ is the probability distribution function of $w$. We also assume that $\int (w\theta)^N e^{-w\theta} dF < \infty$ and $D^2[w] \theta \geq \int \frac{e^{-w\theta}}{1-e^{-w\theta}} dF$.

Unlike the standard Poisson process, where the variance of the process is equal to the mean, the weighted Poisson process features larger variance: $E[N] = \theta \leq D^2[N] = \theta + \theta^2 D^2[w]$. It can thus generate enough variance in consumer’s search results to maintain sufficient competition in the limit. The weighted Poisson setting is also a natural choice to model a random finding process if the search environment changes in a random manner, say, due to unexpected changes in webpage functionality, demand shocks, service breaks, broadband congestion, etc. We observe that when the effect of these random events is sufficiently high, an $DSE$ always exists as long as the firms have many product qualities available.

**Remark 2** If $\{B_N\}_{N=0}^\infty$ is given by the weighted Poisson process with $B_N(w\theta) = \frac{w\theta^N}{N!} e^{-w\theta}$, then (i) $\sum_{N=K+1}^\infty B_N(w\theta) = O\left(\frac{1}{K!}\right)$ as long as $\int (w\theta)^N e^{-w\theta} dF < \infty$ for all $N$ and (ii) $E[N]/D^2[N] \leq 1 - B_0$ as long as $D^2[w] \theta \geq \int \frac{e^{-w\theta}}{1-e^{-w\theta}} dF$. For example, $D^2[w] = 1$ is enough for all $\theta \geq 1$.

As a consequence, as long as the observation rate $\theta$ is sufficiently high and noisy, there is a weighted Poisson process satisfying the conditions of Proposition 3.

**Corollary 4** In a weighted Poisson setting, for any $\theta$ there is a finite cutoff $K^*$ such that $DSE(K, K - 1)$ exists for all $K \geq K^*$. For any $K^2 \geq K^1 \geq K^*$, we have $\Pi(K^2) \geq \Pi(K^1)$.

This is very convenient for the firms because it enables them to undo any profit loss from improving search technology by carrying additional qualities:

**Corollary 5** Consider an improvement from a slower search technology $B(w\theta^1)$ to a faster search technology $B(w\theta^2)$, where $\theta^2 > \theta^1$. Then, for any original number of qualities $K^1$ there exists an extended number of qualities $K^2$ with higher profit, $\Pi(\theta^2, K^2) \geq \Pi(\theta^1, K^1)$, where $K^2 > K^1$.

### Different numbers of product qualities

The number of product qualities $K$ is kept fixed throughout this paper and both firms look initially identical in a sense that they offer the same expected value for money. Particularly, the number of product qualities and the distribution of value for money is the same across the stores. This exercise is well motivated as an analysis of the possible, negative side-effects that a larger number of qualities we observe in the retail markets may have on consumers. However, an extension to different numbers of product qualities would be of obvious interest. So what would high profit equilibria look like in that case? We have no complete analysis for this case but we offer some ideas for future research below.

Suppose that firms carry asymmetric numbers of product qualities, say $K^1$ for firm $i = 1$ and $K^2$ for firm $i = 2$. Proposition 2 suggests that they might not find it optimal to copy as such the equilibrium pricing strategies of $DSE(K^1, K^1 - 1)$ and $DSE(K^2, K^2 - 1)$ as such, respectively, because then consumers would always start from the firm with a lower number of qualities. Because the firm with more products would then only sell to shoppers arriving from the other store its best response would be to carry only discount prices. So this pricing pattern could not arise in an equilibrium because abundant discount prices would encourage consumers
to start from the store with fewer products. This contradicts the assumptions.

The same logic also suggests that, to sustain a high profit equilibrium, where the level of profits exceeds \( B_1/2 \), both stores must offer the same value for money in expectation. Given that the markup distribution depends on the fractions of captives and shoppers that are likely to differ for a firm with \( K^1 \) and for a firm with \( K^2 \) the requirement can be impossible to satisfy except in special cases.

However, we know by Lemma 2 that the low profit equilibrium where stores offer the same randomized value for money for all products always exists. Therefore, if stores can select how many products to carry without costs, then choosing equally many products, \( K^1 = K^2 \), to sustain a discount set equilibrium with higher profit, \( \Pi > B_1/2 \), is obviously an equilibrium if choosing different numbers, \( K^1 = K^2 \), results in the same low profits as carrying one product in a store, \( B_1/2 \).

Generally, Condition (3) suggests that a high profit discount set equilibrium exists as long as the profit gain from captives exceeds the profit loss from shoppers. In practice, firms may apply different instruments to attract captive demand apart from their pricing behavior such as advertising or customer service. Then consumers need not be extremely sensitive to small differences in the number of products in a store and different sized firms may attract some captives.\(^{27}\)

**Scale economies in search technology**

So far we have studied cases where the search technology, \( B \), is independent of the number of products in the stores, \( K \). However, it is also conceivable that in applications search could become easier or more difficult with additional product qualities. To capture this idea in a Poisson setting, we next suppose that an increase in the number of qualities \( K \) modifies the base search rate \( \theta \) by \( \delta(K) \). Thus, the product sampling rate \( \theta(K) = \delta(K)\theta \) becomes a function of \( K \). The multiplier \( \delta(K) \) could be either above one and increasing in \( K \) (for positive economies of scale) or below one and decreasing in \( K \) (for negative economies of scale). To facilitate the exposition, we normalize \( \delta(1) = 1 \).

**Definition 4** (i) There are positive economies of scale in search if \( \delta(K + 1) \geq \delta(K) \geq 1 \) for all \( K \in \mathbb{N} \) and \( \delta(K + 1) > \delta(K) \) for some \( K \in \mathbb{N} \). (i) There are negative economies of scale in search if \( \delta(K + 1) \leq \delta(K) \leq 1 \) for all \( K \in \mathbb{N} \) and \( \delta(K + 1) < \delta(K) \) for some \( K \in \mathbb{N} \).

We find that allowing for scale economies keeps the analysis basically unchanged; the tail probabilities just have to converge fast enough for high values of \( K \):

**Corollary 6** In a weighted Poisson setting, for any \( \theta(K) \) there is a number \( K^* \) such that \( DS(K, K - 1) \) exists for all \( K \geq K^* \) as long as \( \sum_{N=K+1}^{\infty} B_N(w\theta(N)) \rightarrow 0 \) at least at rate \( O(\frac{1}{K}) \).

So, as long as additional products do not drastically improve search efficiency, profits are higher with more products for sufficiently high numbers of products.

Note that it is not obvious whether search has positive or negative scale economies: it can become easier to find an individual product quality when there are more of them but consumers can also get overwhelmed by the larger number. A small number of product qualities can also be displayed in a more compact manner (on one list page) but a larger number maybe needs to be spread around a wider space (on many list pages). Still, the range of \( \delta(2) \) which we think is the most reasonable is \([1, 2]\), where the scale economies are positive but moderate.

To narrow down to this range, suppose that each product has its own independent observation rate \( \phi \), at which it is found on a webpage (representing a store). As a result, (i) if we model a

\(^{27}\)A platform firm like Amazon can also benefit from the commitment opportunities provided by individual retailers’ own pricing choices, which may allow it to carry a more attractive combination of discount prices and regular prices than it could offer on its own.
store with two products as one page with two qualities on it, then the first is found at rate $2\phi$ and the second at rate $\phi$, but, (ii) if we model a store with two products as two pages with one variant in each, then both are found at rate $\phi$. So this provides the estimate that the average finding rate per quality should be between $\phi$ and $2\phi$ for two qualities.\textsuperscript{28,29}

\textbf{PROOF OF REMARK 1}

(i) Since values equal zero, as we have assumed earlier, consumers stay at home. Firms have thus no incentive to offer higher value because no consumer observes it. There exist (almost surely) no other kinds of stay-home equilibria because any tremble in consumers’ beliefs would restore the incentives to search and randomize in prices.

(ii) There is no profitable deviation from $m_1^i$ up to unity as long as $\frac{1-B_i}{2}m_1^i \geq \frac{B_1}{2}m_1^i$ because raising the markup $m_1^i$ will make the consumers continue to the other store so that, instead of selling to half the consumers who observe at least one product, $\frac{1-B_i}{2}$, the firm only sells to consumers who solely observe its own product, $\frac{B_1}{2}$.

\textbf{PROOF OF LEMMA 1}

Our assumptions imply that each store $i$ has some captives $C_i^A$ who observe one if its products $k$ randomly but who observe no products from the other store. They also imply that some consumers $S_{k,i}$ observe, first, some random product $k$ from firm $i$ and, after they switch, some random product $l$ from firm $-i$.

We show first that there exists no equilibrium where firms do not use randomized pricing strategies with random values, $v_k = q_k - p_k$. The proof is by contradiction. Suppose instead that firms apply pure strategies that give a fixed value $v_k = v^i$ for each product quality $k$. Let us consider the products $k = K$ that provide most value for money. Then, if $v^i_K > v^{-i}_K$, firm $i$ has an incentive to decrease $v^i_K$ slightly to extract more revenue from captives $C_i^K$. Also, if $v^i_K = v^{-i}_K$, both firms have an incentive to increase their value offer slightly to capture the shoppers $S_{K,K}$. In conclusion, every store must indeed have at least one product which gives a random value for money. This value must sometimes be lower than unity because, otherwise, firms can increase their payoffs from captives $C_i^K$.

Note that a consumer might respond to a given value observation by switching away the firm depending on her expectations about the value distribution in a store. This might prevent firms from offering certain values in fear of switching. Our assumptions imply that decreasing the value does not stimulate switching.

Next, consider the support of a firm’s lowest markups $\text{supp}(M_i^q)$ and the support of the other firm’s lowest markups $\text{supp}(M_{-i}^q)$. Note particularly that, if these supports would not overlap, i.e. if there would be an interval $S(q,r) = (q - r, q + r)$ such that $S(q,r) \cup \text{supp}(M_i^q) = S(q,r)$ and $S(q,r) \cup \text{supp}(M_{-i}^q)$, then firm $i$ would have a profitable deviation to adjust its markup $m_i^q$ such that all the probability mass in the gap $S(q,r)$ is put on the upper bound of the gap $p = q + r$. This implies that $\text{supp}(M_i^q) \subset \text{supp}(M_{-i}^{-q})$ and $\text{supp}(M_{-i}^q) \subset \text{supp}(M_i^{-q})$ such that the supports of the lowest prices are the same and thus denoted by $\text{supp}(M_i)$.

\textsuperscript{28}Extrapolating this idea, we could of course maintain a constant finding rate $\theta$ for a larger number of product qualities $K$ by replacing the old store, with say just one quality inside the store, by multiple replicas of the old store, each of them with exactly one quality in it. This is just the standard replica argument for constant returns to scale.

\textsuperscript{29}We consider here only constant finding rates $\delta(K)$ for all the qualities in a given store because it is not obvious whether the first ones that a consumer observes should be faster or slower to find than the last ones. Sampling one product could make analyzing another product easier or the products that are the hardest to analyze could be left for last.

29
argument proves that $\text{supp}(M_d)$ cannot have gaps. So they must be intervals $[m, \overline{m}] \subset [0, 1]$. Because the store anyway loses the shoppers $S_{K,K}$ by offering $\overline{m}$, it is optimal to increase $\overline{m}$ up to unity to extract more profit from captives $C_K$. Instead, $\overline{m}$ should be strictly positive because, otherwise, the firm can improve its profits by extracting more profits from captives, $C_K$.

As standard, the supports $M_d$ cannot have atoms unless they trigger a punishment by switching that would not have occurred previously. Otherwise, a firm benefits from reducing its markup slightly below the atom to win over the shoppers when the other store plays the atom. □

**PROOF OF LEMMA 2**

This equilibrium can be constructed like in Varian (1980) or Stahl (1989) if we replace the informed consumers by $S = 1 - B_0 - B_1$ and the uninformed consumers by $2C = B_1$ and note that the sum is less than one, $S + 2C = 1 - R = 1 - B_0 < 1$.

Specifically, by Lemma 1 we know that $\text{supp}(M) = [\underline{m}, \overline{m}]$. A firm’s profit is simply $\Pi(m) = (C + S(1 - M(m)))m$ because captives purchase from it for any markup $m < 1$ but shoppers buy from it only if its markup is lower than the other firm’s markup. This can be evaluated at the upperbound $\overline{p} = 1$ to pin down the profit: $\Pi = C$. The equilibrium price distribution can be obtained by requiring that the profit is the same all over $\text{supp}(M)$: $M$ can be thus solved from $C = (C + S(1 - M(m)))m$. The lowerbound can then be derived by solving $M(m) = 0$.

This symmetric solution is unique if some consumers sample two firms. Since the firms have the same number of products, consumers first approach the one with a lower expected markup or, by our symmetry assumption, choose one firm in random. Asymmetric pricing strategies are thus impossible. Namely, if one store had higher expected values than the other one, it would necessarily attract more captives. But this would imply that the store would actually have higher expected markups because prices always increase with captives. This is a contradiction.30

Deriving the limits is a straightforward calculus exercise. □

**PROOF OF PROPOSITION 2**

**SEARCH PART:**

**Step 1** The probability that the next observed price is a discount price

We derive the conditional probability that a store has a discount price ($l = 1$) and not just regular prices ($l = 0$) assuming that a consumer has observed $r_i^m = N$ regular prices in the store and has not observed one discount price in there.

$$
\Pr (l = 1|r^i_m = N) = \frac{\Pr (l = 1 \text{ and } r^i_m = N)}{\Pr (l = 0 \text{ and } r^i_m = N) + \Pr (l = 1 \text{ and } r^i_m = N)}
$$

$$
= \frac{\Pr (r^i_m = N|l = 1) \Pr (l = 1)}{\Pr (r^i_m = N|l = 0) \Pr (l = 0) + \Pr (r^i_m = N|l = 1) \Pr (l = 1)}
$$

$$
= \frac{K-N}{a + \frac{K-N}{R}(1-a)} = \frac{(K-N)(1-a)}{K a + (K-N)(1-a)},
$$

where

30 Asymmetric pricing strategies are analyzed more in a companion paper Hämäläinen (2018).
Thus, if a consumer has observed \( r^i_m = N \) regular prices (out of \( N \)) in a store, the probability that the next observed price is a discount price is

\[
Pr \left( m^i_{n+1} < 1 | l = 1 \right) Pr \left( l = 1 | r^i_m = N \right) = \frac{1}{K-N} \frac{(K-N)(1-a)}{Ka + (K-N)(1-a)}
\]

whereas, if a consumer has observed \( r^i_m = 0 \) regular prices (out of 0) in a store, the probability that the next observed price is a discount price is

\[
Pr \left( m^i_n < 1 | l = 1 \right) Pr \left( l = 1 | r^i_m = N \right) = \frac{1}{K} (1-a),
\]

where the first term gives the probability of observing a discount price if there is one and the second term gives the probability that there is one. The first term increases with \( N \) but the second term decreases with \( N \). To see which effect dominates, we compare the probabilities of observing a discount price next for \( r^i_m = 0 \) and \( r^i_m = N \).

\[
Pr \left( m^i_{n+1} < 1 | l = 1 \right) Pr \left( l = 1 | r^i_m = N \right) > Pr \left( m^i_n < 1 | l = 1 \right) Pr \left( l = 1 \right)
\]

\[
\frac{1}{K-N} \frac{(K-N)(1-a)}{Ka + (K-N)(1-a)} > \frac{1}{K} (1-a)
\]

\[
\frac{1}{K-N} \frac{K-N}{Ka + (K-N)(1-a)} > \frac{1}{K-N}
\]

\[
Ka + (K-N)(1-a) < K
\]

\[
-N(1-a) < 0
\]

which holds true for all \( N = 1, 2, ..., \) and \( a \in (0, 1) \). In conclusion, the probability of next sampling a discount price is higher in the store where the consumer has earlier sampled a larger number of regular prices \( N \).

**Step 2** Formulating the dynamic optimization problem

We offer a proof that does not require specialized knowledge on continuous time dynamic optimization methods. At each time \( t < d \), a consumer decides in which store to look for the next product knowing the probabilities

\[
B_{t,k} = Pr_t(k) = \begin{cases} 
\frac{(\theta(1-t))^k}{k!} e^{-\theta(1-t)}, & \text{for } k < K - N_1^t - N_2^t, \\
\sum_{i=k}^\infty \frac{(\theta(1-t))^i}{i!} e^{-\theta(1-t)}, & \text{for } k \geq K - N_1^t - N_2^t,
\end{cases}
\]

of observing \( k \) more products before the deadline arrives, which are independent of where the consumer searches.

Because the consumer obviously switches the store after she finds the first discount price, we can represent her strategy by a \( K - N \)-tuple \( \sigma \in \{1, 2\}^K \) of store indices \( i = 1, 2 \), which tells in which store she searches for the remaining \( K - N \) products assuming, that she does not
observe a discount price. For example, \( \sigma = (1, 2, 2) \) for \( K - N = 3 \) captures a strategy where the consumer first tries to look for the next product in store \( i = 1 \). Then, if the product she next observes has a regular price, she tries to look for the following product in store \( i = 2 \) and, if the product she then observes has a regular price, she tries to look for the subsequent product in store \( i = 2 \).

Now, the expected utility of following search plan \( \sigma \) to a consumer who has not observed a discount price by some time \( t < d \) can be decomposed in the following way

\[
V^i_t(N^1_m, N^2_m, \sigma) = \sum_{k=0}^{\infty} B_{t,k} (u(N_d = 1)\phi_k(N_d = 1|\sigma) + u(N_d = 2)\phi_k(N_d = 2, \sigma))
\]

where \( u(N_d = 2) \) denotes the expected utility of finding two discount prices whereas \( u(N_d = 1) \) denotes the expected utility of finding one discount price by the deadline,

\[
\begin{align*}
    u_2 &= E\left[\max\{v^1_K, v^2_K\} | v^1_K > 0 \text{ and } v^2_K > 0\right] > 0 \\
    u_1 &= E\left[v^1_K | v^1_K > 0\right] > 0 \text{ for } i = 1, 2.
\end{align*}
\]

Note that different search strategies \( \sigma \) affect consumer utility only through their effects on the probabilities of finding one and two discount prices, \( \phi_k(u_1|\sigma) \) and \( \phi_k(u_2|\sigma) \). We will next show that \( \phi_k(u_1|\sigma) \) is maximized by continuing in the store where the consumer has observed more discount prices whereas \( \phi_k(u_2|\sigma) \) is independent of where the consumer searches.

The proof constitutes of demonstrating that if the consumers has observed a higher number \( N^1_m \) of regular prices in store \( i = 1 \) and a lower number \( N^2_m \) of regular prices in store \( i = 2 \), then (4) is maximized term-by-term by choosing \( \sigma = (1, ..., K - N^1_m, 2, ..., K - N^2_m) \).

Thus, we want to determine \( \phi_k \)’s as functions of strategy \( \sigma \) for some fixed tuple \( (k, N^1_m, N^2_m) \).

Based on our earlier analysis in Step 1, we introduce the notation

\[
\kappa(N^i_m) = \frac{(1-a)^{K-N^i_m}}{a + (1-a)^K-N^i_m} 
\]

which gives the probability that, if a consumer has so far observed \( N^i_m \) regular prices from a firm \( i \), the next price that she observes from that firm \( i \) is a discount price. Also, we denote by \( \rho(p) \) the probability that a discount price is found on the \( p \)’th draw \( (p \in \{1, ..., k\}) \) and by \( \rho'(p) \) the probability that a competing discount price is afterward observed on the \( p' \)’th draw \( (p' \in \{p + 1, ..., k\}) \).

We can derive the \( \phi_k \)’s from the following equations

\[
\begin{align*}
\phi_k(u_2) &= \rho(1)\rho'(1) + \rho(2)\rho'(2)(1 - \rho(1)) + ... + \rho(k)\rho'(k)(1 - \rho(1)) \cdots (1 - \rho(k-1)) \\
&= \sum_{i=1}^{k} \rho(i)\rho'(i)\prod_{j=1}^{i-1}(1 - \rho(j))
\end{align*}
\]

\[
1 - \phi_k(u_2) - \phi_k(u_1) = (1 - \rho(1)) + (1 - \rho(2)) + ... + (1 - \rho(k))
\]

\[
= \sum_{i=1}^{k} (1 - \rho(i)).
\]

\[\text{31}^3\text{The consumer’s problem is trivial (i) after observing a discount price, } N_d > 0, \text{ and (ii) in the symmetric situation where, } N^1_m = N^2_m, \text{ because we have assumed that, if } t = 0 \text{ the ties are broken in random and, if } t > 0 \text{ they are broken in favor of the store where the consumer is at that particular time. Thus, it suffices to consider asymmetric situations where we can assume without loss that } N^1_m > N^2_m.\]
This shows that vectors \( \rho = (\rho(1), \ldots, \rho(k)) \) and \( \rho' = (\rho'(1), \ldots, \rho'(k)) \) uniquely determine consumer payoffs. Both \( \rho \) and \( \rho' \) depend on \( \sigma \) and \((k, N_m^1, N_m^2)\) and can be constructed from the earlier defined \( \kappa \)'s. However, certain consistency requirements apply because \( N_m^1 (N_m^2) \) obviously increases by one each time the consumer finds a regular price from store \( i = 1 \) \((i = 2)\) and we need to keep track of these dynamics.

Thus, we assume that vector \( \rho \) consists of subvectors \( \rho^1 := (\kappa(N_m^1), \ldots, \kappa(N_m^1 + k^1 - 1)) \) and \( \rho^2 := (\kappa(N_m^1), \ldots, \kappa(N_m^2 + k^2 - 1)) \) where \( k^1 + k^2 = k \). A subvector is obtained by omitting certain elements from a vector but not changing their order.

Another way to put this is saying that the index set \( I = \{1, \ldots, k\} \) of the products the the consumer will still find can be partitioned into disjoint sets, \( I^1 \cup I^2 = I \) and \( I^1 \cap I^2 = \emptyset \), such that \((\rho(i))_{i \in I^1} = \rho \), \((\rho(i))_{i \in I^2} = \rho^1 \), and \((\rho(i))_{i \in I^2} = \rho^2 \).

Additionally, one way to think of the process of building feasible vectors \( \rho \) is that we must construct them element by element, from the first element to the last one, by choosing for every element either the earliest unused element of \( \kappa^1 \) or that of \( \kappa^2 \),

\[
\begin{align*}
\kappa^1 &= (\kappa(N_m^1), \ldots, \kappa(K)) \\
\kappa^2 &= (\kappa(N_m^2), \ldots, \kappa(N_m^1), \ldots, \kappa(K))
\end{align*}
\]

Note that \( \kappa^1 \) and \( \kappa^2 \) are increasing sequences where \( \kappa(N_m) < \kappa((N+1)_m) \) for all \( N_m \in \{N_m^1, \ldots, K-1\} \) as seen from (5). Also, each \( \kappa^1 \) itself can be obtained as a subvector of \( \kappa^2 \) by omitting the first and hence the smallest \( N_m^1 - N_m^2 \) elements of \( \kappa^2 \).

**Step 3** Showing that the probability of observing exactly one discount price before the deadline is larger if the consumer searches where she found more regular prices

Because \( \kappa^1 \) and \( \kappa^2 \) are increasing and \( \kappa^2 \) is a subvector of \( \kappa^1 \), (7) is minimized by selecting

\[
\rho = \begin{cases} 
(k(N_m^1), \ldots, \kappa(N_m^1 + k^1 - 1)), & \text{for } k \leq N_m^1 - K \text{ and } k^1 = k \\
(k(N_m^1), \ldots, \kappa(N_m^1 + k^1 - 1), \kappa(N_m^2), \ldots, \kappa(N_m^2 + k^2 - 1)), & \text{for } N_m^1 - K < k, \text{ and } k^1 = K - N_m^1, \text{ and } k^2 = k - k^1,
\end{cases}
\]

where it is assumed without loss that \( k \leq N_m^1 + N_m^1 - 2K \). Consumer strategy has thus the form

\[
\sigma = (1, 1, \ldots, 1, 2, 2, \ldots, 2).
\]

**Step 4** Showing that the probability of observing exactly two discount prices before the deadline is independent of where the consumer searches for the first discount price

We show that the minimizer of (7) is also a maximizer of (6). First, note that each \( \rho \) with subvectors \( \rho^1 \) and \( \rho^2 \) induces a specific \( \rho' \). To see this, take any \( \rho(i) \) and define the numbers \( n^1(i) \) and \( n^2(i) \) as follows

\[
\begin{align*}
n^1(i) := & \# \{ j \in I^1 | j \leq i \}, \\
n^2(i) := & \# \{ j \in I^2 | j \leq i \}.
\end{align*}
\]

Then we have for the case of \( i \in I^1 \)
\[ \rho'(i) = \begin{cases} 
\sum_{s=N_m^2 + n^2(i)+k-i}^{N_m^2 + n^2(i)+k-i} \kappa(s) \prod_{t=N_m^2 + n^2(i)+k-i}^{s-1} (1 - \kappa(t)), & \text{if } k-i < K - (N_m^2 + n^2(i)) > 0, \\
\sum_{s=N_m^2 + n^2(i)+k-i}^{K} \kappa(s) \prod_{t=N_m^2 + n^2(i)+k-i}^{s-1} (1 - \kappa(t)), & \text{if } k-i \geq K - (N_m^2 + n^2(i)) > 0, \\
0, & \text{if } K - (N_m^2 + n^2(i)) \leq 0.
\end{cases} \]

The case with \( i \in I^2 \) is symmetric.

It is now easy to see that (6) is independent of consumer strategy \( \sigma \) and thus the exact choice of \( \rho \). Intuitively, because the factors \( \kappa(s) \) and \( \kappa(t) \) only change their order, what is gained in the probability of finding the first discount price due to any change in consumer strategy is lost in the probability of finding the second one. Search order thus makes no difference because the consumer must anyway search in both firms until a discount price is found. Thus, the minimizer of (7) is a maximizer of (6) as required. □

**PRICING PART:**

**Step 1 Deriving \( \Pi \) on-path and off-path**

We first derive the profit to a firm who has \( l \) discount prices with \( m < 1 \) and \( K - l \) regular prices with \( m = 1 \). We assume the other firm has \( K \) regular prices and no discount price, with probability \( a \), and \( K - 1 \) regular prices and one discount price, with probability \( 1 - a \). Without loss of generality, let the former firm be firm \( i = 1 \) and let the latter one be \( i = 2 \).

Now consider a consumer who has found \( N \) prices. It is convenient to divide demand of firm \( i = 1 \) from the consumers into six possible consumer segments:

- \( C_{ir}^i(N) \) demand from (captive-like) consumers who start from store \( i = 1, 2 \) and find no discount price,
- \( C_{id}^i(N) \) demand from (captive-like) consumers who start from store \( i = 1, 2 \) and find a discount price only from store \( i \),
- \( S_{id}^i(N) \) demand from (shopper-like) consumers who start from store \( i = 1, 2 \) and find a discount price from both stores.

For concreteness, we assume the perspective of firm \( i = 1 \) next; this is clearly without loss. The profit to firm \( i = 1 \) from consumers who discover \( N \) prices can now be expressed as

\[ \Pi^1(N) = C_{ir}^1(N) + C_{id}^2(N) + (C_{id}^1(N) + C_{id}^2(N)) m + (S_{id}^1(N) + S_{id}^2(N)) (1 - M(m)) m. \]

Firm \( i = 1 \)'s full profit is hence given by the following sum

\[ \Pi^1 = \sum_{N=1}^{\infty} B_N \Pi^1(N). \]

If we write the above expressions of \( C_{ir}, C_{id}^i \) and \( S_{id}^i \) such that \( d \) is kept a free variable, they immediately lend themselves for both on-path (\( l = 0 \) or \( l = 1 \)) and off-path analysis (\( l > 1 \)).

The demands from the defined consumer segments can now be written as
Let us now briefly explain how these are obtained. First, the number two in front of each is there because half the consumers start from firm \(i = 1\) and half of them start from firm \(i = 2\). This implies that, for example, \(C_r^1 + C_d^1 + S_d^1 = \frac{1-B_0}{2}\) because these are those consumers who start from firm \(i = 1\) and then change from \(C_m^1\) to \(C_d^1\) as they find the discount price from firm \(i = 1\) and then from \(C_d^1\) to \(S_d^1\) as they find the discount price from firm \(i = 2\). Of course, not all have time to go through those changes. Thus, we want to calculate how many
end up being \( C^1_{m_i}, C^2_d \) and \( S^1_{d} \) respectively. This depends when they would find each discount price and how many prices they manage to find in total. Consumers \( C^1_d(N) \) start from firm \( i = 1 \) and find only its regular prices. If a consumer finds a total of \( N \) prices, which takes place with probability \( B_N \), the likelihood that none of these is a discount price is \( \frac{(K-d)}{N} \) where \( N \leq K \). Consumers \( C^1_d(N) \) start from firm \( i = 1 \) and find only its discount price. In other words, we require that she does discover a discount price from store \( i = 1 \) but that she does not find a discount price from store \( i = 2 \). Since these events are interlinked because the consumer switches only after she finds the first discount price, we also have to take into account different possible timings as for her observing the first discount price in deriving for instance \( C^1_d \) and \( S^1_{d} \). We hence write \( C^1_d(N) \) as \( \sum_{p=1}^{\min\{N,K\}} \frac{d}{K-p+1} \left( a + (1-a)T(N-p \leq K) \frac{(K-1)}{(N-p)} \right) \), where we sum over all possible finding times \( p \) and \( \frac{d}{K-p+1} \). The likelihood that consumer does not find a discount price before she finds quality \( p \), is \( \frac{(K-d)}{(p-1)} \), the probability that she then finds it is \( \frac{d}{K-p+1} \) and the probability that she does not find another discount price after switching is \( a + (1-a)T(N-p \leq K) \frac{(K-1)}{(N-p)} \). In deriving \( S^1_{d} \) from \( C^1_d \) we can simply change the probability \( a + (1-a)T(N-p \leq K) \frac{(K-1)}{(N-p)} \) (of not finding \( m^{-i} < 1 \) after switching) to the complementary probability \( 1-a - (1-a)T(N-p \leq K) \frac{(K-1)}{(N-p)} \) (of finding \( m^{-i} < 1 \) after switching). For \( C^{-i} \) and \( C^{-i} \) we note that these events arise only if the other firm does not have a discount price, occurring with probability \( a \) and for \( S^{-i} \) that this requires that the other firm does have a discount price, occurring with probability \( 1-a \).

Since unity is in \( supp(M) \) by Lemma 1, profit in \( DSE \) can be obtained by evaluating it for \( l = 1 \) and \( m_d = 1 - \epsilon \) for \( \epsilon \rightarrow 0+ \)

\[
\Pi(K, B | l = 1, m_d < 1) = \sum_{N=1}^{\infty} B_N (C^1_d(N | l = 1) + C^2_d(N | l = 1) + \epsilon \frac{d}{N} + C^2_d(N | l = 1)). \tag{9}
\]

Note that the firm could also raise \( m_d = 1 - \epsilon \) for \( \epsilon \rightarrow 0+ \) a little bit such that \( l = 0 \) and \( m_d = 1 \). That would yield the profit

\[
\Pi(K, B | l = 0, m_d = 1) = \sum_{N=1}^{\infty} B_N (C^1_d(N | l = 0) + C^2_d(N | l = 0) + C^2_d(N | l = 0) + C^2_d(N | l = 0)). \tag{10}
\]

In either case the firm attracts only captives because its markup equals one. The probability of selling to shoppers is negligible. If the firm sets a lower discount price \( m_d < 1 \), however, its profit becomes

\[
\Pi(K, B | l = 0, m_d < 1) = \sum_{N=1}^{\infty} B_N (C^1_d(N | l = 1) + C^2_d(N | l = 1)
\]

\[
+ (C^1_d(N | l = 1) + C^2_d(N | l = 1)) m_d
\]

\[
+ (S^1_d(N | l = 1) + S^2_d(N | l = 1)) (1 - M_d(m_d)) m_d_.. \tag{11}
\]

**Step 2** Deriving \( a, M_d, \) and \( m \)
We can now use these different ways of writing profits (9), (10), and (11) to derive $a$, $M_d$, and $p$. To support randomized pricing strategies, note that all discount prices $m_d \in \text{supp}(M_d)$ should give the firm the same profit. In particular, if it is the case that (10) exceeds (9) for $a = 0$, then the firm has a profitable deviation to using regular prices only unless the other firm also uses regular prices often enough to make it profitable to undercut those higher prices. In that case, $a$ is defined by setting (9) and (10) equal:

$$\sum_{N=1}^{\infty} B_N \left( C_r^1(N | l = 1) + C_r^2(N | l = 1) + C_d^1(N | l = 1) + C_d^2(N | l = 1) \right) =$$

$$\sum_{N=1}^{\infty} B_N \left( C_r^1(N | l = 0) + C_r^2(N | l = 0) + C_d^1(N | l = 0) + C_d^2(N | l = 0) \right).$$

By inserting the expressions for consumers (8) that we derived earlier and by rearranging we get

$$\sum_{N=1}^{\infty} B_N \left( T(N \leq K) + aT(N > K) \right) = \sum_{N=1}^{\infty} B_N \left( T(N \leq K) \frac{K-N}{K} + aT(N > K) + \sum_{p=1}^{\min\{N,K\}} \frac{1}{K} \left( a + (1-a)T(N-p \leq K) \frac{K-N+p}{K} \right) \right) \iff$$

$$\sum_{N=1}^{\infty} B_N \left( \min\left\{ \frac{N}{K}, 1 \right\} \right) = \sum_{N=1}^{\infty} B_N \left( \sum_{p=1}^{\min\{N,K\}} \frac{1}{K} \left( a + (1-a) \max \left\{ \frac{K-N+p}{K}, 0 \right\} \right) \right)$$

$$a = \frac{\sum_{N=1}^{\infty} B_N \left( \min\left\{ \frac{N}{K}, 1 \right\} - \sum_{p=1}^{\min\{N,K\}} \frac{1}{K} \right) + \sum_{p=1}^{\min\{N,K\}} \frac{1}{K} - \sum_{p=1}^{\min\{N,K\}} \frac{1}{K} \max \left\{ \frac{K-N+p}{K}, 0 \right\} )}{\sum_{N=1}^{\infty} B_N \left( \sum_{p=1}^{\min\{N,K\}} \frac{1}{K} \right) \min \left\{ \frac{N-p}{K}, 1 \right\} - T(N > K)}.$$

If the denominator is below zero, then $a = 0$, but if the denominator is above zero, then $a \in (0,1)$.

Next we can derive the equilibrium price distribution $H$ conditional on assumption that $q < 1$ as a function of profits $\Pi$ by requiring that (11) equals (9):

$$\Pi = \sum_{N=1}^{\infty} B_N \left( C_r^1(N | l = 1) + C_r^2(N | l = 1) \right)$$

$$+ (C_d^1(N | l = 1) + C_d^2(N | l = 1)) m_d$$

$$+ (S_d^1(N | l = 1) + S_d^2(N | l = 1)) (1 - M_d(m_d)) m_d.$$
functions of form

A

optimal way to deviate involves choosing the old discount price

m

Deriving the necessary and sufficient condition

prices.

we need not cover those. Essentially the same analysis applies for a larger number of discount

any deviation to (m, m)

Observing that analyzing only deviations to the lower bound suffices

Step 3

such that

\[ m = \frac{\sum_{N=1}^{\infty} B_N \left( C_d^1(N|l = 1) + C_d^2(N|l = 1) \right)}{\sum_{N=1}^{\infty} B_N \left( C_d^1(N|l = 1) + C_d^2(N|l = 1) + S_d^1(N|l = 1) + S_d^2(N|l = 1) \right)} \]  

The lower bound it then the price for which the probability distribution function vanishes

\[ M_d(m) = 0 \]

and for which the firm thus attracts all the shoppers

\[ \sum_{N=1}^{\infty} B_N \left( C_d^1(N|l = 1) + C_d^2(N|l = 1) \right) + (C_d^1(N|l = 1) + C_d^2(N|l = 1)) m + (S_d^1(N|l = 1) + S_d^2(N|l = 1)) m \]

such that

\[ m = \frac{\sum_{N=1}^{\infty} B_N \left( C_d^1(N|l = 1) + C_d^2(N|l = 1) \right)}{\sum_{N=1}^{\infty} B_N \left( C_d^1(N|l = 1) + C_d^2(N|l = 1) + S_d^1(N|l = 1) + S_d^2(N|l = 1) \right)}. \]  

\[ (13) \]

Step 3 Observing that analyzing only deviations to the lower bound suffices

With an extra discount price the firm’s profit is a linear function of form \( A_1 + A_2 m + A_3 m' + A_4 (1 - M_d(m)) m_d + A_5 (1 - M_d(m')) m' \) where \( (1 - M_d(m)) m \) and \( (1 - M_d(m')) m' \) are linear functions of form \( A_6 + A_7 m \); this is easy to see by looking at (11) and (13). As a result, the optimal way to deviate involves choosing the old discount price \( m \) and the new discount price \( m' \) so that they lie on the boundaries of \([m, 1]^2\). We can easily show that the firm never gains by deviating to \((m, m') = (1 - \epsilon, 1 - \epsilon)\) nor by deviating to \((m, m') = (m', 1 - \epsilon)\), where \( \epsilon > 0 \) is a small number. In both cases the firm loses a fraction of its captives but never gains in terms of shoppers starting from the other firm because they almost surely have already found a better price. Therefore, the only relevant deviation to consider is that to \((m, m') = (m, m)\). Obviously, any deviation to \((m, m') \in [0, m]^2\) is dominated by a deviation to \((m, m') = (m, m)\) so that we need not cover those. Essentially the same analysis applies for a larger number of discount prices.

Step 4 Deriving the necessary and sufficient condition

Thus, firm \( i = 1 \) has no profitable deviation from \( DSE \) as long as

\[ \Pi(K, B|l = 1, m_d = m) \geq \Pi(K, B|l = 2, m_{K-1} = m_d = m). \]

This boils down to the necessary and sufficient condition
\[ \sum_{N=1}^{\infty} B_N \left( C^1_r(N|l=1) + C^2_r(N|l=1) \right) \\
+ \left( C^1_d(N|l=1) + C^2_d(N|l=1) \right) m \\
+ \left( S^1_d(N|l=1) + S^2_d(N|l=1) \right) m \geq \\
\sum_{N=1}^{\infty} B_N \left( C^1_r(N|l=2) + C^2_r(N|l=2) \right) \\
+ \left( C^1_d(N|l=2) + C^2_d(N|l=2) \right) m \\
+ \left( S^1_d(N|l=2) + S^2_d(N|l=2) \right) m \]

Note that \( C^1_d + C^1_d + S^1_d = 1 - B_0 \) for both \( l = 1 \) and \( l = 2 \). Since they are thus fixed and both \( C^1_d \) and \( S^1_d \) purchase for certain for \( m \), the loss from this deviation is

\[
\frac{(C^1_r(l=1) - C^1_r(l=2))}{|\Delta C^1_r|} (1-m) + \frac{(C^2_r(l=1) - C^2_r(l=2))}{|\Delta C^2_r|} (1-m),
\]

because those fractions of new consumers now buy for \( m \) whereas before they used to buy for 1. The gain of this deviation is

\[
\frac{(S^2_d(l=2) - S^2_d(l=1))}{|\Delta S^2_d|} m,
\]

because that many new consumers, coming from the other firm after finding its discount price, discover \( m \) earlier and buy for that.

The cases where the firm deviates to a larger number of discount prices \( l > 2 \) must also be covered; the conditions are then basically the same: just replace each \( l = 2 \) with some \( l > 2 \).

\[ \Box \]

**PROOF OF PROPOSITION 1**

This can be obtained directly from the above proof of Proposition 2 by inserting appropriate values of (8) into the expressions (12), (13), (14), and (9) that were derived for \( G \) and \( \Pi \).

**PROOF OF PROPOSITION 3**

So here we analyze only the first and the third terms of (3). The first term can be written as

\[ |\Delta C^1_r| (1-m) = \sum_{N=1}^{\infty} B_N \left( \frac{K-1}{N} \right) \left( \frac{K-2}{N} \right) (1-m), \]

and the third term can be written as

\[ \frac{\sum_{N=1}^{K} B_N}{w(K,N)}. \]

\[ ^{32} \text{We can from now on ignore } T(N \leq K) \text{ and } T(N-p \leq K) \text{ because we are doing this only for taking the limit } K \rightarrow \infty; \text{ both } T \text{'s will equal one for larger values of } K. \]
\[ |\Delta S_d^{-i}| m = (1 - a) \sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} \frac{1}{K} \left( \frac{(K-1)_{N-p}}{N-p} - \frac{(K-2)_{N-p}}{N-p} \right) m. \]

The price ratio can be written as

\[ \frac{1 - \frac{m}{m}}{S_d^i} = \frac{S_d^i}{C_d^i + C_{d-i}} \]

\[ = 2(1 - a) \frac{\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} \frac{1}{K} \left( a + (1 - a)\frac{(K-1)_{N-p}}{(N-p)} \right) + a \sum_{N=K+1}^{\infty} B_N \frac{2K-N}{K}} {\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} \frac{1}{K} g(K, N, p)} + a \sum_{N=K+1}^{\infty} B_N \frac{2K-N}{K}, \]

where we have used a newly defined function \( g(K, N, p) := \frac{(K-1)_{N-p}}{(N-p)} \). It is then easy to calculate that \( w(K, L) = \frac{(K-L)_{L}}{(K-1)K} \) and simplify \( g \) slightly to obtain

\[ w(K, N) = \frac{(K - N)N}{(K - 1)K} \quad \text{and} \quad w(K, N - p) = \frac{(K - N + p)(N - p)}{(K - 1)K}, \]

\[ g(K, N, p) = \frac{N - p}{K} \quad \text{and} \quad 1 - g(K, N, p) = \frac{K - N + p}{K}. \]

We want to show that, for large values of \( K \), the first term of (3) is at least as large as the third term of (3). This requirement can now be stated as

\[ \frac{\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} \frac{K}{N} w(K, N)} {\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} (1 - (1 - a)g(K, N, p)) + aK \sum_{N=K+1}^{\infty} B_N \frac{2K-N}{K}} \geq \frac{\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} w(K, N - p)} {2 \sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} g(K, N, p)}. \]

We have inserted the derived expressions for \( w \) and \( g \), canceled out two \((1-a)\)'s and multiplied both numerators and denominators by \( K \). Next, to take the limit of this, note that functions \( w \) and \( g \) appear both twice in symmetric positions there, after the double summations. These terms can now be rewritten as

\[ 33 \text{We have here replaced } 1 - \frac{(K-1)_{N-p}}{(N-p)} \text{ by } \frac{2K-N}{K}. \]
\[
\frac{K}{N} w(K, N) = \frac{K (K - N) N}{N (K - 1) K} = \frac{1 - \frac{N}{K}}{1 - \frac{1}{K}} \to 1, \text{ as } K \to \infty,
\]
\[
1 - g(K, N, p) = \frac{K - N + p}{K} = 1 - \frac{N}{K} + \frac{p}{K} \to 1, \text{ as } K \to \infty,
\]
\[
K w(K, N - p) = K \frac{(K - N + p)(N - p)}{(K - 1)K} = (N - p) \frac{1 - \frac{N}{K} + \frac{p}{K}}{1 - \frac{1}{K}} \to N - p, \text{ as } K \to \infty,
\]
\[
K g(K, N, p) = K \frac{N - p}{K} = N - p.
\]

This implies that, as long as \(\sum_{N=K+1}^{\infty} B_N = O\left(\frac{1}{K}\right)\) such that \(aK \sum_{N=K+1}^{\infty} B_N \frac{2K-N}{K} \to 0\) as \(K \to \infty\) (we need this convergence) and \(1 - a = O\left(\frac{1}{K}\right)\) (any \(a\) would be just fine here), taking both sides of the inequality to the limit gives

\[
\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} \frac{K}{N} w(K, N) \to 1 \\
\frac{1}{2} \sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} g(K, N, p) \to 1
\]

Note that the same analysis applies for deviations to a larger number of discount prices \(d > 2\) because the limits are just the same for those cases. We just replace \(w(k, N)\) with \(w(k, N, d)\), where \(d\) would then be some larger fixed number:

\[
w(k, N, d) = \frac{(K-1-N)!}{(K-1-d-N)!} \cdot \frac{(K-d)!}{(K)!} \cdot \frac{(K-N)!}{(K-N)!} \\
= \frac{K - N}{K} \left(1 - \frac{(K-1-N)!}{(K-1-d-N)!} \right) \\
= \frac{K - N}{K} \left(1 - \frac{(K-1-N)\cdots(K-d-N+1)}{(K-1)\cdots(K-d+1)}\right) \\
= \frac{K - N}{K} \left(1 - \frac{(K-1-N)(K-d-2)(2-N-K)\cdots(1-d-N-K+1)}{(K-1)(K-d-2)(2-N-K)\cdots(1-d-N-K+1)}\right) \\
\to \frac{N}{K-1} \text{ as } K \to \infty
\]

We also want to make sure that the firm has no profitable deviation to keeping infinite numbers of discount prices. In particular, if the firm had just discount prices the change in its profit from \(C_i\) would be \(\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} \frac{1}{K} \frac{K-N}{K} (1-m)\) and \(S_d^{\prime}\) would be \((1-a)\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} \frac{1}{K} \left(\frac{K-N+p}{K}\right) m\). These can be compared in the similar manner as before.\(^{34}\)

\(^{34}\)Same kind of comparisons can be made for cases where there is a finite number \(N\) of regular prices and an infinite number \(K - N\) discount prices.
\[
\frac{K}{1-a} \frac{1-m}{m} = \frac{2 \sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} (N-p)}{\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} (1 - (1-a)g(K,N,p)) + aK \sum_{N=K+1}^{\infty} B_N \frac{2K-N}{K}} \to \frac{D^2[N]}{1},
\]

because

\[
2 \sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} (N-p) = 2 \sum_{N=1}^{\infty} B_N (N^2 - N) = 2 \sum_{N=1}^{\infty} B_N \left( \frac{N^2}{2} - \frac{N}{2} \right).
\]

\[
= \sum_{N=1}^{\infty} B_N (N^2 - N) = E[N^2] - E[N] = D^2[N],
\]

whereas

\[
\frac{K}{1-a} \frac{\left| \Delta S_d^{-i} \right|}{\left| \Delta C^p_N \right|} = \frac{\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} \left( 1 - \frac{N}{K} + \frac{p}{K} \right)}{\sum_{N=1}^{\infty} B_N \sum_{p=1}^{N} \frac{1}{N} (1 - \frac{N}{K})} \to E[N] \frac{1}{1-B_0}.
\]

So, particularly, this implies that Poisson distribution does not satisfy these requirements as such because for that case \(E[N] = D^2[N] = \theta\). However, a weighted Poisson variable \(N \sim P(w\theta)\) where \(w > 0\) and \(E[w] = 1\) works fine since then \(E[N] = \theta\) but \(D^2[N] = \theta + \theta^2 D^2[w]\). In that case we would also need to require \(E[N]D^2[w] \geq \frac{B_0}{1-B_0}\). □

**References**


