Drip Pricing, After-sales, and Sequential Buying with Behavioral Consumers*

Philipp Brunner§ Christian Zihlmann¶

February 20, 2023

Abstract

A common but controversial practice is to advertise the base product's price up-front, but reveal additional fees for optional add-ons only as the consumer proceeds through the purchase process. We study such an after-sales market with consumers who compare the price of the add-on to the base good price—a behavior that temporarily increases add-on demand. Our model incorporates multiple behavioral micro-foundations to account for this phenomenon. Our findings show that the presence of such behavioral consumers can reduce the surplus of classical consumers and well-intentioned policies aimed at educating consumers may actually decrease consumer surplus.

JEL *classification*: D11, D40, D60, L10, L40

Keywords: Drip pricing, After-sales, Add-on, Heterogeneous consumers, Mental accounting, Relative thinking, Salience

*We are grateful for valuable comments to Jean-Michel Benkert, Berno Büchel, Francisco Gomez Martinez, Paul Heidhues, Holger Herz, Mats Köster, Igor Letina, Marc Möller as well as to seminar and conference participants at DICE, the Swiss Society of Economics and Statistics (2022), the Swiss IO Day (2022), and the University of Bern.

§Department of Economics, University of Bern. Email: philipp.brunner@vwi.unibe.ch

¶Department of Economics, University of Fribourg. Email: christian.zihlmann@unifr.ch.

1 Introduction

Decoupling a product or service into a base and an extra item has become a common business practice. Airlines charge fees for optional services such as seat allocation, while car dealers offer ancillary service contracts. Banks charge late and overdraft fees, and electronic suppliers promote extended warranties. Possibly because it increases the consumer's willingness to pay ("WTP") for the add-on, leading to higher profits for the firm (Morwitz, Greenleaf, and Johnson, 1998; Blake, Moshary, Sweeney, and Tadelis, 2021). As a result, add-on selling and drip pricing[1] have drawn the attention of competition authorities and raised concerns among policymakers.[2] In October 2022, the Biden-Harris administration announced an initiative to tackle potential issues arising from such practices (The White House, 2022).

At the same time, consumers may relate these additional costs to a reference price. Spending \$10 for shipping and handling may feel less significant when the product price is \$100, as opposed to \$20. In the former case, we may perceive that opting for home delivery won't put a dent in the wallet. In the latter case, we may decline the extra service and instead pick up the product in-store. Consumers are indeed more likely to buy a queue-skipping voucher when the price of the ski pass was more expensive (Erat and Bhaskaran, 2012) and to put more effort into redeeming a \$5 discount for a \$25 radio than for a \$500 TV (Thaler, 1980; Bushong, Rabin, and Schwartzstein, 2021).

We integrate this behavior into a model of drip and add-on pricing, which features both classical and behavioral consumers. Our findings reveal that the presence of behavioral consumers motivates firms to increase base good prices, potentially creating an endogenous base good price floor. As a result, behavioral consumers can exert a negative externality on the surplus of classical consumers. Further, regulatory measures like consumer education or exogenous price floors for base goods may have unintended consequences.

In our market, two firms compete in prices with horizontally differentiated base products. Consumers first purchase a base good from a seller. Subsequently, consumers are presented with the seller's offer for an optional ancillary product, which they can reject at

[1] Drip pricing is the sequential presentation of prices and defined as "[...] a pricing technique in which firms advertise only part of a product's price and reveal other charges later as the customer goes through the buying process. The additional charges can be mandatory charges [...] or fees for optional upgrades and add-ons" (Federal Trade Commission, 2012). In this work, we focus on the latter case of optional upgrades and add-ons.

[2] See, for example, the British Competition Market Authority (2022) and the US Federal Trade Commission (2022).

zero costs. The seller enjoys monopoly power in the add-on market. Although there could be multiple sources of market power in the add-on market,[3] this assumption will be relaxed later to show that our results generalize to sequential buying settings with competition in the after-sales market.

Consumers are heterogeneous. Some consumers behave according to traditional theory. For those *classical consumers*, the WTP for the add-on is naturally independent of the price of the previous base good purchase. Other consumers are boundedly rational. Those *behavioral consumers* compare the add-on to the more expensive base good, making them less sensitive to the add-on cost, resulting in temporarily increased demand for the add-on. Our reduced-form model accommodates this behavior without relying on a specific psychological mechanism, and we show that behavioral micro-foundations such as salience, relative thinking, proportional thinking, mental accounting, and anchoring-and-adjustment satisfy our model's assumptions. Therefore, behavioral consumers' WTP for the add-on increases with the price of the base good, leading to an add-on WTP that is strictly greater than that of classical consumers.

Firms face a mixed population of consumers and must decide on their pricing strategy. They know the distribution of consumer types, but they cannot identify an individual's type, making it impossible to engage in price discrimination. Nonetheless, firms can potentially exploit behavioral consumers by increasing the add-on price to their higher add-on WTP. While classical consumers will not buy the add-on anymore, firms can extract a higher add-on mark-up from behavioral consumers. This strategy is referred to as the *exploiting strategy*. Alternatively, firms can choose not to adapt their pricing strategy, resulting in the add-on price being equal to the WTP of the classical consumers. In this case, all base good buyers accept the add-on offer, the *non-exploiting strategy*.

The equilibrium strategy is determined by the proportion of behavioral consumers. With a low share, the exploiting strategy is not optimal as the loss from classical consumers not buying the add-on exceeds the extra revenue from selling it at a higher price to behavioral consumers. The add-on is priced as if all consumers were classical, resulting in the same outcome as in the benchmark economy, which consists of only classical consumers.

The strategy of exploiting behavioral consumers in the after-market is optimal when they become sufficiently frequent. Crucially, pursuing this strategy also affects the optimal

[3] See, for example, Diamond (1971); Shapiro (1994); Ellison (2005); Gabaix and Laibson (2006).

price of the base good due to the following trade-off: Increasing the base good price allows firms to set an even higher add-on price, since the behavioral consumers' WTP for the add-on depends positively on the base good price. Simultaneously, however, demand in the base good market decreases, and with it, also in the after-sales market. The two effects arising from this trade-off determine the optimal prices and the distributional effects between the two consumer types. Indeed, the optimal base good price may be higher or lower than the benchmark price, depending on the proportion of behavioral consumers.

We demonstrate that there exist (unique) equilibria in which classical consumers are worse off: the presence of an intermediate proportion of behavioral consumers creates a negative externality on classical consumers. Yet, higher proportions of behavioral consumers subsidize classical consumers.

Only perfect competition in the base good market offers complete protection for the classical consumer. Due to the (perfectly) competitive pressure, firms cannot increase the price of the base good and the mark-ups from the add-on market must be completely passed on to the base good market. As a result, classical consumers are always weakly better off, benefiting from the cross-subsidization of behavioral consumers. Thus, not only imperfect competition, but also perfect competition may raise important distributional concerns.

Recently, legislation was proposed to state all prices upfront, including those of optional add-ons such as seat allocations (Department of Transportation, 2022). We analyze the effect of such legislation, which, in essence, is an education policy that aims to reduce the share of behavioral consumers. We find that revealing prices upfront can improve total consumer surplus. But, depending on the efficacy of the intervention, it also can harm both consumer types.

Similarly, a government that enacts an exogenous price floor on the base good to prevent loss leading, often also referred to predatory pricing, can actually harm both consumers if the floor is set above the endogenous price floor created by behavioral consumers. In that case, the imposed price floor makes the base good more expensive. As a consequence, firms can take advantage of the behavioral consumer even more, while forgoing some of the competitive pressure to redistribute the mark-up from the after-sales market to the base good market. Further, we discuss the implications of a price cap on the add-on and find similar effects to a education policy.

This paper contributes to the recent debate around drip pricing in economics (Gabaix and Laibson, 2006; Kosfeld and Schüwer, 2016), marketing science (see Ahmetoglu, Furnham, and Fagan, 2014, for a review) and antitrust (see Greenleaf, Johnson, Morwitz, and Shalev, 2016, for a review). A decoupled good consisting of a base and add-on product can lead to increased demand because consumers underweight the add-on price (Hossain and Morgan, 2006; Ellison and Ellison, 2009; Brown, Hossain, and Morgan, 2010; Santana, Dallas, and Morwitz, 2020). In a large field experiment on StubHub.com, Blake et al. (2021) find that drip pricing increases demand in quantity and quality (see also Dertwinkel-Kalt, Köster, and Sutter, 2020). Moreover, drip pricing has been shown to reduce consumer surplus in experimental markets (Rasch, Thöne, and Wenzel, 2020), and there is empirical evidence that add-on purchases are more frequent if base good prices are higher (Xia and Monroe, 2004; Chatterjee, 2010; Erat and Bhaskaran, 2012; Jang and Chung, 2021).

To inform policy, we address these empirical observations and introduce behavioral consumers whose WTP for the add-on positively depends on the base good price into an after-sales market model. The presence of behavioral consumers can lead to the emergence of an endogenous base good price floor, which can reduce competition as firms have less incentive to redistribute profits from add-ons to the base product market. Moreover, we find that education policies, such as stating all prices upfront, may indeed reduce the proportion of behavioral consumers. However and despite good intentions, such policies could do more harm than good.

This paper also relates to the literature on pricing in multi-good settings and loss leading (see Armstrong and Vickers, 2012, for a review). In typical models, firms enjoy ex-post monopoly power over the add-on, allowing them to extract high margins from those after-sales products (Holton, 1957; Diamond, 1971; Lal and Matutes, 1994; Verboven, 1999; Coppi, 2007). However, perfect competition forces firms to redistribute those rents to the base good, which must be sold as a loss-leader to attract consumers ex-ante (Shapiro, 1994; Ellison, 2005). Loss-leading is often seen as a predatory practice that exploits consumers and reduces welfare (Chen and Rey, 2012). For this reason, the issue has gauged the interest of researchers and antitrust agencies alike. For example, 22 U.S. states prohibit the sale of goods below costs, and loss-leading is banned in several countries in the European Union.[4]

In our model, however, a law that enacts a price floor on base good always weakly reduces the surplus of the remaining consumers. Thus, banning loss-leading may actually

[4]See https://www.aeaweb.org/research/loss-leading-bans-retail-competition.

be detrimental for consumers. This result contributes to recent evidence that points towards the potential negative effects of such bans due to other reasons, such as for example a smaller product choice (Johnson, 2017).

Lastly, our insights also contribute to the literature on behavioral industrial organization (Ellison, 2006; Armstrong, 2008; Spiegler, 2011; Armstrong, 2015; Grubb, 2015; Heidhues, Kőszegi, and Murooka, 2016; Heidhues and Kőszegi, 2018; Heidhues, Johnen, and Kőszegi, 2021; Johnen, 2020; Gamp and Krähmer, 2022; Schumacher, 2022). It particularly adds to the literature that investigates heterogeneous consumer populations (Armstrong and Vickers, 2012; Heidhues and Kőszegi, 2017). Our work is related to Ellison (2005) and Gabaix and Laibson (2006), who studied behavioral models in after-sales markets where firms can hide add-on prices. Related is also Rasch et al. (2020), who theoretically and experimentally investigated drip pricing with mandatory add-on purchases. However, in our model, there are no surprise charges: consumers voluntarily purchase the add-on or can reject the seller's offer at no cost. Our article also relates to Michel (2017), where consumers overestimate the value of optional warranties due to underestimating the costs of returning faulty products. But our model is unique in that it considers consumers who differ in whether they evaluate the price of the add-on in reference to the price of the base good, a behavioral foundation that has not yet been introduced in after-sales market models, and is applicable to a variety of add-ons beyond warranties.

Our results show that the presence of behavioral consumers can jeopardize the surplus of classical consumers. This result refines insights from previous literature that usually finds the presence of behavioral consumers does not adversely affect classical consumers, and may even benefit them—the infamous reverse Robin Hood effect (e.g., Armstrong and Vickers, 2012).[5] Because usually the behavioral consumer subsidizes the classical consumer, the discussion has centered on whether to protect behavioral consumers from their own mistakes. Our findings raise the question of whether one should protect the perfectly acting classical consumer from the mistakes and biases of others.

This paper proceeds as follows. Section 2 defines the model set-up. Section 3 provides the equilibrium analysis. Section 4 analyzes policy implications. Consumer behavior is microfounded in Section 5 and Section 6 provides further results. Section 7 concludes. All proofs are presented in appendices.

[5]The issue also received considerable attention in the popular press, see for example `https://thehill.com/opinion/finance/580513-reverse-robin-hood-is-real`.

2 Model

Many products feature add-on components. After purchasing a base product, the consumer is subsequently dripped with the offer for an ancillary product (or service).[6] Formally, we suppose two firms $j \in \{1,2\}$ compete in prices with differentiated base goods, which are imperfect substitutes. Each firm offers a base good at price $p_{1,j}$. There is a continuum of consumers. Firm j faces a weakly concave demand function $D_j(p_{1,j}, p_{1,-j})$, which is twice continuously differentiable, strictly decreasing in its own price and $\lim_{p_{1,j}\to\infty} D_j(\cdot) = 0$. We suppose that the base good demand is (i) supermodular, (ii) the own-price elasticity is stronger than the cross-price elasticity, and (iii) satisfies $\left|\frac{\partial D_j^2(\cdot)}{\partial p_{1,j}^2}\right| \geq \frac{\partial D_j^2(\cdot)}{\partial p_{1,j}\partial p_{1,-j}}$. The last assumption implies that the decrease (increase) of demand is higher when only one firm increases (decreases) prices than when both change prices.[7]

Once a consumer purchased the base good(s), firms offer one unit of an additional good (or service) per base good sold at price $p_{2,j}$. The add-on demand for firm j is thus bounded from above by $D_j(p_{1,j}, p_{1,-j})$. Consumers are locked-in in the aftermarket, which implies monopolistic power for firms. For simplicity, we suppose that add-ons are homogeneous across firms and marginal costs of production for both goods are normalized to zero. The WTP for the add-on of consumer i is given in reduced form by the expression[8]

$$W(v_2, \Delta), \text{ where } \Delta = \beta_i \Delta(p_2, p_1)$$

$W(v_2, \Delta)$ is strictly increasing in both arguments, non-negative, weakly concave and twice continuously differentiable. All consumers receive a gross consumption utility of v_2 for the add-on purchase.[9]

Some consumers may compare the price of the add-on with the price of the base good, leading to a temporarily increased add-on demand. Our model introduces this behavior in a reduced form with $\Delta = \beta_i \Delta(p_2, p_1)$, which is strictly decreasing in p_2, strictly increasing in the reference price p_1 and $\Delta(p_2, p_1) = 0$ when $p_2 = p_1$. In Section 5, we formally show that such behavior can be micro-founded by salience, relative thinking, proportional

---

[6]Our model also applies to tentative purchases of the base product, assuming that consumers search little once tentatively committed to a base good, as demonstrated by Rasch et al. (2020).

[7]For example, the linear demand function derived in Singh and Vives (1984) satisfies these assumptions. In section 6.2, we show that our results also hold for unit demand à la Hotelling (1929).

[8]To ease notation, we will suppress the firm index j when not necessary. Thus, a single subscript indicates whether it is the price of the base good or add-on.

[9]The gross utility v_2 also captures any potential complementary effects between the base product and the add-on.

thinking, mental accounting or anchoring and adjustment. The parameter β_i captures the strength of the behavioral mechanism. $\beta_i = 0$ characterizes a classical consumer who is not subject to any of the above discussed behaviors. The classical consumer's WTP for the add-on is independent of the base good price and constant $W(v_2, 0) = W(v_2)$.

A β_i larger than zero characterizes a behavioral consumer who puts the add-on price p_2 in relation to the price of the base good p_1. The argument $\Delta(p_2, p_1)$ captures this behavioral mechanism. It follows that behavioral consumers have a higher WTP for the add-on than classical consumers when $p_1 > p_2$.[10] Note that we do not specify a utility function for behavioral consumers. Since none of our results depend on a specification, we remain agnostic about whether the behavioral mechanism increases the utility of behavioral consumers or not, and thus, whether they receive a utility of $W(v_2, \Delta)$ or $W(v_2)$ when consuming the add-on.[11] In the following, we focus on the case of a more expensive base good than the add-on, such that $\Delta > 0$ in any equilibrium.[12] In Appendix C.4, we investigate the case in which the base good costs less than the add-on.

We analyze an economy that potentially consists of both types, classical and behavioral consumers, i.e. $\beta_i \in \{0, \beta\}$ with $i = \{c, b\}$ and $\beta \in (0, 1]$.[13] The share of behavioral consumers in the population is denoted with $\alpha \in [0, 1]$. Firms know the distribution of the types of consumers but cannot identify an individual's type. It follows that firms cannot price discriminate and need to offer the same prices p_1 and p_2 to all consumers. The timing of the game is as follows:

- Period 0: Firms choose the prices p_1 and p_2 simultaneously.
- Period 1: Consumers observe the base good price p_1, choose a seller, and buy the base good(s).
- Period 2: Each firm offers an add-on to its base good consumers. Consumers observe the add-on offer and either accept or reject it.

[10]Note that the mechanism could also work through a price rather than WTP distortion. That is, behavioral consumers may misperceive the price of the add-on and perceive it as cheaper than it actually is. This affects the incentive constraint to buy the add-on similarly to a WTP distortion and thus, does not change our analysis.

[11]Proposition 3 deals with welfare of behavioral consumers. We show in the proof that the result is independent of the welfare specification.

[12]We rule out the corner solution $p_1 = p_2$, which implies $\Delta = 0$ and an identical WTP for classical and behavioral consumers, $W(v_2) = W(v_2, \Delta)$. Therefore, we focus on interior solutions and consider only equilibria with $p_1 > p_2$.

[13]Since β reflects the strength of the behavioral mechanism of an individual, it is restricted to the unit interval.

Note that we assume that the add-on does and can not affect consumer choice in the base good market. They choose a firm only because of the surplus provided by the base good. This assumption is reasonable in a number of settings. For example, the add-on price may be truly unobservable at the time of the base good purchase: many firms reveal prices of add-ons only after a (tentative) base good purchase, a practice known as drip pricing (Competition Market Authority, 2022). Firms may not need to commit to the add-on price ex ante or firms may not advertise and shroud add-on prices (see Gabaix and Laibson, 2006; Spiegler, 2006; Gamp, 2015; Spiegler, 2016).[14] Finally, add-on prices may be too expensive to learn ex ante before arriving at a point of sale (Ellison, 2005; Heidhues et al., 2021).[15]

3 Equilibrium Analysis

We solve the game for subgame perfect Nash equilibria in pure strategies by means of backward induction.

3.1 Aftermarket

In period 2, after the purchase of the base good, consumers with WTP $W(v_2, \Delta)$ can buy an add-on at price p_2. Classical consumers ($\beta = 0$) buy the add-on when $W(v_2) \geq p_2$. Behavioral consumers ($\beta \in (0, 1]$) buy when $W(v_2, \Delta) \geq p_2$. Therefore, the demand for the add-on of firm j is given by

$$Q_j(p_{2,j}, D_j(p_{1,j}, p_{1,-j})) = \begin{cases} D_j(p_{1,j}, p_{1,-j}) & \text{if } p_{2,j} \leq W(v_2), \\ \alpha D_j(p_{1,j}, p_{1,-j}) & \text{if } W(v_2) < p_{2,j} \leq W(v_2, \Delta), \\ 0 & \text{if } p_{2,j} > W(v_2, \Delta). \end{cases}$$

Observe that the add-on demand also depends indirectly on p_1, because only base good buyers can purchase the add-on.

[14]When prices are unobservable, in line with Gabaix and Laibson (2006), we suppose consumers form Bayesian posteriors about the add-on with the beliefs that firms set monopolistic add-on prices since they are profit-maximizing. These rational expectations are identical across firms and, thus, do not affect consumer's choice problem.

[15]See Spiegler (2011); Heidhues and Kőszegi (2018) for a review of behavioral models in add-on pricing.

3.2 Firms' Problem

The profit function of firm j is given by

$$\pi_j(p_{1,j}, p_{1,-j}, p_{2,j}) = p_{1,j} D_j(p_{1,j}, p_{1,-j}) + p_{2,j} Q_j(p_{2,j}, D_j(\cdot)). \tag{1}$$

Firms have monopolistic power in the aftermarket and extract the entire rent and will make one of the two consumer types indifferent. Two possible prices emerge in equilibrium, implicitly defined by $p_2^* \in \{W(v_2), W(v_2, \Delta)\}$. If $p_2^* = W(v_2)$, firms do not exploit behavioral consumers. All consumers accept the additional offer. We refer to this as the *non-exploiting strategy.* If $p_2^* = W(v_2, \Delta)$, however, firms do exploit the behavioral consumer and price the add-on at behavioral consumers' WTP. Consequently, classical consumers do not accept the add-on offer. We refer to this as the *exploiting strategy.*

Selecting one strategy determines $p_{2,j}$ and $Q_j(p_{2,j}, D_j(\cdot))$ in Equation (1). Given a chosen strategy, firm j maximizes its profits by choosing the base good price $p_{1,j}$, which yields the implicitly defined best response functions for any $p_{1,-j}$. Depending on the rival's actions, we obtain either (i) symmetric non-exploiting prices p_1^n and profits π^n, (ii) symmetric exploiting prices $p_1^e(\alpha)$ and profits π^e or (iii) an asymmetric outcome, where the non-exploiting firm sets $\tilde{p}_1^n(\alpha)$ and gets $\tilde{\pi}^n$, and the exploiting firm sets $\tilde{p}_1^e(\alpha)$ and receives $\tilde{\pi}^e$, where

$$\begin{aligned}
\pi^n &= \pi(p_1^n, p_1^n, W(v_2)) \\
\pi^e &= \pi(p_1^e(\alpha), p_1^e(\alpha), W(v_2, \Delta)) \\
\tilde{\pi}^n &= \pi\left(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha), W(v_2)\right) \\
\tilde{\pi}^e &= \pi\left(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha), W(v_2, \Delta)\right).
\end{aligned}$$

The full derivation of all prices and profits are characterized in Appendix A. Lemma 3 in Appendix A.3 shows that the exploiting profits π^e and $\tilde{\pi}^e$ are strictly increasing in α. This is because behavioral consumers—who can be exploited by firms—become more frequent when α increases. The symmetric non-exploiting profit π^n is independent of the share of behavioral consumers because firms set the add-on price to the WTP of the classical consumer, implying that all consumers accept the add-on offer. The asymmetric non-exploiting profit $\tilde{\pi}^n$ is either increasing or decreasing in α.[16]

[16]See Appendix A for details and explanation.

3.3 Equilibrium

The formal equilibrium derivation is provided in Appendix A.4. The share of behavioral consumers α crucially determines which equilibrium arises. When behavioral consumers are particularly frequent, then both firms exploit in equilibrium by choosing $p_1^* = p_1^e(\alpha)$, $p_2^* = W(v_2, \Delta)$. When the share of behavioral consumers is low, then neither firm exploits and both set $p_1^* = p_1^n$, $p_2^* = W(v_2)$ in equilibrium. For a wide range of α, the *symmetric non-exploiting equilibrium* and the *symmetric exploiting equilibrium*, respectively, are unique. Only for an intermediate share of behavioral consumers multiple equilibria exists. Either the best response implies to do the same as the rival and both, the symmetric non-exploiting and symmetric exploiting equilibrium exist, or the best response is to do the opposite and multiple asymmetric equilibria arise.

The equilibrium structure is intuitive. Firms face a trade-off between a higher aftermarket demand versus a higher mark-up. When the share of behavioral consumers is low, the demand effect dominates. The income from selling a high-priced add-on to only a few behavioral consumers cannot compensate for the demand loss arising from classical consumers who decline the additional offer. Accordingly, firms do not exploit and set the add-on price at the classical consumers' WTP $W(v_2)$. When the share of behavioral consumers is large, both firms exploit behavioral consumers by setting $p_2^* = W(v_2, \Delta)$. The demand loss from not serving classical consumers in the aftermarket is (over)compensated by the higher add-on mark-up because sufficiently many behavioral consumers are in the population.

3.4 The base good price

In classical aftersales models, the only incentive firms face is to lower the base good price in order to lock in more consumers (Diamond, 1971). An important consequence of introducing consumers that follow our behavioral pattern, is that it gives firms a countervailing incentive to increase the base good price. A more expensive base good increases behavioral consumers' WTP for the add-on. This allows firms to extract a higher mark-up in the add-on market, increasing the value of the aftermarket. Yet, a higher base good price leads to a lower demand in the base good market, which also implies less demand for the add-on. Hence, firms that exploit in the aftermarket face a trade-off when setting the optimal base good price $p_1^e(\alpha)$ or $\tilde{p}_1^e(\alpha)$, and consequently also $\tilde{p}_1^n(\alpha)$, since base good prices are strategic complements. This trade-off is captured by the relationship of the two semi-elasticities

$$\epsilon_D = \frac{-\partial D(\cdot)/\partial p_{1,j}}{D(\cdot)} \quad \text{and} \quad \epsilon_W = \frac{\partial W(v_2,\Delta)/\partial p_{1,j}}{W(v_2,\Delta)}.$$

The base good demand semi-elasticity, ϵ_D, denotes the demand effect of a price change in the base good market and thus, the anount of consumers in the the add-on market. The second semi-elasticity, ϵ_W, captures the effect of a base good price change on the add-on WTP of behavioral consumers. Depending on which effect dominates, the optimal base good price of an exploiting firm is either a decreasing or increasing function in the share of behavioral consumers α. When $\epsilon_D > \epsilon_W$, the demand effect is stronger and the optimal prices $p_1^e(\alpha)$, $\tilde{p}_1^e(\alpha)$ and $\tilde{p}_1^n(\alpha)$ are decreasing in α. In this case, it is profitable to attract and lock-in more consumers by lowering the base good price. In contrast, when $\epsilon_D < \epsilon_W$, the optimal base good prices are increasing in the share of behavioral consumers. This is the case when the base good demand is relatively inelastic and a price change has little effect on the sold quantity of base goods. We show in Lemma 4 in Appendix A that the order of the semi-elasticities, ϵ_D and ϵ_W, is monotonic in α and thus, also the price functions.[17]

Crucially, this implies that the base good price in the symmetric exploiting equilibrium and asymmetric equilibrium depends on the share of behavioral consumers in the population. As a consequence, the presence of behavioral consumers affects classical consumers in the base good market.

To analyze the effect of behavioral consumers on classical buyers, we consider an economy consisting of classical consumers only ($\alpha = 0$) as the benchmark. The benchmark base good price is given by p_1^b. In equilibrium, firms sell the add-on to all consumers at $p_2 = W(v_2)$. Thus, the benchmark outcome is identical to the symmetric non-exploiting equilibrium, and we obtain $p_1^b = p_1^n$, which implies that in any symmetric non-exploiting equilibrium, firms price as if there were only classical consumers. It is immediate to see that a low share of behavioral consumers has no effect on the surplus of classical consumers.

Whether the base good in a symmetric exploiting or asymmetric equilibrium is cheaper or more expensive than in the benchmark depends crucially on α and whether ϵ_D or ϵ_W is

[17] That is for specific functions $D(\cdot)$ and $W(\cdot)$, it is either $\epsilon_D \geq \epsilon_W$ for all α or $\epsilon_D \leq \epsilon_W$ for all α.

stronger. We implicitly define the critical price threshold

$$\bar{\alpha}_p = \begin{cases} \dfrac{W(v_2)}{W(v_2,\Delta) + \dfrac{\frac{\partial W(v_2,\Delta)}{\partial p_{1,j}} D(p^*_{1,j}, p^*_{1,-j})}{\frac{\partial D(p^*_{1,j}, p^*_{1,-j})}{\partial p_{1,j}}}}, & \text{for } \epsilon_D \neq \epsilon_W \\ \infty, & \text{for } \epsilon_D = \epsilon_W. \end{cases}$$

When $\alpha = \bar{\alpha}_p$, then the base good costs the same in any equilibrium, $p^n_1 = p^b_1 = p^e_1(\bar{\alpha}_p) = \tilde{p}^n_1(\bar{\alpha}_p) = \tilde{p}^e_1(\bar{\alpha}_p)$. Observe that, since $D(p_{1,j}, p_{1,-j})$ is decreasing in $p_{1,j}$, the denominator of $\bar{\alpha}_p$ is not necessarily positive, but depends on the relationship of the semi-elasticities. The price threshold $\bar{\alpha}_p$ is positive when $\epsilon_D > \epsilon_W$ and negative when $\epsilon_D < \epsilon_W$.[18] Lemma 1 captures when the base good is cheaper or more expensive than in the benchmark economy.

Lemma 1.

(i) Suppose $\epsilon_D > \epsilon_W$. If $\alpha \in (\min\{\bar{\alpha}, \hat{\alpha}\}, \bar{\alpha}_p)$, then the base good is more expensive in any symmetric exploiting or asymmetric equilibrium than in the benchmark. If $\alpha > \bar{\alpha}_p$, then the base good is cheaper in any symmetric exploiting equilibrium.

(ii) Suppose $\epsilon_D < \epsilon_W$. The base good is always more expensive in any symmetric exploiting or asymmetric equilibrium than in the benchmark.

For $\epsilon_D > \epsilon_W$, the price functions $p^e_1(\alpha)$, $\tilde{p}^n_1(\alpha)$, and $\tilde{p}^e_1(\alpha)$ are all decreasing in α. Hence, when the share of behavioral consumers is sufficiently low ($\alpha < \bar{\alpha}_p$), then the base good, offered by an exploiting firm (and by the non-exploiting firm in the asymmetric case), is more expensive compared to the benchmark case. Note that symmetric exploiting and asymmetric equilibria exist only when $\alpha > \min\{\bar{\alpha}, \hat{\alpha}\}$.[19] Otherwise, for a large share $\alpha > \bar{\alpha}_p$, the base good is cheaper in any symmetric exploiting equilibrium.[20] For $\epsilon_D < \epsilon_W$, the price functions are increasing in α, and the price threshold $\bar{\alpha}_p$ is negative. Hence, for any share of behavioral consumers, the base good of an exploiting firm (and of the non-exploiting firm in the asymmetric case) is more expensive than in the benchmark economy.

[18] The inequality $W(v_2, \Delta) > \frac{\partial W(v_2,\Delta)}{\partial p_{1,j}} \frac{D(p^*_{1,j}, p^*_{1,-j})}{\frac{\partial D(p^*_{1,j}, p^*_{1,-j})}{\partial p_{1,j}}}$ can be rearranged to $\epsilon_D > \epsilon_W$.

[19] We define the profit thresholds $\bar{\alpha}$ and $\hat{\alpha}$ in Appendix A.4. They are necessary to formalize the equilibrium characterization.

[20] Asymmetric equilibria do not exist when $\alpha > \bar{\alpha}_p$.

3.5 The Surplus of Classical Consumers

We turn now to the central part of our analysis and main result. Combining the results from Lemma 1 and the equilibrium characterization (Lemma 5) identifies that classical consumers can be harmed by the presence of behavioral consumers. Importantly, when $\epsilon_D > \epsilon_W$, then the price threshold is always larger than the profit thresholds. That is $\bar{\alpha}_p > \max\{\bar{\alpha}, \hat{\alpha}\}$. Hence, there exists an interval, in which α is such that a symmetric exploiting or asymmetric equilibrium exists and the base good price in these equilibria is larger than in the benchmark economy.[21] When $\epsilon_D < \epsilon_W$, then the base good is always more expensive in a symmetric exploiting or asymmetric equilibrium. Proposition 1 states the conditions when classical consumers benefit or are harmed by the presence of behavioral consumers.

Proposition 1 (The effect on the surplus of classical consumers)**.**

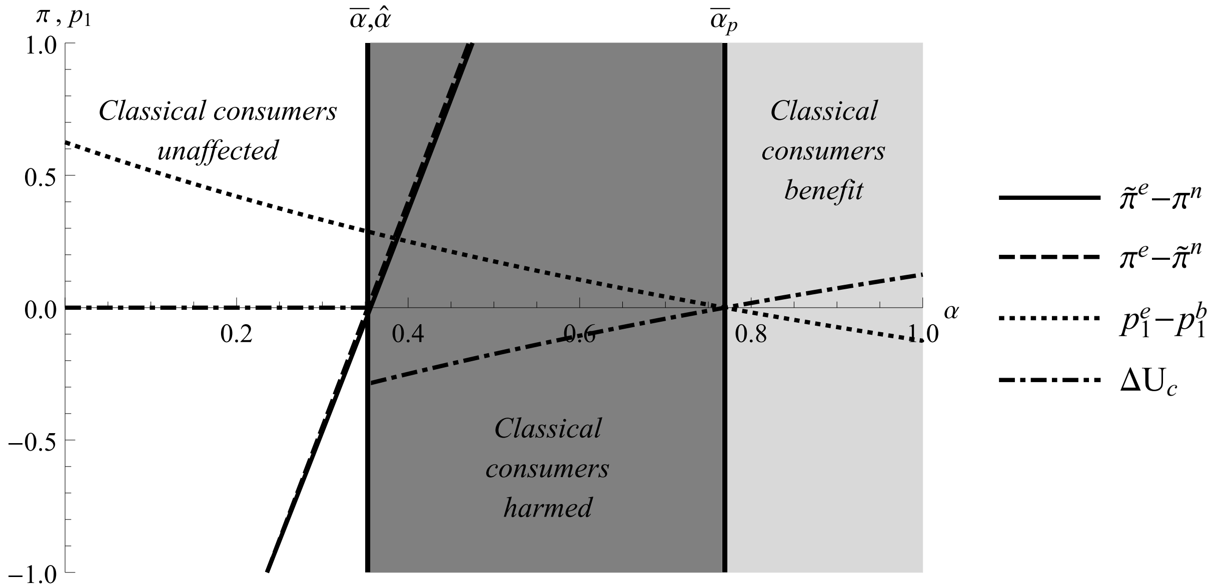
(a) Suppose $\epsilon_D > \epsilon_W$. Then the presence of behavioral consumers: (*i*) *harms classical consumers in any symmetric exploiting equilibrium if $\alpha < \bar{\alpha}_p$ and benefits them otherwise,* (*ii*) *harms classical consumers in any asymmetric equilibrium,* (*iii*) *has no effect in any symmetric non-exploiting equilibrium.*

(b) Suppose $\epsilon_D < \epsilon_W$. Then the presence of behavioral consumers: (*i*) *harms classical consumers in any symmetric exploiting or asymmetric equilibrium,* (*ii*) *has no effect in any symmetric non-exploiting equilibrium.*

When the share of behavioral consumers is low as in (aiii) and (bii), there is no effect on classical consumers because no firm exploits, which leads to the same outcome as in the benchmark economy. The presence of behavioral consumers, however, may harm classical consumers once there are sufficient behavioral buyers in the economy. For a relatively elastic base good demand ($\epsilon_D > \epsilon_W$), this is the case when the share of behavioral consumers is intermediate, such that exploiting is optimal, but the base good is more expensive than in the benchmark economy. Then, the presence of behavioral consumers harms classical consumers because they have to pay more for the desired base good than when all consumers would be classical and not some subject to a behavioral mechanism. However, since the optimal base good price is a decreasing function when $\epsilon_D > \epsilon_W$, classical

[21] When $\alpha > \max\{\bar{\alpha}, \hat{\alpha}\}$, then the unique symmetric exploiting equilibrium exists. When $\hat{\alpha} < \alpha < \bar{\alpha}$, then the asymmetric equilibria exist. When $\bar{\alpha} < \alpha < \hat{\alpha}$, then the multiple symmetric equilibria exist.

Figure 1: Proposition 1 (a) with symmetric prices.



Note: The figure depicts Proposition 1 (a) with $\epsilon_D > \epsilon_W$ and $\bar{\alpha} < \hat{\alpha}$, using $W(v_2, \Delta) = v_2\beta(1 + p_1 - p_2)$ with $\beta = 1$. The base good demand function is adopted from Singh and Vives (1984): $D_j(p_{1,j}, p_{1,-j}) = \frac{v_1}{1+d} - \frac{p_{1,j}}{1-d^2} + \frac{dp_{1,-j}}{1-d^2}$. The parameter specifications are $v_1 = 9$, $v_2 = 1$, $d = 0.4$ and marginal costs $c = 2.6$. Notation:
π^n, π^e: symmetric non-exploiting and exploiting profits;
$\tilde{\pi}^n, \tilde{\pi}^e$: profits in asymmetric outcomes;
p_1^e, p_1^b: symmetric exploiting and benchmark economy prices;
$\hat{\alpha}$: profit threshold such that $\pi^n = \tilde{\pi}^e$; $\bar{\alpha}$: profit threshold such that $\tilde{\pi}^n = \pi^e$;
ΔU_c: Classical consumer surplus change.

consumers are better off when the share is large ($\alpha > \bar{\alpha}_p$) because the base good is cheaper than in the benchmark economy. These three different outcomes of Proposition 1 (a) are depicted Figure 1. For a relatively inelastic demand for the base good ($\epsilon_D < \epsilon_W$), classical consumers are always harmed when at least one firm exploits because then the base good is always more expensive than in the benchmark.

Our finding implies that consumers, who behave perfectly according to classical economic theory, do not always benefit from behavioral or "naive" consumers. In our model, classical consumers may be harmed by the presence of behavioral consumers. This enables a distinctive evaluation of the results from the previous literature, who usually report that classical or "sophisticated" consumers can only benefit from the presence of behavioral consumers (Spiegler, 2011; Heidhues and Kőszegi, 2018). This has major policy implications, which we will discuss in more detail in Section 4.

So far, we have focused only on the surplus of classical consumers. Economists, however, should be concerned about total consumer surplus. The results in Proposition 2 state how exploitation by firms affects the surplus of behavioral consumers and total consumer welfare. The variable $CS^E(\alpha)$, which is increasing in α when$\epsilon_D > \epsilon_W$, captures the consumer surplus in the base good market when firms exploit in equilibrium, and, similarly, CS^{NE} denotes the consumer surplus when no firm exploits.

Proposition 2 (Consumer surplus)**.**

(a) Suppose $\epsilon_D > \epsilon_W$

(i) Behavioral consumers are always worse off when they are exploited.

(ii) Suppose at least one firm exploits in equilibrium. Then, the total consumer surplus is strictly lower when $\alpha D(\cdot)[W(v_2, \Delta) - W(v_2)] > CS^E(\alpha) - CS^{NE}$. The condition is always satisfied for $\alpha \leq \bar{\alpha}_p$.

(b) Suppose $\epsilon_D < \epsilon_W$. Then behavioral consumer and total consumer surplus is strictly lower when at least one firm exploits in equilibrium.

First, behavioral consumers always have a lower surplus in an exploiting equilibrium than in a symmetric non-exploiting equilibrium. This is independent of whether the behavioral mechanism increases the received add-on utility or not. The effect on total consumer surplus is mixed when $\epsilon_D > \epsilon_W$. On the one hand, exploitation can decrease the base good price below the price in a symmetric non-exploiting (or benchmark) equilibrium. This leads to an increased demand for the base good, but it also implies that more behavioral consumers are exploited in the aftermarket. This trade-off is captured by the inequality $\alpha D(\cdot)[W(v_2, \Delta) - W(v_2)] > CS^E(\alpha) - CS^{NE}$.[22] On the left-hand side, we have the surplus effect of behavioral consumers buying the add-on, which is strictly positive. The right-hand side displays the change in consumer surplus in the base good market, which can be positive or negative. When $\alpha \leq \bar{\alpha}_p$, by Lemma 1, the base good is more expensive when at least one firm exploits compared to a symmetric non-exploiting (or benchmark) equilibrium. This implies less demand, and it must be that $CS^E(\alpha) \leq CS^{NE}$. Hence, when classical consumers are harmed according to Proposition 1, then the total consumer surplus is always lower because of exploitation.[23] For $\alpha > \bar{\alpha}_p$, the effect of exploitation

[22]Note that we need to adjust the condition slightly for asymmetric equilibria since not all behavioral consumers are exploited. However, since asymmetric equilibria exist only for $\alpha < \bar{\alpha}_p$, the adjusted condition is always satisfied.

[23]Note that consumer surplus is also in the extreme case of $\alpha = 1$ always lower. This is because Proposition 2 states that behavioral consumers always have a lower surplus, and with $\alpha = 1$, only behavioral consumers exist.

is unclear on consumer surplus. In this case, the base good is cheaper, which increases the surplus in the base good market $CS^E(\alpha) > CS^{NE}$. Therefore, it depends on whether the positive effect in the base good market dominates the negative impact in the after-sales market. When $\epsilon_D < \epsilon_W$, all consumers are worse off when firms exploit, and thus, consumer surplus is always lower.

3.6 Monopoly and Perfect Competition

In this section, we discuss the outcomes of the two extreme cases of competition, monopoly and perfect competition. The findings in Proposition 1 and 2 are robust to a monopoly setting, while perfect competition eliminates the harmful effect on classical consumers. We defer the details of the formal analysis and results to Appendix B.4. First, the analysis and findings in a setting with only one firm in the base good market are identical to the imperfect case, except that asymmetric equilibria do not exist.[24] This result is non-trivial since the cross-subsidization result vanishes with monopolistic competition in the existing literature (Heidhues and Kőszegi, 2018).

Under perfect price competition, that is when base goods are perfect substitutes, then only the positive effect of behavioral consumers on classical consumers survives. The intuitive reason is that due to competitive pressure, firms cannot increase the base good price above the benchmark level. Otherwise, firms would face zero demand. Thus, when firms exploit in equilibrium, then the base good price must be strictly lower, which benefits classical consumers. Hence, under perfect competition, classical consumers can never be harmed by the presence of behavioral consumers.

4 Policy Implications

In this section, we apply our main result stated in Proposition 1 and analyze how different policies affect the individual welfare of consumers. First, we consider a policy that educates behavioral consumers and thus reduces their frequency in the population, such as for example revealing all prices (separately) up-front. Second, we analyze the effect of a price

[24] When base goods are perfectly differentiated, then each firm is a monopolist in its respective base good market. Compared to the imperfect competition case, we need much less structure on the base good demand function. We simply impose that $D(p_1)$ is strictly decreasing, twice continuously differentiable, $\lim_{p_1 \to \infty} D(p_1) = 0$ and satisfies $D(p_1)D''(p_1) < 2D'(p_1)^2$, which, for instance, holds for log-concave but also CES demand functions

floor regulation on the base good, which is a common tool used by policymakers to prevent loss-leading and predatory pricing. Lastly, we discuss the impact of a price cap on the add-on.

4.1 Educating behavioral consumers

The Department of Transportation (2022) proposed a rule to require airlines to reveal the full price of a ticket up-front, including ancillary services such as checked baggage. This leads to so-called partitioned pricing, where all prices are shown up-front but still separately for the base good and ancillary services. This is in contrast to drip pricing, where prices are presented sequentially, and all-inclusive pricing, where consumers see one total price. Compared to drip pricing, partitioned pricing leads to lower consumer demand for the add-on (Robbert and Roth, 2014). Compared to all-inclusive pricing, however, partitioned pricing still leads to misperceptions among consumers, pushing up their WTP for the add-on. Therefore, we assume that such a policy intervention, as proposed by the Department of Transportation (2022), reduces the frequency of behavioral consumers, because it educates some of them who then become a classical consumer after the intervention.

Suppose there exists an instrument or technology for policymakers to reduce the share of behavioral consumers in the population, and that the ex-ante share of behavioral consumers is such that the unique symmetric exploiting equilibrium exists.[25] To characterize the policy impact, we need to distinguish between effective and ineffective instruments. An *effective* policy leads to a sufficiently large reduction of behavioral consumers such that after the intervention, exploiting is not optimal anymore for either firm. Hence, the measure results in the unique symmetric non-exploiting equilibrium. In contrast, an *ineffective* policy reduces the share of behavioral consumers only by a bit such that both firms still exploit in equilibrium.

Definition 1.

(i) An effective education policy reduces α sufficiently large such that the symmetric non-exploiting equilibrium emerges ex-post.

[25] The intuition for ex-ante asymmetric equilibria is similar.

(ii) An ineffective education policy reduces α only a little such that both firms still exploit ex-post.

Further, we need to distinguish whether classical consumers benefit or are harmed by the presence of behavioral consumers prior to the policy implementation. We call any equilibrium, in which classical consumers benefit, according to Proposition 1, *beneficial equilibrium.* Otherwise, when classical consumers are harmed by the presence of behavioral consumers, we have a *harmful equilibrium.*

Proposition 3 (Education)**.**

(a) Suppose $\epsilon_D > \epsilon_W$.

(i) Any ineffective education policy makes all consumers worse off.

(ii) Any effective education policy benefits behavioral consumers. Classical consumers benefit when the ex-ante equilibrium was harmful and are worse off if the ex-ante equilibrium was beneficial.

(iii) Consumer surplus (i) decreases when the policy was ineffective, (ii) increases when the policy was effective and the ex-ante equilibrium was harmful, (iii) impact is ambiguous when the policy was effective and the ex-ante equilibrium was beneficial.

(b) When $\epsilon_D < \epsilon_W$, all consumers benefit from any policy, which increases consumer surplus always.

When the base good demand is relatively elastic as in Proposition 3 (a), then the optimal base good price under symmetric exploiting $p_1^e(\alpha)$ is decreasing in α. Since an ineffective policy reduces the number of behavioral consumers only by a little, the intervention does not change firms' behavior but strictly increases $p_1^e(\alpha)$. Hence, firms still exploit in equilibrium, and consumers must pay more for the base good, which makes any consumer clearly worse off. In addition, the add-on also becomes more expensive.

In contrast, an effective policy affects a firm's behavior, and there is no exploitation ex-post. The intervention results in the symmetric non-exploiting equilibrium. This always benefits behavioral consumers, independently of whether the behavioral mechanism increases add-on consumption utility or not. In any case, the add-on surplus increases, while the policy can affect the base good surplus positively or negatively, depending on whether the base good is cheaper or more expensive ex-post. We show in the proof of Proposition 3 that the positive surplus change in the add-on market always dominates a negative effect on the base good surplus.

For classical consumers, it depends on whether the ex-ante equilibrium was harmful or beneficial. In the former case, classical consumers benefit from an effective policy because the base good is cheaper ex-post. In the latter case, they are hurt, because an effective policy prevents the cross-subsidization from behavioral consumers, which results in a more expensive base good ex-post.[26]

Part *(iii)* states the implications on total consumer surplus. When a policy affects classical and behavioral consumers similarly, then the welfare effects are clear. An ineffective policy decreases total consumer surplus, while an effective intervention increases total consumer surplus when the ex-ante equilibrium was harmful. Only when the ex-ante equilibrium is beneficial, then the impact of an effective policy is ambiguous. The intervention prevents the so-called reverse Robin Hood exercise, the cross-subsidization of behavioral consumers to classical consumers. The effect on total consumer surplus depends on which group is stronger affected by the policy.

When the base good demand is relatively inelastic as in Proposition 3 (b), then the optimal base good price under symmetric exploiting $p_1^e(\alpha)$ is increasing in α. This implies that any reduction in the share of behavioral consumers lowers the base good price, which benefits all consumers. Hence, any policy is beneficial and increases consumer surplus.

The results in Proposition 3 provide important insights for policymakers. Not every education policy improves the welfare of consumers. On the contrary, they may even hurt consumers. When demand effects are relatively strong, then policymakers need to be careful with imposing regulations and interventions as they may worsen the situation for consumers.

4.2 Price Floor

Loss-leading is a controversial practice that raises concerns over anti-competitive effects. For that reason, predatory pricing is banned in many US States and some European countries.[27] Policymakers impose a price floor on goods by prohibiting pricing below costs with the aim of protecting consumers. The literature finds mixed results on the

[26] The add-on surplus for classical consumers is still zero ex-post.

[27] See for example https://www.aeaweb.org/research/loss-leading-bans-retail-competition.

effectiveness of this policy (e.g., Chen and Rey, 2012; Johnson, 2017). In our model, a binding price floor can only yield negative effects on consumers in an exploiting equilibrium.

We impose a price floor that does not affect the benchmark economy, $\underline{p}_1 \leq p_1^b$.[28] This price floor is binding only in an exploiting equilibrium with $\epsilon_D > \epsilon_W$, where the base good is cheaper than in the benchmark economy.[29] This is the case when the share of behavioral consumers is sufficiently large. Intuitively, firms decrease base good prices to attract more behavioral consumers that are willing to buy the overpriced add-on and, thus, can be exploited.

Proposition 4 (Price floor)**.** *Suppose $\epsilon_D > \epsilon_W$ and $\underline{p}_1 \leq p_1^b$. A binding price floor $\underline{p}_1$ (i) increases the add-on price and (ii) reduces the base good demand. Remaining consumers in the market are strictly worse off by a binding regulation.*

Since firms must offer the base good at a higher price than in equilibrium, the demand for the base good declines and the add-on becomes more expensive with a binding price floor. This is because an exploiting firm sets $p_2 = W(v_2, \Delta)$ and $W(v_2, \Delta)$ is increasing in p_1. This clearly harms consumers that remain in the market as the prices for both goods increase. This is also the case for classical consumers that dropped out of the market because of the regulation. Without a price floor, they would buy the base good and obtain a positive surplus. The effects on behavioral consumers that left the market are ambiguous. Similar to classical consumers, they lose a positive surplus from the base good, but do not receive a weakly negative add-on surplus.[30]

Hence, the overall effect on consumer surplus is unclear. A binding price floor reduces the number of consumers that buy the expensive add-on. In other words, the regulation limits the number of consumers that firms can exploit, which positively affects some behavioral consumers. But simultaneously, it harms all other consumers in the population.

4.3 Add-on price cap

In his 2023 State of the Union speech, Biden called for a \$8 cap on credit card late fees (The White House, 2023), with the intention of extending such a policy to add-on fees

[28] A binding price floor in the benchmark economy affects only the base good market. It leads to a higher base good price and consequently to fewer quantities sold. The add-on price is unaffected.

[29] When $\epsilon_D < \epsilon_W$, the base good price in an exploiting equilibrium is always larger than in the benchmark and a price floor is never binding.

[30] Whether the add-on surplus is negative or zero depends on the welfare specification for behavioral consumers.

offered by concert and sports promoters.

We discuss the impact of this regulation on the outcome of our model intuitively. First, with a price cap $\bar{p}_2 < W(v_2)$, only the non-exploiting equilibrium exists because firms cannot set the add-on price above the WTP of classical consumers.[31] Thus, a price cap prevents the exploitation of behavioral consumers completely. But this comes at efficiency costs. Since firms redistribute revenues from add-on selling, the price cap reduces those earnings, which increases the base good price. This, in turn, lowers the base-good demand and lowers firms' profits. Therefore, a price cap $\bar{p}_2 < W(v_2)$ solves the problem of exploitation, but the effect on consumer surplus is unclear as fewer base goods are sold at a higher price.

When $W(v_2) < \bar{p}_2 < W(v_2, \Delta)$, exploiting equilibria may still exist ex-post. Similar to an education policy, we need to distinguish between inefficient and efficient price caps.[32] Analogous to Definition 1, with an ineffective price cap, firms still exploit behavioral consumers after the regulation, while an effective price cap prevents exploitation and the non-exploiting equilibrium emerges ex-post. Whether a price cap is effective depends on how strongly it affects the add-on revenue. An effective price cap limits exploitation sufficiently enough such that selling the add-on to all consumers is more profitable. The impact of an effective price cap is similar to the result in Proposition 3 (b ii). It benefits classical consumers when the ex-ante equilibrium was harmful and hurts them when the ex-ante equilibrium was beneficial. The intuition is analogous to the education policy. Depending on the ex-ante equilibrium, classical consumers pay a higher or lower base good price ex-post, while still receiving zero surplus from the add-on. Behavioral consumers are always better off. Thus, when the ex-ante equilibrium was harmful, an effective price cap regulation unambiguously increases consumer surplus. When the ex-ante equilibrium was beneficial, the impact on consumer surplus is unclear.

When the price cap is ineffective, then firms still exploit behavioral consumers and set a higher base good price because they earn less add-on revenue. This clearly harms classical consumers. Behavioral consumers enjoy a lower add-on price, but also must pay more for the base good. Thus, the impact on them is unclear.

In general, a price cap on the add-on limits the extent to which firms can exploit behavioral consumers. When exploiting is still optimal (or $\bar{p}_2 < W(v_2)$), however, then

[31] When $\bar{p}_2 = W(v_2)$, the price cap leads to the identical outcome of the non-exploiting (or benchmark) equilibrium without a price cap.

[32] We consider only binding price caps. Otherwise, the regulation has no effect on the outcome.

the regulation leads to more expensive base goods. The implications are similar to an education policy. However, instead of reducing the share of behavioral consumers directly, a price cap tackles the profits that firms can make from exploitation, which increases the profit thresholds. In other words, a larger share of behavioral consumers is required that exploitation is profitable. Graphically, a price cap $W(v_2) < \bar{p}_2 < W(v_2, \Delta)$ implies that the lines denoted with $\bar{\alpha}$ and $\hat{\alpha}$ in Figure 1 shift to the right and the region where classical consumers are unaffected increases.

Our discussion shows that a price regulation in the add-on market leads to non-trivial effects. It has a mixed impact on individual welfare, and it is unclear whether consumer surplus increases or decreases. A formal analysis, however, is beyond the scope of this paper, and we encourage future research to investigate on this topic.

5 Microfoundations

For the behavioral consumer, the base good price increases the WTP for the add-on, as empirically documented in a range of studies (Erat and Bhaskaran, 2012; Xia and Monroe, 2004; Chatterjee, 2010). In this section, we discuss several mechanisms that may microfound such behavior, which is captured in our reduced-form model through the function $W(v_2, \Delta)$.

Relative thinking. Relative thinking has been shown to be an important determinant in individual decision-making (Thaler, 1980; Jacowitz and Kahneman, 1995). In Bushong et al. (2021), 48% of participants are willing to accept a 30-minute drive to save \$25 for a \$1000 laptop, while 73% of participants are willing to do so to save the same monetary amount when shopping for \$100 headphones. Somerville (2022) experimentally shows that more than two-thirds of the participants are better characterized as relative thinkers than as standard utility maximizers.

In Bushong et al. (2021), consumers put a relative weight $w(\Delta_k)$ on each consumption dimension $k = v, p$; where $\Delta_k = \max k_s - \min k_s$ for $s = 1, 2$ and $w(\Delta_p)$ is a differentiable and decreasing function on $(0, \infty)$. Adapting the model to our setting, the behavioral consumer is a relative thinker regarding the price dimension. To focus on this channel, we set to $w(\Delta_v) = 1$. Behavioral consumers' incentive constraint for the purchase of the add-on can be written as $v_2 - w(\Delta_p)p_2 \geq 0$. We employ the parameterized example of their model: $w(\Delta_p) = (1 - \rho) + \frac{\rho}{\Delta_p + \xi}$ where $\rho \in [0, 1)$ and $\xi \in (0, \infty)$. Rearranging the incentive constraint yields $W(v_2,\Delta) = \frac{v_2}{(1-\rho)+\frac{\rho}{\Delta_p+\xi}} \geq p_2$.[33] Therefore, whenever the base good is more expensive than the add-on, the model of Bushong et al. (2021) satisfies our assumptions on $W(v_2,\Delta)$.

Somerville (2022) provides a similar parameterized function for relative thinking. The incentive constraint is characterized by $v_2 - (\Delta_p)^y \cdot p_2 \geq 0$ with $y \in (-1,0)$ and $\Delta_p = \max p_k - \min p_k$ for $k = 1,2$. Rearranging yields $W(v_2,\Delta) = \frac{v_2}{(\Delta_p)^y} \geq p_2$. Again, whenever $p_1 > p_2$, the micro-foundation of Somerville (2022) satisfies our assumption of consumer behavior.

Proportional thinking. Closely related is proportional thinking (Thaler, 1980). In Tversky and Kahneman's (1981) famous jacket-calculator example, a person is willing to exert more effort to save \$10 when the relative amount of money saved is higher (see also the replications by Mowen and Mowen, 1986; Frisch, 1993; Ranyard and Abdel-Nabi, 1993). Azar (2011) shows that consumers are willing to pay more for the same constant improvement in quality when the good's price is higher. In a field experiment, Blake et al. (2021) document a lower proportional price ($\frac{p_1}{p_2}$) boosts add-on sales.[34] Formally, a behavioral consumer perceives the add-on prices as $\frac{p_2}{p_1}$ (or $\frac{p_2}{p_1+p_2}$), which implies an incentive constraint of $v_2 - \frac{p_2}{p_1} \geq 0$. Rearranging shows that the WTP $W(v_2,p_1)$ is increasing in p_1.[35]

Salience. Consumers may devote more attention to product attributes that are more salient. For example, it is documented that consumers underreact to taxes when those are not salient (Chetty, Looney, and Kroft, 2009; Feldman and Ruffle, 2015; Taubinsky and Rees-Jones, 2018). Also, when prices become less salient, demand substantially increases (Finkelstein, 2009; Sexton, 2015). In a large field experiment on StubHub.com, Blake et al. (2021) show that drip pricing strategies increase demand due to the additional fee appearing less salient for consumers (see also Brown et al., 2010; Hossain and Morgan, 2006; Dertwinkel-Kalt et al., 2020).

Bordalo, Gennaioli, and Shleifer (2022) formalize salience theory. In their model, the surplus function for behavioral consumers is $\hat{V} = \sum_k w_k \pi_k a_k$ for a good with k attributes,

[33]Note that $\frac{\partial W(v_2,\Delta)}{\partial p_1} > 0$ and $\frac{\partial^2 W(v_2,\Delta)}{\partial p_1^2} < 0$ when $p_1 > p_2$ and $\frac{\partial W(v_2,\Delta)}{\partial v_2} > 0$ and $\frac{\partial^2 W(v_2,\Delta)}{\partial v_2^2} = 0$.

[34]"An old selling trick is to quote a low price for a stripped-down model and then coax the consumer into a more expensive version in a series of increments each of which seems small relative to the entire purchase" (Thaler, 1980, p. 51).

[35]See also Azar (2007) who develops a model of add-on pricing with mixed consumers in which behavioral consumers' add-on WTP is given by $w(P_L) = dP_L^{\alpha\beta}$, where d is a constant capturing utility, P_L the price of the base good, $\alpha \in [0,1]$ captures the extent of proportional thinking of a consumer, and $\beta \in [0,1]$ reflects the extent of relative thinking inherent in a certain decision context. Setting $d = v_2$, $P_L = p_1$, $\alpha = \beta_i$ and $\beta = 1$, leads directly to our reduced form $W(v_2,p_1) = v_2 p_2^{\beta_i}$ with $\beta_i \in \{0,\beta\}$.

where w_k is a weighting function capturing bottom-up attention to salient attributes, π_k is the decision weight attached to attribute k, and a_k denotes the attribute's value (see also Bordalo, Gennaioli, and Shleifer, 2012, 2013, 2020). Suppose that in our setup, consumers consider the base good and add-on only as a single bundle product only once they are confronted with the add-on offer.[36] In that case, the product has four attributes: base good quality $a_{v_1} = v_1$ and price $a_{p_1} = p_1$; and add-on quality $a_{v_2} = v_2$ and price $a_{p_2} = p_2$. We suppose that $w_{v_1} = w_{v_2} = 1$, and for ease of exposition, we also assume $\pi_{v_2} = 1$, $\pi_{p_2} = -1$. Firms can use a high and salient base good price p_1 to draw attention away from the add-on price p_2. Importantly, w_k is increasing in the salience of k and weakly decreasing in other attributes $-k$ salience. Thus, increasing p_1 makes the base good price more salient, with the consequence of p_2 becoming less salient. This in turn decreases $w_{p_2}(\Delta)$ with the consequence that behavioral consumers put less weight on the add-on price. Therefore, we suppose that $w_{p_2}(\Delta)$ depends on the (relative) difference of p_1 and p_2, which is captured by Δ. A behavioral consumer buys the add-on when $v_2 - w_{p_2}(\Delta)p_2 > 0$. Rearranging the incentive constraint yields $W(v_2, \Delta) = \frac{v_2}{w_{p_2}(\Delta)} \geq p_2$, where $w_{p_2}(\Delta)$ is decreasing in Δ. Hence, $W(v_2, \Delta)$ is increasing in both arguments. Note that the salience effect is of diminishing sensitivity (Bordalo et al., 2022), which corresponds to our concavity assumption.

Mental accounting. Because consumers are mental accountants " [...] sellers have a distinct advantage in selling something if its cost can be added on to another larger purchase" Thaler (1985, p. 209). See also Ranyard and Abdel-Nabi (1993); Moon, Keasey, and Duxbury (1999); Erat and Bhaskaran (2012).

The transaction utility theory from Thaler (1985) is a two-stage process. First, there is a judgment process, where consumers evaluate potential transactions. The total utility is defined as $w(z, p, p^*) = v(\bar{p} - p) + v(-p : -p^*)$, where $\bar{p}$ is the valuation for a good z with price p, reference price p^*, and $v(\cdot)$ is a concave function. The term $v(\bar{p} - p)$ captures the acquisition utility, which is simply the net utility accrued by the trade and corresponds to the add-on net utility of classical consumers.[37] The transaction utility (or reference outcome) is captured by $v(-p : -p^*)$, which depends on the add-on price and the reference price. Note that $v(-p : -p) = 0$, $v(-p : -p^*) > 0$ when $p < p^*$, and

[36]Note that by assumption, in our after-sales market, the add-on is simply not on offer before the base good purchase. Thus, since consumers observe the add-on only after purchasing the base good, we suppose that salience does not play a role in the base good market.

[37]We use directly the notation $v(\bar{p} - p)$ instead of $v(\bar{p}, -p)$, since Thaler (1985) argues that acquisition utility will generally be coded as integrated outcome.

$v(-p : -p^*)$ is increasing in p^*. Intuitively, when the reference price exceeds the market price, then it affects the value of good z positively. The size of the effect depends on the difference between p and p^*. Second, there is a decision process, where consumers (dis-)approve each potential transaction. A behavioral consumer will buy a good z if $\frac{w(z,p,p^*)}{p} > k$, where k is a constant. We interpret $k = 0$ as the outside option of not buying the add-on. Supposing $v(\bar{p} - p) = W(\bar{p}) - p$ and setting $\bar{p} = v_2$, $p = p_2$, and $p^* = p_1$ leads to the incentive constraint $\frac{W(v_2)-p_2+v(-p_2:-p_1)}{p_2} \geq 0$. Assuming $p_1 > p_2$, then $W(v_2, v(-p_2 : -p_1)) = W(v_2) + v(-p_2 : -p_1) \geq p_2$ implies $\frac{\partial v(-p_2:-p_1)}{\partial p_1} > 0$, and, since $v(\cdot)$ is concave, the assumptions on $W(v_2, \Delta)$ with $\Delta = v(-p_2 : -p_1)$ are satisfied. Therefore, consumers subject to mental thinking can be characterized as behavioral consumers in our model.[38]

Reference point dependence and anchoring-and-adjustment. A large amount of experimental evidence documents the importance of reference points in individual decision-making, starting with Kahneman and Tversky (1979); Tversky and Kahneman (1974); Jacowitz and Kahneman (1995); Tversky and Kahneman (1981). Arbitrary high anchors have been shown to increase the WTP for a variety of goods (Ariely, Loewenstein, and Prelec, 2003; Bergman, Ellingsen, Johannesson, and Svensson, 2010; Fudenberg, Levine, and Maniadis, 2012; Maniadis, Tufano, and List, 2014; Alevy, Landry, and List, 2015; Yoon, Fong, and Dimoka, 2019; Ioannidis, Offerman, and Sloof, 2020). Also, it is documented that the price observed in previous market periods affects subsequent bids of market participants (Tufano, 2010; Beggs and Graddy, 2009; Ferraro, Messer, Shukla, and Weigel, 2021). The former (Office of Fair Trading, 2013, p. 8) suggests: "For example, consumers may use a heuristic called 'anchoring and adjustment', in which case consumers will anchor on the base price and insufficiently adjust for the surcharge" (see Furnham and Boo, 2011, for a literature review on the heuristic). Therefore, we argue that anchoring and adjustment is a suitable explanation for our reduced form function $W(v_2, \Delta)$. Formally, we incorporate the distance between p_2 and the reference price p_1 as the behavioral mechanism into the incentive constraint, $u - p_i + \gamma(\tilde{p} - p_i) \geq 0$, where $\gamma(\cdot)$ captures loss aversion (Wenner,

[38] Our reduced-form model also accommodates Erat and Bhaskaran (2012), who provide a mental accounting model in the context of add-on selling. The behavioral mechanism is defined as a mental book value $BV = p - V$, where p is the paid base good price and V is the cumulative benefit a consumer has obtained so far from using the base good, which increases over time. Thus, BV is maximal just after the base good purchase occurred. Further, a consumer buys the add-on if and only if $p_A \leq u_A + \gamma u_A BV$. Setting $p = p_1$, $p_A = p_2$ and $u_A = v_2$ translates immediately to our reduced form incentive constraint $W(v_2, BV(p_1)) = v_2(1 + \gamma BV(p_1)) \geq p_2$, where $BV(p_1)$ is strictly increasing in p_1.

2015). Setting $u = v_2$, $p_i = p_2$ and $\tilde{p} = p_1$ yields immediately $W(v_2, \Delta) = v_2 + \gamma(\Delta) \geq p_2$ with $\Delta = p_1 - p_2$.

6 Further Results

6.1 After-sales Competition (Sequential Buying)

We relax the lock-in assumption and allow for competition in the after-sales market. We suppose the same setup as in the baseline model, but a fraction $\rho \in (0, 1)$ of base good buyers search for the cheapest add-on, while the fraction $(1 - \rho)$ stays loyal and purchases the add-on from the same company. Firms know the distribution of loyal consumers but cannot price discriminate. They choose prices p_1 and p_2 simultaneously and can commit to add-on prices.[39]

In equilibrium, firms still choose between the non-exploiting strategy ($p_2 \leq W(v_2)$) and the exploiting strategy ($W(v_2) < p_2 \leq W(v_2, \Delta)$), but mix over the choice of add-on prices. To see this, consider the symmetric equilibria given in Lemma 5. Since searching consumers buy the add-on from the cheapest seller, a firm can profitably deviate by setting a slightly lower price and capturing all non-loyal customers. The other extreme, marginal cost pricing and earning zero after-sales profits, is also not optimal. Firms can always just sell the add-on to the loyal consumers at the WTP of either classical or behavioral consumers and make positive after-sales profits. Thus, there is mixing in add-on prices, where firms must be indifferent between mixing, and potentially attracting some consumers from the rival, or setting $p_2 \in \{W(v_2), W(v_2, \Delta)\}$, and sell the add-on to only loyal (non-) classical consumers. This result resembles the findings of Baye and Morgan (2001).

Therefore, the expected profit from the add-on market of a non-exploiting firm is given by $(1-\rho)D_j(\cdot)W(v_2)$. Similarly, an exploiting firm expects to earn $\alpha(1-\rho)D_j(\cdot)W(v_2, \Delta)$ in the aftermarket. These can be substituted into the profit function (1), which simplifies the maximization problem greatly, and we can proceed as in the baseline model. Note that the introduction of searching consumers is simply a rescaling and does not change the analysis qualitatively. Relaxing the lock-in assumption lowers not only the (expected) profits of an exploiting firm but also of non-exploiting firms. Hence, the equilibria characterization is

[39] This assumption is merely for simplicity. The results do not change when p_1 and p_2 are chosen sequentially. In any equilibria, firms mix their choice of p_2, independently of simultaneous or sequential price setting.

identical to Lemma 5 with following exception for p_2^*:

Lemma 2.

(i) A non-exploiting firm draws an add-on price p_2^n from a continuous and atomless price distribution $F^n(p_2^n)$ with $p_2^n \in (\underline{p}_2^n, W(v_2))$.

(ii) An exploiting firm draws prices an add-on price p_2^e from a continuous and atomless price distribution $F^e(p_2^e)$ with $p_2^e \in (\max\{\underline{p}_2^e, W(v_2)\}, W(v_2, \Delta))$.

Importantly, the semi-elasticities ϵ_D and ϵ_W, and the price threshold $\bar{\alpha}_p$ are (qualitatively) unchanged and we still have $\bar{\alpha}_p > \max\{\bar{\alpha}, \hat{\alpha}\}$. Thus, the results from Lemma 1 and Proposition 1 follow immediately. Therefore, our central finding that the presence of behavioral consumers can harm classical consumers does not rely on the lock-in assumption.

6.2 Unit Demand

We apply our framework to a model with unit base good demand and horizontal differentiation, which is commonly used in the literature (e.g., Ellison, 2005; Gabaix and Laibson, 2006; Armstrong and Vickers, 2012; Heidhues and Kőszegi, 2017). The full analysis is provided in Appendix C.2.

We find similar results as in the baseline model with imperfect competition. We show that the equilibrium is alike as characterized in Lemma 5 and the optimal base good price p_1 behaves similar to Lemma 1. Crucially, the main findings stated in Proposition 1 hold and are not affected by the different demand structure. When firms exploit behavioral consumers, this can benefit or harm classical consumers.

7 Conclusion

We study an after-sales market with behavioral consumers who are subject to an effect that temporarily increases the WTP for the add-on. We show that several well-known and studied mechanisms, such as relative thinking, proportional thinking, salience, mental accounting or anchoring-and-adjustment, can motivate our reduced-form model. When confronted with such behavior, firms face an incentive to increase the base good price.

We provide a novel result in the context of add-on selling and drip pricing: behavioral consumers may harm the surplus of classical consumers when firms compete imperfectly in the base good market. These findings lead to the question of whether one should protect the classical consumer from the mistakes and biases of others. Behavioral consumers exert quite ambiguous effects on the surplus of classical consumers. While a relatively equal mixture of the two types makes the classical worse off than in the benchmark economy, a considerable proportion of behavioral consumers is necessary to increase the surplus of the classical at the expense of behavioral consumers.

We use our model to assess the impact of some potential policies. Education policies that reduce the share of the behavioral consumer can increase consumer surplus but could also decrease it. Similarly, for exogenous price floors imposed in the base good market, originally implemented with the intention to prevent loss leading, harms all consumers remaining in the market. While classical consumers' surplus is not jeopardized by the behavioral consumer in a perfectly competitive market, competition comes along with a distributional effect in which the behavioral type subsidizes the classical consumer.

Our model and findings could also be relevant in labor markets, organizational settings and for gift-exchange (Akerlof, 1982; Fehr, Kirchsteiger, and Riedl, 1993; Hart and Moore, 2008; Azar, 2019). Experimental evidence suggests that wages that are low relative to past wages decrease labor supply (Bracha, Gneezy, and Loewenstein, 2015). Similarly, the flat wage agreed-upon in a contract might represent the base good and serve as reference. Consequently, workers might differently perceive and reciprocate the very same $1000 year-end bonus. A similar effect may play a role in meal allowances or other additional employee benefits.

Our model comes along with some limitations. For example, we presume that firms have no choice in first offering a base good and then an add-on. Variations of our model could investigate firms' optimal pricing strategies and its consequences on welfare in product design, such as Apffelstaedt and Mechtenberg (2021) do with context-sensitive consumers. For example, due to behavioral consumers, decoupling a bundle product into a base and extra good may increase demand, such as various studies have empirically shown (Morwitz et al., 1998; Blake et al., 2021). Moreover, our model does not capture a setting with mandatory add-ons, such as unavoidable surcharge fee's (Rasch et al., 2020). Future research could address this and extend our model to accommodate mandatory surcharges. We further assume that behavioral consumers' WTP for the add-on is monotonously increasing in the price of the base good. A disproportionately high-priced add-on, however, may be perceived as unfair (Herz and Taubinsky, 2017; Rabin, 1993; Robbert and Roth, 2014). Fairness effects may impose an upper limit for the add-on price. Finally, it would be valuable if future research can provide clean and causal empirical evidence for our theoretical implications.

References

Ahmetoglu, G., A. Furnham, and P. Fagan (2014): "Pricing practices: A critical review of their effects on consumer perceptions and behaviour," *Journal of Retailing and Consumer Services*, 21, 696–707.

Akerlof, G. A. (1982): "Labor contracts as partial gift exchange," *The quarterly journal of economics*, 97, 543–569.

Alevy, J. E., C. E. Landry, and J. A. List (2015): "Field experiments on the anchoring of economic valuations," *Economic Inquiry*, 53, 1522–1538.

Apffelstaedt, A. and L. Mechtenberg (2021): "Competition for context-sensitive consumers," *Management Science*, 67, 2828–2844.

Ariely, D., G. Loewenstein, and D. Prelec (2003): "Coherent Arbitrariness: Stable Demand Curves Without Stable Preferences," *The Quarterly Journal of Economics*, 118, 73–106.

Armstrong, M. (2008): "Interactions between competition and consumer policy," *Competition Policy International*, 4, 97–147.

——— (2015): "Search and ripoff externalities," *Review of Industrial Organization*, 47, 273–302.

Armstrong, M. and J. Vickers (2012): "Consumer protection and contingent charges," *Journal of Economic Literature*, 50, 477–93.

Azar, O. H. (2007): "Relative thinking theory," *The Journal of Socio-Economics*, 36, 1–14.

——— (2011): "Do people think about absolute or relative price differences when choosing between substitute goods?" *Journal of Economic Psychology*, 32, 450–457.

——— (2019): "Do fixed payments affect effort? Examining relative thinking in mixed compensation schemes," *Journal of Economic Psychology*, 70, 52–66.

Baye, M. R. and J. Morgan (2001): "Information gatekeepers on the internet and the competitiveness of homogeneous product markets," *American Economic Review*, 91, 454–474.

Beggs, A. and K. Graddy (2009): "Anchoring Effects: Evidence from Art Auctions," *American Economic Review*, 99, 1027–39.

Bergman, O., T. Ellingsen, M. Johannesson, and C. Svensson (2010): "Anchoring and cognitive ability," *Economics Letters*, 107, 66–68.

Blake, T., S. Moshary, K. Sweeney, and S. Tadelis (2021): "Price salience and product choice," *Marketing Science*, 40, 619–636.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012): "Salience theory of choice under risk," *The Quarterly journal of economics*, 127, 1243–1285.

——— (2013): "Salience and consumer choice," *Journal of Political Economy*, 121, 803–843.

——— (2020): "Memory, attention, and choice," *The Quarterly journal of economics*, 135, 1399–1442.

——— (2022): "Salience," *Annual Review of Economics*, 14, 521–44.

Bracha, A., U. Gneezy, and G. Loewenstein (2015): "Relative pay and labor supply," *Journal of labor economics*, 33, 297–315.

Brown, J., T. Hossain, and J. Morgan (2010): "Shrouded attributes and information suppression: Evidence from the field," *The Quarterly Journal of Economics*, 125, 859–876.

Bushong, B., M. Rabin, and J. Schwartzstein (2021): "A model of relative thinking," *The Review of Economic Studies*, 88, 162–191.

Chatterjee, P. (2010): "Consumer response to promotions in the presence of surcharge: implications for online retailing," *Journal of Customer Behaviour*, 9, 117–134.

Chen, Z. and P. Rey (2012): "Loss leading as an exploitative practice," *American Economic Review*, 102, 3462–82.

Chetty, R., A. Looney, and K. Kroft (2009): "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99, 1145–77.

Competition Market Authority (2022): "Evidence Review of Online Choice Architecture and Consumer and Competition Harm," *UK Government Report*.

Coppi, L. (2007): "Aftermarket Monopolization: The Emerging Consensus in Economics," *The Antitrust Bulletin*, 52, 53–71.

Department of Transportation (2022): "Enhancing Transparency of Airline Ancillary Service Fees," *Federal Register*, 2022-22214, 63718–63738.

Dertwinkel-Kalt, M., M. Köster, and M. Sutter (2020): "To buy or not to buy?

Price salience in an online shopping field experiment," *European Economic Review*, 130, 103593.

DIAMOND, P. A. (1971): "A model of price adjustment," *Journal of economic theory*, 3, 156–168.

ELLISON, G. (2005): "A model of add-on pricing," *The Quarterly Journal of Economics*, 120, 585–637.

——— (2006): "Bounded rationality in industrial organization," *Econometric Society Monographs*, 42, 142.

ELLISON, G. AND S. F. ELLISON (2009): "Search, obfuscation, and price elasticities on the internet," *Econometrica*, 77, 427–452.

ERAT, S. AND S. R. BHASKARAN (2012): "Consumer mental accounts and implications to selling base products and add-ons," *Marketing Science*, 31, 801–818.

FEDERAL TRADE COMMISSION (2012): "The Economics of Drip Pricing," Tech. rep., U.S. Government.

——— (2022): "Unfair or Deceptive Fees Trade Regulation Rule," *Federal Register*, 87 FR 67413, 67413 – 67424.

FEHR, E., G. KIRCHSTEIGER, AND A. RIEDL (1993): "Does fairness prevent market clearing? An experimental investigation," *The quarterly journal of economics*, 108, 437–459.

FELDMAN, N. E. AND B. J. RUFFLE (2015): "The impact of including, adding, and subtracting a tax on demand," *American Economic Journal: Economic Policy*, 7, 95–118.

FERRARO, P., K. D. MESSER, P. SHUKLA, AND C. WEIGEL (2021): "Behavioral Biases among Producers: Experimental Evidence of Anchoring in Procurement Auctions," .

FINKELSTEIN, A. (2009): "E-ztax: Tax salience and tax rates," *The Quarterly Journal of Economics*, 124, 969–1010.

FRISCH, D. (1993): "Reasons for framing effects," *Organizational behavior and human decision processes*, 54, 399–429.

FUDENBERG, D., D. K. LEVINE, AND Z. MANIADIS (2012): "On the robustness of anchoring effects in WTP and WTA experiments," *American Economic Journal: Microeconomics*, 4, 131–45.

FURNHAM, A. AND H. C. BOO (2011): "A literature review of the anchoring effect," *The journal of socio-economics*, 40, 35–42.

Gabaix, X. and D. Laibson (2006): "Shrouded attributes, consumer myopia, and information suppression in competitive markets," *The Quarterly Journal of Economics*, 121, 505–540.

Gamp, T. (2015): "Search, differentiated products, and obfuscation," *Beitraege zur Jahrestagung des Vereins für Socialpolitik 2015: Oekonomische Entwicklung - Theorie und Politik - Session: Monopoly and Oligopoly No. D24-V2*.

Gamp, T. and D. Krähmer (2022): "Competition in search markets with naive consumers," *The RAND Journal of Economics*.

Greenleaf, E. A., E. J. Johnson, V. G. Morwitz, and E. Shalev (2016): "The price does not include additional taxes, fees, and surcharges: A review of research on partitioned pricing," *Journal of Consumer Psychology*, 26, 105–124.

Grubb, M. D. (2015): "Behavioral consumers in industrial organization: An overview," *Review of Industrial Organization*, 47, 247–258.

Hart, O. and J. Moore (2008): "Contracts as reference points," *The Quarterly journal of economics*, 123, 1–48.

Heidhues, P., J. Johnen, and B. Kőszegi (2021): "Browsing versus studying: A pro-market case for regulation," *The Review of Economic Studies*, 88, 708–729.

Heidhues, P. and B. Kőszegi (2017): "Naivete-based discrimination," *The Quarterly Journal of Economics*, 132, 1019–1054.

——— (2018): "Behavioral industrial organization," *Handbook of Behavioral Economics: Applications and Foundations 1*, 1, 517–612.

Heidhues, P., B. Kőszegi, and T. Murooka (2016): "Inferior products and profitable deception," *The Review of Economic Studies*, 84, 323–356.

Herz, H. and D. Taubinsky (2017): "What Makes a Price Fair? An Experimental Study of Transaction Experience and Endogenous Fairness Views," *Journal of the European Economic Association*, 16, 316–352.

Holton, R. H. (1957): "Price discrimination at retail: the supermarket case," *The Journal of Industrial Economics*, 6, 13–32.

Hossain, T. and J. Morgan (2006): "... plus shipping and handling: Revenue (non) equivalence in field experiments on ebay," *The BE Journal of Economic Analysis & Policy*, 6.

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

Ioannidis, K., T. Offerman, and R. Sloof (2020): "On the effect of anchoring on valuations when the anchor is transparently uninformative," *Journal of the Economic Science Association*, 6, 77–94.

Jacowitz, K. E. and D. Kahneman (1995): "Measures of anchoring in estimation tasks," *Personality and Social Psychology Bulletin*, 21, 1161–1166.

Jang, S. and J. Chung (2021): "What drives add-on sales in mobile games? The role of inter-price relationship and product popularity," *Journal of Business Research*, 124, 59–68.

Johnen, J. (2020): "Dynamic competition in deceptive markets," *The RAND Journal of Economics*, 51, 375–401.

Johnson, J. P. (2017): "Unplanned purchases and retail competition," *American Economic Review*, 107, 931–65.

Kahneman, D. and A. Tversky (1979): "Prospect theory: An analysis of decision under risk," *Econometrica*, 47, 363–391.

Kosfeld, M. and U. Schüwer (2016): "Add-on Pricing in Retail Financial Markets and the Fallacies of Consumer Education*," *Review of Finance*, 21, 1189–1216.

Lal, R. and C. Matutes (1994): "Retail pricing and advertising strategies," *Journal of business*, 345–370.

Maniadis, Z., F. Tufano, and J. A. List (2014): "One Swallow Doesn't Make a Summer: New Evidence on Anchoring Effects," *American Economic Review*, 104, 277–90.

Michel, C. (2017): "Market regulation of voluntary add-on contracts," *International Journal of Industrial Organization*, 54, 239–268.

Moon, P., K. Keasey, and D. Duxbury (1999): "Mental accounting and decision making: The relationship between relative and absolute savings," *Journal of Economic Behavior & Organization*, 38, 145–153.

Morwitz, V. G., E. A. Greenleaf, and E. J. Johnson (1998): "Divide and prosper: consumers' reactions to partitioned prices," *Journal of marketing research*, 35, 453–463.

Mowen, M. M. and J. C. Mowen (1986): "An empirical examination of the biasing effects of framing on business decisions," *Decision Sciences*, 17, 596–602.

Office of Fair Trading (2013): "Partitioned Pricing Research: The economics of partitioned pricing," *UK Government Report*.

Rabin, M. (1993): "Incorporating fairness into game theory and economics," *The American economic review*, 1281–1302.

Ranyard, R. and D. Abdel-Nabi (1993): "Mental accounting and the process of multiattribute choice," *Acta Psychologica*, 84, 161–177.

Rasch, A., M. Thöne, and T. Wenzel (2020): "Drip pricing and its regulation: Experimental evidence," *Journal of Economic Behavior & Organization*, 176, 353–370.

Robbert, T. and S. Roth (2014): "The flip side of drip pricing," *Journal of Product & Brand Management*, 23, 413–419.

Santana, S., S. K. Dallas, and V. G. Morwitz (2020): "Consumer reactions to drip pricing," *Marketing Science*, 39, 188–210.

Schumacher, H. (2022): "Competitive Markets and Boundedly Rational Expectations," .

Sexton, S. (2015): "Automatic bill payment and salience effects: Evidence from electricity consumption," *Review of Economics and Statistics*, 97, 229–241.

Shapiro, C. (1994): "Aftermarkets and Consumer Welfare: Making Sense of Kodak Symposium: Post-Chicago Economics," *Antitrust Law Journal*, 63, 483.

Singh, N. and X. Vives (1984): "Price and quantity competition in a differentiated duopoly," *The Rand journal of economics*, 546–554.

Somerville, J. (2022): "Range-Dependent Attribute Weighting in Consumer Choice: An Experimental Test," *Econometrica*, 90, 799–830.

Spiegler, R. (2006): "Competition over agents with boundedly rational expectations," *Theoretical Economics*, 1, 207–231.

——— (2011): *Bounded rationality and industrial organization*, Oxford University Press.

——— (2016): "Choice complexity and market competition," *Annual Review of Economics*, 8, 1–25.

Taubinsky, D. and A. Rees-Jones (2018): "Attention variation and welfare: theory and evidence from a tax salience experiment," *The Review of Economic Studies*, 85, 2462–2496.

Thaler, R. (1980): "Toward a positive theory of consumer choice," *Journal of economic behavior & organization*, 1, 39–60.

——— (1985): "Mental accounting and consumer choice," *Marketing science*, 4, 199–214.

The White House (2022): "The President's Initiative on Junk Fees and Related Pricing Practices," Tech. rep., U.S. Government.

——— (2023): "Remarks of President Joe Biden – State of the Union Address as Prepared for Delivery," Tech. rep., U.S. Government.

Tufano, F. (2010): "Are true preferences revealed in repeated markets? An experimental demonstration of context-dependent valuations," *Experimental Economics*, 13, 1–13.

Tversky, A. and D. Kahneman (1974): "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185, 1124–1131.

——— (1981): "The Framing of Decisions and the Psychology of Choice," *Science*, 211, 453–458.

Verboven, F. (1999): "Product line rivalry and market segmentation with an application to automobile optional engine pricing," *The Journal of Industrial Economics*, 47, 399–425.

Wenner, L. M. (2015): "Expected prices as reference points—theory and experiments," *European Economic Review*, 75, 60–79.

Xia, L. and K. B. Monroe (2004): "Price partitioning on the internet," *Journal of Interactive marketing*, 18, 63–73.

Yoon, S., N. M. Fong, and A. Dimoka (2019): "The robustness of anchoring effects on preferential judgments." *Judgment & Decision Making*, 14.

A Auxiliary Results

To ease notation, we denote the first and second derivative of the base-good demand function with respect to its own price with $D_j'(p_{1,j}, p_{1,-j}) = \frac{\partial D_j(p_{1,j}, p_{1,-j})}{\partial p_{1,j}}$ and $D_j''(p_{1,j}, p_{1,-j}) = \frac{\partial^2 D_j(p_{1,j}, p_{1,-j})}{\partial p_{1,j}^2}$.

A.1 Non-exploiting strategy

Suppose firm j does not exploit and sets $p_{2,j} = W(v_2)$. This implies $Q_j(p_{2,j}, D_j(\cdot)) = D_j(p_{1,j}, p_{1,-j})$ and the profit function (1) reduces to

$$\pi_j(p_{1,j}, p_{1,-j}, W(v_2)) = \big[p_{1,j} + W(v_2)\big] D_j(p_{1,j}, p_{1,-j}). \tag{2}$$

Note that the optimization problem in the benchmark economy ($\alpha = 0$) is identical to (2). Maximizing this expression with respect to $p_{1,j}$ yields the first-order condition

$$\begin{aligned} D_j(p_{1,j}, p_{1,-j}) + D_j'(p_{1,j}, p_{1,-j})[p_{1,j} + W(v_2)] = 0 \\ \Leftrightarrow \quad p_{1,j} = \frac{-D_j(p_{1,j}, p_{1,-j})}{D_j'(p_{1,j}, p_{1,-j})} - W(v_2). \end{aligned}$$

Substituting $p_{1,j}$ in expression (2) leads to

$$\pi_j(p_{1,j}, p_{1,-j}, W(v_2)) = \frac{-D_j(p_{1,j}, p_{1,-j})^2}{D_j'(p_{1,j}, p_{1,-j})}.$$

Whether firm j sets p_1^n or $\tilde{p}_1^n(\alpha)$ depends on the action of firm $-j$. First, suppose firm $-j$ does not exploit. Then, both firms set

$$p_1^n = \frac{-D(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)} - W(v_2)$$

and obtain

$$\pi^n = \pi(p_1^n, p_1^n, W(v_2)) = \frac{-D(p_1^n, p_1^n)^2}{D'(p_1^n, p_1^n)}.$$

Observe that neither p_1^n nor π^n depend on α. Therefore, the symmetric non-exploiting outcome is independent of the share of behavioral consumers and thus, identical to the benchmark outcome. That is $p_1^n = p_1^b$ and $\pi^n = \pi^b$. Now suppose firm $-j$ exploits and sets $\tilde{p}_1^e(\alpha)$. Then, firm j sets

$$\tilde{p}_1^n(\alpha) = \frac{-D(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha))}{D'(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha))} - W(v_2)$$

and obtains

$$\tilde{\pi}^n = \pi\left(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha), W(v_2)\right) = \frac{-D(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha))^2}{D'(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha))}.$$

A.2 Exploiting strategy

Suppose firm j exploits and sets $p_{2,j} = W(v_2, \Delta)$. This implies $Q_j(p_{2,j}, D_j(\cdot)) = \alpha D_j(p_{1,j}, p_{1,-j})$ and the profit function (1) reduces to

$$\pi_j(p_{1,j}, p_{1,-j}, W(v_2, \Delta)) = \left[p_{1,j} + \alpha W(v_2, \Delta)\right] D_j(p_{1,j}, p_{1,-j}). \tag{3}$$

Maximizing this expression with respect to $p_{1,j}$ yields the first-order condition

$$\begin{aligned} &[1 + \alpha W'(v_2, \Delta)] D_j(p_{1,j}, p_{1,-j}) + D_j'(p_{1,j}, p_{1,-j})[p_{1,j} + \alpha W(v_2, \Delta)] = 0 \\ \Leftrightarrow \quad & p_{1,j} = \frac{-[1 + \alpha W'(v_2, \Delta)] D_j(p_{1,j}, p_{1,-j})}{D_j'(p_{1,j}, p_{1,-j})} - \alpha W(v_2, \Delta), \end{aligned}$$

where $W'(v_2, \Delta) = \frac{\partial W(v_2, \Delta)}{\partial p_{1,j}}$. Substituting $p_{1,j}$ in expression (3) leads to

$$\pi_j(p_{1,j}, p_{1,-j}, W(v_2, \Delta)) = \frac{-[1 + \alpha W'(v_2, \Delta)] D_j(p_{1,j}, p_{1,-j})^2}{D_j'(p_{1,j}, p_{1,-j})}.$$

Whether firm j sets $p_1^e(\alpha)$ or $\tilde{p}_1^e(\alpha)$ depends on the action of firm $-j$. First, suppose firm $-j$ exploits. Then, both firms set

$$p_1^e(\alpha) = \frac{-[1 + \alpha W'(v_2, \Delta)] D(p_1^e(\alpha), p_1^e(\alpha))}{D'(p_1^e(\alpha), p_1^e(\alpha))} - \alpha W(v_2, \Delta)$$

and obtain

$$\pi^e = \pi(p_1^e(\alpha), p_1^e(\alpha), W(v_2, \Delta)) = \frac{-[1 + \alpha W'(v_2, \Delta)] D(p_1^e(\alpha), p_1^e(\alpha))^2}{D'(p_1^e(\alpha), p_1^e(\alpha))}.$$

Now suppose firm $-j$ does not exploit and sets $\tilde{p}_1^n(\alpha)$. Then, firm j sets

$$\tilde{p}_1^e(\alpha) = \frac{-[1 + \alpha W'(v_2, \Delta)] D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))} - \alpha W(v_2, \Delta)$$

and obtains

$$\tilde{\pi}^e = \pi\left(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha), W(v_2, \Delta)\right) = \frac{-[1+\alpha W'(v_2,\Delta)]D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))^2}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}.$$

A.3 Derivatives

It is crucial for our analysis to understand how the base-good prices and profits react to changes in α. Recall that $D_j(p_{1,j}, p_{1,-j})$ is a concave function and is strictly decreasing in $p_{1,j}$, which implies $D'(\cdot) < 0$ and $D''(\cdot) \leq 0$. Further, we have $\frac{\partial D_j(p_{1,j},p_{1,-j})}{\partial p_{1,-j}} \geq 0$ and $\frac{\partial^2 D_j(p_{1,j},p_{1,-j})}{\partial p_{1,j}\partial p_{1,-j}} \geq 0$, since base goods are strategic complements and demand is supermodular, and stronger own price elasticity implies $\left|D_j'(p_{1,j}, p_{1,-j})\right| > \left|\frac{\partial D_j(p_{1,j},p_{1,-j})}{\partial p_{1,-j}}\right|$. Finally, since $W(v_2, \Delta)$ is strictly increasing in all arguments and weakly concave, we have $W'(v_2,\Delta) = \frac{\partial W(v_2,\Delta)}{\partial p_{1,j}} > 0$ and $W''(v_2,\Delta) = \frac{\partial^2 W(v_2,\Delta)}{\partial p_{1,j}^2} \leq 0$. Given the assumptions on $D(\cdot)$ and $W(v_2, \Delta)$, Lemma 3 characterizes how profits and prices react to a change of α.

Lemma 3.

(a) p_1^n and π^n are constant in α.

(b) π^e and $\tilde{\pi}^e$ are strictly increasing in α.

(c) $p_1^e(\alpha)$, $\tilde{p}_1^e(\alpha)$, $\tilde{p}_1^n(\alpha)$ and $\tilde{\pi}^n$ are (i) strictly decreasing in α if $\epsilon_D > \epsilon_W$, (ii) strictly increasing in α if $\epsilon_D < \epsilon_W$, and (iii) constant in α if $\epsilon_D = \epsilon_W$.

Proof. (a)

$$\begin{aligned}
\frac{\partial p_1^n}{\partial \alpha} &= \left[\frac{-D'(p_1^n,p_1^n)^2}{D'(p_1^n,p_1^n)^2} + \frac{D(p_1^n,p_1^n)D''(p_1^n,p_1^n)}{D'(p_1^n,p_1^n)^2}\right]\frac{\partial p_1^n}{\partial \alpha} \\
\Leftrightarrow \quad \frac{\partial p_1^n}{\partial \alpha} &\left[2 - \frac{D(p_1^n,p_1^n)D''(p_1^n,p_1^n)}{D'(p_1^n,p_1^n)^2}\right] = 0, \\
\frac{\partial \pi_1^n}{\partial \alpha} &= \frac{-2D(p_1^n,p_1^n)D'(p_1^n,p_1^n)^2 + D(p_1^n,p_1^n)^2 D''(p_1^n,p_1^n)}{D'(p_1^n,p_1^n)^2}\underbrace{\frac{\partial p_1^n}{\partial \alpha}}_{=0} = 0.
\end{aligned}$$

(b) First, we state $\frac{\partial p_1^e(\alpha)}{\partial \alpha}$, which will be needed to derive $\frac{\partial \pi_1^e}{\partial \alpha}$.

$$
\begin{aligned}
\frac{\partial p_1^e(\alpha)}{\partial \alpha} &= -\left[1 + 2\alpha W'(v_2,\Delta) + \frac{\alpha W''(v_2,\Delta)D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))}\right]\frac{\partial p_1^e(\alpha)}{\partial \alpha}\\
&\quad + (1+\alpha W'(v_2,\Delta))\frac{D(p_1^e(\alpha),p_1^e(\alpha))D''(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))^2}\frac{\partial p_1^e(\alpha)}{\partial \alpha}\\
&\quad - W(v_2,\Delta) - \frac{W'(v_2,\Delta)D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))}\\
\Leftrightarrow \quad \frac{\partial p_1^e(\alpha)}{\partial \alpha} &= \frac{-W(v_2,\Delta) - \frac{W'(v_2,\Delta)D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))}}{(1+\alpha W'(v_2,\Delta))\left(2 - \frac{D(p_1^e(\alpha),p_1^e(\alpha))D''(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))^2}\right) + \frac{\alpha W''(v_2,\Delta)D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))}}.
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \pi_1^e}{\partial \alpha} &= D(p_1^e(\alpha),p_1^e(\alpha))\Bigg[-\frac{W'(v_2,\Delta)D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))} - \frac{\partial p_1^e(\alpha)}{\partial \alpha}\Bigg[(1+\alpha W'(v_2,\Delta))\\
&\quad \left(2 - \frac{D(p_1^e(\alpha),p_1^e(\alpha))D''(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))^2}\right) + \frac{\alpha W''(v_2,\Delta)D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))}\Bigg]\Bigg]\\
&= D(p_1^e(\alpha),p_1^e(\alpha))W(v_2,\Delta) > 0,
\end{aligned}
$$

where the second equality follows from substituting $\frac{\partial p_1^e(\alpha)}{\partial \alpha}$. Before we can derive $\frac{\partial \tilde{\pi}_1^e}{\partial \alpha}$, we need to state $\frac{\partial \tilde{p}_1^n(\alpha)}{\partial \alpha}$ and $\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}$:

$$
\frac{\partial \tilde{p}_1^n(\alpha)}{\partial \alpha} = \underbrace{\left[\frac{\frac{D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))\frac{\partial^2 D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}{\partial \tilde{p}_1^n(\alpha)\partial \tilde{p}_1^e(\alpha)}}{D'(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))^2} - \frac{\frac{\partial D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}{\partial \tilde{p}_1^e(\alpha)}}{D'(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}}{2 - \frac{D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))D''(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}{D'(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))^2}}\right]}_{=A}\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}.
$$

Note that $A \geq 0$ since $D(\cdot)$ is concave, supermodular, strictly decreasing in the first argument and increasing in the second argument. Taking the derivative of $\tilde{p}_1^e(\alpha)$ with respect to α yields

$$
\begin{aligned}
&\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}\Bigg[(1+\alpha W'(v_2,\Delta))\left[2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2}\right]\\
&+ \frac{\alpha W''(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}\Bigg] = -W(v_2,\Delta) - \frac{W'(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}
\end{aligned}
$$

$$+ (1+\alpha W'(v_2,\Delta)) \underbrace{\left[\frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))\frac{\partial^2 D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{\partial \tilde{p}_1^e(\alpha)\partial \tilde{p}_1^n(\alpha)}}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - \frac{\frac{\partial D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{\partial \tilde{p}_1^n(\alpha)}}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}\right]}_{=B} \frac{\partial \tilde{p}_1^n(\alpha)}{\partial \alpha}.$$

Note that $B \geq 0$. Substituting $\frac{\partial \tilde{p}_1^n(\alpha)}{\partial \alpha}$ yields

$$\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha} = \frac{-W(v_2,\Delta) - \frac{W'(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}}{(1+\alpha W'(v_2,\Delta))\left[2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - AB\right] + \frac{\alpha W''(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}}.$$

Now we can take the derivative of $\tilde{\pi}_1^e$ with respect to α.

$$\begin{aligned}
\frac{\partial \tilde{\pi}_1^e}{\partial \alpha} &= D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))\Bigg[-\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}\Bigg[(1+\alpha W'(v_2,\Delta)) \\
&\left[2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2}\right] + \frac{\alpha W''(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}\Bigg] \\
&+ \frac{\partial \tilde{p}_1^n(\alpha)}{\partial \alpha}(1+\alpha W'(v_2,\Delta)) \underbrace{\left[\frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))\frac{\partial^2 D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{\partial \tilde{p}_1^e(\alpha)\partial \tilde{p}_1^n(\alpha)}}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - \frac{2\frac{\partial D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{\partial \tilde{p}_1^n(\alpha)}}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}\right]}_{=C} \\
&- \frac{W'(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}\Bigg] \\
&= D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))\Bigg[-\frac{W'(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))} - \frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}\Bigg[(1+\alpha W'(v_2,\Delta)) \\
&\left[2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - AC\right] + \frac{\alpha W''(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}\Bigg] \\
&= D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))\Bigg[\left(W(v_2,\Delta) + \frac{W'(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}\right) \cdot \\
&\frac{(1+\alpha W'(v_2,\Delta))\left[2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - AC\right] + \frac{\alpha W''(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}}{(1+\alpha W'(v_2,\Delta))\left[2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - AB\right] + \frac{\alpha W''(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}} \\
&- \frac{W'(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}\Bigg] > 0
\end{aligned}$$

The inequality follows since $B < C$ and $2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - AC > 0$ because $\left|\frac{\partial D_j^2(\cdot)}{\partial p_{1,j}^2}\right| \geq \frac{\partial D_j^2(\cdot)}{\partial p_{1,j}\partial p_{1,-j}}$. Thus, the fraction on the second last line is strictly positive and strictly smaller than 1. Since $\frac{W'(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))} < 0$ the result follows.

(c) First observe that, given our assumptions, the denominators of $\frac{\partial p_1^e(\alpha)}{\partial \alpha}$ and $\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}$ are strictly positive. Whether the nominators are positive or negative depends on whether ϵ_D or ϵ_W dominates.

$$\begin{aligned}
\frac{\partial p_1^e(\alpha)}{\partial \alpha} &\leq 0 \\
-W(v_2,\Delta) - \frac{W'(v_2,\Delta)D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))} &\leq 0 \\
-\frac{D'(p_1^e(\alpha),p_1^e(\alpha))}{D(p_1^e(\alpha),p_1^e(\alpha))} &\geq \frac{W'(v_2,\Delta)}{W(v_2,\Delta)} \\
\epsilon_D &\geq \epsilon_W.
\end{aligned}$$

Hence, it follows $\epsilon_D \geq \epsilon_W \Leftrightarrow \frac{\partial p_1^e(\alpha)}{\partial \alpha} \leq 0$. Thus, $p_1^e(\alpha)$ is strictly decreasing if $\epsilon_D > \epsilon_W$, strictly increasing if $\epsilon_D < \epsilon_W$ and constant if $\epsilon_D = \epsilon_W$.

The argument for $\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}$ and thus, $\tilde{p}_1^e(\alpha)$, is analogous. Observe that $\frac{\partial \tilde{p}_1^n(\alpha)}{\partial \alpha} \leq 0$ if $\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha} \leq 0$ and strictly positive otherwise. Thus, the result for $\tilde{p}_1^n(\alpha)$ follows immediately.

Finally,

$$\frac{\partial \tilde{\pi}_1^n}{\partial \alpha} = \underbrace{-D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))\frac{\frac{\partial D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}{\partial \tilde{p}_1^e(\alpha)}}{D'(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}}_{>0}\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}.$$

Hence, $\tilde{\pi}_1^n$ is strictly decreasing if $\epsilon_D > \epsilon_W$, strictly increasing if $\epsilon_D < \epsilon_W$ and constant if $\epsilon_D = \epsilon_W$.

□

Lemma 4 shows that the prices $p_1^e(\alpha)$, $\tilde{p}_1^e(\alpha)$ and $\tilde{p}_1^n(\alpha)$, and thus $\tilde{\pi}_1^n$, are monotonic in α. For a given $D(\cdot)$ and $W(v_2,\Delta)$, the price functions are either increasing or decreasing for all $\alpha \in [0,1]$.

Lemma 4. *Fix $D(\cdot)$ and $W(v_2,\Delta)$. The base-good prices $p_1^e(\alpha)$, $\tilde{p}_1^e(\alpha)$ and $\tilde{p}_1^n(\alpha)$ are monotonic in the share of behavioral consumers α.*

Proof. We provide the proof for $p_1^e(\alpha)$. The argument for $\tilde{p}_1^e(\alpha)$ and $\tilde{p}_1^n(\alpha)$ are analogous. Observe that

$$\frac{\partial \epsilon_D}{\partial p_1^e(\alpha)} = \frac{-D''(p_1^e(\alpha),p_1^e(\alpha))D(p_1^e(\alpha),p_1^e(\alpha)) + D'(p_1^e(\alpha),p_1^e(\alpha))^2}{D(p_1^e(\alpha),p_1^e(\alpha))^2} > 0$$

since $D(p_{1,j}, p_{1,-j})$ is concave, and

$$\frac{\partial \epsilon_W}{\partial p_1^e(\alpha)} = \frac{W''(v_2, \Delta)W(v_2, \Delta) - W'(v_2, \Delta)^2}{W(v_2, \Delta)^2} < 0$$

since $W(v_2, \Delta)$ is concave.

First, suppose $\epsilon_D > \epsilon_W$ at an initial share of behavioral consumers $\alpha_0 \in [0, 1]$. By Lemma 3, we have $\frac{\partial p_1^e(\alpha)}{\partial \alpha} < 0$. This implies, since $\frac{\partial \epsilon_D}{\partial p_1^e(\alpha)} > 0$ and $\frac{\partial \epsilon_W}{\partial p_1^e(\alpha)} < 0$, that ϵ_D and ϵ_W are converging for $\alpha > \alpha_0$ and diverging for $\alpha < \alpha_0$.

Since ϵ_D and ϵ_W are converging for an increasing α, there exists a threshold value $\bar{\alpha} > \alpha_0$ such that $\epsilon_D = \epsilon_W$. Note that $\bar{\alpha} > 1$ is possible. By Lemma 3, we have $\frac{\partial p_1^e(\alpha)}{\partial \alpha} = 0$ when $\epsilon_D = \epsilon_W$. Hence, a further increase $\alpha > \bar{\alpha}$ does not change the optimal base-good price $p_1^e(\alpha)$. But then it must be $\epsilon_D = \epsilon_W$ for all $\alpha \geq \bar{\alpha}$ and thus, $p_1^e(\alpha)$ is constant in α for all $\alpha \geq \bar{\alpha}$ and strictly decreasing in α for all $\alpha \in [\alpha_0, \bar{\alpha})$.

Since ϵ_D and ϵ_W are diverging for a decreasing α, $p_1^e(\alpha)$ is a strictly decreasing function for all $\alpha \in [0, \alpha_0]$. Hence, $p_1^e(\alpha)$ is strictly decreasing in the domain $\alpha \in [0, \bar{\alpha})$ and constant in α for all $\alpha \geq \bar{\alpha}$, which implies that $p_1^e(\alpha)$ is monotonic for $\alpha \in [0, 1]$ if $\epsilon_D > \epsilon_W$ at α_0.

Now, suppose that $\epsilon_D < \epsilon_W$ at an initial share of behavioral consumers $\alpha_0 \in [0, 1]$. By Lemma 3, we have $\frac{\partial p_1^e(\alpha)}{\partial \alpha} > 0$. This implies again that ϵ_D and ϵ_W are converging for $\alpha > \alpha_0$ and diverging for $\alpha < \alpha_0$. Thus, we can apply the same argument as above. This implies that $p_1^e(\alpha)$ is a strictly increasing function for all $\alpha \in [0, \bar{\alpha})$ and constant in α for all $\alpha \geq \bar{\alpha}$, which implies that $p_1^e(\alpha)$ is monotonic for $\alpha \in [0, 1]$ if $\epsilon_D < \epsilon_W$ at α_0.

Observe that the argument does not depend on the specific value of α_0 and the statements are true for any $\alpha_0 \in [0, 1]$. Therefore, $p_1^e(\alpha)$ must be monotonic in α. The argument for $\tilde{p}_1^e(\alpha)$ and $\tilde{p}_1^n(\alpha)$ follows immediately by replacing $p_1^e(\alpha)$. □

A.4 Equilibrium

Lemma 5 characterizes the Nash equilibria in pure strategies. We implicitly define the unique profit threshold $\hat{\alpha}$ such that $\pi^n = \tilde{\pi}^e$ when $\alpha = \hat{\alpha}$. When $\epsilon_D > \epsilon_W$, we can also implicitly define the unique threshold $\bar{\alpha}$ such that $\tilde{\pi}^n = \pi^e$ when $\alpha = \bar{\alpha}$.[40]

Lemma 5 (Equilibrium).

(a) Suppose $\epsilon_D > \epsilon_W$.

[40]When $\epsilon_D < \epsilon_W$, then both profits, $\tilde{\pi}^n$ and π^e, are strictly increasing in α.

(i) *If* $\alpha < \min\{\bar{\alpha}, \hat{\alpha}\}$, *then both firms do not exploit and set* $p_1^* = p_1^n$ *and* $p_2^* = W(v_2)$.

(ii) *If* $\alpha > \max\{\bar{\alpha}, \hat{\alpha}\}$, *then both firms exploit and set* $p_1^* = p_1^e(\alpha)$ *and* $p_2^* = W(v_2, \Delta)$.

(iii) *If* $\bar{\alpha} < \alpha < \hat{\alpha}$, *then either both firms do not exploit or both firms exploit.*

(iv) *If* $\hat{\alpha} < \alpha < \bar{\alpha}$, *then firm* j *does not exploit and sets* $p_{1,j}^* = \tilde{p}_1^n(\alpha)$ *and* $p_{2,j}^* = W(v_2)$, *and firm* $-j$ *exploits and sets* $p_{1,-j}^* = \tilde{p}_1^e(\alpha)$ *and* $p_{2,-j}^* = W(v_2, \Delta)$.

(b) *Suppose* $\epsilon_D < \epsilon_W$.

(i) *If* $\alpha < \hat{\alpha}$, *then only symmetric equilibria exist.*

(ii) *If* $\alpha > \hat{\alpha}$, *then symmetric exploiting and asymmetric equilibria exist.*

In the case of (a), $\epsilon_D > \epsilon_W$, the symmetric non-exploiting equilibrium in (i) and the symmetric exploiting equilibrium in (ii) are unique.[41] In (iii), the best response of a firm is to do the same as the rival, and in (iv), the best response is to do the opposite.[42] Thus, for intermediate values of α, we observe either multiple symmetric equilibria or multiple asymmetric equilibria.

In part (b), when $\epsilon_D < \epsilon_W$, we observe a similar pattern of equilibria, but we cannot characterize when a unique symmetric equilibrium emerges. For a low share of behavioral consumers (i), either both firms do not exploit (when $\tilde{\pi}^n > \pi^e$) or there exists multiple symmetric equilibria like in case $(aiii)$. For a large α (ii), either both firms exploit, or an asymmetric outcome emerges like in case (aiv).

A.5 Proof of Lemma 5

We will first prove two intermediate result.

Lemma 6 (Unique thresholds)**.**

(i) *The critical threshold* $\hat{\alpha}$ *is the unique solution to* $\pi^n = \tilde{\pi}^e$ *and* $\alpha < \hat{\alpha} \Leftrightarrow \pi^n > \tilde{\pi}^e$.

[41] If $\pi^n > \tilde{\pi}^e$ and $\tilde{\pi}^n > \pi^e$, non-exploiting is the dominant strategy for both firms. Similarly, if $\pi^n < \tilde{\pi}^e$ and $\tilde{\pi}^n < \pi^e$, then exploiting is the dominant strategy.

[42] Lemma 5 $(a)(iii)$ also applies, when $\bar{\alpha} = \alpha < \hat{\alpha}$ or $\bar{\alpha} < \alpha = \hat{\alpha}$. When $\hat{\alpha} = \alpha < \bar{\alpha}$, then, next to the asymmetrica equilibria, there exist also the symmetric non-exploiting equilibrium. Similarly, when $\hat{\alpha} < \alpha = \bar{\alpha}$, then, next to the asymmetrica equilibria, there exist also the symmetric exploiting equilibrium. In the special case of $\alpha = \hat{\alpha} = \bar{\alpha}$, any strategy is optimal since $\pi^n = \pi^e = \tilde{\pi}^n = \tilde{\pi}^e$.

(ii) Suppose $\epsilon_D > \epsilon_W$. The critical threshold $\bar{\alpha}$ is the unique solution to $\tilde{\pi}^n = \pi^e$ and $\alpha < \bar{\alpha} \Leftrightarrow \tilde{\pi}^n > \pi^e$.

Proof. (i) By Lemma 3, π^n is constant in α and $\tilde{\pi}^e$ is strictly increasing in α. Thus, there exists a unique solution solved for α such that $\pi^n = \tilde{\pi}^e$ and $\alpha < \hat{\alpha} \Leftrightarrow \pi^n > \tilde{\pi}^e$.

(ii) When $\epsilon_D > \epsilon_W$, then, by Lemma 3, π^e is strictly increasing in α and $\tilde{\pi}^n$ is decreasing in α. Thus, there exists a unique solution solved for α such that $\tilde{\pi}^n = \pi^e$ and $\alpha < \bar{\alpha} \Leftrightarrow \tilde{\pi}^n > \pi^e$.

□

Lemma 7 (Dominant strategies)**.**

(i) Non-exploiting is the dominant strategy for both firms if $\pi^n > \tilde{\pi}^e$ and $\tilde{\pi}^n > \pi^e$.

(ii) Exploiting is the dominant strategy for both firms if $\pi^n < \tilde{\pi}^e$ and $\tilde{\pi}^n < \pi^e$.

Proof. (i) First, suppose that firm $-j$ does not exploit. The best response of firm j is to not exploit since $\pi^n > \tilde{\pi}^e$. Now suppose that firm $-j$ does exploit. The best response of firm j is to not exploit since $\tilde{\pi}^n > \pi^e$. Hence, in any case, the best response is to not exploit and thus, the dominant strategy. The best response of firm $-j$ is similarly.

(ii) First, suppose that firm $-j$ does not exploit. The best response of firm j is to exploit since $\pi^n < \tilde{\pi}^e$. Now suppose that firm $-j$ does exploit. The best response of firm j is to exploit since $\tilde{\pi}^n < \pi^e$. Hence, in any case, the best response is to exploit and thus the dominant strategy. The best response of firm $-j$ is similarly.

□

Now, we can proof the statements in Lemma 5.

(a) (i) By Lemma 6, we have $\pi^n > \tilde{\pi}^e$ and $\tilde{\pi}^n > \pi^e$ if $\alpha < \min\{\bar{\alpha}, \hat{\alpha}\}$. Hence, by Lemma 7, it is optimal for both firms to not exploit behavioral consumers, and set $p_1^* = p_1^n$ and $p_2^* = W(v_2)$.

(ii) By Lemma 6, we have $\pi^n < \tilde{\pi}^e$ and $\tilde{\pi}^n < \pi^e$ if $\alpha > \max\{\bar{\alpha}, \hat{\alpha}\}$. Hence, by Lemma 7, it is optimal for both firms to exploit behavioral consumers, and set $p_1^* = p_1^e(\alpha)$ and $p_2^* = W(v_2, \Delta)$.

(iii) By Lemma 6, we have $\pi^n > \tilde{\pi}^e$ and $\tilde{\pi}^n < \pi^e$ if $\bar{\alpha} < \alpha < \hat{\alpha}$. Suppose that firm $-j$ does not exploit. The best response of firm j is to not exploit since $\pi^n > \tilde{\pi}^e$. Now suppose that firm $-j$ does exploit. The best response of firm j is to exploit since $\tilde{\pi}^n < \pi^e$. Hence, the best response of firm j is to do the same as firm $-j$. The best response of firm $-j$ is similarly. Thus, there exists two Nash equilibria in pure strategies {(not exploit, not exploit),(exploit,exploit)}.

(iv) By Lemma 6, we have $\pi^n < \tilde{\pi}^e$ and $\tilde{\pi}^n > \pi^e$ if $\hat{\alpha} < \alpha < \bar{\alpha}$. Suppose that firm $-j$ does not exploit. The best response of firm j is to exploit since $\pi^n < \tilde{\pi}^e$. Now suppose that firm $-j$ does exploit. The best response of firm j is to not exploit since $\tilde{\pi}^n > \pi^e$. Hence, the best response of firm j is to do the opposite as firm $-j$. The best response of firm $-j$ is similarly. Thus, there exists two Nash equilibria in pure strategies {(not exploit, exploit),(exploit, not exploit)}.

(b) (i) By Lemma 6, we have $\pi^n > \tilde{\pi}^e$. If $\tilde{\pi}^n > \pi^e$, then by Lemma 7, it is optimal for both firms to not exploit behavioral consumers. Thus, the unique symmetric non-exploiting equilibrium emerges. Otherwise, if If $\tilde{\pi}^n < \pi^e$, case $(a)(iii)$ arises and the best response of firm j is to do the same as firm $-j$. Thus, multiple symmetric equilibria emerge. In either case, only symmetric equilibria exist.

(ii) By Lemma 6, we have $\pi^n < \tilde{\pi}^e$. If $\tilde{\pi}^n < \pi^e$, then by Lemma 7, it is optimal for both firms to exploit behavioral consumers. Thus, the unique symmetric exploiting equilibrium emerges. Otherwise, if If $\tilde{\pi}^n > \pi^e$, case $(a)(iv)$ arises and the best response of firm j is to do the opposite as firm $-j$. Thus, multiple asymmetric equilibria emerge. The symmetric non-exploiting equilibrium does not exist if $\alpha > \hat{\alpha}$.

□

B Proofs Main Results

B.1 Proof of Lemma 1

First,

$$p_1^n = p_1^e(\alpha) \quad \Leftrightarrow \quad \alpha = \frac{W(v_2)}{W(v_2,\Delta) + \frac{\frac{\partial W(v_2,\Delta)}{\partial p_{1,j}^*} D(p_{1,j}^*, p_{1,-j}^*)}{D'(p_{1,j}^*, p_{1,-j}^*)}} = \bar{\alpha}_p.$$

Further, observe that $p_1^n = \tilde{p}_1^n(\alpha)$ implies $D(p_1^n, p_1^n) = D(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha))$. But then it must be $p_1^n = \tilde{p}_1^e(\alpha)$ when $p_1^n = \tilde{p}_1^n(\alpha)$. Similarly, $p_1^e(\alpha) = \tilde{p}_1^e(\alpha)$ implies $D(p_1^e(\alpha), p_1^e(\alpha)) = D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))$. But then it must be $p_1^e(\alpha) = \tilde{p}_1^n(\alpha)$ when $p_1^e(\alpha) = \tilde{p}_1^e(\alpha)$. Combining these two statements implies $p_1^n = \tilde{p}_1^n(\alpha) = \tilde{p}_1^e(\alpha) = p_1^e(\alpha)$, which is true when $\alpha = \bar{\alpha}_p$.

In the benchmark economy ($\alpha = 0$) only the non-exploiting strategy is possible, which implies $p_1^b = p_1^n$. Hence, $p_1^b = p_1^n = p_1^e(\bar{\alpha}_p) = \tilde{p}_1^n(\bar{\alpha}_p) = \tilde{p}_1^e(\bar{\alpha}_p)$.

(i) $\epsilon_D > \epsilon_W$ implies $\bar{\alpha}_p > 0$. Further, by Lemma 3, the prices $p_1^e(\alpha)$, $\tilde{p}_1^n(\alpha)$ and $\tilde{p}_1^e(\alpha)$ are decreasing in α when $\epsilon_D > \epsilon_W$ and $p_1^b = p_1^n$ are constant in α. Hence, for any $\alpha \in (\min\{\bar{\alpha}, \hat{\alpha}\}, \bar{\alpha}_p)$, it follows $p_1^* \in \{p_1^e(\alpha), \tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha)\} > p_1^b$, and for any $\alpha > \bar{\alpha}_p$ it follows $p_1^* \in \{p_1^e(\alpha), \tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha)\} < p_1^b$.

(ii) $\epsilon_D < \epsilon_W$ implies $\bar{\alpha}_p < 0$. By Lemma 3, the prices $p_1^e(\alpha)$, $\tilde{p}_1^n(\alpha)$ and $\tilde{p}_1^e(\alpha)$ are increasing in α when $\epsilon_D < \epsilon_W$ and $p_1^b = p_1^n$ are constant in α. Hence, for any $\alpha > 0 > \bar{\alpha}_p$, it follows $p_1^* \in \{p_1^e(\alpha), \tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha)\} > p_1^b$.

□

B.2 Proof of Proposition 1

Lemma 8. *Suppose $\epsilon_D > \epsilon_W$. The price threshold is larger than any profit threshold, $\max\{\hat{\alpha}, \bar{\alpha}\} < \bar{\alpha}_p$.*

Proof. Suppose $\alpha = \bar{\alpha}_p$. Hence, $p_1^n = p_1^e(\bar{\alpha}_p) = \tilde{p}_1^n(\bar{\alpha}_p) = \tilde{p}_1^e(\bar{\alpha}_p)$. It follows

$$\pi^n = \frac{-D(p_1^n, p_1^n)^2}{D'(p_1^n, p_1^n)} < \frac{-[1 + \alpha W'(v_2, \Delta)] D(\tilde{p}_1^e(\bar{\alpha}_p), \tilde{p}_1^n(\bar{\alpha}_p))^2}{D'(\tilde{p}_1^e(\bar{\alpha}_p), \tilde{p}_1^n(\bar{\alpha}_p))} = \tilde{\pi}^e,$$

since $p_1^n = \tilde{p}_1^n(\bar{\alpha}_p) = \tilde{p}_1^e(\bar{\alpha}_p)$ and $W'(v_2, \Delta) > 0$. By Lemma 6, it must be $\alpha > \hat{\alpha}$ when $\pi^n < \tilde{\pi}^e$. Thus, $\bar{\alpha}_p > \hat{\alpha}$.

Further, we have

$$\tilde{\pi}^n = \frac{-D(\tilde{p}_1^n(\bar{\alpha}_p), \tilde{p}_1^e(\bar{\alpha}_p))^2}{D'(\tilde{p}_1^n(\bar{\alpha}_p), \tilde{p}_1^e(\bar{\alpha}_p))} < \frac{-[1 + \alpha W'(v_2, \Delta)]D(p_1^e(\bar{\alpha}_p), p_1^e(\bar{\alpha}_p))^2}{D'(p_1^e(\bar{\alpha}_p), p_1^e(\bar{\alpha}_p))} = \pi^e,$$

since $p_1^e(\bar{\alpha}_p) = \tilde{p}_1^n(\bar{\alpha}_p) = \tilde{p}_1^e(\bar{\alpha}_p)$ and $W'(v_2, \Delta) > 0$. By Lemma 6, it must be $\alpha > \bar{\alpha}$ when $\tilde{\pi}^n < \pi^e$. Thus, $\bar{\alpha}_p > \bar{\alpha}$. Hence, $\max\{\hat{\alpha}, \bar{\alpha}\} < \bar{\alpha}_p$. □

We denote the utility consumers receive from the base good with v_1. The surplus of a classical consumer in the benchmark economy ($\alpha = 0$) is given by $v_1 - p_1^b + W(v_2) - p_2 = v_1 - p_1^b$ since $p_2 = W(v_2)$ in any benchmark (and symmetric non-exploiting) equilibrium. Hence, not consuming the add-on does not decrease the surplus of a classical consumer. Thus, a classical consumer benefits, compared to the benchmark, from the presence of behavioral consumers when $p_1^* < p_1^b$. Otherwise, when $p_1^* > p_1^b$, classical consumers are harmed.

(a) (*i*) By Lemma 5, there exists symmetric exploiting equilibria with $p_1^* = p_1^e(\alpha)$. By Lemma 1, we have $p_1^e(\alpha) > p_1^b$ for $\alpha \in (\min\{\bar{\alpha}, \hat{\alpha}\}, \bar{\alpha}_p)$, which reduces a classical consumer's surplus compared to the benchmark. Thus, classical consumers are harmed by the presence of behavioral consumers. Lemma 8 proofs that $\alpha \in (\max\{\hat{\alpha}, \bar{\alpha}\}, \bar{\alpha}_p)$ exists. By Lemma 1, we have $p_1^e(\alpha) < p_1^b$ for all $\alpha > \bar{\alpha}_p$, which increases a classical consumer's surplus compared to the benchmark. Thus, classical consumers benefit by the presence of behavioral consumers.

(*ii*) By Lemma 5, asymmetric equilibria exist only if $\hat{\alpha} < \alpha < \bar{\alpha}$. Therefore, by Lemma 8, we have $\alpha < \bar{\alpha}_p$ in any asymmetric equilibrium, which implies, by Lemma 1, $\tilde{p}_1^n(\alpha) > p_1^b$ and $\tilde{p}_1^e(\alpha) > p_1^b$. Hence, regardless from which firm classical consumers buy the base good, their surplus is lower compared to the benchmark. Thus, classical consumers are harmed by the presence of behavioral consumers in any asymmetric equilibrium.

(*iii*) We have $p_1^n = p_1^b$ in any symmetric non-exploiting equilibrium. Hence, the surplus of a classical consumer is the same as in the benchmark and they are unaffected by the presence of behavioral consumers.

(b) (*i*) By Lemma 5, there exists symmetric exploiting equilibria and asymmetric equilibria. By Lemma 1, we have $p_1^* \in \{p_1^e(\alpha), \tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha)\} > p_1^b$ for all α. Hence, a classical consumer's surplus is lower compared to the benchmark in any symmetric exploiting equilibrium or asymmetric equilibrium. Thus, classical consumers are harmed by the presence of behavioral consumers.

(ii) We have $p_1^n = p_1^b$ in any symmetric non-exploiting equilibrium. Hence, the surplus of a classical consumer is the same as in the benchmark and they are unaffected by the presence of behavioral consumers.

□

B.3 Proof of Proposition 2

(a) The total surplus of a behavioral consumer, when the behavioral effect does not increase the add-on utility, is given by $U_b = v_1 - p_1 + W(v_2) - p_2$, where v_1 denotes the gross utility received from the base good. The total surplus of a behavioral consumer, when the behavioral effect increases the add-on utility, is given by $\tilde{U}_b = v_1 - p_1 + W(v_2, \Delta) - p_2$.

(i) The condition that behavioral consumers are worse of by exploitation is independent of whether U_b or $\tilde{U}_b$ applies.

$$\begin{aligned}
U_b^{NE} = v_1 - p_1^n + W(v_2) - p_2 &> v_1 - p_1^e(\alpha) + W(v_2) - p_2 = U_b^E \\
\Leftrightarrow \quad & p_1^e(\alpha) > p_1^n + W(v_2) - W(v_2, \Delta) \\
\tilde{U}_b^{NE} = v_1 - p_1^n + W(v_2, \Delta) - p_2 &> v_1 - p_1^e(\alpha) + W(v_2) - p_2 = \tilde{U}_b^E \\
\Leftrightarrow \quad & p_1^e(\alpha) > p_1^n + W(v_2) - W(v_2, \Delta)
\end{aligned}$$

Observe that the term $W(v_2, \Delta)$ is different in U_b^E and $\tilde{U}_b^{NE}$ because in the former, the reference price is $p_1^e(\alpha)$ and in the latter p_1^n. However, we show that the condition is satisfied for any reference price. The condition $p_1^e(\alpha) > p_1^n + W(v_2) - W(v_2, \Delta)$ holds for all $\alpha \in [0, 1]$. It is immediate to see that the condition is satisfied when $\alpha \leq \bar{\alpha}_p$, which implies $p_1^e(\alpha) \geq p_1^n$ by Lemma 1, and since $W(v_2) < W(v_2, \Delta)$.

When $\alpha > \bar{\alpha}_p$, which implies $p_1^e(\alpha) < p_1^n$, we need an intermediate step. Observe that

$$
\begin{aligned}
\frac{[1+\alpha W'(v_2,\Delta)]D(p_1^e(\alpha),p_1^e(\alpha))}{-D'(p_1^e(\alpha),p_1^e(\alpha))} &> \frac{D(p_1^n,p_1^n)}{-D'(p_1^n,p_1^n)} \\
\alpha > \frac{\frac{D(p_1^n,p_1^n)}{-D'(p_1^n,p_1^n)}\frac{-D'(p_1^e(\alpha),p_1^e(\alpha))}{D(p_1^e(\alpha),p_1^e(\alpha))}-1}{W'(v_2,\Delta)} &= \frac{\frac{\epsilon_{D(E)}}{\epsilon_{D(NE)}}-1}{W'(v_2,\Delta)} \\
\alpha \geq 0 &> \frac{\frac{\epsilon_{D(e)}}{\epsilon_{D(NE)}}-1}{W'(v_2,\Delta)},
\end{aligned}
$$

where $D(s) = D(p_1^s, p_1^s)$ for $s = n, e$. The last inequality follows from the fact that $\frac{\partial \epsilon_D}{\partial p_1} > 0$ when $\epsilon_D > \epsilon_W$ by the proof of Lemma 4. Thus, we have $\epsilon_{D(n)} > \epsilon_{D(e)}$ when $p_1^e(\alpha') < p_1^n$, which implies $\frac{\epsilon_{D(e)}}{\epsilon_{D(n)}} - 1 {<} 0$.

Now, we show that $p_1^e(\alpha) < p_1^n + W(v_2) - W(v_2,\Delta)$ never holds for $\alpha \in [0,1]$.

$$
\begin{aligned}
p_1^e(\alpha) &< p_1^n + W(v_2) - W(v_2,\Delta) \\
\frac{[1+\alpha W'(v_2,\Delta)]D(p_1^e(\alpha),p_1^e(\alpha))}{-D'(p_1^e(\alpha),p_1^e(\alpha))} - \alpha W(v_2,\Delta) &< \frac{D(p_1^n,p_1^n)}{-D'(p_1^n,p_1^n)} - W(v_2,\Delta) \\
\alpha &> \underbrace{\frac{\frac{[1+\alpha W'(v_2,\Delta)]D(p_1^e(\alpha),p_1^e(\alpha))}{-D'(p_1^e(\alpha),p_1^e(\alpha))} - \frac{D(p_1^n,p_1^n)}{-D'(p_1^n,p_1^n)}}{W(v_2,\Delta)}}_{>0} + 1 > 1,
\end{aligned}
$$

which is a contradiction for any $\alpha \in [0,1]$. Hence, it must be $p_1^e(\alpha) > p_1^n + W(v_2) - W(v_2,\Delta)$ for any $\alpha \in [0,1]$, which implies that behavioral consumers are always better off in a non-exploiting equilibrium (or benchmark economy) than in an exploiting equilibrium.

(ii) Total consumer surplus is larger in a symmetric non-exploiting equilibrium when

$$
\begin{aligned}
CS^{NE} &> CS^E(\alpha) - \alpha D(\cdot)[W(v_2,\Delta) - W(v_2)] \\
\alpha D(\cdot)[W(v_2,\Delta) - W(v_2)] &> CS^E(\alpha) - CS^{NE}
\end{aligned}
$$

By Lemma 1 we have $p_1^e(\alpha) \geq p_1^n$ when $\alpha \leq \bar{\alpha}_p$, which implies $D(p_1^e(\alpha), p_1^e(\alpha)) \leq D(p_1^n, p_1^n)$. Thus, it must be $CS^E(\alpha) \leq CS^{NE}$ which implies $\alpha D(\cdot)[W(v_2,\Delta) - W(v_2)] > 0 \geq CS^E(\alpha) - CS^{NE}$ since $W(v_2,\Delta) > W(v_2)$.

(b) When $\epsilon_D < \epsilon_W$, then by Lemma 1 we have $p_1^e(\alpha) > p_1^n$. Following the argument of part (a), behavioral consumers are always better off in a non-exploiting equilibrium when $p_1^e(\alpha) > p_1^n$, which is always the case when firms exploit.

Further, $p_1^e(\alpha) > p_1^n$ implies again $CS^E(\alpha) < CS^{NE}$ and thus, total consumer surplus is always lower under exploitation □

B.4 Monopoly and Perfect Competition

Suppose that base goods are perfectly differentiated, then each firm is a monopolist in its respective base-good market. Further, suppose that $D(p_1)$ is strictly decreasing, twice continuously differentiable, $\lim_{p_1 \to \infty} D(p_1) = 0$ and satisfies $D(p_1)D''(p_1) < 2D'(p_1)^2$. Observe that the monopolist's maximization problem is identical to equation (1), and yields $\pi^n = \pi(p_1^n, W(v_2))$ when choosing the non-exploiting strategy and $\pi^e = \pi(p_1^e(\alpha), W(v_2, \Delta))$ when choosing the exploiting strategy. Observe that the profits and prices are similar to the symmetric outcomes with two firms. Therefore, we can directly apply Lemma 3 and Lemma 4, which implies that π^n is constant in α and π^e strictly increasing in α. Define the profit threshold $\hat{\alpha}$ such that $\pi^n = \pi^e$.

Lemma 9.

(i) If $\alpha < \hat{\alpha}$, then the monopolist does not exploit and sets $p_1^* = p_1^n$ and $p_2^* = W(v_2)$.

(ii) If $\alpha > \hat{\alpha}$, then the monopolist exploits and sets $p_1^* = p_1^e(\alpha)$ and $p_2^* = W(v_2, \Delta)$.

Proof. The proof is analogous to the proof of Lemma 5. □

The remainder of the analysis is similar to the baseline model with two firms. The critical price threshold $\bar{\alpha}_p$ is unchanged. Therefore, Lemma 1 without asymmetric prices follows immediately. Further, analogous to Lemma 8, we have $\hat{\alpha} < \bar{\alpha}_p$. Thus, Proposition 5 (a) and (b) below follow and is analogous to Proposition 1 and 2.

Proposition 5 (Monopoly and perfect competition)**.**

(a) Under a monopolist, the presence of behavioral consumers harms classical consumers in any exploiting equilibrium except if $\alpha > \bar{\alpha}_p$ and $\epsilon_D > \epsilon_W$. Classical consumers are unaffected in any non-exploiting equilibrium.

(b) Behavioral consumers are always worse off when a monopolist exploits them. For $\epsilon_D > \epsilon_W$, total consumer surplus is strictly lower when $\alpha D(\cdot)[W(v_2, \Delta) - W(v_2)] >$

$CS^E(\alpha) - CS^{NE}$. The condition is always satisfied for $\alpha \leq \bar{\alpha}_p$. For $\epsilon_D < \epsilon_W$, total consumer surplus is strictly lower when a monopolist exploits.

(c) *Classical consumers are never harmed by the presence of behavioral consumers under perfect price competition. Classical consumers benefit in any symmetric exploiting equilibrium and are unaffected in any symmetric non-exploiting equilibrium.*

Proof. (a) The proof is analogous to the proof to Proposition 1.

(b) The proof is analogous to the proof to Proposition 2. Firms must earn zero profits under perfect competition, which implies $p_{1,j}D(\cdot) = -p_2Q(\cdot)$. Further, they must offer the lowest price given the zero profit constraint. Otherwise, firms would face zero demand. Therefore, it must be $p_1^* = \min\{-W(v_2), -\alpha W(v_2, \Delta)\}$. Hence, it is optimal to exploit behavioral consumers only if $\alpha W(v_2, \Delta) > W(v_2)$. The unique symmetric exploiting equilibrium exists if and only if $\alpha > \frac{W(v_2)}{W(v_2,\Delta)}$. Otherwise, when $\alpha < \frac{W(v_2)}{W(v_2,\Delta)}$, the unique symmetric non-exploiting equilibrium exists. In the benchmark economy with $\alpha = 0$, firms choose $p_2 = W(v_2)$ and $p_1^b = -W(v_2)$. Thus, in any symmetric non-exploiting equilibrium, firms set $p_1^n = p_1^b = -W(v_2)$ and classical consumers are unaffected by the presence of behavioral consumers. In any exploiting equilibrium, it must be $p_1^e = -\alpha W(v_2, \Delta) < -W(v_2) = p_1^b$. Hence, classical consumers have to pay strictly less in any exploiting equilibrium than in the benchmark and thus, benefit. Lastly, there exist no profitable deviations for firms. Changing p_2 leads to less add-on revenues and thus, a higher p_1 and zero base-good demand. Increasing p_1 leads to zero demand and thus zero profits. Decreasing p_1 would lead to negative profits.

□

B.5 Proof of Proposition 3

(a) The total surplus of a classical consumer is given by $v_1 - p_1 + W(v_2) - p_2$, where v_1 denotes the gross utility received from the base good. The total surplus of a behavioral consumer, when the behavioral effect does not increase the add-on utility, is given by $U_b = v_1 - p_1 + W(v_2) - p_2$. The total surplus of a behavioral consumer, when the behavioral effect increases the add-on utility, is given by $\tilde{U}_b = v_1 - p_1 + W(v_2, \Delta) - p_2$.

(i) Observe that $\frac{\partial p_1^e(\alpha)}{\partial \alpha} < 0$ when $\epsilon_D > \epsilon_W$ by Lemma 3. By definition, the ex-post equilibrium is identical to the ex-ante equilibrium when the policy is ineffective and firms still set $p_1^e(\alpha)$. Denote with α' the share of behavioral consumers ex-ante and with $\alpha'' < \alpha'$ the share ex-post. Since $\alpha'' < \alpha'$ and $\frac{\partial p_1^e(\alpha)}{\partial \alpha} < 0$, we have $p_1^e(\alpha'') > p_1^e(\alpha')$. Since $\frac{\partial W(v_2,\Delta)}{\partial p_1} > 0$, the add-on price $p_2^* = W(v_2, \Delta)$ also increases ex-post. Observe that the add-on surplus is either unaffected or worse ex-post and the base-good surplus strictly lower ex-post since v_1 is unchanged and p_1 strictly larger. Hence, all consumers are worse off by an ineffective policy.

(ii) We first proof the result on classical consumers. By definition, an effective policy leads to a non-exploiting equilibrium ex-post with $p_1^* = p_1^n = p_1^b$ and $p_2^* = W(v_2)$. The add-on surplus remains at zero since $p_2^* = W(v_2)$. Hence, a classical consumer benefits from an effective policy when $p_1^e(\alpha') > p_1^b$, which, by Proposition 1, is the case when the ex-ante equilibrium was harmful. Similarly, when the ex-ante equilibrium was beneficial, which implies $p_1^e(\alpha') < p_1^b$, then a classical consumer is worse off ex-post.

The condition that behavioral consumers benefit from an effective policy is independent of whether U_b or $\tilde{U}_b$ applies.

$$
\begin{aligned}
U_b^B = v_1 - p_1^b + W(v_2) - p_2 &> v_1 - p_1^e(\alpha') + W(v_2) - p_2 = U_b^E \\
\Leftrightarrow \quad & p_1^e(\alpha') > p_1^b + W(v_2) - W(v_2, \Delta) \\
\tilde{U}_b^B = v_1 - p_1^b + W(v_2, \Delta) - p_2 &> v_1 - p_1^e(\alpha') + W(v_2) - p_2 = \tilde{U}_b^E \\
\Leftrightarrow \quad & p_1^e(\alpha') > p_1^b + W(v_2) - W(v_2, \Delta)
\end{aligned}
$$

Since we obtain the identical condition, we can remain agnostic about the exact welfare function for behavioral consumers. They benefit from the effective policy when the total surplus in the ex-post equilibrium is larger than in the ex-ante exploiting equilibrium, which is the case when $p_1^e(\alpha') > p_1^b + W(v_2) - W(v_2, \Delta)$. We now proof that the condition $p_1^e(\alpha') > p_1^b + W(v_2) - W(v_2, \Delta)$ holds for all $\alpha \in [0, 1]$. It is immediate to see that the condition is satisfied when the ex-ante equilibrium was harmful, which implies $p_1^e(\alpha') > p_1^b$, since $W(v_2) < W(v_2, \Delta)$.

When $p_1^e(\alpha') < p_1^b$, we need an intermediate step. Observe that

$$\begin{aligned}
\frac{[1+\alpha W'(v_2,\Delta)]D(p_1^e(\alpha),p_1^e(\alpha))}{-D'(p_1^e(\alpha),p_1^e(\alpha))} &> \frac{D(p_1^b,p_1^b)}{-D'(p_1^b,p_1^b)} \\
\alpha > \frac{\frac{D(p_1^b,p_1^b)}{-D'(p_1^b,p_1^b)}\frac{-D'(p_1^e(\alpha),p_1^e(\alpha))}{D(p_1^e(\alpha),p_1^e(\alpha))}-1}{W'(v_2,\Delta)} &= \frac{\frac{\epsilon_{D(e)}}{\epsilon_{D(b)}}-1}{W'(v_2,\Delta)} \\
\alpha \geq 0 &> \frac{\frac{\epsilon_{D(e)}}{\epsilon_{D(b)}}-1}{W'(v_2,\Delta)},
\end{aligned}$$

where $D(s) = D(p_1^s, p_1^s)$ for $s = b, e$. The last inequality follows from the fact that $\frac{\partial \epsilon_D}{\partial p_1} > 0$ when $\epsilon_D > \epsilon_W$ by the proof of Lemma 4. Thus, we have $\epsilon_{D(b)} > \epsilon_{D(e)}$ when $p_1^e(\alpha') < p_1^b$, which implies $\frac{\epsilon_{D(e)}}{\epsilon_{D(b)}} - 1 < 0$.

Now, we show that $p_1^e(\alpha') < p_1^b + W(v_2) - W(v_2,\Delta)$ never holds for $\alpha \in [0,1]$.

$$\begin{aligned}
p_1^e(\alpha') &< p_1^b + W(v_2) - W(v_2,\Delta) \\
\frac{[1+\alpha W'(v_2,\Delta)]D(p_1^e(\alpha),p_1^e(\alpha))}{-D'(p_1^e(\alpha),p_1^e(\alpha))} - \alpha W(v_2,\Delta) &< \frac{D(p_1^b,p_1^b)}{-D'(p_1^b,p_1^b)} - W(v_2,\Delta) \\
\alpha &> \underbrace{\frac{\frac{[1+\alpha W'(v_2,\Delta)]D(p_1^e(\alpha),p_1^e(\alpha))}{-D'(p_1^e(\alpha),p_1^e(\alpha))} - \frac{D(p_1^b,p_1^b)}{-D'(p_1^b,p_1^b)}}{W(v_2,\Delta)}}_{>0} + 1 > 1,
\end{aligned}$$

which is a contradiction for any $\alpha \in [0,1]$. Hence, it must be $p_1^e(\alpha') > p_1^b + W(v_2) - W(v_2,\Delta)$ for any $\alpha \in [0,1]$, which implies that the total surplus in the benchmark economy (or non-exploiting equilibrium) is always larger than in the exploiting equilibrium. Therefore, any effective policy benefits behavioral consumers.

(iii) Since all consumers are individually worse off by an ineffective policy, it must be that also total consumer surplus decreases. Similarly, when all consumers are better off, then consumer surplus must increase. As shown above, this is the case when the policy is effective and the ex-ante equilibrium was harmful. Lastly, when the policy is effective and the ex-ante equilibrium was beneficial, then behavioral consumer surplus increases and classical consumer surplus decreases. Hence, the effect on total consumer surplus is unclear.

(b) When $\epsilon_D > \epsilon_W$, then by Lemma 3, $p_1^e(\alpha)$ is increasing in α. Thus, any decrease in α

reduces base-good and add-on prices. Therefore, any policy is beneficial.

□

B.6 Price Floor

We impose a price floor $\underline{p}_1 \leq p_1^b = p_1^n$ on the base good. Observe that this does not affect the benchmark equilibrium. Further, $\underline{p}_1$ is never binding when $\epsilon_D < \epsilon_W$ since, by Lemma 1, this implies $p_1^* \in \{p_1^e(\alpha), \tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha)\} > p_1^b \geq \underline{p}_1$ for any α. Thus, we focus on the case of $\epsilon_D > \epsilon_W$. We define $\underline{\alpha}$ such that $\alpha = \underline{\alpha} \Leftrightarrow p_1^e(\alpha) = \underline{p}_1$.

Proposition 6 (Price floor)**.** *Suppose $\epsilon_D > \epsilon_W$ and $\underline{p}_1 \leq p_1^b$. A price floor (i) increases the base-good price, (ii) increases the add-on price, and (iii) reduces the base-good demand if and only if $\alpha > \underline{\alpha}$. classical consumers are strictly worse off and behavioral consumers are weakly worse off by a binding regulation.*

Proof. By Lemma 3, $p_1^* \in \{p_1^e(\alpha), \tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha)\}$ is decreasing in α when $\epsilon_D > \epsilon_W$. Therefore, the price floor is binding, $p_1^* < \underline{p}_1$, if and only if $\alpha > \underline{\alpha}$. Thus, the base-good price increases by the regulation if and only if $\alpha > \underline{\alpha}$. Next, since $\frac{\partial W(v_2,\Delta)}{\partial p_1} > 0$ and $p_2^* = W(v_2, \Delta)$ for any exploiting firm, the add-on price increases. Lastly, $p_1^* < \underline{p}_1$ implies $D(p_{1,j}^*, p_{1,-j}^*) > D(\underline{p}_1, \max\{p_{1,-j}^*, \underline{p}_1\})$.

Classical buyers still only purchase the base good, but now at a higher price than without regulation, $p_1^* < \underline{p}_1$. Thus, an individual classical consumer is worse off when the regulation is binding and unaffected when not binding.

The effect on individual behavioral buyers in the base-good market is identical. The effect on the surplus in the add-on market depends on how we define the actual utility that a behavioral consumer enjoys from purchasing the add-on. That is whether a behavioral consumer receives the whole WTP $W(v_2, \Delta)$ or only $W(v_2)$ as utility. In the former case, the add-on surplus remains zero, and the regulation has no effect. In the latter case, the add-on surplus decreases as the utility is fixed while the add-on price increases. Thus, a behavioral buyer is weakly worse off in the add-on market by the regulation.

□

C Further Results

C.1 Sequential Buying

Observe that the equilibrium entails mixing of p_2 with searching consumers. Depending on the chosen strategy, setting the monopolistic add-on price $p_2 \in \{W(v_2), W(v_2, \Delta)\}$ with probability 1 is not optimal. A firm can profitably deviate by setting a slightly lower add-on price and capture the add-on demand of all searching consumers. The expected profit of a non-exploiting firm is given by

$$\mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, p_2^n) = p_{1,j}D_j(\cdot) + (1-\rho)D_j(\cdot)p_2^n + \rho[1 - F^n(p_2^n)][D_j(\cdot) + D_{-j}(\cdot)]p_2^n.$$

The term $1 - F^n(p_2^n)$ denotes the probability to set a lower add-on price than the competitor. The expected profit of an exploiting firm is given by

$$\mathbb{E}\pi_j^e(p_{1,j}, p_{1,-j}, p_2^e) = p_{1,j}D_j(\cdot) + \alpha(1-\rho)D_j(\cdot)p_2^e + \alpha\rho[1 - F^e(p_2^e)][D_j(\cdot) + D_{-j}(\cdot)]p_2^e.$$

However, firms can always obtain positive add-on profits by selling the add-on to loyal consumers at $p_2 \in \{W(v_2), W(v_2, \Delta)\}$, and earn $(1-\rho)D_j(\cdot)W(v_2)$ and $\alpha(1-\rho)D_j(\cdot)W(v_2, \Delta)$, respectively, in the aftermarket. Therefore, firms must be indifferent between mixing and just selling to loyal consumers at the monopolistic price. We show in the proof of Lemma 2 below that $F^n(W(v_2)) = 1$ and $F^e(W(v_2, \Delta)) = 1$. This allows us to rewrite the expected profits accordingly

$$\begin{aligned}
\mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, W(v_2)) &= [p_{1,j} + (1-\rho)W(v_2)]D_j(\cdot), \\
\mathbb{E}\pi_j^e(p_{1,j}, p_{1,-j}, W(v_2, \Delta)) &= [p_{1,j} + \alpha(1-\rho)W(v_2, \Delta)]D_j(\cdot).
\end{aligned}$$

Observe that the maximization problems are very similar to the baseline model and identical when $\rho = 0$. Thus, we can proceed like in the baseline model and derive the base-good prices and profits in the three different outcomes. The results of Lemma 3 and Lemma 4 are very similar, we only need to adjust properly for the term $(1 - \rho)$. Further, the equilibrium structure is identical to Lemma 5, with the only difference that firms mix over p_2 instead of setting an add-on price with probability 1, which we will prove below. The result of Lemma 1 is unchanged and we still have $\max\{\hat{\alpha}, \bar{\alpha}\} < \bar{\alpha}_p$. Therefore, Proposition 1 follows immediately. The derivations and proofs are available on request.

C.1.1 Proof of Lemma 2

(i) A non-exploiting firm must be indifferent between mixing over p_2 and setting $p_2 = W(v_2)$. Thus, we can derive the equilibrium price distribution $F^n(p_2^n)$

$$\begin{aligned}
\mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, p_2^n) &= \mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, W(v_2)) \\
(1-\rho)D_j(\cdot)p_2^n + \rho[1 - F^n(p_2^n)][D_j(\cdot) + D_{-j}(\cdot)]p_2^n &= (1-\rho)D_j(\cdot)W(v_2) \\
F^n(p_2^n) &= 1 - \frac{(1-\rho)D_j(\cdot)[W(v_2) - p_2^n]}{\rho[D_j(\cdot) + D_{-j}(\cdot)]p_2^n}.
\end{aligned}$$

The upper bound is given by $W(v_2)$

$$F^n(W(v_2)) = 1 - \frac{(1-\rho)D_j(\cdot)[W(v_2) - W(v_2)]}{\rho[D_j(\cdot) + D_{-j}(\cdot)]W(v_2)} = 1.$$

Set $F^n(p_2^n) = 0$ to obtain the lower bound $\underline{p}_2^n$

$$\begin{aligned}
F^n(\underline{p}_2^n) &= 0 \\
(1-\rho)D_j(\cdot)[W(v_2) - \underline{p}_2^n] &= \rho[D_j(\cdot) + D_{-j}(\cdot)]\underline{p}_2^n \\
\underline{p}_2^n &= \frac{(1-\rho)D_j(\cdot)W(v_2)}{D_j(\cdot) + \rho D_{-j}(\cdot)}.
\end{aligned}$$

We can easily verify that $\mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, \underline{p}_2^n) = \mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, W(v_2))$, which implies that firms obtain the same expected profit for all prices on the equilibrium support. The price distribution $F^n(p_2^n)$ is continuous and atomless since $D(\cdot)$ is continuous, $W(v_2)$ is constant, and $\frac{\partial F^n(p_2^n)}{\partial p_2^n} > 0$. For a detailed proof see Baye and Morgan (2001).

(ii) The proof is analogous to part (i). We simply have to replace p_2^n with p_2^e and $W(v_2)$ with $W(v_2, \Delta)$. Note that an exploiting firm must set an add-on price $p_2^e > W(v_2)$. Therefore, the lower bound is given by $\max\{\underline{p}_2^e, W(v_2)\}$. It is easily verifiable that $\mathbb{E}\pi_j^e(p_{1,j}, p_{1,-j}, W(v_2)) = \mathbb{E}\pi_j^e(p_{1,j}, p_{1,-j}, W(v_2, \Delta))$ when $\underline{p}_2^e < W(v_2)$.

□

C.2 Unit Demand

We use a Hotelling model to analyze the unit demand case with classical and behavioral consumers, which are uniformly distributed on the interval $[0, 1]$. Consumers buy at most one unit of the base good with valuation v_1 at price p_1. We suppose that v_1 is sufficiently large. Two firms are located at each extreme, $l \in \{0,1\}$. They sell identical main products and add-ons, and produce at similar marginal costs c and zero, respectively. Without loss of generality, assume that firm j is located at $l=0$ and firm $-j$ at $l=1$. Buying a good imposes transportation costs t on the consumer. The rest of the setup is identical to the baseline model in Section 2, but we use an explicit WTP function $W(v_2,\Delta) = v_2\beta_i(1+p_1-p_2)$ with $\beta_i \in \{0,1\}$.

C.2.1 Aftermarket

In the last stage, after buying the base good, consumers can buy an add-on with valuation v_2 at price p_2. A classical consumer ($\beta=0$) buys the add-on when $v_2 \geq p_2$ and a behavioral consumer ($\beta=1$) buys when $\frac{v_2(1+p_1)}{1+v_2} \geq p_2$. Similar to the baseline model, firms extract the entire rent and choose $p_2^* \in \{v_2, \frac{v_2(1+p_1)}{1+v_2}\}$ in equilibrium. Therefore, the add-on demand is given by $Q_j(p_{2,j}, D_j(p_{1,j}, p_{1,-j})) = \{D_j(\cdot), \alpha D_j(\cdot)\}$.

C.2.2 Firm's Problem

The base-good demand of either firm is determined by the indifferent consumer $\bar{x}$, who is located at

$$\bar{x} = \frac{1}{2} + \frac{p_{1,-j} - p_{1,j}}{2t}.$$

The demand functions are

$$\begin{aligned} D_j(p_{1,j}, p_{1,-j}) &= \bar{x} = \frac{1}{2} + \frac{p_{1,-j} - p_{1,j}}{2t}, \\ D_{-j}(p_{1,-j}, p_{1,j}) &= 1 - \bar{x} = \frac{1}{2} + \frac{p_{1,j} - p_{1,-j}}{2t}. \end{aligned}$$

The profit function of firm j is given by

$$\pi_j(p_{1,j}, p_{1,-j}, p_{2,j}) = [p_{1,j} - c]\left[\frac{1}{2} + \frac{p_{1,-j} - p_{1,j}}{2t}\right] + Q_j(p_{2,j}, D_j(.))p_{2,j}.$$

The base-good prices and firm profits in the symmetric non-exploiting and symmetric exploiting outcome are given by

$$p_1^n = t + c - v_2$$

$$\pi^n = \pi(p_1^n, p_1^n, v_2) = \frac{t}{2}$$

$$p_1^e = t + \frac{c - \frac{\alpha v_2}{1+v_2}}{1 + \frac{\alpha v_2}{1+v_2}}$$

$$\pi^e = \pi\left(p_1^e, p_1^e, \frac{v_2(1+p_1^e)}{1+v_2}\right) = \frac{t}{2}\left(1 + \frac{\alpha v_2}{1+v_2}\right)$$

We can observe immediately that $\pi^e > \pi^n$ for all $\alpha > 0$.[43] The reason for this is the covered market assumption, which is often used in Hotelling models. However, possible asymmetric strategies enable the existence of symmetric non-exploiting equilibria.

The prices and profits in asymmetric outcomes are

$$\tilde{p}_1^n = t + \frac{2(c - v_2)}{3} + \frac{c - \frac{\alpha v_2}{1+v_2}}{3(1 + \frac{\alpha v_2}{1+v_2})}$$

$$\tilde{p}_1^e = t + \frac{c - v_2}{3} + \frac{2(c - \frac{\alpha v_2}{1+v_2})}{3(1 + \frac{\alpha v_2}{1+v_2})}$$

$$\tilde{\pi}^n = \pi\left(\tilde{p}_1^n, \tilde{p}_1^e, v_2\right) = \frac{1}{2t}\left[t + \frac{v_2(1 + v_2 + \alpha(v_2 - 1 - c))}{3(1 + v_2(1+\alpha))}\right]^2$$

$$\tilde{\pi}^e = \pi\left(\tilde{p}_1^e, \tilde{p}_1^n, \frac{v_2(1+\tilde{p}_1^e)}{1+v_2}\right) = \frac{\left[t(1 + v_2(1+\alpha)) - \frac{1}{3}v_2(1 + v_2 + \alpha(v_2 - 1 - c))\right]^2}{2t(1+v_2)(1 + v_2(1+\alpha))}.$$

Note that demands under asymmetric strategies can be negative. We focus on interior solutions and assume that $D(\tilde{p}_1^n, \tilde{p}_1^e) > 0$ and $D(\tilde{p}_1^e, \tilde{p}_1^n) > 0$.

C.3 Equilibrium

The equilibria characterization is similar to Lemma 5.[44]

Lemma 10.

(i) If $\alpha < \min\{\bar{\alpha}, \hat{\alpha}\}$, then both firms do not exploit and set $p_1^* = p_1^n$ and $p_2^* = v_2$.

(ii) If $\alpha > \max\{\bar{\alpha}, \hat{\alpha}\}$, then both firms exploit and set $p_1^* = p_1^e$ and $p_2^* = \frac{v_2(1+p_1^e)}{1+v_2}$.

(iii) If $\bar{\alpha} < \alpha < \hat{\alpha}$, then either both firms do not exploit symmetrically or both firms exploit symmetrically.

(iv) If $\hat{\alpha} < \alpha < \bar{\alpha}$, then firm j does not exploit and sets $p_1^* = \tilde{p}_1^n$ and $p_2^* = v_2$, and firm $-j$ exploits and sets $p_1^* = \tilde{p}_1^e$ and $p_2^* = \frac{v_2(1+p_1^e)}{1+v_2}$.

[43] It can be shown that the introduction of behavioral consumers does not affect the optimal location of a firm.

[44] If $D(\tilde{p}_1^n, \tilde{p}_1^e) = 0$ or $D(\tilde{p}_1^e, \tilde{p}_1^n) = 0$, only symmetric equilibria exists.

Proof. The proof is analogeous to the proof of Lemma 5. Note that $\frac{\partial \tilde{\pi}^n}{\partial \alpha} < 0$. Thus, the threshold $\bar{\alpha}$ exists. Further, we have $\pi^n > \tilde{\pi}^e$ and $\tilde{\pi}^n > \pi^e$ when $\alpha = 0$. Since $\frac{\partial \pi^n}{\partial \alpha} = 0$, $\frac{\partial \pi^e}{\partial \alpha} > 0$, $\frac{\partial \tilde{\pi}^e}{\partial \alpha} > 0$, and $\frac{\partial \tilde{\pi}^n}{\partial \alpha} < 0$, the thresholds $\hat{\alpha}$ and $\bar{\alpha}$ must be unique. □

The critical price threshold is given by $\bar{\alpha}_p = \frac{1+v_2}{1+c-v_2}$. This leads to the following result similar to 1.

Lemma 11.

(i) Suppose $1 + c > v_2$. The base good is more expensive in any symmetric exploiting or asymmetric equilibrium when the share of behavioral consumers is sufficiently low, $\alpha < \bar{\alpha}_p \Leftrightarrow p_1^* \in \{p_1^e, \tilde{p}_1^n, \tilde{p}_1^e\} > p_1^b$.

(ii) Suppose $1 + c < v_2$. The base good is always more expensive in any symmetric exploiting or asymmetric equilibrium, $p_1^* \in \{p_1^e, \tilde{p}_1^n, \tilde{p}_1^e\} > p_1^b$ for all α.

Proof. (i)

$$\begin{aligned}
\alpha &< \frac{1+v_2}{1+c-v_2} \\
c + \alpha v_2(1+c-v_2) &< c + v_2(1+v_2) \\
(c-v_2)(1+\frac{\alpha v_2}{1+v_2}) &< c - \frac{\alpha v_2}{1+v_2} \\
t + c - v_2 &< t + \frac{c - \frac{\alpha v_2}{1+v_2}}{1+\frac{\alpha v_2}{1+v_2}} \\
p_1^b &< p_1^e
\end{aligned}$$

$$\begin{aligned}
\alpha &< \frac{1+v_2}{1+c-v_2} \\
c + \alpha v_2(1+c-v_2) &< c + v_2(1+v_2) \\
(c-v_2)(1+\frac{\alpha v_2}{1+v_2}) &< c - \frac{\alpha v_2}{1+v_2} \\
\frac{1}{3}(c-v_2) &< \frac{c - \frac{\alpha v_2}{1+v_2}}{3(1+\frac{\alpha v_2}{1+v_2})} \\
t + c - v_2 &< t + \frac{2(c-v_2)}{3} + \frac{c - \frac{\alpha v_2}{1+v_2}}{3(1+\frac{\alpha v_2}{1+v_2})} \\
p_1^b &< \tilde{p}_1^n
\end{aligned}$$

$$
\begin{aligned}
\alpha &< \frac{1+v_2}{1+c-v_2} \\
2c+2\alpha v_2(1+c-v_2) &< 2c+2v_2(1+v_2) \\
2(c-v_2)(1+\frac{\alpha v_2}{1+v_2}) &< 2\left(c-\frac{\alpha v_2}{1+v_2}\right) \\
2(c-v_2) &< \frac{2(c-\frac{\alpha v_2}{1+v_2})}{3(1+\frac{\alpha v_2}{1+v_2})} \\
t+c-v_2 &< t+\frac{c-v_2}{3}+\frac{2\left(c-\frac{\alpha v_2}{1+v_2}\right)}{3(1+\frac{\alpha v_2}{1+v_2})} \\
p_1^b &< \tilde{p}_1^e
\end{aligned}
$$

(ii)

$$
\begin{aligned}
p_1^b &< p_1^e \\
t+c-v_2 &< t+\frac{c-\frac{\alpha v_2}{1+v_2}}{1+\frac{\alpha v_2}{1+v_2}} \\
\alpha(1+c-v_2) &< 1+v_2 \\
\alpha > 0 &> \frac{1+v_2}{1+c-v_2}
\end{aligned}
$$

Since $1+c-v_2<0$, the direction of inequality reverses when dividing. The proof for $p_1^b < \tilde{p}_1^n$ and $p_1^b < \tilde{p}_1^e$ when $1+c-v_2<0$ is analogeous.

□

Similar to Lemma 8, the price threshold is always larger than the profit thresholds when $1+c-v_2>0$.

Lemma 12. *Suppose* $1+c>v_2$. *Then* $\max\{\hat{\alpha},\tilde{\alpha}\}<\bar{\alpha}_p$.

Proof. Suppose $\alpha=\bar{\alpha}_p=\frac{1+v_2}{1+c-v_2}$. Then

$$
\begin{aligned}
\tilde{\pi}^e &> \pi^n \\
\frac{t(1+c)}{2(1+c-z)} &> \frac{t}{2} \\
0 &> -\frac{tv_2}{2}
\end{aligned}
$$

By Lemma 10, it must be $\alpha > \hat{\alpha}$ when $\pi^n < \tilde{\pi}^e$. Thus, $\bar{\alpha}_p > \hat{\alpha}$.

Further, when $\alpha = \bar{\alpha}_p = \frac{1+v_2}{1+c-v_2}$, then

$$\pi^e > \tilde{\pi}^n$$
$$\frac{t(1+c)}{2(1+c-z)} > \frac{t}{2}$$
$$0 > -\frac{tv_2}{2}$$

Note that $\pi^n = \tilde{\pi}^n$ and $\pi^e = \tilde{\pi}^e$ when $\alpha = \bar{\alpha}_p = \frac{1+v_2}{1+c-v_2}$. By Lemma 10, it must be $\alpha > \bar{\alpha}$ when $\tilde{\pi}^n < \pi^e$. Thus, $\bar{\alpha}_p > \bar{\alpha}$. Hence, $\max\{\hat{\alpha}, \tilde{\alpha}\} < \bar{\alpha}_p$. □

Given the results of Lemma 10, Lemma 11 and Lemma 12, Proposition 7 follows immediately, which analogeous to the main finding in the baseline model stated by Proposition 1.

Proposition 7 (Unit demand)**.**

(a) Suppose $1 + c > v_2$. Then the presence of behavioral consumers: (i) harms classical consumers in any symmetric exploiting equilibrium if $\alpha < \bar{\alpha}_p$ and benefits otherwise, (ii) harms classical consumers in any asymmetric equilibrium, (iii) has no effect in any symmetric non-exploiting equilibrium.

(b) Suppose $1 + c < v_2$. Then the presence of behavioral consumers: (i) harms classical consumers in any symmetric exploiting or asymmetric equilibrium, (ii) has no effect in any symmetric non-exploiting equilibrium.

Proof. The proof is analogous to the proof of Proposition 1, where $1 + c > v_2$ corresponds to the case of $\epsilon_D > \epsilon_W$ and $1 + c < v_2$ corresponds to $\epsilon_D < \epsilon_W$. □

C.4 Cheaper Base Good than Add-on

In the baseline model, we focused on the case that the behavioral mechanism affects the add-on WTP positively by restricting Δ to be positive. Let us now consider the opposite when the add-on is more expensive than the base good. Then, the behavioral mechanism decreases the add-on WTP. We suppose the same setup as in the baseline model but allow $\Delta = \beta_i(p_1 - p_2)$ to be negative and focus on the case of cheap base goods and expensive add-ons such that $\Delta < 0$ in any equilibrium. Further, for simplicity, we consider monopolistic base good markets.[45] Crucially, behavioral consumers now have a lower WTP for the add-on, which has several implications. The add-on demand is given by

$$Q(p_2, D(p_1)) = \begin{cases} D(p_1) & \text{if } p_2 \leq W(v_2, \Delta), \\ (1-\alpha)D(p_1) & \text{if } W(v_2, \Delta) < p_2 \leq W(v_2), \\ 0 & \text{if } p_2 > W(v_2). \end{cases}$$

Contrary to the baseline model, all consumers buy the add-on if it is priced at the WTP of behavioral consumers, while only the fraction $(1-\alpha)$ accepts the add-on offer when $p_2 = W(v_2)$. The profit function, adjusted for the monopoly case, is still given by Equation (1) and the firm chooses between the non-exploiting strategy ($p_2^* = W(v_2)$) and the exploiting strategy ($p_2^* = W(v_2, \Delta)$). Due to the negative behavioral effect, exploiting implies now lowering the add-on price below the WTP of classical consumers and selling the add-on to all. The non-exploiting profit $\pi^n = \pi(p_1^n(\alpha), W(v_2))$ is strictly decreasing in α, while the exploiting profit $\pi^e = \pi(p_1^e, W(v_2, \Delta))$ is independent of the share of behavioral consumers. Thus, we can define the profit threshold $\alpha = \hat{\alpha} \Leftrightarrow \pi^n = \pi^e$. Similarly to Lemma 5, for a share below the threshold, the monopolist does not exploit behavioral consumers and sets $p_2^* = W(v_2)$. There are only a few behavioral consumers that do not buy the add-on. When α is sufficiently large, the monopolist selects the exploiting strategy as the missed revenue in the aftermarket would be too high otherwise.

Interestingly, none of the results with $\Delta < 0$ depend on the semi-elasticities ϵ_D and ϵ_W. The optimal non-exploiting price $p_1^n(\alpha)$ is strictly increasing in α. The optimal base good price when all consumers purchase the add-on (p_1^e) is, like in the baseline model, independent of the share of behavioral consumers. Further, the outcome of the benchmark economy with $\alpha = 0$ is now different from both, the exploiting and non-exploiting equilibrium. The base good is always the cheapest in the benchmark, $p_1^b < \min\{p_1^n(\alpha), p_1^e\}$. The simple reason for this is that firms in after-sales markets redistribute add-on earnings to lower the base good price to attract more consumers.[46]

In the benchmark economy, all consumers purchase the add-on at the price $W(v_2) > W(v_2, \Delta)$, which clearly yields higher add-on profits than in any (non-)exploiting equilib-

[45] The analysis can easily be extended to the two firms' case. The differences in results between monopolistic and imperfect competition are similar to the baseline model with $\Delta > 0$.

[46] For this reason, $p_1^n(\alpha)$ is increasing in α as the add-on earnings decline with more behavioral consumers in the population.

rium.[47] This implies that even a few behavioral consumers already affect the economy. Because not everyone buys the add-on in the non-exploiting equilibrium, the base good becomes more expensive consequently. Hence, by not accepting the additional offer, behavioral consumers indirectly increase the base good price. In the exploiting equilibrium, all consumers purchase the add-on but at a lower price than in the benchmark economy. Therefore, in contrast to Proposition 1, the presence of behavioral consumers always affects classical consumers: When behavioral consumers with $\Delta < 0$ are present, in any equilibrium, the base good is more expensive than in the benchmark case.

Proposition 8 (Cheaper base good than add-on)**.** *Suppose* $\Delta < 0$*.*

(i) The presence of behavioral consumers harms a classical consumer in any non-exploiting equilibrium for all $\alpha > 0$*.*

(ii) If $p_1^e - p_1^b > W(v_2) - W(v_2, \Delta)$*, then the presence of behavioral consumers harms a classical consumer in any exploiting equilibrium. Otherwise, a classical consumer benefits.*

Importantly, classical consumers are harmed when the monopolist *does not exploit* behavioral consumers. In any non-exploiting equilibrium, classical consumers pay the same for the add-on as in the benchmark economy but strictly more for the base good when $\alpha > 0$. Thus, they are clearly worse off. The impact in an exploiting equilibrium is ambiguous. Compared to the benchmark, classical consumers have to pay more for the base good but less for the add-on. Which effect dominates determines whether classical consumers benefit or are harmed by the presence of behavioral consumers.

C.5 Proofs

In the setup with $\Delta < 0$, the exploiting strategy still implies to price the add-on at the WTP of behavioral consumers, $p_2 = W(v_2, \Delta)$. However, this is now below the WTP of classical consumers. Therefore, exploiting implies to sell the add-on to all customers and non-exploiting implies to sell the add-on only to classical consumers. The monopolist's

[47] When $\alpha = 0$, the benchmark outcome and non-exploiting equilibrium coincide.

maximization problems given a chosen strategy are

$$\max_{p_1} \pi^n(p_1, W(v_2)) = \max_{p_1}[p_1 + (1-\alpha)W(v_2)]D(p_1),$$
$$\max_{p_1} \pi^e(p_1, W(v_2,\Delta)) = \max_{p_1}[p_1 + W(v_2,\Delta)]D(p_1).$$

Maximizing each expression with respect to p_1 yields the prices and profits given the monopolist exploits or not

$$\begin{aligned}
p_1^n(\alpha) &= \frac{-D(p_1^n(\alpha))}{D'(p_1^n(\alpha))} - (1-\alpha)W(v_2),\\
\pi^n(p_1^n(\alpha), W(v_2)) &= \frac{-D(p_1^n(\alpha))^2}{D'(p_1^n(\alpha))},\\
p_1^e &= \frac{-[1+W'(v_2,\Delta)]D(p_1^e)}{D'(p_1^e)} - W(v_2,\Delta),\\
\pi^e(p_1^e, W(v_2,\Delta)) &= \frac{-[1+W'(v_2,\Delta)]D(p_1^e)^2}{D'(p_1^e)}.
\end{aligned}$$

In contrary to the baseline model, the non-exploiting base-good price and profit depend on α, while the exploiting base-good price and profit are independent of α.

Lemma 13.

(a) $p_1^n(\alpha)$ is strictly increasing in α and π^n is strictly decreasing in α.

(b) p_1^e and π^e are constant in α.

Proof. (a)

$$\begin{aligned}
\frac{\partial p_1^n(\alpha)}{\partial\alpha} &= \frac{W(v_2)}{2 - \frac{D(p_1^n(\alpha))D''(p_1^n(\alpha))}{D'(p_1^n(\alpha))^2}} > 0\\
\frac{\partial \pi^n}{\partial\alpha} &= -D(p_1^n(\alpha))\left[2 - \frac{D(p_1^n(\alpha))D''(p_1^n(\alpha))}{D'(p_1^n(\alpha))^2}\right]\frac{\partial p_1^n(\alpha)}{\partial\alpha} = -D(p_1^n(\alpha))W(v_2) < 0.
\end{aligned}$$

(b) Taking $\frac{\partial p_1^e}{\partial\alpha}$ and rearranging yields

$$\begin{aligned}
\frac{\partial p_1^e}{\partial\alpha}\left[(1+W'(v_2,\Delta))\left(2 - \frac{D(p_1^e)D''(p_1^e)}{D'(p_1^e)^2}\right) + \frac{W''(v_2,\Delta)D(p_1^e)}{D'(p_1^e)}\right] &= 0,\\
\frac{\partial\pi^e}{\partial\alpha} = -(1+W'(v_2,\Delta))D(p_1^e)\left[2 - \frac{D(p_1^e)D''(p_1^e)}{D'(p_1^e)^2}\right]\underbrace{\frac{\partial p_1^e}{\partial\alpha}}_{=0} &= 0.
\end{aligned}$$

□

A crucial difference to the baseline model is that $p_1^n(\alpha)$ is strictly increasing in α and does not depend on the semi-elasticities. In fact, ϵ_D and ϵ_W are not important when $\Delta < 0$. Since π^n is strictly decreasing and π^e is constant in α, we can define the profit threshold $\hat{\alpha}$ and characterize the equilibria.

Lemma 14.

(i) If $\alpha < \hat{\alpha}$, then the monopolist does not exploit and sets $p_1^* = p_1^n$ and $p_2^* = W(v_2)$.

(ii) If $\alpha > \hat{\alpha}$, then the monopolist exploits and sets $p_1^* = p_1^e(\alpha)$ and $p_2^* = W(v_2, \Delta)$.

Proof. The proof is analogous to the proof to Lemma 5. □

The equilibrium structure is identical to the baseline model. Non-exploiting is optimal when α is low. Otherwise, exploiting is optimal. This is because when the share of behavioral consumers sufficiently large, then the loss of not selling the add-on to them is too high.

In the benchmark economy, the monopolist prices the add-on at $W(v_2)$ and sells it to all customers. Therefore, no equilibria with behavioral consumers is identical to the benchmark outcome. In the non-exploiting equilibrium, the add-on demand is strictly lower, and in the exploiting equilibrium, the add-on price $p_2 = W(v_2, \Delta) < W(v_2)$ is strictly smaller. This is a second major difference to the baseline model with $\Delta > 0$. This implies that the base good is always the cheapest in the benchmark, $p_1^b = \frac{-D(p_1^b)}{D'(p_1^b)} - W(v_2) < \min\{p_1^n(\alpha), p_1^e\}$.

Lemma 15. $p_1^b < \min\{p_1^n(\alpha), p_1^e\}$ *for all* $\alpha > 0$.

Proof. First, observe that $p_1^b = p_1^n(\alpha) \Leftrightarrow 0 = \alpha W(v_2)$ is feasible only when $\alpha = 0$. By Lemma 13, $p_1^n(\alpha)$ is strictly increasing in α. Hence, since p_1^b is independent of α, it must follow that $p_1^b < p_1^n(\alpha)$ when $\alpha > 0$.

We prove $p_1^b < p_1^e$ in several steps. First, observe that $p_1^b \neq p_1^e$ for all $\alpha \in \mathbb{R}$ because

$$\begin{aligned}
& p_1^b = p_1^e \\
\Leftrightarrow \quad & -W(v_2) = \frac{-W'(v_2, \Delta) D(p_1^e)}{D'(p_1^e)} - W(v_2, \Delta) \\
\Leftrightarrow \quad & \frac{W'(v_2, \Delta) D(p_1^e)}{D'(p_1^e)} = W(v_2) - W(v_2, \Delta),
\end{aligned}$$

which is a contradiction since $\frac{W'(v_2,\Delta)D(p_1^b)}{D'(p_1^b)} < 0$ and $W(v_2) - W(v_2, \Delta) > 0$. Hence, since p_1^b and p_1^e are both constant in α, it must be either $p_1^b < p_1^e\ \forall\alpha$ or $p_1^b > p_1^e\ \forall\alpha$.

Next, observe that $p_1^e = p_1^n(\alpha)$ when

$$\alpha = \bar{\alpha}_p = 1 - \underbrace{\frac{W'(v_2,\Delta)D(p_1^e)}{D'(p_1^e)W(v_2)}}_{<0} - \underbrace{\frac{W(v_2,\Delta)}{W(v_2)}}_{<1} > 0.$$

Since $p_1^b < p_1^n(\alpha)$ for all $\alpha > 0$ and $p_1^e = p_1^n(\alpha)$ when $\alpha = \bar{\alpha}_p > 0$, it follows that $p_1^b < p_1^e$ for $\alpha \geq \bar{\alpha}_p$. But since p_1^b and p_1^e are both constant in α, it must be $p_1^b < p_1^e$ for any α. $\square$

Now can we prove the statements in Proposition 8.

C.5.1 Proof of Proposition 8

The proof follows closely the proof of Proposition 1.

(i) By Lemma 14 and Lemma 15, we have $p_1^b < p_1^n(\alpha)$ and $p_2 = W(v_2)$ in any non-exploiting equilibrium. Hence, classical consumers pay the same as in the benchmark economy for the add-on, but strictly more for the base good, which reduces a classical consumer's surplus compared to the benchmark. Thus, classical consumers are harmed by the presence of behavioral consumers.

(ii) By Lemma 14 and Lemma 15, we have $p_1^b < p_1^e$ and $p_2 = W(v_2, \Delta) < W(v_2)$ in any exploiting equilibrium. Hence, compared to the benchmark, classical consumers pay strictly less ($W(v_2) - W(v_2, \Delta) > 0$) for the add-on and strictly more for the base good ($p_1^b - p_1^e < 0$). If $p_1^e - p_1^b > W(v_2) - W(v_2, \Delta)$, the negative effect dominates, which reduces a classical consumer's surplus compared to the benchmark. Thus, classical consumers are harmed by the presence of behavioral consumers. Otherwise, if $p_1^e - p_1^b < W(v_2) - W(v_2, \Delta)$, the positive effect dominates, which increases a classical consumer's surplus compared to the benchmark. Thus, classical consumers benefit by the presence of behavioral consumers.