Competition for Context-Sensitive Consumers

Apffelstaedt, Arno and Mechtenberg, Lydia

University of Cologne, University of Hamburg

15 November 2016

Competition for Context-Sensitive Consumers*

Arno Apffelstaedt† Lydia Mechtenberg‡

This Version: November, 2016

Abstract: When preferences are sensitive to context, firms may try to influence purchase decisions by designing the environment of consumption choices. Confirming anecdotal evidence on retailer marketing tricks, we show that competitive retailers exploit context-sensitivity by designing environments that drive a wedge between preferences before and after entering a store. This wedge induces naïve consumers to switch preference from a loss-leader product that attracts consumers into the store to a more profitable product. Depending on the quality preferences and budget of consumers, markets are either in an "up-selling" or "down-selling" equilibrium. In the former, firms attract consumers with low-quality products, compete on prices, and design context to ultimately sell a product of higher price. In the latter, firms attract consumers with high-price products, compete on quality, and design context to ultimately sell a product of lower quality. When modeling context-sensitivity according to the theories of *Salience*, *Focusing*, or *Relative Thinking*, designing context comes down to the introduction of a single decoy product. This decoy draws consumer attention at the store to the favorable attributes of the product the firm aims to sell.

Keywords: Choice Context, Retailer Competition, Up-Selling, Down-Selling, Decoys, Naïveté

JEL Codes: D03, D11, D41

---

*We thank Leonie Baumann, Pedro Bordalo, Tom Cunningham, Markus Dertwinkel-Kalt, Nicola Gennaioli, Michael D. Grubb, Francesco Nava, Joshua Schwartzstein, Adam Szeidl, and, especially, Paul Heidhues, Botond Kőszegi, and Andrei Shleifer, for very insightful discussions and suggestions. We also thank audiences at several conferences, seminars and workshops for valuable comments.

†University of Cologne. Email: apffelstaedt@wiso.uni-koeln.de

‡University of Hamburg. Email: lydia.mechtenberg@uni-hamburg.de

I. Introduction

A search for marketing tricks in blogs and online newspapers yields an abounding number of tipps (to firms) and warnings (to consumers) about consumer manipulation. A recurring story is about the "anchoring decoy": When US kitchenware retailer Williams-Sonoma saw their breadmaker at \$279 flopping, they offered, next to it, a larger model priced at \$429, and sales on the \$279 model picked up. Since then, putting an over-priced decoy next to a product with similar features to make customers believe the latter to be a bargain has become a (reportedly) widely used marketing strategy.[1] Similar strategies make use of what is sometimes labelled the "middling effect" (customers shy away from extreme prices and go for the middle-priced option) or the "everyone-loves-a-deal effect" (add a crossed-out price on the price tag that exceeds the offer price and make the product look as if on sale, i.e., a "bargain"). Taking a more general viewpoint, the marketing literature speaks of "atmospheric effects" when it reports how the grouping of merchandise, color schemes, lighting, price displays, music, scents and other variables at the point of purchase affect shopping behavior.[2] Although brandished under a variety of labels, such marketing tricks boil down to the usage of one habit of mind: to assess the properties of products relative to the immediate environment in which they are presented. In short, context matters.

At first sight, such marketing tricks look intuitive. However, even accepting that purchase decisions are affected by context, it is far from straightforward that such manipulations survive competition. We show that such tricks play a role even in perfectly competitive retail markets because consumers typically need to enter the store of a particular firm before being able to buy a product: An internet shopper, for instance, can browse and compare the product-lines of different online stores using price comparison websites such as PriceGrabber and Google Shopping, but ultimately has to enter a store—say, amazon.com—to purchase a product. This structure gives individual retailers (even perfectly competitive ones) considerable power over the design of the final choice environment. If consumers are context-sensitive and underestimate the degree to which their preferences are malleable, firms will use this power to establish choice environments that induce consumers to purchase more profitable products than they intended to buy when entering the store.

To show formally how retail firms use context to manipulate consumer choice we develop a model that introduces context-sensitivity into a standard competitive retail market. In our model, there is only one good, but this good can be differentiated in quality and price.

[1] This story occurs, for example, on https://www.quora.com/What-psychological-tricks-do-retailers-use-to-get-people-to-spend-more-money, accessed 20-09-2016.

[2] For a review of experimental marketing research on the effects of "atmosheric variables" on shopping behavior, see, for example, Turley and Milliman (2000).

Market supply of the good is competitive: There is a large number of (ex-ante symmetric) firms that each can offer any menu of quality-price combinations—a *product-line*—at a firm-specific store (for example, a website or a supermarket). When making a purchase decision, consumers first compare the product-lines on offer and then enter the store of one firm to purchase a product. We introduce context-sensitivity by assuming that the choice environment at any given store can lead the consumer to overvalue by a factor β either the quality of a product (leading to the distorted value function $\beta q_j - p_j$) or its price (leading to the distorted value function $q_j - \beta p_j$).

We show that when a firm has the power to manipulate the choice-environment of its store and consumers underestimate the extent to which their preferences are malleable, the firm will attract consumers with a competitive *attraction product* but design a store-context that induces them to switch to a more profitable *target*. Such strategies are profitable at the margin because they allow firms to compete for consumers on a set of preferences that is different from the preferences at the moment of purchase. Our model thereby predicts the widespread marketing practice of using variations in the presentation of options to "up-sell" or "down-sell" a customer, i.e., making the customer purchase a product of higher price (up-sale) or lower cost (down-sale). While the incentives to manipulate context are strong, fierce competition for consumer entry leads firms to compete away the extra profits earned on such manipulation. As a result, the attraction product becomes a so-called "loss leader", i.e., a product priced below marginal costs.

To make more explicit predictions with regard to *how* firms manipulate the choice environment, we embed and compare three recently suggested theories of choice-set dependent preferences: *Salience* (Bordalo, Gennaioli and Shleifer, 2013, henceforth BGS), *Focusing* (Kőszegi and Szeidl, 2013, henceforth KS), and *Relative Thinking* (Bushong, Rabin and Schwartzstein, 2016, henceforth BRS). These theories have in common that they make preferences at the store dependent on whether quality or price of a product stands out in contrast to the firm's entire product-line. While the three frameworks are very different in their particular definition of how the choice-set affects consumer attention and thus preferences, we show that all of them allow firms to profitably manipulate preferences by inflating the product-line with unattractive products. The characteristics of these products mirror the construction of so-called *decoys* in classical experiments on context-effects (see e.g., Huber, Payne and Puto, 1982; Simonson, 1989). Under all three specifications, firms in equilibrium use product-lines that include exactly one such decoy. This decoy is necessary and sufficient to shift preferences at the store from the attraction product to the target.

We also show that sophisticated agents and agents who overestimate their response to contextual clues will not be up- or down-sold in equilibrium. Firms react to these consumers as if their preferences were stable and undistorted. Our predictions do not change when consumers who are prone to profitable deception and those who are not co-exist in the population. In equilibrium, the market perfectly separates the former from the latter: While consumers who underestimate the malleability of their preferences will always be deceived by firms that use context to up- or down-sell, there also exist firms that offer undistorted product-lines where types who "see through" the up- and down-selling regimes of deceptive firms can purchase a product that is aligned with their outside preference.

Related Literature. Theoretical contributions dealing with the question of how firms react to context-sensitivity in market settings are rare. Kamenica (2008) considers a monopolistic seller and shows that consumers with choice-set dependent preferences need not necessarily exhibit a behavioral "mistake". In particular, if a consumer is imperfectly informed about the quality of a product, then it can be rational to update her valuation once she faces the product-line. This is because the popularity of products can reveal something about their quality (given that there are also better informed consumers in the market). Kamenica (2008) shows that, given that there is also uncertainty about the production cost, a monopolist may be able to "manipulate" the quality-perception of rational, uninformed consumers by adding decoy products to the product-line. While this is an important result that sheds new light on the importance of consumer inference, it is definitely not the end of the story. Context-effects have been found in experimental settings with no explanatory room for inference, see, e.g., Herne (1999), Ariely, Loewenstein and Prelec (2003), Mazar, Kőszegi and Ariely (2014) and Jahedi (2011). Moreover, the conjecture that context-sensitive shopping behavior is largely irrational seems corroborated by the extensive on-line discussion of marketing techniques that all seem to "manipulate" or "trick" consumers into purchase decisions. Thus, while it is plausible that context-sensitivity *evolutionary originates* from rational inference—representing a heuristic that is often true and time-efficient (see also Wolfe and Horowitz, 2004)—, it is even more prevalent in situations where it cannot be rationalized. We therefore believe that a *behavioral* model of firms using context to exploit naïve consumers is required to round out the picture of how context-sensitivity affects market outcomes. Our paper provides one such model which is, in addition, sufficiently general to nest three popular ways of modeling context-sensitivity.

The paper closest to ours is "Competition for Attention" by Bordalo, Gennaioli and Shleifer (2016). In this paper, the authors assume that consumers have context-dependent preferences according to the Salience theory and study the price-quality competition of *single-product* duopolists on a "direct" market, i.e., where consumers buy products *without entering* the store of a particular firm. The main results of our paper, on the other hand, stem from the possibility of firms to offer *multiple products* and the observation that preferences may change *after entering* the store of a particular firm. We consider the Salience theory, but also other approaches to modeling context-sensitivity. An approach to context that is structurally different to ours is followed by Piccione and Spiegler (2012) and Spiegler (2014). In these papers the authors assume that the way how firms describe their products (*when* attracting consumers) can affect the ability of consumers to compare the products of different firms. We are interested, on the other hand, in how firms present their products at the moment of purchase (i.e., *after* consumers have been attracted). In our framework, consumers are always able to compare the products of all firms, but this comparison may be defective when—due to store-context—consumers have wrong expectations about their preferences over the product-lines of some stores.

There are other, less immediate relations to the existing literature. Due to its 2-period choice structure with preference inconsistencies our model is (from a technical standpoint) related to models of contracting with time-inconsistent consumers (see, e.g., DellaVigna and Malmendier, 2004, 2006; Gul and Pesendorfer, 2001; Heidhues and Kőszegi, 2010; Gottlieb, 2008; Eliaz and Spiegler, 2006). Some of our findings parallel theirs. For example, we find that (context-induced) time-inconsistency influences market outcomes if and only if consumers *underestimate* the effect of this bias on their future choices. This mirrors results of Heidhues and Kőszegi (2010). Our result that competitive firms attract naïve, time-inconsistent consumers with loss-leader products mirrors results by DellaVigna and Malmendier (2004, 2006). Importantly, while this literature considers an exogeneous source of time-inconsistency, it arises endogenously in our model from the equilibrium decision of firms to manipulate context at the point of purchase. It is interesting to see that the common retail market structure allows for a similar exploitation of incorrect consumer expectations as a contract environment.

Organization of the paper. The remainder of the paper is organized as follows. We introduce the formal model in the next section. In section III. we first derive the rational benchmark and then carve out the major impact of assuming context-sensitivity in retail markets, which is the possibility of firms to "fool" (i.e., up- or down-sell) naïve consumers. We also show in this section how the feasibility and profitability of such strategies depends on the naïveté of consumers and the type of choice-environment that the firm selects. Section IV. addresses the questions of *what* environment firms will construct in equilibrium and *how* such context is constructed when it is a function of the choice-set, as suggested by the theories of Salience (BGS), Focusing (KS) and Relative Thinking (BRS). Section V. concludes and relates our results to anecdotal and empirical evidence from actual retail markets. All proofs are in the appendix.

II. The Model

II.A. The Market

A unit mass of consumers has demand for a good that can be differentiated in quality $q \in \mathbb{R}$ and price $p \in \mathbb{R}$, where quality and price are both measured in dollars. There is a minimum quality $\underline{q} > 0$ and a maximum price $b > 0$ agents are willing to accept. Each consumer demands one good. There is a large number K of firms in the market. Each firm k owns a store. To purchase from firm k, a consumer has to enter its store. At the store, the firm can offer any menu of products J^k. Each product $j \in J^k$ implements the good at some level of quality $q_j \in \mathbb{R}$ and price $p_j \in \mathbb{R}$. The set $M^k = ((q_j, p_j))_{j \in J^K}$ is called the *product-line* of firm k. Each firm k also chooses *how* to present its product-line to consumers who enter its store. This choice is represented by the variable Θ^k, which we call the *store-context* of firm k. (We define Θ^k in detail further below). Instead of entering a store and purchasing a product, consumers can select an outside option of no purchase. The sequence of events is as follows.

- Firms simultaneously commit to a product-line M^k and a store-context Θ^k.
- Each consumer then moves in two stages:
  - Stage 1: The consumer observes the product-lines M^k of all firms and then decides to enter one store to make a purchase *or* to exercise the outside option and leave the market without purchase.
  - Stage 2: If the consumer has entered store k, she selects a product $j \in J^k$ in context Θ^k.

II.B. Context-Sensitive Consumers

When evaluating products *outside* of a firm-specific store, consumers value a product of given quality and price at all firms equally. Without loss of generality (henceforth w.l.o.g.), let this "global" surplus function be given by

$$u_j = q_j - p_j. \tag{1}$$

We assume (w.l.o.g.) that the outside option of no purchase yields surplus $u_0 = 0$. *Inside* a firm-specific store, the "local" valuation of firm k's products may differ from Equation (1) due to the consumer now being exposed to the local context of the store: Let Θ^k be a vector that has as many entries as the firm has products in the product-line (i.e., $|\Theta^k| = |J^k|$).

Element $\theta_j^k \in \Theta^k$ identifies the effect of local context at store k—for instance, the color of price-tags or the relative position of product j in the product-line—on the valuation of product j. We assume, in particular, that store-context can either increase the perceived quality ($\theta_j^k = Q$) or the perceived price ($\theta_j^k = P$) of a product, thereby leading to an inflation or deflation of product value relative to Equation (1). If the local context at store k has no influence on the valuation of product j, we write $\theta_j^k = N$. With a given context Θ^k, consumers then value products $j \in J^k$ inside store k with the surplus function

$$\hat{u}_j^k = \begin{cases} q_j - p_j & \text{if } \theta_j^k = N, \\ \beta q_j - p_j & \text{if } \theta_j^k = Q, \\ q_j - \beta p_j & \text{if } \theta_j^k = P, \end{cases} \tag{2}$$

where $\beta \geq 1$ measures the size of contextual distortions; the possibility of $\beta = 1$ nests the rational model.[3]

When making their entry decision in stage 1, consumers observe the product-lines of all stores and form an expectation about their purchase in stage 2. This expectation depends, of course, on the consumer's awareness of possible preference distortions at the store. We allow for different types. A perfectly *sophisticated* type knows Θ^k and β, and will therefore always predict her behavior correctly. On the other end is a perfectly *naïve* type who is either (completely) unaware of context-effects or (falsely) believes that her valuations are consistent across different contexts. We capture these two extremes as well as their convex hull by assuming that all consumers are aware of the environment at firm k (and thus of Θ^k), but are heterogeneous regarding their belief about the size of parameter β. Specifically, each consumer has a point-belief $E(\beta) = \tilde{\beta}$ and predicts herself to value products inside of store k with the surplus function

$$E_{\tilde{\beta}}\left[\hat{u}_j^k\right] = \hat{u}_j^k|_{\beta=\tilde{\beta}}. \tag{3}$$

The distribution of types in the population is $F(\tilde{\beta})$, with density $f(\tilde{\beta})$. We assume $f(\tilde{\beta}) = 0$ for any $\tilde{\beta} \leq 1$, which implies that agents may mispredict the size of contextual distortions, but not the direction. The lower bound $\tilde{\beta} = 1$ identifies the perfectly naïve type. Note that any type $\tilde{\beta} \neq \beta$ is *partially* naïve. We subdivide naïves into *optimists* ($\tilde{\beta} < \beta$), who *under*estimate the effect of context on their choice and *pessimists* ($\tilde{\beta} > \beta$), who *over*estimate

[3] Note, importantly, that we do *not* claim that the outside assessment of products is free of distortions. The crucial element of our model is not the particular form of Equation (1), but that—once that consumers have entered the store—preferences may change *relative* to this outside assessment. Our results go through for any limitation of "local" context effects to small values, i.e., for any β arbitrarily close to 1.

this effect. This categorization will play an important effect for market supply in equilibrium.

II.C. Firms

Each firm maximizes its profit π^k by choosing a product-line and a context for its store. For large parts of the paper it will be sufficient to consider a reduced form model, where the firm chooses the distortion Θ^k (i.e., whether a product is quality- or price-inflated inside the store) directly. This allows us to capture the effect of environmental variables on consumer choice in a very general manner. When solving the model, we first consider the direct choice of Θ^k under different technological restrictions (section IV.A.) and then consider an extended model where context Θ^k is a function of the product-line M^k (section IV.B.).

Throughout the paper, firms simultaneously commit to a (finite) menu of products M^k (with $(q_j, p_j) \in \mathbb{R}^2$) and a context Θ^k before consumers move. We assume that firms—e.g., through historical observations of consumer choice—have perfect knowledge of the context-sensitivity parameter β and of the distribution of beliefs $F(\tilde{\beta})$, but cannot observe the type of individual consumers. Firms have symmetric cost-functions. When a consumer purchases a good of quality q from firm k, the firm incurs a cost $c(q)$ that we assume is strictly convex increasing in the quality delivered, $c'(q) > 0$, $c''(q) > 0$, and satisfies $c(0) = c'(0) = 0$. These standard (Inada) conditions imply that for a given (context-dependent) surplus function (see Equation (2)) there exists a unique, strictly positive quality q_j^c that is cost-efficient. In particular,

$$q_j^c = \begin{cases} q^* := \arg\max_q [q - c(q)] & \Leftrightarrow c'(q^c) = 1 \quad \text{if } \theta_j^k = N, \\ q^Q := \arg\max_q [\beta q - c(q)] & \Leftrightarrow c'(q^c) = \beta \quad \text{if } \theta_j^k = Q, \\ q^P := \arg\max_q [q - \beta c(q)] & \Leftrightarrow c'(q^c) = \dfrac{1}{\beta} \quad \text{if } \theta_j^k = P. \end{cases}$$

Note that $q^Q > q^* > q^P > 0$. We concentrate on interior results by demanding that minimum quality $\underline{q}$ is sufficiently low and maximum willingness to pay b sufficiently high that consumers do not *per-se* reject buying cost-efficient quality q_j^c at cost. This is true with any context-effect $\theta_j^k \in \{N, Q, P\}$ iff $\underline{q} \leq q^P$ and $b \geq c(q^Q)$, which we assume henceforth. Finally, we assume that any distortion of context (choosing a store context other than $\theta_j^k = N$ $\forall j \in J^k$) entails a positive but infinitely small cost as does the inclusion of one additional product in the product-line.[4]

[4]This assumption sustains the results of an analysis without set-up costs but requires that firms distort context or add products only if this has non-zero marginal effect on profits. This helps us to define equilibria that pin down store-context and the size of product-lines exactly.

II.D. Solution Concept

We analyze market supply in the competitive equilibrium, where the latter is defined as as a tuple $(\boldsymbol{M}, \boldsymbol{\Theta})$, $\boldsymbol{M} := (M^k)_{k=1,...,K}$, $\boldsymbol{\Theta} := (\Theta^k)_{k=1,...,K}$, with the following properties:

1. (Nash Equilibrium) Firms play mutual best responses. For every $k \in \{1, ..., K\}$, $\pi^k((M^k, \Theta^k), (\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})) \geq \pi^k((M^{k'}, \Theta^{k'}), (\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})\ \forall (M^{k'}, \Theta^{k'}) \neq (M^k, \Theta^k)$.

2. (Competitive Market) For every $k \in \{1, ..., K\}$, $\pi^k((M^k, \Theta^k), (\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})) = 0$.

To resolve possible tie-breaks, we make two assumptions. First, whenever indifferent, a consumer chooses each surplus-maximizing option with positive probability. Second, we assume that there exists a smallest monetary unit $\delta > 0$, which we take to be positive but infinitesimally small.[5] This is equivalent to assuming that a firm, when best-responding, can resolve tie-breaks in favor of the strictly more profitable product. We will exploit this equivalence when solving the model.

III. Setting the Stage

In this section, we first provide a rational benchmark against which to compare our later results by solving the model for the case of $\beta = 1$. We then work-out the main impact of assuming context-sensitivity ($\beta > 1$) on the store-level: Firms can use the design of a store to sell naïve consumers a product that is different to the product they were attracted with. We show that such a "fooling strategy" is only profitable when attracting naïves who prior to entering the store underestimate the effect of store-context on their choice (i.e., optimists). We also detail what type of distortion Θ^k (quality- vs. price-inflating) is profitable in such a case (Lemma 3). When no optimistic consumers exists who can be profitably fooled, the market outcome does not change from the rational benchmark (Proposition 1).

III.A. Rational Benchmark

When consumers are not sensitive to store context, our set-up yields a standard Bertrand outcome. Assume that $\beta = 1$. This is equivalent to assuming that valuations are undistored ($\theta_j^k = N$ for all j and k): Preferences are stable and for both stages (outside and inside stores) uniquely defined by Equation (1). Neither the two-step choice of consumers nor potential naïveté is relevant in such a case because every consumer perfectly predicts her

[5] Formally, let $\delta = \frac{1}{10^z}$ where $z \in \mathbb{Z}$ is an integer. Firms then choose qualities and prices from a discretized set of real numbers $R_z = \{r \in \mathbb{R} | (r \cdot 10^z) \in \mathbb{Z}\}$. In the limit $z \to \infty$ (i.e., $\delta \to 0^+$) this set is equal to $\mathbb{R}$.

behavior in stage 2: The choice of a store is equivalent with the choice of a final product. A market so defined generates standard Bertrand incentives: Firms compete by offering a single product with cost-efficient quality q^* at a price that marginally undercuts the price of other competitors. The firms with the lowest price win all consumers. In equilibrium, at least two firms will offer q^* at marginal cost $c(q^*)$ and share the market. In this situation, the only other choice that yields zero profit (by avoiding any positive costs) is to choose $(M^k, \Theta^k) = \emptyset$. It follows:

Lemma 1 (Rational Benchmark) *If consumers are* not *context-sensitive ($\beta = 1$), a competitive equilibrium exists. In any such equilibrium, at least 2 firms share the market, each firm offering 1 product with quality q^* at marginal cost $c^* = c(q^*)$. All other firms choose $M^k = \emptyset$.*

III.B. Attraction and Fooling

For the rest of the paper, let $\beta > 1$ such that preferences are sensitive to the context in which products are presented at the point of purchase. Having consumers first select a store and then a product may now have important consequences for market supply. To see this, note that consumers are attracted to a store by the product they *expect* to purchase. If a consumer is naïve regarding future preference changes, this product must not necessarily conform to the product the consumer will ultimately purchase. We therefore define:

Definition 1 (Attraction Product) *We call product $j \in J^k$ an attraction product $a(\tilde{\beta})$ of firm k if and only if a consumer of type $\tilde{\beta}$ expects to purchase product j when entering store k.*

Definition 2 (Target) *We call product $j \in J^k$ the target t of firm k if and only if a consumer who enters store k purchases product j.*

Of course, sophisticated consumers perfectly foresee their behavior at the store implying that for these consumers the firm's target is also the attraction product. In particular, these consumers enter store k iff target t is feasible ($p_t \leq b$ and $q_t \geq \underline{q}$) and provides at least as high (undistorted) surplus as any other firm's target. If a consumer is naïve, however, she is prone to mispredicting her choice at a store where preferences are distorted by local context. In this case, the consumer might be attracted to a store by a product that is not the target. If a firm designs a store that attracts type $\tilde{\beta} \neq \beta$ with a product that is not the target, we say that the firm *fools* the consumer.

Definition 3 (Fooling) *Firm k* fools *type $\tilde{\beta}$ if and only if $a(\tilde{\beta}) \neq t$. If firm k fools type $\tilde{\beta}$,*

$$\text{(IC)} \qquad \hat{u}_t^k \geq \hat{u}_{a(\tilde{\beta})}^k$$

$$\text{(PCC)} \qquad E_{\tilde{\beta}}\left[\hat{u}_t^k\right] \leq E_{\tilde{\beta}}\left[\hat{u}_{a(\tilde{\beta})}^k\right]$$

with at least one of the inequalities being strict.

In this definition, condition (IC) is a standard incentive compatibility constraint: Inside store k, the consumer weakly prefers the target over the attraction product. Condition (PCC), on the other hand, is what we call a *perceived choice* constraint: When entering store k, a consumer with expectation $E(\beta) = \tilde{\beta} \neq \beta$ (falsely) expects to weakly prefer the attraction product over the target. Before we move to the analysis of fooling in detail, the following two points deserve mention:

Remark 1 (Rational Benchmark) *If consumers are* not *context-sensitive ($\beta = 1$), valuations are consistent across stages. Then no type $\tilde{\beta}$ can be fooled.*

Remark 2 (The Sole Function of Store-Context is Fooling) *If $a(\tilde{\beta}) = t$ for all $\tilde{\beta} \in$* $\text{supp}[f(\tilde{\beta})]$*, then the firm cannot improve by distorting preferences.*

To see why Remark 2 must be true, assume that firm k does not fool. Then $a(\tilde{\beta}) = t$ for all $\tilde{\beta} \in \text{supp}[f(\tilde{\beta})]$. The entire consumer population then *correctly* expects to purchase the target and receive (undistorted) surplus $u_t = q_t - p_t$ at store k, even if store-context Θ^k may distort valuations ex-post. It follows that the profit-maximal choice of quality and price is already determined by the outside valuation of products, implying that store-context cannot affect the best response.

III.C. Who is susceptible to fooling attempts?

It is clear that any naïve consumer can be fooled by a suitable choice of product-line M^k and preference distortion Θ^k. However, fooling may not always be more profitable for a firm than the standard Bertrand strategy of undercutting competitors with a single competitively priced product. We want to understand under which circumstances a fooling strategy indeed yields higher profits and may therefore be used in equilibrium. As a first step, in the next lemma we highlight important structural differences between a store-design that fools consumers who *over*estimate their sensitivity to context (pessimists) and a store-design that fools consumer who *under*estimate this sensitivity (optimists).

Lemma 2 (Naïveté and Fooling) *Let $\beta > 1$. Let product t be the target of firm k. Assume that the firm fools consumers of type $\tilde{\beta}^0$ by attracting them with product $a \neq t$. Then it is true that*

- *Store-context distorts valuations of at least one of the two products: $(\theta_a, \theta_t) \neq (N, N)$.*
- *The fooled consumer is (partially) naïve regarding contextual distortions: $\tilde{\beta}^0 \neq \beta$.*

a) *Pessimists and optimists cannot be fooled by the same pair of products (a, t):*

   If the fooled consumer is pessimistic *($\tilde{\beta}^0 > \beta$), other pessimistic types may be attracted by product $a \neq t$, but optimistic consumer always correctly expect to prefer the target. In particular, $\forall \tilde{\beta} < \beta$, $E_{\tilde{\beta}}\left[\hat{u}_t^k\right] > E_{\tilde{\beta}}\left[\hat{u}_a^k\right]$.*

   If the fooled consumer is optimistic *($\tilde{\beta}^0 < \beta$), other optimistic types may be attracted by product $a \neq t$, but pessimistic consumers always correctly expect to prefer the target. In particular, $\forall \tilde{\beta} > \beta$, $E_{\tilde{\beta}}\left[\hat{u}_t^k\right] > E_{\tilde{\beta}}\left[\hat{u}_a^k\right]$.*

b) *Being fooled increases undistorted surplus for pessimists, but lowers it for optimists:*

   If the fooled consumer is pessimistic *($\tilde{\beta}^0 > \beta$), she receives* higher *undistorted surplus than expected. In particular, $u_t > u_a$.*

   If the fooled consumer is optimistic *($\tilde{\beta}^0 < \beta$), she receives* lower *undistorted surplus than expected. In particular, $u_t < u_a$.*

Above lemma presents two important auxiliary results: Part a) implies that one group of naïves (pessimists or optimists) always "see through" a deceptive strategy that is meant to fool the other group. Unless a firm designs separate attraction products for each type (say, $a_1 \neq t$ for optimists and $a_2 \neq t$ for pessimists), there is always one group of consumers that is not fooled but correctly expects to purchase the target when entering the store. This result will be important later when we characterize the equilibrium for "mixed" populations where both groups co-exist. Part b), on the other hand, relates the sign of naïveté (pessimist or optimist) to the question of whether fooling is profitable: When fooled, consumers who underestimate their sensitivity to context purchase a product of lower undistorted surplus than expected ($u_t < u_a$). Intuition suggests that firms may be able to extract some of this "lost" surplus in the form of positive fooling profits. On the other hand, when pessimistic consumers ($\tilde{\beta} > \beta$) are fooled, $u_t > u_a$, suggesting that the deception of those consumers is not profitable. The next lemma confirms this intuition: Part a) establishes that firms must always lower the selling price of a target in order to fool pessimists. It follows that a standard Bertrand strategy (without fooling the consumer) is more profitable when selling to this group of naïves. Conversely, part b) establishes that fooling can yield *higher* profits than a Bertrand strategy when selling to optimists. However, not every distortion of preferences yields a profitable fooling outcome: To be able to sell a given target at a higher price than with a standard Bertrand strategy, the distortion Θ^k must increase the perceived surplus of the target relative to the attraction product. Only then will the consumer be induced to switch away from a competitively priced attraction product and be willing to purchase the more profitable target. We show in Lemma 3 part b) what this implies for the choice of distortions θ_a and θ_t (the effect of context on the valuation of the attraction product and the target) and how this relates to whether the consumer will be up-sold ($q_t > q_a$) or down-sold ($q_t < q_a$) as a result.

Lemma 3 (Profitable Fooling) *Let $\beta > 1$. Assume w.l.o.g. that $b \to \infty$. Fix any target quality $q_t \geq \underline{q}$. If firm k does <u>not</u> fool, the maximum price the firm can sell quality q_t to consumers of type $\tilde{\beta}^0$ is*

$$p_t^0 := q_t - \bar{u}(\tilde{\beta}^0),$$

where $\bar{u}(\tilde{\beta}^0) \geq 0$ is the highest undistorted surplus that a consumer with belief $\tilde{\beta}^0$ expects to receive when not *shopping at store k. Compare this to a fooling strategy where firm k attracts type $\tilde{\beta}^0$ with a different product $a \neq t$. Then it is true that:*

a) If type $\tilde{\beta}^0$ is pessimistic ($\tilde{\beta}^0 > \beta$), fooling her is unprofitable*: Conditional on selling target $t \neq a$, the firm must charge a price p_t that is strictly lower than p_t^0.*

b) If type $\tilde{\beta}^0$ is optimistic ($\tilde{\beta}^0 < \beta$), fooling her is profitable*:*

- *If store-context inflates qualities, $(\theta_a, \theta_t) = (Q, Q)$* and *the target has higher quality than the attraction product, $q_t > q_a$ (i.e., the firm up-sells), or*
- *If store-context inflates prices, $(\theta_a, \theta_t) = (P, P)$* and *the target has lower price than the attraction product, $p_t < p_a$ (i.e., the firm down-sells), or*
- *If store-context asymmetrically distorts surplus in favor of the target, $(\theta_a, \theta_t) \in \{(P, Q), (P, N), (N, Q)\}$,*

the firm can sell target $t \neq a$ at a price p_t that is strictly higher than p_t^0.

Moreover, fooling with other store-contexts is unprofitable. In particular, if store-context asymmetrically distorts surplus in favor of the attraction product, $(\theta_a, \theta_t) \in \{(Q, P), (Q, N), (N, P)\}$, the consumer cannot be fooled to purchase target t at any $p_t \geq 0$.

We end this section by characterizing the equilibrium for populations that consist entirely of consumers that are either impossible to fool or do not lend themselves to *profitable* fooling strategies: sophisticated and pessimistic consumers. When faced with such populations, firms maximize profit by behaving identical to standard Bertrand competitors, i.e., by offering a single, cost-efficient product at the lowest competitive price. In equilibrium, context-sensitivity will then be rendered inconsequential for market supply:

Proposition 1 (Sophistication and pessimism induces the rational outcome) *If consumers are context-sensitive ($\beta > 1$), but (weakly) overestimate their sensitivity to context ($\tilde{\beta} \geq \beta$ for all consumers), market supply is identical to the rational benchmark.*

Lemma 3 implies, on the other hand, that consumers who underestimate the malleability of their preferences ($\tilde{\beta} < \beta$) will be fooled if firms have the possibility to design a store-context that distorts preferences. We turn to the question of how such context may be constructed and define the equilibrium for populations that are profitable to fool in the next section.

IV. Equilibrium with Naïve Populations

Consider a naïve consumer that lends himself to profitable fooling. A store may try to influence the consumer in many different ways. For instance, it may try to use the color of price-tags to influence the perception and thus, the valuation of prices. Or, it may spray a certain perfume at the store entrance and hope that this leads the consumer to spend more money on quality. Meyers-Levy, Zhu and Jiang (2010) report, for example, that tactile properties of the flooring of a store (e.g., whether it is hard or soft) can affect the quality-assessment of products. While we desire to keep our model as general as possible, we also wish to be more specific about the construction of context. In particular, we aim at findings for a range of recent, much-attended models suggesting that consumers overweight that attribute of a product that "stands out" in the choice-set: the *Salience* model by BGS, the *Focusing* model by KS, and the model of *Relative Thinking* by BRS. These models make consumer attention and thus, distortions θ_j^k, a function of the product-line of firm k.

Our approach is as follows: In a first step, we keep the assumption that Θ^k is a direct choice-variable of the firm. The technology to manipulate context is thus assumed independent of the product-line and might as well be the color of price-tags, the spraying of perfume, or decisions about the tactile properties of flooring. Under the assumption that firms choose Θ^k directly, in section IV.A. we explore two possible assumptions regarding technological restrictions: (1) Contextual distortions are *store-wide*: For any two products j, j' at a given firm k, $\theta^k_j = \theta^k_{j'} = \theta^k$ and firms can choose $\theta^k \in \{Q, P, N\}$ and (2) Contextual distortions are *product-specific*: Firms can choose $\theta^k_j \in \{Q, P, N\}$ for each product individually. We show that under both assumptions, the unique equilibrium strategy when consumers underestimate contextual distortions ($\tilde{\beta} < \beta$) is a fooling strategy that uses context to up- or down-sell. In a second step, we nest the models of Salience (BGS), Focusing (KS) and Relative Thinking (BRS) in our framework and make Θ^k a function of the product-line. We show that in this case, firms achieve the profit-maximizing distortion by adding one additional product to their product-line. This *decoy* shifts consumer attention at the store to the the advantageous attribute of the target and thereby enables the fooling outcome. We show in detail where this decoy needs to be located in price-quality space according to each of the three models.

We assume for most of this section that *all* consumers are optimistic ($\tilde{\beta} < \beta$) and thus, are susceptible to profitable fooling attempts by firms. We return to the more general case of a "mixed" population at the end of this section where we define the equilibrium for societies where optimistic as well as sophisticated/pessimistic consumers co-exist. Outcomes and market-supply for each group of consumers will be unaffected by this co-existence: Our results for an entirely optimistic population (which we present next) will therefore continue to hold (for the optimistic part of society) in the general case.

IV.A. Firms Choose Context Directly

We present the central result of our paper below. The proposition characterizes the equilibrium when firms have an unspecified technology at hand that lets them choose the context of their store Θ^k. We show in a later proposition that this result also characterizes market supply when store-context is endogenous to the product-line and follows the Salience (BGS), Focusing (KS) and Relative Thinking (RT) frameworks.

Proposition 2 (Fooling Equilibrium) *Assume that consumers are context-sensitive ($\beta > 1$), and underestimate their sensitivity to context ($\tilde{\beta} < \beta$ for all consumers).*

a) **(Store-Wide Distortions.)** *Assume that for any two products j, j' at a given firm k, $\theta^k_j = \theta^k_{j'} = \theta^k$ and firms choose $\theta^k \in \{Q, P, N\}$. A competitive equilibrium exists. In any such equilibrium, at least 2 firms share the market, each firm offering 2 products, t and $a \neq t$. All other firms choose $(M^k, \Theta^k) = \emptyset$. Let*

$$\nu^{(Q,Q)} := [q^Q - c(q^Q)] + (\beta - 1)(q^Q - \underline{q}), \text{ and}$$
$$\nu^{(P,P)} := [q^P - c(q^P)] + (\beta - 1)[b - c(q^P)].$$

- *If $\nu^{(Q,Q)} > \nu^{(P,P)}$, a firm with strictly positive demand chooses $\theta^k = Q$, sells quality $q_t = q^Q > q^*$ at marginal cost $p_t = c(q^Q)$, but attracts with quality $q_a = \underline{q} < q_t$ at price $p_a = c(q^Q) - \beta(q^Q - \underline{q}) < p_t$. The firm* up-sells. *Note that $p_a < c(q_a)$.*
- *If $\nu^{(Q,Q)} < \nu^{(P,P)}$, a firm with strictly positive demand chooses $\theta^k = P$, sells quality $q_t = q^P < q^*$ at marginal cost $p_t = c(q^P)$, but attracts with quality $q_a = q^P + \beta[b - c(q^P)] > q_t$ at price $p_a = b > p_t$. The firm* down-sells. *Note that $p_a < c(q_a)$.*
- *If $\nu^{(Q,Q)} = \nu^{(P,P)}$, a firm with strictly positive demand chooses θ^k, (q_t, p_t) and (q_a, p_a) according to one of the two cases described above.*

b) **(Product-Specific Distortions)** *Assume that firms choose $\theta_j^k \in \{Q, P, N\}$ for each product $j \in J^k$ individually. A competitive equilibrium exists. In any such equilibrium, at least 2 firms share the market, each firm offering 2 products, t and $a \neq t$. All other firms choose $(M^k, \Theta^k) = \emptyset$. A firm with strictly positive demand chooses $\Theta^k = (\theta_a, \theta_t) = (P, Q)$, sells quality $q_t = q^Q > q^*$ at marginal cost $p_t = c(q^Q)$, but attracts with quality $q_a = \beta(q^Q + b) - c(q^Q) > q_t$ at price $p_a = b > p_t$. The firm* down-sells. *Note that $p_a < c(q_a)$.*

In the following, we provide a comprehensive intuition for this result and highlight important observations and predictions. Note first that Lemma 3 implies that the equilibrium must involve fooling: When consumers underestimate context effects ($\tilde{\beta} < \beta$), fooling is more profitable than any truthful sales strategy. If firms *can* distort context, they *will* do so and fool. We now explain the structure of *optimal* (i.e., equilibrium) fooling strategies.

Understanding consumers: Context-induced switching surplus. For a firm that fools, profit-maximization implies that perceived surplus of the target at the store is identical to the perceived surplus of the attraction product at the store, i.e., $\hat{u}_a^k = \hat{u}_t^k$.[6] This condition can be transformed into $u_a = u_t + [(\hat{u}_t^k - u_t) - (\hat{u}_a^k - u_a)]$, or

$$(4) \qquad u_a \equiv \nu^{(\theta_a, \theta_t)} := \underbrace{u_t}_{\text{undistorted surplus of target}} + \underbrace{(\beta - 1)\begin{cases} (q_t - q_a) & \text{if } (\theta_a, \theta_t) = (Q, Q), \\ (p_a - p_t) & \text{if } (\theta_a, \theta_t) = (P, P), \\ (q_t + p_a) & \text{if } (\theta_a, \theta_t) = (P, Q). \end{cases}}_{\text{context-dependent (virtual) surplus from switching to target}}$$

Expression (4) has an intriguing interpretation: Because u_a is the expected surplus at store k, a fooled consumer behaves *as if* she was maximizing undistorted surplus u_t plus a

[6]If this was not the case, the firm could increase the target's price p_t or decrease its quality q_t and thereby increase profits.

term that measures the context-dependent subjective benefit of "changing her mind" when switching from product a to product t. If $(\theta_a, \theta_t) = (Q, Q)$, this relative advantage comes from switching to a product of higher quality. If $(\theta_a, \theta_t) = (P, P)$, the consumer perceives a "bargain effect" when switching from the expensive product to the cheaper option. Finally, if the distortion is asymmetric, $(\theta_a, \theta_t) = (P, Q)$, the effect is a combination of perceived bargain (*not* purchasing product a) and quality-increase (purchasing t). Most intuitively, this context-induced (virtual) surplus—and therefore the incentive of firms to fool—increases in the size of the context-sensitivity parameter β.

Marketing and the choice of attraction products. For a given target, maximizing profit means choosing an attraction product that maximizes the virtual surplus of switching to this target at the store. A different contextual distortion might therefore call for a different "marketing" (i.e., attraction) strategy. If $(\theta_a, \theta_t) = (Q, Q)$, a firm achieves the highest profit when attracting the consumer with the lowest quality possible, $q_a = \underline{q}$, while competing with other firms on price p_a. If $(\theta_a, \theta_t) \in \{(P, P), (P, Q)\}$, the bargain effect is maximized by fixing p_a at the maximum acceptable price b and competing with other firms on quality q_a. The lower is $\underline{q}$ or the higher is b, the more extreme are the products with which firms attract consumers to their store.

Cost-efficiency and the choice of targets. While the optimal attraction product is either a "top-of-the-range" (price $p_a = b$) or "budget" (quality $q_a = \underline{q}$) option, the optimal target is a product that lies somewhere in the middle. In particular, when the target is quality-inflated—i.e., if $(\theta_a, \theta_t) \in \{(Q, Q), (P, Q)\}$—, selling quality $q_t = q^Q := \arg\max_q[\beta q - c(q)]$ is cost-efficient, while if the price of the target is inflated, selling $q_t = q^P := \arg\max_q[q - \beta c(q)]$ is cost-efficient. Note that $q^Q > q^* > q^P$ and that the differences between these qualities increase in the sensitivity to context, that is, in β.

Up-selling vs. down-selling. The previous two paragraphs show that up- and down-selling attempts (from extreme attraction products to "mid-range" targets) are an integral part of any fooling equilibrium. Whether firms engage in up- *or* down-selling depends on the profitability of each regime. Condition (4) clearly shows that the asymmetric distortion $(\theta_a, \theta_t) = (P, Q)$ weakly dominates the symmetric distortions in virtual surplus and hence, in profit.[7] When firms have access to a technology that allows for product-specific choice of θ_j^k, constructing $(\theta_a, \theta_t) = (P, Q)$ and *down*-selling the consumer is therefore always profit-maximizing. With store-wide distortions, on the other hand, it depends on parameter values whether the virtual surplus $\nu^{(\theta_a, \theta_t)}$ is larger in the case of a quality-inflating context ($\theta^k = Q$) or a price-inflating context ($\theta^k = P$) and hence, whether firms use up- or down-selling

[7] Not shown in Expression 4 are other asymmetric distortions which are also dominated by $(\theta_a, \theta_t) = (P, Q)$. If $(\theta_a, \theta_t) = (N, Q)$, the switching surplus is $(\beta - 1)q_t$, while for $(\theta_a, \theta_t) = (P, N)$, it is $(\beta - 1)p_a$.

strategies in equilibrium. In particular, because virtual surplus increases in the difference between expected and realized purchase, the characteristics of the attraction product play an important role for profitability: Up-selling entails a high switching surplus and is particularly profitable if $\underline{q}$ is low. Similarly, down-selling is profitable if b is high. Accordingly, up-selling attempts should be prevalent in markets where consumers have a limited budget (b low) and can be attracted with a low-quality product ($\underline{q}$ low), while down-selling attempts are likely to be found in markets where consumers expect a high minimum quality ($\underline{q}$ high) and do not shy away from high prices (b high).

Zero-profit equilibrium and loss-leader products. Profits are zero in equilibrium even if consumers are (infinitely) heterogeneous in their beliefs $\tilde{\beta}$ and the number of firms is limited (to any number $K \geq 2$). This is a result of the optimality condition $\hat{u}_a^k = \hat{u}_t^k$, which implies that *any* consumer who strictly underestimates the effect of contextual distortions—i.e, holds belief $\tilde{\beta} < \beta$—(wrongly) expects to prefer product a over t at the store: If $\hat{u}_a = \hat{u}_t$, and fooling is successful according to Lemma 2, then $E_{\tilde{\beta}}\left[\hat{u}_a^k\right] > E_{\tilde{\beta}}\left[\hat{u}_t^k\right]$ for any $\tilde{\beta} < \beta$. Because the best-response does not generate heterogenous expectations among consumers with $\tilde{\beta} < \beta$, consumers are treated as if they were perfectly homogeneous and competition is Bertrand-like despite the extra fooling profits. In equilibrium, the context-dependent mark-ups on the target are offset by negative mark-ups on the attraction product, $p_a < c(q_a)$. In other words, the attraction product is a "loss leader".

Incentives to fool. With no "fooling profits" left in equilibrium, one might be tempted to think that the incentives to fool are small. To show that this is not the case, we make these incentives explicit. A simple transformation of Expression (4) yields a general expression for the maximum profits of a firm that attracts consumers who have an outside option $\bar{u} \geq 0$:

$$\pi_t = \underbrace{q_t - c(q_t) - \bar{u}}_{\text{max. profits without fooling}} + \underbrace{(\beta - 1)\begin{cases}(q_t - q_a) & \text{if } (\theta_a, \theta_t) = (Q, Q),\\ (p_a - p_t) & \text{if } (\theta_a, \theta_t) = (P, P),\\ (q_t + p_a) & \text{if } (\theta_a, \theta_t) = (P, Q).\end{cases}}_{\text{context-dependent fooling profits}} \tag{5}$$

Clearly, the higher the degree of context-sensitivity β, the higher are the incentives to fool and deviate from the rational benchmark.[8]

Welfare. There are two aspects of the equilibrium that seem important for a welfare analysis. The first aspect is the apparent difference in products purchased by rational and

[8]Note that Expression (5) also nests information on profits in less competitive market situations. The maximum surplus available outside of the firm ($\bar{u} \geq 0$) can be used to parameterize the degree of competition in the market. For example, setting $\bar{u} = 0$ yields the monopoly profits and shows that for a monopolist, higher context-sensitivity directly translates into higher profits.

fooled consumers. If—for instance—welfare is defined by undistorted surplus generated by trade occurring in equilibrium, the exploitation of context sensitivity leads to an efficiency loss (relative to the rational benchmark) that increases in the context sensitivity parameter β. Second, introspection suggests that the practice of fooling itself—i.e., the deliberate generation of false expectations—may affect "experienced utility" and thus welfare of the consumer.[9] Taking this point of view, it is the "degree" of fooling that may be important for the policy maker. The degree of fooling may be defined, for example, as the difference between the expected and received (undistorted) surplus $u_a - u_t$. By Expression (4), the degree of fooling is then nothing else than the context-dependent virtual surplus experienced when switching products at the store. In equilibrium:

$$u_a - u_t = \begin{cases} 0 & \text{if consumers are rational or sophisticated,} \\ (\beta - 1)(q^Q - \underline{q}) & \text{if consumers are fooled and } (\theta_a, \theta_t) = (Q, Q), \\ (\beta - 1)[b - c(q^Q)] & \text{if consumers are fooled and } (\theta_a, \theta_t) = (P, P), \\ (\beta - 1)(q^Q + b) & \text{if consumers are fooled and } (\theta_a, \theta_t) = (P, Q). \end{cases}$$

Following this definition, the degree of fooling increases in β as well as in the possibility of firms to attract consumers with 'extreme' products having very low quality $\underline{q}$ or very high price b. Note also that the degree of naïveté, expressible by the difference between actual and expected context sensitivity $\beta - \tilde{\beta}$, does not affect outcomes and thus, leaves welfare completely unaffected. This is due to all optimists, regardless of their degree of naïveté, having identical (yet false) expectations about their purchase in equilibrium.

IV.B. Salience, Focusing, and Relative Thinking

We now introduce three specific models of contextual distortion to our framework that have recently been suggested in the literature. These models make context a function of the product-line—assuming that consumers overweight attributes to which *relative comparisons* between products draw their attention.

Salience. Bordalo, Gennaioli and Shleifer (2013, BGS) assume that consumers attach disproportionally high weight to *salient* attributes, where "[a]n attribute is salient for a good when it stands out among the good's attribute relative to that attribute's average level in the choice set" (BGS, abstract, p. 803). We apply the original salience definition by BGS

[9] We follow Chetty (2015, p.2) by acknowledging that "Behavioral biases [...] often generate differences between welfare from a policy maker's perspective, which depends on an agent's experienced utility (his actual well-being), and the agent's decision utility (the objective the agent maximizes when making choices)." We see our formal model as only representing "decision utility", making an additional definition of "experienced utility" for welfare analysis necessary and valid.

(BGS, Definition 1 and Assumption 1) to a choice set equal to the product line of store k:

Assumption S (Salience) *Let z_R^k be the* average *level of attribute $z \in \{q,p\}$ at store k. The salience of attribute z_j, $z \in \{s,p\}$ at store k is given by a symmetric and continuous (real-valued) function $\sigma\left(z_j, z_R^k\right)$ that satisfies ordering and homogeneity of degree zero.*[10] *Then*

$$\theta_j^k = \begin{cases} Q & \textit{iff } \sigma\left(q_j, q_R^k\right) > \sigma\left(p_j, p_R^k\right) \\ P & \textit{iff } \sigma\left(q_j, q_R^k\right) < \sigma\left(p_j, p_R^k\right) \\ N & \textit{otherwise.} \end{cases}$$

The properties of the salience function listed in Assumption S are in analogy to key features of sensory perception. Ordering says that salience increases in contrast: the higher the difference to the reference, the more salient is an attribute. Homogeneity of degree zero captures the concept of diminishing sensitivity, implying that relative differences instead of absolute differences matter in comparisons. For $z_j > 0$ and $z_R^k > 0$, $\sigma(z_j, z_R^k) = \frac{|z_j - z_R^k|}{(z_j + z_R^k)}$ is an example of a salience function that satisfies these properties.

Focusing. Kőszegi and Szeidl (2013, KS) argue "that a person focuses more on, and hence overweights, attributes in which her options differ more" (KS, abstract, p. 53). We implement the central assumption of KS (Assumption 1) in the following way:

Assumption F (Focusing) *Let Δ_z^k be the spread of attribute $z \in \{q,p\}$ at store k, $\Delta_z^k := \max_{j \in J^k} z_j - \min_{j \in J^k} z_j$, and let $\kappa_F \geq 0$ be some (exogenously defined) threshold. Then*

$$\theta_j^k = \begin{cases} Q & \textit{iff } \Delta_q^k - \Delta_p^k > \kappa_F \\ P & \textit{iff } \Delta_p^k - \Delta_q^k > \kappa_F \\ N & \textit{otherwise.} \end{cases}$$

Note that for any two products $\{j,i\} \in J^k$, $\theta_j^k = \theta_i^k = \theta^k$.

Assumption F says that a consumer focuses more on, and hence overweights, quality (price) iff the spread in qualities (prices) at store k is larger than the spread in prices (qualities) by some threshold $\kappa_F \geq 0$. This is a straightforward adaption of the central assumption of KS to our rank-dependent framework where "focus-weights" can only take values of 1 or

[10] These are defined as follows. (1) Ordering: Let $\mu = sgn\left(z_k - z_R^k\right)$. Then, for any $\varepsilon, \varepsilon' \geq 0$ with $\varepsilon + \varepsilon' > 0$, $\sigma\left(z_j + \mu\varepsilon, z_R^k - \mu\varepsilon'\right) > \sigma\left(z_j, z_R^k\right)$. (2) Homogeneity of degree zero: $\sigma\left(\alpha z_j, \alpha z_R^k\right) = \sigma\left(z_j, z_R^k\right) \forall \alpha > 0$. These definitions are valid for $z_j > 0$ and $z_R^k > 0$. In order to work with nonpositive arguments, we would need to formulate additional properties, see BGS. For our results it is however sufficient to have salience defined in the positive domain.

$\beta \geq 1$. The exogeneous threshold $\kappa_F \geq 0$ measures the level of stimulus necessary for such a preference distortion.

Relative Thinking. Bushong, Rabin and Schwartzstein (2016, BRS) model the idea that "[f]ixed differences loom smaller when compared to large differences" (BRS, abstract, p.1). The consumer thus "weighs a given change along a consumption dimension by less when it is compared to bigger changes along that dimension" (ibid.). We base our implementation on the central norming assumptions N0-N2 in BRS.

Assumption RT (Relative Thinking) *Let Δ_z^k be the spread of attribute $z \in \{q, p\}$ at store k, $\Delta_z^k := \max_{j \in J^k} z_j - \min_{j \in J^k} z_j$, and let $\kappa_{RT} \geq \beta$ be some (exogenously defined) threshold. Then*

$$\theta_j^k = \begin{cases} Q & \textit{iff } \dfrac{\Delta_p^k}{\Delta_q^k} > \kappa_{RT} \\ P & \textit{iff } \dfrac{\Delta_q^k}{\Delta_p^k} > \kappa_{RT} \\ N & \textit{otherwise.} \end{cases}$$

Note that for any two products $\{j, i\} \in J^k$, $\theta_j^k = \theta_i^k = \theta^k$.

Assumption RT says that fixed differences loom smaller when compared to large attribute-spreads Δ_z^k at the store. When the ratio $\frac{\Delta_p^k}{\Delta_q^k}$ is sufficiently large, any given price-difference looms (relatively) small compared to the overall price-spread while any given quality-difference looms (relatively) large compared to the overall quality-spread, leading the consumer to disproportionally concentrate on quality in pairwise comparisons of products. The right-hand-side of the condition ($\kappa_{RT} \geq \beta$) is due to norming assumption N2 in BRS which "assures that people show some sensitivity to absolute consumption utility differences" (BRS, p.8). A higher threshold κ_{RT} implies stronger contextual stimuli are required to distort preferences from the baseline.[11]

[11] Assumption RT implements norming assumptions N0-N2 in the following way. Let $w(\cdot)$ denote the weight function that attaches weight $w_z^k \in \{1, \beta\}$ to attribute $z \in \{q, p\}$. N0 is simply the assumption that $w(\cdot)$ is a function of the attribute spread Δ_z^k. Now suppose that quality has a higher weight than price, i.e. $w_q^k = \beta$ and $w_p^k = 1$. According to our framework, $\theta_j^k = Q$ for all products $j \in J^k$. By N1,

$$w(\Delta_q^k) > w(\Delta_p^k) \Rightarrow \Delta_q^k < \Delta_p^k.$$

But N2 makes a more restrictive assumption, namely,

$$w(\Delta_q^k) > w(\Delta_p^k) \wedge \Delta_q^k < \Delta_p^k \Rightarrow w(\Delta_q^k)\Delta_s^k < w(\Delta_p^k)\Delta_p^k \Leftrightarrow \beta\Delta_q^k < \Delta_p^k,$$

which is identical to our implementation by Assumption RT if $\kappa_{RT} = \beta$. The possibility of $\kappa_{RT} > \beta$ captures cases where stronger stimulus is required. An analogous statement establishes the case of $\theta_j^k = P$.

Before we formulate the equilibrium when Θ^k is defined according to one of the theories introduced above, note that the Salience specification (Assumption S) is the only specification that allows distortions to be *product-specific*. Under Assumptions F and RT, contextual distortions affect all products equally and are therefore always store-wide. Furthermore, the set-up of all three theories implies that consumer preferences between any two products can be manipulated by a third product that is added to the choice-set. When context is defined according to one of the three theories, we can therefore make the following observation.

Proposition 3 (Fooling with Decoys: Salience, Focusing, and Relative Thinking) *Assume that consumers are context-sensitive ($\beta > 1$), and underestimate their sensitivity to context ($\tilde{\beta} < \beta$ for all consumers). Under Assumption S, F, or RT, a competitive equilibrium exists. In any such equilibrium, at least 2 firms share the market, each firm offering 3 products, t, $a \neq t$, and $d \notin \{t, a\}$. The firm sells product t, but attracts with product a. The sole function of product d is to manipulate preferences: Product d is a* decoy*. Moreover,*

a) Under Assumption F or RT, a firm with strictly positive demand chooses (θ_a, θ_t), (q_t, p_t), and (q_a, p_a) according to Proposition 2, Part a).

b) Under Assumption S, a firm with strictly positive demand chooses (θ_a, θ_t), (q_t, p_t), and (q_a, p_a) according to Proposition 2, Part b).

The important take-away from Proposition 3 is that with each of the three specifications the product-line itself becomes a technology that enables firms to distort consumer preferences after entering the store. In particular, any profit-maximizing distortion of attraction product and target can be realized by adding just *one* additional product to the product-line. The function of this *decoy* is to manipulate the "reference points" of consumers, making the target look disproportionally attractive compared to the attraction product at the store. This technology enables firms to virtually *choose* θ_j^k and fool according to Proposition 2. All three specifications predict a single decoy in equilibrium. To achieve incentive compatibility, the decoy must necessarily be an unattractive option (neither preferred outside nor inside the store). The location of this product in quality-price-space however depends on which specification is employed.

Decoy-Positions: Focusing and Relative Thinking. Figure 1 shows decoy-positions under Assumptions F and RT in equilibrium. Under these assumptions, firms are restricted to store-wide distortions. There are two cases: (1) the firm up-sells, $q_t > q_a$, by inflating perceived qualities: $(\theta_a, \theta_t) = (Q, Q)$ (left figure), or (2) the firm down-sells, $q_t < q_a$, by inflating perceived prices: $(\theta_a, \theta_t) = (P, P)$ (right figure). To achieve the profit-maximizing distortion without violating incentive compatibility, the firm has to add a decoy to the product-line that resides within the boundaries of the grey shaded areas in Figure 1.[12] A product located in this area is "unattractive" itself but manipulates the attribute spreads Δq and Δp—which constitute *reference points* of consumer attention according to the theories of Focusing and Relative Thinking—in favor of the target. Note that the shaded areas for Assumption F and RT do not overlap, implying that decoys can help identifying the two models from data: According to the theory of Focusing (KS), a consumer overweights the attribute in which her options differ more. A successful decoy therefore serves to extend the spread of that attribute on which the target is *better* than the attraction product (quality when up-selling, price when downselling). The theory of Relative Thinking (BRS), on the other hand, makes a somewhat opposite assumption, namely that a fixed difference looms smaller when compared to a large spread. A successful decoy therefore has to extend the spread of that attribute on which the target is *worse* than the attraction product (price when up-selling, quality when down-selling). Under both theories, a choice that resonates with experimental literature and anecdotal evidence on so-called *decoy effects* is to construct a decoy that copies the target in one attribute, but is strictly worse along the other dimension.[13] The white markers in Figure I illustrate such a choice.

Decoy-Positions: Salience. When attention is modeled according to BGS' model of Salience (Assumption S), contextual distortions of quality and price may be product-specific and choosing distortion $(\theta_a, \theta_t) = (P, Q)$ is profit-maximizing. Figure II illustrates how the firm can construct this distortion by manipulating the position of the reference point (q_R, p_R) using a single decoy.[14] The figure depicts the case when, as in equilibrium, $q_a > q_t$ and $p_a > p_t$. The construction (which we also use to prove Proposition 3) exploits two central implications of the Salience framework: (1) If product $j \in J^k$ neither dominates nor is dominated by the reference point, i.e., $\left(q_j - q_R^k\right)\left(p_j - p_R^k\right) > 0$, then the "advantageous" attribute of product j—higher quality *or* lower price relative to the average—is overweighted

[12]The shaded areas show decoy positions for minimum thresholds $\kappa_F = 0$ and $\kappa_{RT} = \beta$. Larger thresholds demand decoys that are located further away from products t and a.

[13]See, e.g., Huber, Payne and Puto (1982); Doyle et al. (1999); Herne (1999) for experimental literature on asymmetrically dominated decoys. It is interesting to see that *optimal* firm behavior produces entire product-lines that very strongly resemble experimental tests as well as anecdotal evidence on decoy effects. While the characteristics of the decoy (for a given set of competitor and target) derive from the particular specification of consumer attention, it is our model that explains (and predicts) the *existence* of all three products—decoy, competitor and target—at one firm. We are not aware of another theory with a similar prediction.

[14]Recall that the reference point according to the theory of Salience is a "fictitious" product with quality q_R and price p_R conforming to the average quality and price, respectively, in the choice-set.

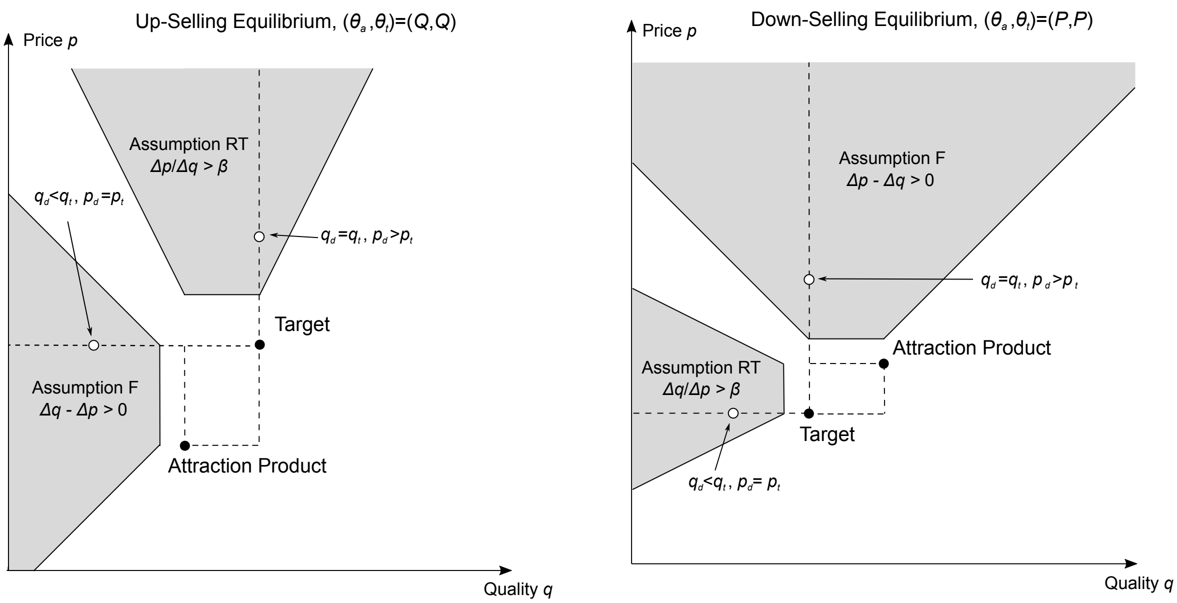

Figure I
Decoy positions (shaded areas) under Assumption F or Assumption RT

if and only if the product has better-than-average quality-to-price ratio, that is, $\frac{q_j}{p_j} > \frac{q_R^k}{p_R^k}$.[15] (2) If one attribute of product $j \in J^k$ is average while the other is not (e.g., $q_j = q_R^k$, but $p_j \neq p_R^k$), then the latter is "outstanding" and thus salient. Following these rules, the firm can implement the distortion $(\theta_a, \theta_t) = (P, Q)$ by constructing a reference point (q_R^k, p_R^k) that is either clearly dominated by the target ($p_R = p_t$, but $q_R < q_t$) or by the attraction product ($q_R = q_a$, but $p_R > p_t$). Which of the two constructions is feasible depends on whether the target or the attraction product has a higher quality-to-price ratio (see the left-hand-side and right-hand-side of Figure II, respectively). In both cases, such a reference point can always be constructed—using a single, unattractive decoy—without violating incentive compatibility (as exemplarily shown in Figure II).

IV.C. Mixed Populations

Our propositions have so far stated results for the polar cases of either an entirely sophisticated/pessimistic consumer population (Proposition 1, inducing the rational outcome) or an entirely optimistic consumer population (Propositions in this section, inducing a fooling outcome). The next proposition shows that equilibrium outcomes for both optimistic ($\tilde{\beta} < \beta$) and non-optimistic ($\tilde{\beta} \geq \beta$) consumers—as well as market supply for each group—

[15] Also, iff $\frac{q_j}{p_j} < \frac{q_R^k}{p_R^k}$, then the "disadvantageous" attribute of product j is overweighted, while iff $\frac{q_j}{p_j} = \frac{q_R^k}{p_R^k}$, consumers weigh both attributes equally.

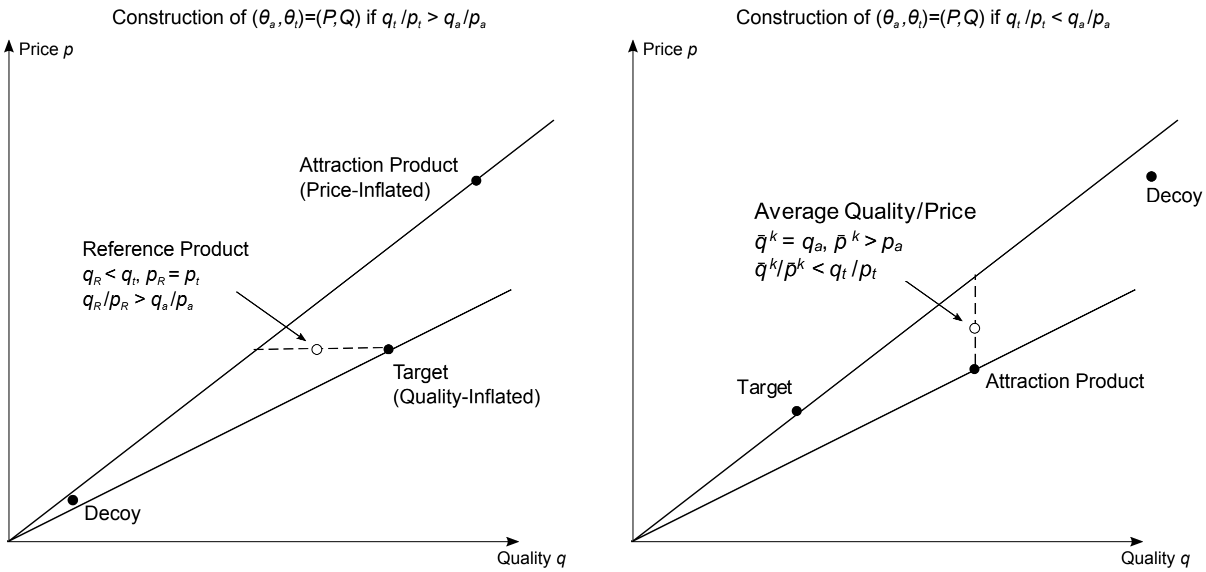

Figure II
Construction of distortion $(\theta_a, \theta_t) = (P, Q)$ (with one decoy) under Assumption S

are unchanged when the consumer population is mixed, i.e., when these groups co-exist in society.

Proposition 4 (Equilibrium with Fooling and Non-Fooling Firms) *Assume that consumers are context sensitive ($\beta > 1$) and firms either directly choose θ_j^k (with store-wide or product-specific distortions, following the assumptions in Proposition 2), or Assumption S, F, or RT is satisfied (firms can manipulate θ_j^k indirectly using decoy products). A competitive equilibrium exists. Moreover,*

a) If there exist consumers who are sophisticated or overestimate their sensitivity to context ($\tilde{\beta} \geq \beta$), then there exist at least 2 non-fooling ('truthful') firms that each offer 1 product with quality q^* at marginal cost $c(q^*)$ and all consumers with $\tilde{\beta} \geq \beta$ buy at one of these firms.

b) If there exist consumers who underestimate their sensitivity to context ($\tilde{\beta} < \beta$), then there exist at least 2 fooling firms that each offer 2 or 3 products, according to Propositions 2 and 3, respectively, and all consumers with $\tilde{\beta} < \beta$ buy at one of these firms.

c) Firms that do not supply the market (according to cases a or b), choose $(M^k, \Theta^k) = \emptyset$.

Intuition for this result follows from the fact that sophisticated and pessimistic consumers necessarily "see through" any fooling schemes that firms design to exploit optimistic consumers (see also Lemma 2). Knowing that they would buy the inefficient target when entering such a store, these consumers can be attracted by 'truthful' firms that offer commitment devices in the form of single-product stores (where fooling attempts are impossible). In equilibrium, competition leads these truthful firms to sell cost-efficient quality q^* at marginal cost c^*. Optimistic consumers, on the other hand, can be profitably fooled even if their outside option is to buy q^* at price c^*. Competition for this group of consumers leads to the equilibrium product-lines defined in Proposition 2 (and confirmed under Assumption S, F, or RT in Proposition 3).

V. Discussion and Conclusion

We have presented a theory of firms that use contextual variations in the presentation of options to enable a deceptive marketing strategy able to exploit the naïvité of consumers. In this final section, we first turn to the question of how to confront our theory with anecdotal and empirical evidence and then shortly discuss welfare implications and possible regulatory interventions. We concentrate throughout on our predictions regarding market supply for consumers who underestimate their sensitivity to context, as this is where market supply deviates from the predictions of the rational benchmark.[16]

Store-context as a marketing strategy. We predict that firms design context to drive a wedge between preferences before and after entering a store in order to compete with a product that is different to the product that they ultimately aim to sell. Anecdotal evidence in the form of consumer and marketing blogs is ample for such up- and down-selling attempts to be wide-spread in actual retail markets. With regard to up-selling, we present just one of many possible examples in Figure III. The figure shows an online advertisement as well as the product-line of the *Dollar Shave Club*, an online retailer of razors.[17] In this example, the firm attracts consumers with "great razors" priced at \$1, but once consumers have entered the online-store, tries up-selling from this budget option (now unattractively called "the humble twin") to a more expensive razor. Note that the example matches the description of a product-line that uses an attraction product ('The Humble Twin"), a target ("The 4X"–"Member favorite") and a decoy ("The Executive"). Other popular examples include up-selling attempts in airline ticket and rental car markets—where firms typically attract consumers with low priced-products via price-comparison websites but then try up-selling to more expensive options after the consumer has entered the website of a particular firm—or

[16]As we have shown in the main part of the paper, predictions for context-sensitive consumers who do *not* underestimate their bias is identical to a standard model without context-sensitivity.

[17]We have borrowed this example from the website https://econsultancy.com/blog/66879-10-powerful-examples-of-upselling-online/ (accessed 09-20-2016), which presents many more.

the famous *The Economist* example by Dan Ariely (Ariely, 2008).

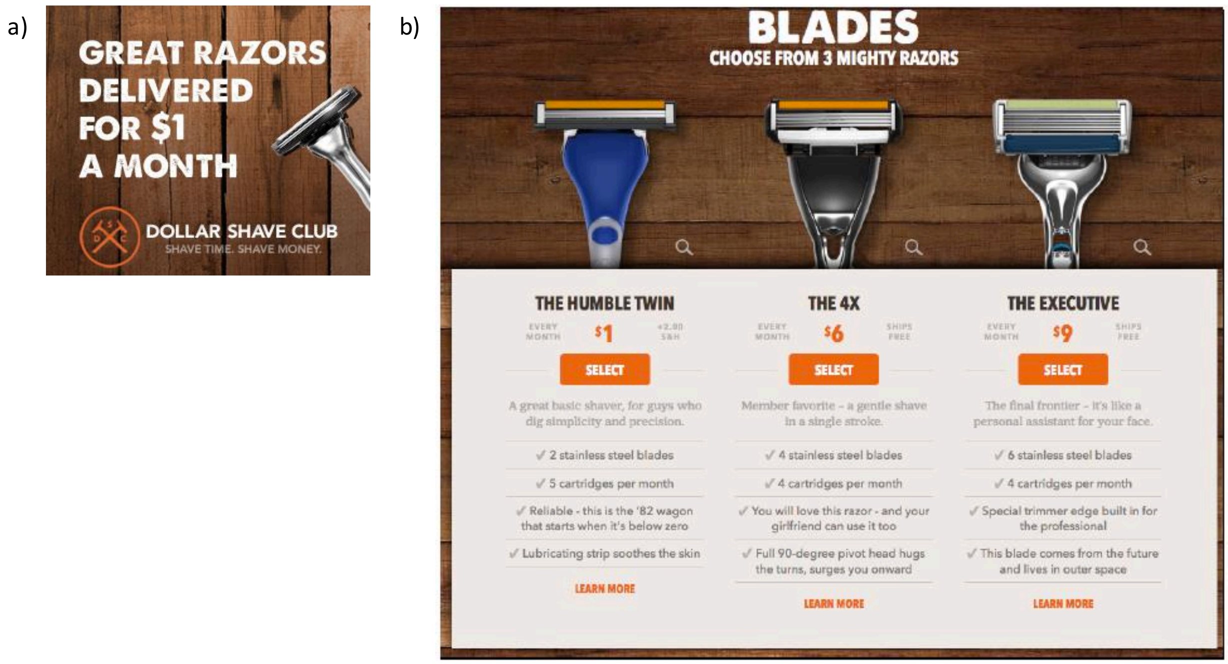

Figure III
Example for an (alleged) up-selling strategy on the internet: a) Advertised Product of the *Dollar Shave Club*, b) Product-Line at the web store of the *Dollar Shave Club*

Source: a) http://www.razorpedia.com/blog/dollar-shave-club-business-update, b) https://econsultancy.com/blog/66879-10-powerful-examples-of-upselling-online/, both websites accessed on 09-20-2016.

With regard to *down-selling*, similar examples can be found. For instance, CBS reporter Jeremy Quittner writes on “How to Boost Sales by Down-Selling Your Customers” and gives an illustrative example from the fashion industry: “Take handbag company Oovoo Designs in Alexandria, Va. In 2008, Pauline Lewis, founder of the $500,000 company, saw sales slipping about 15 percent for her high-end bags [...]. So Lewis took immediate action. [...] From her research she came away with a new idea: a line of $35 cell-phone holders with some of the same hand-embroidered flourishes as the pricier items. [...] ‘Younger women can’t afford the $250 bag, but they love the brand so much [...],’ she adds. The strategy ultimately offset losses for the more up-market bags. She sold about 1,500 of the cell-phone holders in 2008, and then about 2,500 the following year.”[18] Note that in this example, the attraction product

[18]http://www.cbsnews.com/news/how-to-boost-sales-by-down-selling-your-customers/, accessed 09-20-2016.

is some high-end handbag (that is ultimately *not* sold) while actual profits are made on a cheaper, lower-quality item, the cell-phone holder.

Up-selling vs down-selling. Our model predicts that firms can use both up- and down-selling regimes to exploit context-sensitive consumers. To the best of our knowledge, our theory is the first to explain both up-selling and down-selling attempts of firms within one framework. Which of the two strategies is actually chosen by a firm depends on its technological ability to distort context as well as on market characteristics: We predict up-selling to be more likely in “no-frills ”markets where consumers can be attracted with low-quality products at low prices—like in the razor-example above—and down-selling to be more likely in “up-scale” markets where consumers can be attracted by luxurious products at high prices—conforming to the handbag-example above. While it is easier to find specific examples and blogs covering up-selling attempts, marketing methods in up-scale markets seem to have much in common with a down-selling strategy. This is true for the fashion industry but also for car brands like Mercedes or BMW. These firms advertise quality, not price, when showing expensive sports cars and putting up slogans like “sheer driving pleasure” or “driving DNA” to attract consumers, while making the bulk of their final sales on cheaper cars of considerably less sportiness.

Loss leaders as attraction products. Our model predicts that in sufficiently competitive markets the product with which consumers are attracted to the store is a loss-leader. Chevalier, Kashyap and Rossi (2003) provide empirical support for existing loss-leading models like Lal and Matutes (1994) and Hosken, Matsa and Reiffen (2000). Using price and sales data on a large grocery store in Chicago, they find that it is the advertised product varieties that are discounted in times of high (idiosyncratic) demand. This finding is consistent with our theory. However, it does not separate it from existing models that treat loss leading in isolation from up- or down-sell strategies and that assume, other than we do, that loss leaders are meant as much for sale as other products. Some evidence for this missing piece is provided by Ellison and Ellison (2009). Using data on online markets for computer memory, the authors find that firms attract consumer with a loss-leader product via the price-search engine *Pricewatch.* Once attracted, consumers are manipulated into purchasing a product with a higher profit-margin. Hence, Ellison and Ellison (2009) provide evidence for the use of loss-leaders in combination with an up-selling strategy.

Context Effects. The evidence and examples presented by Ellison and Ellison (2009) are very much in line with the predictions of our theory. Still, there is as yet no systematic evidence that it is the *context* at the final point of purchase that induces consumers to switch products. An extensive experimental literature documents effects of exogenous context on consumer choice that are consistent with our theory and match the design of choice environments in many examples of up- and down-selling attempts. This pertains both to studies of "atmospheric effects" (context created by such things as the color scheme, lighting, and flooring of a store, see Turley and Milliman, 2000, for a review) and choice-set related *decoy-effects* (see, e.g., Huber, Payne and Puto, 1982; Simonson, 1989; Simonson and Tversky, 1992; Herne, 1999; Doyle et al., 1999). The latter evidence is particularly in line with our model when nesting the frameworks suggested by BGS, KS, and BRS, which predict that firms use decoys similar to the experimental evidence to manipulate preferences at the store. We are however not aware of clean evidence on consumer sensitivity to or firms' strategic choice of store-context in competitive markets. This gap in the empirical literature opens up to a promising new avenue for future research.

We end with a note on welfare and policy. Welfare statements—and for this reason, policy recommendations—are inherently difficult in situations where consumers make time-inconsistent choices. To borrow two terms recently suggested by Chetty (2015), making conclusive statements about welfare in the presence of such biases requires the modeler to define what is the "experienced utility" of the consumer (his actual well-being) and how this utility compares to his "decision utility" (the objective the agent maximizes when making choices). The question of how experienced utility should be exactly defined is ultimately an empirical one. Irrespective of the finer functional details of such a definition we believe that there are two aspects that seem important for a welfare analysis of the equilibrium we have described. First, the possibility of firms to manipulate preferences by manipulating the environment of a choice leads competitive firms to sell products that may not be cost-efficient from a welfare perspective. Because firms in equilibrium use both, up- and down-selling strategies, and therefore may sell products that have quality both higher or lower against any stable benchmark, such efficiency losses are likely to occur for any definition of experienced utility that is non-identical with decision-utility at the moment of purchase. Second, introspection suggests that the systematic practice of up- and down-selling *itself* may affect experienced utility and thus welfare of the consumer because it builds on the deliberate generation of false expectations. Taking this point of view, it is the "degree" of fooling—i.e., the difference between expected and received utility—that is important for the policy maker. Notably in this case, welfare does not only depend on the characteristics of the product purchased, but also on the characteristics of the product with which the consumer is attracted to the store. In particular, attraction products that are more "extreme" in terms of quality and price (for example, advertisements with "ultra-low-budget" or "top-of-the-range" products) are likely to entail more extreme up- and down-selling attempts and produce larger welfare effects.

Above reasoning suggests that policy makers might want to curb the power of firms to use store-context as a marketing tool. In fact, the equilibrium practice of up- and down-selling we have described in this paper resembles the Federal Trade Commission (FTC) definition of a *Bait-and-Switch*: "Bait advertising is an alluring but insincere offer to sell a product or service which the advertiser in truth does not intend or want to sell. Its purpose is to switch consumers from buying the advertised merchandise, in order to sell something else, usually at a higher price or on a basis more advantageous to the advertiser" (16 C.F.R. §238.0). The conventional way the FTC deals with bait-and-switch schemes is to require that any offers a firm makes in order to attract consumers are *bona fide*, enforced in practice by not allowing firms to prevent the consumer from buying the product if they will, for example, by "refusal to show, demonstrate, or sell the product offered in accordance with the terms of the offer", "failure to have available at all outlets listed in the advertisement a sufficient quantity of the advertised product", or "use of a sales plan or method of compensation for salesmen or penalizing salesmen, designed to prevent or discourage them from selling the advertised product" (16 C.F.R. §238.3). However, because store-context induces consumers to switch products *at their own will* once they have entered the store, such conventional policy does not have any bite in our set-up and is ineffective in helping the naïve consumer population. More than that, directly regulating store-context seems a futile attempt due to the vast number of environmental variables that may affect preferences at the point of purchase. A more promising approach is to curb the "context-related" power of retailers by increasing the possibilities for consumers to purchase products without having to enter the store of a particular firm. While this is difficult to achieve in bricks-and-mortar markets, it is easy in online-markets. Policy makers could, for instance, require that firms advertising on product comparison websites such as Google Shopping or PriceCrawler must allow consumers to purchase offers *directly* on the comparison website. The viability of such a policy is demonstrated by the fact that the option to purchase an airline ticket without leaving the Google Flights search engine is already available in some cases.[19] Besides the apparent low cost of such a policy intervention, demanding that retailers make their products available on competitive platforms has other desirable features: First, the regulation is effective independently of the technology that is used to manipulate consumer preferences at the store. In fact, it is effective in cases where there is a "rational" (i.e., search cost) explanation for deceptive up-and down-selling stategies at the store level (see, e.g., Ellison and Wolitzky, 2012). In that sense, the proposed regulation is "model-neutral". Second, the regulation is "non-paternalistic" in that it helps naïve consumers without altering market supply for rational or sophisticated consumers. Third, the existence of an option to purchase products on a competitive platform without forbidding consumers to proceed to a firm-specific

[19]See https://support.google.com/flights/answer/6236307 (accessed 10-21-2016).

store for purchase (such as realized on the Google Flights search engine) gives researchers the possibility to study purchase decision in competitive as well as *in-store* situations and thus enables a better understanding of the effect the current retail market structure has on market outcomes.

References

Ariely, Dan. 2008. *Predictably Irrational.* New York:Harper Collins.

Ariely, Dan, George Loewenstein, and Drazen Prelec. 2003. "Coherent Arbitrariness: Stable Demand Curves Without Stable Preferences." *Quarterly Journal of Economics*, 118(1): 74–105.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2013. "Salience and Consumer Choice." *Journal of Political Economy*, 121(5): 803–843.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2016. "Competition for Attention." *Review of Economic Studies*, 83: 481–513.

Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein. 2016. "A Model of Relative Thinking." mimeo (This version: March 31, 2016).

Chetty, Ray. 2015. "Behavioral Economics and Public Policy: A Pragmatic Perspective." National Bureau of Economic Research Working Paper No. 20928, February 2015.

Chevalier, Judith A., Anil K. Kashyap, and Peter E. Rossi. 2003. "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data." *American Economic Review*, 93(1): 15–37.

DellaVigna, Stefano, and Ulrike Malmendier. 2004. "Contract Design and Self-Control: Theory and Evidence." *The Quarterly Journal of Economics*, 119(2): 353–402.

DellaVigna, Stefano, and Ulrike Malmendier. 2006. "Paying Not to Go to the Gym." *American Economic Review*, 96(3): 694–719.

Doyle, John R., David J. O'Connor, Gareth M. Reynolds, and Paul A. Bottomley. 1999. "The Robustness of the Asymmetrically Dominated Effect: Buying Frames, Phantom Alternatives, and In-Store Purchases." *Psychology and Marketing*, 16(3): 225–243.

Eliaz, Kfir, and Ran Spiegler. 2006. "Contracting with Diversly Naive Agents." *Review of Economic Studies*, 73: 689–714.

Ellison, Glen, and Sarah Fisher Ellison. 2009. "Search, Obfuscation, and Price Elasticities on the Internet." *Econometrica*, 77(2): 427–452.

Ellison, Glenn, and Alexander Wolitzky. 2012. "A Search Cost Model of Obfuscation." *RAND Journal of Economics*, 43(3): 417–441.

Gottlieb, Daniel. 2008. "Competition over Time-Inconsistent Consumers." *Journal of Public Economic Theory*, 10(4): 673–684.

Gul, Faruk, and Wolfgang Pesendorfer. 2001. "Temptation and Self-Control." *Econometrica*, 69(6): 1403–1435.

Heidhues, Paul, and Botond Kőszegi. 2010. "Exploiting Naivete about Self-Control in the Credit Market." *American Economic Review*, 100(5): 2279–2303.

Herne, Kaisa. 1999. "The Effects of Decoy Gambles on Individual Choice." *Experimental Economics*, 2: 31–40.

Hosken, Daniel, David Matsa, and David Reiffen. 2000. "How do Retailers Adjust Prices?: Evidence from Store-Level Data." Federal Trade Commission Bureau of Economics Working Paper No. 230, January 2000.

Huber, Joel, John W. Payne, and Christopher Puto. 1982. "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis." *Journal of Consumer Research*, 9(1): 90–98.

Jahedi, Salar. 2011. "A Taste for Bargains." mimeo.

Kamenica, Emir. 2008. "Contextual Inference in Markets: On the Informational Content of Product Lines." *American Economic Review*, 98(5): 2127–2149.

Kőszegi, Botond, and Adam Szeidl. 2013. "A Model of Focusing in Economic Choice." *The Quarterly Journal of Economics*, 128(1): 53–104.

Lal, Rajiv, and Carmen Matutes. 1994. "Retail Pricing and Advertising Strategies." *Journal of Business*, LXVII(1994): 345–370.

Mazar, Nina, Botond Kőszegi, and Dan Ariely. 2014. "True Context-Dependent Preferences? The Causes of Market-Dependent Valuations." *Journal of Behavioral Decision Making*, 27(4): 200–208.

Meyers-Levy, Joan, Rui Zhu, and Lan Jiang. 2010. "Context Effects from Bodily Sensations: Examining Bodily Sensations Induced by Flooring and the Moderating Role of Product Viewing Distance." *Journal of Consumer Research*, 37(1): 1057–1091.

Piccione, Michele, and Ran Spiegler. 2012. "Price Competition under Limited Comparability." *Quarterly Journal of Economics*, 127(1): 97–135.

Simonson, Itamar. 1989. "Choice Based on Reasons: The Case of Attraction and Compromise Effects." *Journal of Consumer Research*, 16(2): 158–174.

Simonson, Itamar, and Amos Tversky. 1992. "Choice in Context: Tradeoff Contrast and Extremeness Aversion." *Journal of Marketing Research*, 29(3): 281–295.

Spiegler, Ran. 2014. "Competitive Framing." *American Economic Journal: Microeconomics*, 6(3): 35–58.

Turley, L.W., and Ronald E. Milliman. 2000. "Atmospheric Effects on Shopping Behavior: A Review of the Experimental Evidence." *Journal of Business Research*, 49: 193–211.

Wolfe, Jeremy M., and Todd S. Horowitz. 2004. "What Attributes Guide the Development of Visual Attention and How Do They Do It?" *Nature: Reviews Neuroscience*, 5: 495–501.

Appendix: Proofs to the Results in the Main Text

We use the following method throughout all proofs to find market supply in the competitive equilibrium: First, we derive the best response of some firm k to a fixed competitor offer $(\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})$ *conditional* on attracting a positive share of consumers under the assumption that the maximum price b consumers are able to pay is arbitrarily large, i.e., $b \to \infty$. In general, this best response will be unique and continuous in $(\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})$. Due to this characteristic, in a second step, we can find the competitive market supply by searching for the competitor offer $(\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})$ that equates the profits of this response to zero. At this point, firms that supply the market will sell a cost-efficient quality (q^*, q^Q, or q^P) at cost, making zero profit. When we drop the assumption $b \to \infty$, consumers will always buy such a product if $b \geq c(q^Q) > c(q^*) > c(q^P)$, which holds by section II.C. of this paper. The (interior) solution we define using this method is thus valid without the assumption $b \to \infty$. Moreover, firms who do not supply the market must always choose $(M^k, \Theta^k) = \emptyset$, because this is the only response that avoids any costs and yields nonnegative profits. While supplying the market at cost and choosing $(M^k, \Theta^k) = \emptyset$ both yield zero profits and are thus best responses, in equilibrium, at least 2 firms must choose to supply the market. Otherwise there would exist some firm k that faced only competitors choosing $(M^k, \Theta^k) = \emptyset$, making a deviation to monopoly profits possible. In general, we therefore have a range of competitive equilibria that all result in the same market supply: At least 2 firms share the market and sell at cost, while all other firms choose $(M^k, \Theta^k) = \emptyset$.

Proof of Lemma 1 (Rational Benchmark). Let $\beta = 1$. This implies that consumers are homogeneous and have time-consistent surplus function $u_j = q_j - p_j$. Context leaves valuations unaffected, $\theta_j^k = N$ for all j and k.

Consider some firm k and fix the competitor offer $\boldsymbol{M}^{-k}$. Let $\bar{u} \geq 0$ be the maximum surplus attainable outside of firm k (this surplus is implicitly defined by $\boldsymbol{M}^{-k}$ and the outside option of no purchase). Let $b \to \infty$ and consider the best response *conditional* on attracting a positive share of consumers. Fix some quality $q_j \geq \underline{q}$. The firm can sell q_j to all consumers at price $p_j = \lim_{\delta \to 0}(q_j - \bar{u} - \delta) = q_j - \bar{u}$. At this price, the firm offers just enough surplus to let consumers marginally improve over the highest surplus available elsewhere, thereby winning all consumers. For given quality q_j, no other price can achieve higher profits: A higher price implies the loss of all consumers, a lower price cannot attract more. This price implies profit $\pi^k = q_j - \bar{u} - c(q_j)$ and thus, the profit-maximizing quality to sell is $q^* := \arg\max[q - c(q)]$, or $c'(q^*) = 1$. Note that $q^* > \underline{q}$ by assumption, making this interior solution valid. Offering additional products is costly and cannot increase profits. It follows: Conditional on attracting a positive share of consumers, the unique best response is the product-line $M^k = ((q^*, q^* - \bar{u}))$. Note that the best response so defined is unique and continuous in $\bar{u}$. Market supply in the competitive equilibrium can thus be found by searching for $\bar{u}$ where this response yields zero profits. This unique point exists at $\bar{u} = q^* - c(q^*)$, implying marginal cost pricing, $p_j = c(q^*)$ and the product-line $M^* = ((q^*, c(q^*))$. This solution is valid by our model assumption $b > c(q^*)$, such that we can drop the assumption $b \to \infty$.

Given that some firm offers $M^* = ((q^*, c(q^*))$, other firms face $\bar{u} = q^* - c(q^*)$. There are two best responses: (1) Sell $M^* = ((q^*, c(q^*))$ as well, which yields zero profits, (2) Offer nothing, $M^k = \emptyset$, which is the only response avoiding all costs and also yields zero profits. In any equilibrium, at least 2 firms must offer the product-line M^*: If no firm offered M^*, then any firm would face an outside option $\bar{u} = 0 < q^* - c(q^*)$ and there would exist a deviation incentive to monopoly profits. If only one firm offered M^*, then, similarly, this firm could earn monopoly profits by deviating. We thus have a range of competitive equilibria that all result in the same market supply: At least 2 firms share the market and offer M^*, while all other firms choose $M^k = \emptyset$.

■

Proof of Lemma 2 (Fooling Regimes). Assume that the firm fools the consumer, selling product $t \in J^k$, but attracting the consumer with another product $a \in J^k$, $a \neq t$. Then

$$\text{(by IC)} \qquad \hat{u}_t^k \geq \hat{u}_a^k$$

$$\text{(by PCC)} \qquad E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] \leq E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right]$$

with at least one inequality strict: incentive compatibility (IC) requires that the consumer weakly prefers the target at the store, the perceived choice constraint (PCC) requires that the consumer *expects* to weakly prefer the attraction product at the store. To induce the consumer to switch products, at least one inequality needs to be strict. Fix θ_a and θ_t. If $(\theta_a, \theta_a) = (N, N)$, $E_{\tilde{\beta}^0}\left[\hat{u}_j^k\right] = \hat{u}_j^k = u_j$ for $j \in \{a, b\}$ and any $\tilde{\beta}^0$ such that both, IC and PCC hold with equality, a contradiction. For at least one inequality to be strict, $(\theta_a, \theta_a) \neq (N, N)$: if the consumer is fooled, store-context distorts the valuation of at least one of the products. Also, if $\tilde{\beta}^0 = \beta$, then $E_{\tilde{\beta}^0}\left[\hat{u}_j^k\right] = \hat{u}_j^k$ for any distortion (θ_a, θ_t). Again, IC and PCC cannot hold with one inequality being strict. Thus, if the consumer is fooled, she must be naïve regarding contextual distortions, that is, $\tilde{\beta}^0 \neq \beta$.

Assume for the rest of the proof that $(\theta_a, \theta_t) \neq (N, N)$ and $\tilde{\beta}^0 \neq \beta$. Define the function $v_j(\gamma) := \hat{u}_j^k|_{\beta=\gamma}$. Note that $v_j(1) = u_j$, $v_j(\beta) = \hat{u}_j^k$ and $v_j(\tilde{\beta}^0) = E_{\tilde{\beta}^0}\left[\hat{u}_j^k\right]$. We can rewrite the fooling constraints as

$$\text{(by IC)} \qquad \hat{u}_t^k \geq \hat{u}_a^k \Leftrightarrow v_t(\beta) \geq v_a(\beta)$$

$$\text{(by PCC)} \qquad E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] \leq E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right] \Leftrightarrow v_t(\tilde{\beta}^0) \leq v_a(\tilde{\beta}^0).$$

The two conditions imply that if the consumer is fooled, $\exists \gamma^0 \in [\min\{\tilde{\beta}^0, \beta\}, \max\{\tilde{\beta}^0, \beta\}]$ s.t. $v_t(\gamma^0) - v_a(\gamma^0) = 0$ (a point where products a and t yield identical surplus). We first want to show that this crossing is unique. For this, note that for a given distortion θ_j, $\partial v_j / \partial \gamma = const$. Thus, $\partial[v_t(\gamma) - v_a(\gamma)]/\partial\gamma = const$. Because at least one inequality is strict, $\partial[v_t(\gamma) - v_a(\gamma)]/\partial\gamma \neq 0$. Thus, if a crossing exists, it must be unique. It follows:

a) If the consumer is *pessimistic* ($\tilde{\beta}^0 > \beta$) and fooled, then $\exists! \gamma^0 \in [\beta, \tilde{\beta}^0]$ s.t. $v_t(\gamma^0) - v_a(\gamma^0) = 0$. By IC, $v_t(\beta) - v_a(\beta) \geq 0$ and thus, $v_t(\gamma^0) - v_a(\gamma^0) < 0 \; \forall \gamma > \gamma^0$ and $v_t(\gamma^0) - v_a(\gamma^0) > 0 \; \forall \gamma < \gamma^0$.

   This implies that

   - $\forall \tilde{\beta} > \gamma^0$, $v_t(\tilde{\beta}) - v_a(\tilde{\beta}) < 0 \Leftrightarrow E_{\tilde{\beta}}\left[\hat{u}_t^k\right] - E_{\tilde{\beta}}\left[\hat{u}_a^k\right] < 0$: all *pessimistic* agents with $\tilde{\beta} > \gamma^0$ (falsely) expect to prefer product a over product t at the store and are therefore *also* fooled by the pair (a, t). If $\gamma^0 = \beta$, *all* pessimistic agents are fooled.
   - $\forall \tilde{\beta} < \beta \leq \gamma^0$, $v_t(\tilde{\beta}) - v_a(\tilde{\beta}) > 0 \Leftrightarrow E_{\tilde{\beta}}\left[\hat{u}_t^k\right] - E_{\tilde{\beta}}\left[\hat{u}_a^k\right] > 0$: all *optimistic* agents (correctly) expect to prefer product t over product a at the store and are therefore *not* fooled by the pair (a, t).
   - $v_t(1) - v_a(1) > 0 \Leftrightarrow u_t - u_a > 0$ by $\gamma^0 > 1$: the target yields higher undistorted surplus than the attraction product.

b) If the consumer is *optimistic* ($\tilde{\beta}^0 < \beta$) and fooled, then $\exists! \gamma^0 \in [\tilde{\beta}^0, \beta]$ s.t. $v_t(\gamma^0) - v_a(\gamma^0) = 0$. By IC, $v_t(\beta) - v_a(\beta) \geq 0$, and thus, $v_t(\gamma^0) - v_a(\gamma^0) > 0 \; \forall \gamma > \gamma^0$ and $v_t(\gamma^0) - v_a(\gamma^0) < 0 \; \forall \gamma < \gamma^0$.

   This implies that

   - $\forall \tilde{\beta} < \gamma^0$, $v_t(\tilde{\beta}) - v_a(\tilde{\beta}) < 0 \Leftrightarrow E_{\tilde{\beta}}\left[\hat{u}_t^k\right] - E_{\tilde{\beta}}\left[\hat{u}_a^k\right] < 0$: all optimistic agents with $\tilde{\beta} < \gamma^0$ (falsely) expect to prefer product a over t at the store and are therefore *also* fooled by the pair (a, t). If $\gamma^0 = \beta$, *all* optimistic agents are fooled.
   - $\forall \tilde{\beta} > \beta \geq \gamma^0$, $v_t(\tilde{\beta}) - v_a(\tilde{\beta}) > 0 \Leftrightarrow E_{\tilde{\beta}}\left[\hat{u}_t^k\right] - E_{\tilde{\beta}}\left[\hat{u}_a^k\right] > 0$: all *pessimistic* agents (correctly) expect to prefer product t over product a at the store and are therefore *not* fooled by the pair (a, t).

- $v_t(1) - v_a(1) < 0 \Leftrightarrow u_t - u_a < 0$ by $\gamma^0 > 1$: the target yields lower undistorted surplus than the attraction product.

■

Proof of Lemma 3 (Profitable Fooling). Fix $(\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})$. $(\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})$ implies an outside option with surplus $\bar{u}(\tilde{\beta}^0) \geq 0$ for a consumer of type $\tilde{\beta}^0$. Assume $\underline{q} \to 0$ and $b \to \infty$. Fix any target quality $q_t = q_t^0 \geq 0$. If the firm does not fool, the consumer correctly expects to purchase the target when entering firm k. Thus, conditional on not fooling, the maximum selling price for quality q_t^0 is

$$p_t^0 := q_t^0 - \bar{u}(\tilde{\beta}^0).$$

For example, the firm could only offer product t (and no other product). Then consumers of type $\tilde{\beta}^0$ enter the store of firm k if $u_t \geq \bar{u}(\tilde{\beta}^0)$. Given quality $q_t = q_t^0$ and price p_t^0, this condition holds with equality. Formally, the firm can achieve that type $\tilde{\beta}^0$ enter the store with certainty by choosing $p_t = \lim_{\delta \to 0}[q_t^0 - \bar{u}(\tilde{\beta}^0) - \delta] = p_t^0$.

Part a) (If type $\tilde{\beta}^0$ is pessimistic ($\tilde{\beta}^0 > \beta$), fooling her is *unprofitable*). Assume that firm k fools type $\tilde{\beta}^0$. We will first show that fooling a pessimistic type ($\tilde{\beta}^0 > \beta$) is unprofitable. For this, we will derive an upper bound on the price for a given target quality q_t^0, $\bar{p}_t(q_t^0)$ (conditional on fooling the consumer and selling her target $t \neq a$), and show that this bound is lower than the price p_t^0.

Consider stage 2, i.e., the decision of the consumer of what product $j \in J^k$ to purchase after she has entered the store of firm k. A lower bound on the (context-dependent) surplus of the target is given by $\hat{u}_t^k = \hat{u}_a^k$: Lowering $\hat{u}_t^k$ by charging a higher price p_t or offering a lower quality q_t will make the consumer choose product a over t, violating incentive compatibility. Rewriting $\hat{u}_t^k = \hat{u}_a^k$ as $(\hat{u}_t^k - u_t) + u_t = (\hat{u}_a^k - u_a) + u_a \Leftrightarrow (\hat{u}_t^k - u_t) + q_t - p_t = (\hat{u}_a^k - u_a) + u_a$ and solving this expression for p_t and q_t, respectively, yields as an upper bound on price:

$$p_t = q_t - u_a + (u_a - u_t), \tag{6}$$

Now consider stage 1. Condition $\hat{u}_t^k = \hat{u}_a^k$ implies that if the consumer is fooled, $E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] < E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right]$: the consumer expects to *strictly* prefer product a over t at store k. She enters the store iff the *expected* purchase—i.e., product a—yields undistorted surplus that is as least as high as her outside option, i.e., $u_a \geq \bar{u}(\tilde{\beta}^0)$. Consider the bound on p_t as defined in Equations (6): Clearly, this bound is maximized if the participation constraint $u_a \geq \bar{u}(\tilde{\beta}^0)$ binds, i.e., if $u_a = \bar{u}(\tilde{\beta}^0)$. We conclude: conditional on fooling and selling target $t \neq a$ to a consumer of type $\tilde{\beta}^0$, an upper bound on the price for given target quality q_t^0 is given by

$$\bar{p}_t(q_t) := q_t - \bar{u}(\tilde{\beta}^0) + (u_a - u_t).$$

It is now easy to see that fooling pessimistic types is not profitable. By Lemma 2, if a pessimistic type is fooled, $u_a < u_t$: It follows from $\bar{p}_t(q_t)$ and $\underline{q}_t(p_t)$ that if the firm sells $t \neq a$, then it must charge a lower price $p_t < p_t^0$ for quality q_t^0. This concludes the proof for part a).

Part b) (If type $\tilde{\beta}^0$ is optimistic ($\tilde{\beta}^0 < \beta$), fooling her is *profitable*). Turning to the case of an optimistic type, note first that if $\tilde{\beta}^0 < \beta$, then by Lemma 2, $(u_t - u_a) < 0 \Leftrightarrow (u_a - u_t) > 0$ and thus, $\bar{p}_t(q_t^0) > p_t^0$. This suggests that fooling a pessimistic type may be profitable. We will now show that this is indeed the case iff store-context distorts the surplus of products a and t as described in the lemma. Given a distortion (θ_a, θ_t), we can rewrite IC and PCC as conditions on the attributes of products a and t:

- Assume that $(\theta_a, \theta_t) = (Q, Q)$. If $\tilde{\beta}^0 < \beta$, the following two statements are equivalent:

  1. If the consumer is fooled,

     $$\begin{aligned} &\text{(by IC)} && \hat{u}_t^k \geq \hat{u}_a^k \Leftrightarrow \beta(q_t - q_a) \geq p_t - p_a \\ &\text{(by PCC)} && E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] \leq E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right] \Leftrightarrow \tilde{\beta}^0(q_t - q_a) \leq p_t - p_a, \end{aligned}$$

     with at least one inequality strict.

  2. If the consumer is fooled, $q_t > q_a$ and $p_t > p_a$ (the firm up-sells).

- Assume that $(\theta_a, \theta_t) = (P, P)$. If $\tilde{\beta}^0 < \beta$, the following two statements are equivalent:

  1. If the consumer is fooled,

     $$\begin{aligned} &\text{(by IC)} && \hat{u}_t^k \geq \hat{u}_a^k \Leftrightarrow \beta(p_a - p_t) \geq q_a - q_t \\ &\text{(by PCC)} && E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] \leq E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right] \Leftrightarrow \tilde{\beta}^0(p_a - p_t) \leq q_a - q_t, \end{aligned}$$

     with at least one inequality strict.

  2. If the consumer is fooled, $q_t < q_a$ and $p_t < p_a$ (the firm down-sells).

- Assume that $(\theta_a, \theta_t) = (P, Q)$. If $\tilde{\beta}^0 < \beta$, the following two statements are equivalent:

1. If the consumer is fooled,

$$\text{(by IC)} \qquad \hat{u}_t^k \geq \hat{u}_a^k \Leftrightarrow \beta(q_t + p_a) \geq q_a + p_t$$
$$\text{(by PCC)} \qquad E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] \leq E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right] \Leftrightarrow \tilde{\beta}^0(q_t + p_a) \leq q_a + p_t,$$

   with at least one inequality strict.

2. If the consumer is fooled, $q_t + p_a > 0$ and $q_a + p_t > 0$.

- Assume that $(\theta_a, \theta_t) = (P, N)$. If $\tilde{\beta}^0 < \beta$, the following two statements are equivalent:

1. If the consumer is fooled,

$$\text{(by IC)} \qquad \hat{u}_t^k \geq \hat{u}_a^k \Leftrightarrow \beta p_a \geq q_a - q_t + p_t$$
$$\text{(by PCC)} \qquad E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] \leq E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right] \Leftrightarrow \tilde{\beta}^0 p_a \leq q_a - q_t + p_t,$$

   with at least one inequality strict.

2. If the consumer is fooled, $p_a > 0$ and $q_a - q_t + p_t > 0$.

- Assume that $(\theta_a, \theta_t) = (Q, P)$. If $\tilde{\beta}^0 < \beta$, the following two statements are equivalent:

1. If the consumer is fooled,

$$\text{(by IC)} \qquad \hat{u}_t^k \geq \hat{u}_a^k \Leftrightarrow \beta(q_a + p_t) \leq q_t + p_a$$
$$\text{(by PCC)} \qquad E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] \leq E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right] \Leftrightarrow \tilde{\beta}^0(q_a + p_t) \geq q_t + p_a,$$

   with at least one inequality strict.

2. If the consumer is fooled, $q_a + p_t < 0$ and $q_t + p_a < 0$.

- Assume that $(\theta_a, \theta_t) = (Q, N)$. If $\tilde{\beta}^0 < \beta$, the following two statements are equivalent:

1. If the consumer is fooled,

$$\text{(by IC)} \qquad \hat{u}_t^k \geq \hat{u}_a^k \Leftrightarrow \beta q_a \leq q_t - p_t + p_a$$
$$\text{(by PCC)} \qquad E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] \leq E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right] \Leftrightarrow \tilde{\beta}^0 q_a \geq q_t - p_t + p_a,$$

   with at least one inequality strict.

2. If the consumer is fooled, $q_a < 0$ and $q_t - p_t + p_a < 0$.

- Assume that $(\theta_a, \theta_t) = (N, P)$. If $\tilde{\beta}^0 < \beta$, the following two statements are equivalent:

1. If the consumer is fooled,

$$\text{(by IC)} \qquad \hat{u}_t^k \geq \hat{u}_a^k \Leftrightarrow \beta p_t \leq q_t - q_a + p_a$$
$$\text{(by PCC)} \qquad E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] \leq E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right] \Leftrightarrow \tilde{\beta}^0 p_t \geq q_t - q_a + p_a,$$

   with at least one inequality strict.

2. If the consumer is fooled, $p_t < 0$ and $q_t - q_a + p_a < 0$.

It is obvious from the conditions derived above that any distortion $(\theta_a, \theta_t) \in \{(Q,P),(Q,N),(N,P)\}$ cannot lead to a profitable fooling outcome. In particular, fooling with any one of these distortions requires that either, $q_a < 0$, or $p_t < 0$, or both. But if $q_a < 0$, the attraction product has below minimum quality $\underline{q}$ and can therefore not attract consumers to the store, while if $p_t < 0$, the firm would sell the target strictly below marginal cost and make negative profit. We conclude: If store-context asymmetrically distorts context in favor of the attraction product, $(\theta_a, \theta_t) \in \{(Q,P),(Q,N),(N,P)\}$, the consumer cannot be fooled to purchase target t at any $p_t \geq 0$.

It remains to be shown that fooling is profitable if $(\theta_a, \theta_t) \in \{(Q,Q),(P,P),(P,Q),(P,N),(E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right] N, Q)\}$. In particular, we will show that with any one of these distortions, the firm can indeed sell quality q_t^0 at price $\bar{p}_t(q_t^0) \geq p_t^0$. W.l.o.g, we assume for the rest of the proof that either a and t are the only products that firm k offers, or that other existing products do not violate IC and PCC, that is, $\forall j \in J^k, j \notin \{a,t\}$, $\hat{u}_j^k < \hat{u}_t^k$ and $E_{\tilde{\beta}^0}\left[\hat{u}_j^k\right] < E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right]$. Fix some $q_t^0 > 0$ and $p_t = \bar{p}_t(q_t^0)$. Recall that the construction of $\bar{p}_t(q_t^0)$ implies that IC is satisfied with equality, i.e, $\hat{u}_t^k = \hat{u}_a^k$. Formally, the firm chooses q_j and p_j from the discretized set of real numbers $\bar{\mathbb{R}}(\delta)$ where $\delta \to 0$ and achieves that the consumer purchases the target with certainty by choosing $\hat{u}_t^k$ arbitrarily close but above $\hat{u}_a^k$, i.e., $\hat{u}_t^k = \lim_{\delta\to 0}(\hat{u}_a^k + \delta) = \hat{u}_a^k$. Similarly, the construction of $\bar{p}_t^0$ implies that $u_a = \bar{u}(\tilde{\beta}^0)$: If the consumer is fooled, she expects to receive surplus identical to her outside option $\bar{u}(\tilde{\beta}^0)$. Again, by choosing u_a arbitrarily close but above $\bar{u}(\tilde{\beta}^0)$, $u_a = \lim_{\delta\to 0}(u_a + \delta) = \bar{u}(\tilde{\beta})$, the firm can guarantee that the consumer enters its store with certainty. To prove that the firm can sell q_t at $p_t = \bar{p}_t(q_t^0)$ it remains be shown that—given distortion (θ_a, θ_t)—there exists an attraction product with $q_a > 0$ s.t. $E_{\tilde{\beta}^0}\left[\hat{u}_t^k\right] < E_{\tilde{\beta}^0}\left[\hat{u}_a^k\right]$: the consumer expects to strictly prefer product a over t. Note that with q_t and p_t being fixed, $\hat{u}_t^k = \hat{u}_a^k$ determines a one-to-one function between q_a and p_a. We are thus left with only one degree of freedom. Consider the three possible cases listed in the lemma:

- Assume that $(\theta_a, \theta_t) = (Q,Q)$. PCC holds with strict inequality iff $q_t > q_a$ and $p_t > p_a$. Pick $q_a \in (0, q_t^0)$, which exists by construction. For example, choose $q_a = \underline{q}$. Then

$q_a < q_t$ and by $\hat{u}_t^k = \hat{u}_a^k \Leftrightarrow p_a = p_t - \beta(q_t - q_a) < p_t$: PCC holds with strict inequality. (q.e.d.)

- Assume that $(\theta_a, \theta_t) = (P, P)$. PCC holds with strict inequality iff $p_t < p_a$ and $q_t < q_a$. Pick $p_a > \bar{p}_t(q_t^0)$, which exists by construction. For example, choose $p_a = b$. Then $p_a > p_t$ and by $\hat{u}_t^k = \hat{u}_a^k \Leftrightarrow q_a = q_t + \beta(p_a - p_t) > q_t$: PCC holds with strict inequality. (q.e.d.)

- Assume that $(\theta_a, \theta_t) \in \{(P, Q), (P, N), (N, Q)\}$.

  1. If $(\theta_a, \theta_t) = (P, Q)$, PCC holds with strict inequality iff $q_t + p_a > 0$ and $q_a + p_t > 0$. Pick $p_a > 0$ sufficiently large, s.t. by $\hat{u}_t^k = \hat{u}_a^k \Leftrightarrow q_a = \beta(q_t^0 + p_a) - \bar{p}_t(q_t^0) > 0$. For example, choose $p_a = b$. Then $q_t + p_a > 0$ and $q_a + p_t > 0$ by construction: PCC holds with strict inequality. (q.e.d.)
  2. If $(\theta_a, \theta_t) = (P, N)$, PCC holds with strict inequality iff $p_a > 0$ and $q_a - q_t + p_t > 0$. Pick $p_a > 0$ sufficiently large, s.t. by $\hat{u}_t^k = \hat{u}_a^k \Leftrightarrow q_a = \beta p_a + q_t - \bar{p}_t(q_t^0) > 0$. For example, choose $p_a = b$. Then $p_a > 0$ and $q_a - q_t + p_t = \beta p_a > 0$ by construction: PCC holds with strict inequality. (q.e.d.)
  3. If $(\theta_a, \theta_t) = (N, Q)$, PCC holds with strict inequality iff $q_t > 0$ and $q_a - p_a + p_t > 0$. Pick $p_a > 0$ sufficiently large, s.t. by $\hat{u}_t^k = \hat{u}_a^k \Leftrightarrow q_a = p_a + \beta q_t - \bar{p}_t(q_t^0) > 0$. For example, choose $p_a = b$. Then $q_t > 0$ and $q_a - p_a + p_t = \beta q_t > 0$ by construction: PCC holds with strict inequality. (q.e.d.)

This concludes the proof for part b). ■

Proof of Proposition 1 (Sophistication/pessimism induces the rational outcome). Let $\beta > 1$ (consumers are context-sensitive). Assume that $\tilde{\beta} \geq \beta$ for all consumers.

We first derive the (unique) best response for a generic firm k *conditional* on attracting a positive share of consumers under the assumption that $b \to \infty$. Fix the competitor offer $(\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})$ and let $\bar{u}(\tilde{\beta}) \geq 0$ be type $\tilde{\beta}$'s *expected* maximum surplus attainable outside of firm k. By Lemma 3, fooling pessimistic types $\tilde{\beta} > \beta$ is not profitable: The firm can sell any quality $q_t \geq \underline{q}$ at a strictly higher price if $a^k(\tilde{\beta}) = t^k$ for all $\tilde{\beta}$. Also, sophisticated consumers cannot be fooled. It follows that if $\tilde{\beta} \geq \beta$ for all consumers, a firm can sell any target t^k at a strictly higher profit if it does not fool. So assume that the firm does not fool. Then all consumers correctly expect to buy the target when entering firm k and the demand of firm k depends entirely on the characteristics of this target, q_t and p_t. Let $D(q_t, p_t) \in [0, 1]$ be the corresponding demand function of firm k. The profit of firm k is then $\pi^k(q_t, p_t) = D(q_t, p_t)[p_t - c(q_t)]$ and depends only on the characteristics of product t:

Offering more than this product is unnecessary yet costly and cannot be part of the best response. Fix quality q_t and p_t at a strictly positive demand $\bar{D} = D(\bar{q}_t, \bar{p}_t)$. Note that $\bar{D} = D(\bar{q}_t, \bar{p}_t) = D(\bar{q}_t + \Delta q, \bar{p}_t + \Delta q) > 0$, for any increment in quality $\Delta q \in \mathbb{R}$, because $\bar{u}_t = \bar{q}_t + \Delta q - (\bar{p}_t + \Delta q) = \bar{q}_t - \bar{p}_t$. Solving for the profit-maximizing Δq, $\arg\max_{\Delta q} \pi^k(\Delta q) = \bar{D}[\bar{p}_t + \Delta q - c(\bar{q}_t + \Delta q)]$, yields the condition $c'(\bar{q}_t + \Delta q) = 1$. In other words, for any positive demand $\bar{D}$, the profit-maximizing quality to sell is defined by $c(q_t) = 1$, i.e., $q_t = q^*$. This interior solution is valid by assumption that $\underline{q} < q^*$. Because profits (conditional on not fooling) depend on the outside valuation (undistorted surplus) u_t, distorting context at the store is unnecessary yet costly and cannot be part of the best response. It follows that the unique best response is to offer one undistorted product with quality $q_t = q^*$, $\theta_t = N$ at price $p_t = \arg\max_{p \in \mathbb{R}} D(q^* - p)[p - c(q^*)]$ (and no other products). Market supply in any equilibrium must follow this rule: If a firm with a positive market share would choose differently,—by the uniqueness of the best response derived above—there would exist a deviation incentive. The only other response that can be profit-maximizing is to choose $(M^k, \Theta^k) = \emptyset$, i.e., to not supply any products, which yields zero profits. Note that best response behavior is *near-identical* to the rational benchmark: Firms behave *as if* consumers were rational but (possibly) heterogeneous in their outside options $\bar{u}(\tilde{\beta})$. However, at any point of mutual best response, $\bar{u}(\tilde{\beta}) = \bar{u} \; \forall \tilde{\beta} \geq \beta$: If no firm fools, all consumers must expect to yield the same maximum surplus. Once $\bar{u}$ is unique, the unique best response conditional on attracting a positive share of consumers collapses to $(M^k, \Theta^k) = ((q^*, q^* - \bar{u}), (\theta_t^k = N))$— identical to the best response in the rational benchmark. Hence, the competitive equilibrium must conform to the equilibrium derived in Lemma 1. The remainder of the proof is identical to the second part of the proof of Lemma 1 and is therefore omitted. ■

Proof of Proposition 2 (Fooling Equilibrium).

Part a) (Store-Wide Distortions). We begin the proof by considering a perfectly homogeneous, optimistic consumer population with unique type $\tilde{\beta}^0 < \beta$. Consider a generic firm k. Fix the competitor offer $(\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})$ and let $\bar{u} = \bar{u}(\tilde{\beta}^0) \geq 0$ be type $\tilde{\beta}^0$'s *expected* maximum surplus attainable outside of firm k. Assume that for any two products j, i at firm k, $\theta_j^k = \theta_i^k = \theta^k$ and the firm chooses $\theta^k \in \{Q, P, N\}$. Assume (for now) that $b \to \infty$. Consider the best response *conditional* on attracting a positive share of consumers. By Lemma 3, the best response will involve fooling and the distortion of context. This yields strictly higher profits than not fooling and choosing $\theta^k = N$. Hence, the best response will involve choosing either $\theta^k = Q$ or $\theta^k = P$. We will now derive the two equilibrium candidates that derive from assuming either $\theta^k = Q$ or $\theta^k = P$.

- Assume that $\theta^k = Q$. The maximum price the firm can sell any target quality q_t is given by the upper bound $\bar{p}_t(q_t)$ which we have derived in the proof of Lemma 3. To achieve $\bar{p}_t(q_t)$, offering a second product $a \neq t$ is necessary and sufficient. Holding more than 2 products is unnecessary yet costly and can thus not be part of the best response. If $\theta^k = Q \Rightarrow (\theta_a, \theta_t) = (Q, Q)$, by (the proof of) Lemma 3, the firm sells q_t at $p_t = \bar{p}_t(q_t)$ iff it chooses $q_a < q_t$ and $p_a < p_t$. If $(\theta_a, \theta_t) = (Q, Q)$, $\bar{p}_t(q_t)$ can be rewritten as

$$\bar{p}_t(q_t, q_a, p_a) = \beta(q_t - q_a) + p_a,$$

under the condition $q_a - p_a = \bar{u}$ (the participation constraint binds). To find the best response, we need to choose q_t and (q_a, p_a) such that profit at this price is maximized. Consider the choice of q_t first. Because quality q_t is inflated by a factor β when $(\theta_a, \theta_t) = (Q, Q)$, it is easy to see that the cost-efficient quality to sell is

$$q_t = q^Q := \arg\max_q [\beta q - c(q)] \Leftrightarrow c'(q^Q) = \beta.$$

This interior solution is valid by assumption $q^Q > \underline{q}$. We are left with the choice of the attraction product (q_a, p_a). Maximizing profit for any q_t implies maximizing $\bar{p}_t(q_t, q_a, p_a)$ under the constraint $q - p_a = \bar{u}$. There are 2 opposing forces: Minimizing q_a and maximizing p_a. The profit-maximizing choice is to minimize q_a: Because quality q_a is inflated at the store, the positive effect on profits of decreasing quality q_a is larger than the positive effect of increasing price p_a. The unique profit-maximizing choice is therefore to choose $q_a = \underline{q}$, which implies $p_a = \underline{q} - \bar{u}$. Note that this choice satisfies the fooling conditions $q_a < q_t$ and $p_a < p_t$ for any $q_t > 0$. For later reference, note the "marketing" implications of this best-response: To attract consumers, the firm fixes attraction quality $q_a = \underline{q}$ and competes with other firms on the price of this low-quality product. We conclude: Conditional on $\theta^k = Q$, the best response in the domain of positive profits is unique and continuous: the firm offers 2 products, t and $a \neq t$, with $(q_t, p_t) = (q^Q, \bar{p}_t(q^Q))$ and $(q_a, p_a) = (\underline{q}, \underline{q} - \bar{u})$.

- Assume that $\theta^k = P$. Analogously to the case of $\theta^k = Q$, we find the best response by maximizing profit at price $\bar{p}_t(q_t)$ which we can now express as

$$\bar{p}_t(q_t, q_a, p_a) = p_a - \frac{1}{\beta}(q_a - q_t)$$

under the condition $q_a - p_a = \bar{u}$ (the participation constraint binds). 2 products, t and $a \neq t$ are necessary and sufficient to yield this maximum price for any target quality q_t. Holding more products cannot be part of a best response. By (the proof of) Lemma 3, the firm sells q_t at $p_t = \bar{p}_t(q_t)$ iff it chooses $q_a > q_t$ and $p_a > p_t$. With price of the target being inflated at the store, the cost-efficient quality to sell is

$$q_t = q^P := \arg\max_q \left[q - \beta c(q)\right] \Leftrightarrow c'(q^P) = \frac{1}{\beta}.$$

This interior solution is valid by assumption $q^P \geq \underline{q}$. Maximizing profit for any q_t implies maximizing $\bar{p}_t(q_t, q_a, p_a)$ under the constraint $q_a - p_a = \bar{u}$. There are 2 opposing forces: Minimizing q_a and maximizing p_a. Contrary to the case of $\theta^k = Q$, the profit-maximizing choice now is to maximize p_a: Because price p_a is inflated at the store, the positive effect on profits of increasing price p_a is larger than the positive effect of decreasing quality q_a. The unique profit-maximizing choice is therefore to choose $p_a = b$, which implies $q_a = b + \bar{u}$. Note that this choice satisfies the fooling conditions $p_a > p_t$ and $q_a > q_t$ for any $p_t < b$. For later reference, note the "marketing" implications of this best-response: To attract consumers, the firm fixes attraction price $p_a = b$ and competes with other firms on the quality of this high-price product. We conclude: Conditional on $\theta^k = P$, the best response in the domain of positive profits is unique and continuous: the firm offers 2 products, t and $a \neq t$, with $(q_t, p_t) = (q^P, \bar{p}_t(q^P))$ and $(q_a, p_a) = (b + \bar{u}, b)$.

Note that the best response in both cases is independent of the degree of naïveté of type $\tilde{\beta}^0 < \beta$: Due to the optimality condition $\hat{u}_t^k = \hat{u}_a^k$ (the IC binds), any consumer with belief $\tilde{\beta} < \beta$ (falsely) believes to purchase product a with certainty. The best response does *not* generate heterogeneous expectations among a purely optimistic consumer population. If firms play mutual best responses, any heterogeneity in types $\tilde{\beta}$ is therefore rendered unimportant for market supply: Uniqueness of the best response (given a distortion θ^k) implies that firms generating positive demand must choose according to it; otherwise, there would exists a strict deviation incentive. This response does not generate heterogeneous expectations. Firms not generating positive demand, on the other hand, choose $(M^k, \Theta^k) = \emptyset$ to avoid positive costs and thus negative profits. These firms do not generate heterogeneous expectations either. It follows that in any equilibrium, $\bar{u}(\tilde{\beta}) = \bar{u}\ \forall \tilde{\beta} < \beta$: the outside option is a unique value. We can find market supply in the competitive equilibrium by searching for $\bar{u}$ that equates the best response profits to zero. This yields the following two candidates for equilibrium market supply:

$$(Q^*) \qquad \theta^k = Q,\ (q_t, p_t) = (q^Q, c(q^Q)),\ (q_a, p_a) = (\underline{q}, c(q^Q) - \beta(q^Q - \underline{q}))$$

$$(P^*) \qquad \theta^k = P,\ (q_t, p_t) = (q^P, c(q^P)),\ (q_a, p_a) = (q^P + [b - c(q^P)], b)$$

By reasoning analogous to the second part of the proof of Lemma 1, at least 2 firms must provide a product-line according to (Q^*) or (P^*). These firms share the market. All other firms choose $(M^k, \Theta^k) = \emptyset$. Fix an equilibrium where at least one firm chooses (M^k, θ^k) according to (Q^*). Then there must be at least one other firm that provides the same expected surplus $\bar{u} = u_a = \underline{q} - c(q^Q) + \beta(q^Q - \underline{q})$. Otherwise, the firm would have a deviation incentive to strictly positive profits. What remains to be checked is a deviation towards the other regime $\theta^k = P$, where the maximum profit is given by the unique best response defined above. In other words, the firm may offer 2 products, t and $a \neq t$, with $(q_t, p_t) = (q^P, \bar{p}_t(q_t))$ and $(q_a, p_a) = (b + \bar{u}, b)$. There exists a strict deviation incentive iff, under this formulation, $q_t - p_t > 0$. Rearranging, this is the case iff $\nu^{(Q,Q)} < \nu^{(P,P)}$, where

$$\begin{aligned}
\nu^{(Q,Q)} &:= (q^Q - c(q^Q)) + (\beta - 1)(q^Q - \underline{q}), \text{ and} \\
\nu^{(P,P)} &:= (q^P - c(q^P)) + (\beta - 1)(b - c(q^P)).
\end{aligned}$$

Analogously, in an equilibrium where at least one firm plays according to (P^*), firms have a deviation incentive towards $\theta^k = Q$ iff $\nu^{(Q,Q)} > \nu^{(P,P)}$.

We conclude: A competitive equilibrium exists. In any such equilibrium, at least 2 firms share the market. These firms offer 2 products, t and $a \neq t$. All other firms choose $(M^k, \Theta^k) = \emptyset$. The characteristics of t and a as well as θ^k are uniquely defined by (Q^*) iff $\nu^{(Q,Q)} > \nu^{(P,P)}$ and by (P^*) iff $\nu^{(Q,Q)} < \nu^{(P,P)}$. If $\nu^{(Q,Q)} = \nu^{(P,P)}$, any firm that supplies the market chooses t and a according to either (Q^*) or (P^*). As a final step, we can drop the assumption that $b \to \infty$. In particular, our characterization is valid for any $b \geq c(q^Q) > c(q^P)$ as assumed in the model section of this paper. This concludes the proof for part a) (Store-Wide-Distortions).

Part b) (Product-Specific Distortions). Assume that firms choose $\theta_j^k \in \{Q, P, N\}$ for each product $j \in J^k$ individually. The proof works similarly as the proof for part a). We start again with the assumption of a homogeneous population with unique type $\tilde{\beta}^0 < \beta$ and determine the best response *conditional* on attracting a positive market share under the assumption that $b \to \infty$. Fix the competitor offer $(\boldsymbol{M}^{-k}, \boldsymbol{\Theta}^{-k})$ and let $\bar{u} = \bar{u}(\tilde{\beta}^0) \geq 0$ be type $\tilde{\beta}^0$'s *expected* maximum surplus attainable outside of firm k. Fix some quality $q_t \geq \underline{q}$. By the proof of Lemma 3, the maximum price the firm can sell any q_t is $\bar{p}_t(q_t)$. To sell at this price, a second product $a \neq t$ is necessary and sufficient. Offering more products is unnecessary yet costly and hence, cannot be part of the best response. Moreover, the (in-store) valuation of at least one product $j \in \{a, t\}$ must be distorted, in particular, with any distortion $(\theta_a, \theta_t) \in \{(Q, Q), (P, P), (P, Q), (P, N), (N, Q)\}$, a strictly higher price than without fooling can be realized. It follows that the best response must involve one of these distortions. It is easy to see that choosing $(\theta_a, \theta_t) = (P, Q)$ strictly dominates any other choice of (θ_a, θ_t): No other distortion yields an overvaluation of the target relative to the attraction product that is as extreme. This is of course reflected in $\bar{p}_t(q_t)$, which under the condition that $q_a - p_a = \bar{u}$ (the participation constraint binds) can be rewritten as

$$\bar{p}_t(q_t, q_a, p_a) = \begin{cases} \beta q_t - q_a + \beta p_a & \text{if } (\theta_a, \theta_t) = (P, Q) \\ q_t - q_a + \beta p_a & \text{if } (\theta_a, \theta_t) = (P, N) \\ \beta q_t - q_a + p_a & \text{if } (\theta_a, \theta_t) = (N, Q) \\ \beta q_t - \beta q_a + p_a & \text{if } (\theta_a, \theta_t) = (Q, Q) \\ \dfrac{1}{\beta} q_t - \dfrac{1}{\beta} q_a + p_a & \text{if } (\theta_a, \theta_t) = (P, P). \end{cases}$$

If $(\theta_a, \theta_t) \in \{(P, P), (P, Q), (P, N), (N, Q)\}$ (all except $(\theta_a, \theta_t) = (Q, Q)$), $\bar{p}_t(q_t, q_a, p_a)$ is maximized by choosing $p_a = b$ (which implies $q_a = b + \bar{u}$). This choice satisfies all fooling constraints: the consumer indeed enters the store of firm k and buys q_t (see also the proof of Lemma 3). Clearly, $(\theta_a, \theta_t) = (P, Q)$ yields the highest price. To see the dominance of $(\theta_a, \theta_t) = (P, Q)$ over $(\theta_a, \theta_t) = (Q, Q)$, note that $\bar{p}_t(q_t, q_a, p_a)$ is maximized under $(\theta_a, \theta_t) = (Q, Q)$ by choosing $q_a = \underline{q}$ (which implies $p_a = q_a - \bar{u}$). This is a feasible choice also when $(\theta_a, \theta_t) = (P, Q)$, which yields a strictly higher price. Hence, only $(\theta_a, \theta_t) = (P, Q)$ can be part of the best response.

Given 2 products t and $a \neq t$ as well as distortion $(\theta_a, \theta_t) = (P, Q)$, we need to define the profit-maximizing choice of (q_t, p_t) and (q_a, p_a). For any q_t, the profit-maximizing price is $p_t = \bar{p}_t(q_t, q_a, p_a)$ as defined above. With quality being inflated at the store, the cost-efficient choice of q_t is

$$q_t = q^Q := \arg\max_q [\beta q - c(q)] \Leftrightarrow c'(q^Q) = \beta.$$

This interior solution is valid by assumption $q^Q > \underline{q}$. We have already noted above that $\bar{p}_t(q_t, q_a, p_a)$ is maximized by choosing $p_a = b$ and thus, $q_a = b + \bar{u}$. While there are 2 opposing forces when maximizing $\bar{p}_t(q_t, q_a, p_a)$—minimizing q_a and maximizing p_a—, maximizing p_a is the dominant choice: Due to price p_a being inflated at the store, a marginal increase in price (accompanied by a marginal increase in quality) always yields a higher marginal effect on profits than the equivalent decrease in quality. The "marketing" implication of this choice is identical to the case of a purely price-inflated store (see case a), $\theta^k = P$): Competition outside the store is on quality q_a and not on price. It follows: Conditional on attracting a positive share of consumers, the unique best response of firm k is to offer 2 products, t and $a \neq t$, choose distortion $\Theta^k = (\theta_a, \theta_t) = (P, Q)$, and product-characteristics $(q_t, p_t) = (q^Q, \bar{p}_t(q^Q))$ and $(q_a, p_a) = (b + \bar{u}, b)$.

Identical to part a) of this proof, the best response is independent of the degree of naïveté of type $\tilde{\beta}^0 < \beta$. Also, firms not generating positive demand choose $(M^k, \Theta^k) = \emptyset$ to avoid positive costs and thus negative profits. By analogous statements as those in part a) it follows that at any point of mutual best response, any heterogeneity in types $\tilde{\beta}$ is rendered unimportant for market supply. It follows that in any equilibrium, $\bar{u}(\tilde{\beta}) = \bar{u}\ \forall \tilde{\beta} < \beta$: the outside option is a unique value. We can find market supply in the competitive equilibrium by searching for $\bar{u}$ that equates the profits in the best response defined above to zero. We conclude: A competitive equilibrium exists. In any such equilibrium, at least 2 firms share the market. These firms offer 2 products, t and $a \neq t$. All other firms choose $(M^k, \Theta^k) = \emptyset$. For the firms that share the market, Θ^k and the characteristics of products a and t are uniquely defined by

$$(PQ^*) \qquad \Theta^k = (\theta_a, \theta_t) = (P, Q),\ (q_t, p_t) = (q^Q, c(q^Q)),\ (q_a, p_a) = (\beta(q^Q + b) - c(q^Q), b).$$

As a final step, we can drop the assumption that $b \to \infty$. In particular, our characterization is valid for any $b \geq c(q^Q)$ as assumed in the model section of this paper. This concludes the proof for part b). ■

Proof of Proposition 3 (Fooling with Salience, Focusing, or Relative Thinking). The proof is constructed as follows: We will first consider Assumption F (Focusing) and Assumption RT (Relative Thinking). Both of these assumptions imply that distortions are store-wide, i.e., for any two products j, i in J^k, $\theta^k_j = \theta^k_i = \theta^k$. Following Proposition 2, if firms have a technology that allows for store-wide distortions, they will either want to fool with $(\theta_a, \theta_t) = (Q, Q)$ or $(\theta_a, \theta_t) = (P, P)$. We will show that if context is a function of the product-line and follows Assumption F or Assumption RT, a firm can construct $(\theta_a, \theta_t) = (Q, Q)$ and $(\theta_a, \theta_t) = (P, P)$ and fool according to the best response defined in the proof of Proposition 2 if and only if it introduces a third product to the product-line. In other words, one *decoy* is necessary and sufficient to fool according to Proposition 2, part a). We will then turn to Assumption S (Salience). This assumption allows firms to construct *product-specific* distortions. Proposition 2, part b) has shown that if firms can choose θ^k_j for each product individually, they will want to fool with $(\theta_a, \theta_t) = (P, Q)$. Again, we will show that under Assumption S, the firm can construct such distortion and fool according to the best response defined in the proof of Proposition 2 if and only if it introduces a third product to the product-line. In other words, one *decoy* is necessary and sufficient to fool according to Proposition 2, part b).

Assumption F (Focusing).

Step 1: Fooling is not possible without a third product (a decoy is necessary). Assume that Assumption F holds and that firm k offers only two products a and t, $a \neq t$. The firm may either fool with $(\theta_a, \theta_t) = (Q, Q)$ or $(\theta_a, \theta_t) = (P, P)$.

- Assume that the firm fools with $(\theta_a, \theta_t) = (Q, Q)$. We have shown in the proofs of Lemma 2 and Lemma 3 that fooling an optimistic consumer ($\tilde{\beta} < \beta$) with $(\theta_a, \theta_t) = (Q, Q)$ implies $u_t < u_a$, $q_t > q_a$ and $p_t > p_a$ (the firm up-sells). Note that $u_t < u_a \Leftrightarrow q_t - q_a < p_t - p_a$. By Assumption F this implies that $\Delta_q^k < \Delta_p^k$, and thus, $\theta^k \in \{P, N\}$, a contradiction.
- Assume that the firm fools with $(\theta_a, \theta_t) = (P, P)$. We have shown in the proofs of Lemma 2 and Lemma 3 that fooling an optimistic consumer ($\tilde{\beta} < \beta$) with $(\theta_a, \theta_t) = (P, P)$ implies $u_t < u_a$, $q_t < q_a$ and $p_t < p_a$ (the firm down-sells). Note that $u_t < u_a \Leftrightarrow p_a - p_t < q_a - q_t$. By Assumption F this implies that $\Delta_q^k > \Delta_p^k$, and thus, $\theta^k \in \{Q, N\}$, a contradiction.

This concludes the proof of step 1. Note that this result is not an artefact of our rank-based implementation of KS, but a generic characteristic of the Focusing framework, which requires, by Assumption 1 in KS, that whenever preferences are shifted towards a product that is dominant in one attribute (z), but not in the other ($-z$), $\Delta_z^k > \Delta_{-z}^k$, which is in contradiction to fooling condition $u_t < u_a$.

Step 2: Fooling is always possible with a third product (a single decoy is sufficient). Assume that Assumption F holds and that firm k offers three products, a, t, and d. The firm may either fool with $(\theta_a, \theta_t) = (Q, Q)$ or $(\theta_a, \theta_t) = (P, P)$.

- Assume that the firm wants to fool with $(\theta_a, \theta_t) = (Q, Q)$. Fix any characteristics (q_a, p_a) and (q_t, p_t) that imply that the consumer is fooled if $(\theta_a, \theta_t) = (Q, Q)$. Then $u_t < u_a$, $q_t > q_a$, and $p_t > p_a$. (For example, choose the characteristics defining the best response in the proof of Proposition 2.) Choose $p_d = p_t$ and $q_d < q_t - (p_t - p_a) - \kappa_F$.[20] Then by Assumption F, $\Delta_q^k - \Delta_p^k > \kappa_F \Leftrightarrow \theta^k = Q \Rightarrow (\theta_a, \theta_t) = (Q, Q)$. Note that product d is strictly dominated by target t and thus, does not violate incentive compatibility of the fooling regime: The firm can sell product t according to the best response defined in the proof of Proposition 2.

[20] Recall from Assumption F that $\kappa_F \geq 0$ is some (exogenously defined) threshold that measures the level of stimulus necessary for a preference distortion.

- Assume that the firm wants to fool with $(\theta_a, \theta_t) = (P, P)$. Fix any characteristics (q_a, p_a) and (q_t, p_t) that imply that the consumer is fooled if $(\theta_a, \theta_t) = (P, P)$. Then $u_t < u_a$, $q_t > q_a$, and $p_t > p_a$. (For example, choose the characteristics defining the best response in the proof of Proposition 2.) Choose $q_d = q_t$ and $p_d > p_t + (q_a - q_t) + \kappa_F$.[21] Then by Assumption F, $\Delta_p^k - \Delta_q^k > \kappa_F \Leftrightarrow \theta^k = P \Rightarrow (\theta_a, \theta_t) = (P, P)$. Note that product d is strictly dominated by target t and thus, does not violate incentive compatibility of the fooling regime: The firm can sell product t according to the best response defined in the proof of Proposition 2.

This concludes the proof of step 2. Again, notice that this a result generic to the Focusing framework and does not depend on the rank-based formulation of preferences that we have assumed for our model. The Focusing framework assumes utility weights to be a function of the attribute spread Δ_z^k. Most naturally, such spreads are open to manipulation by a single option, i.e., a single decoy.

It follows that under Assumption F, the characterization of products a and t corresponds to the equilibrium defined in Proposition 2, part a). Holding more than three products is unnecessary yet costly which implies that the fooling equilibrium of Proposition 2 will be realized with exactly three products of which one is a decoy.

Assumption RT (Relative Thinking).

Step 1: Fooling is not possible without a third product (a decoy is necessary). We show that norming assumptions N1 and N2 in BRS imply that fooling is impossible with only 2 products. The result then readily extends to Assumption RT. Attention weights in BRS are a function of the spread of an attribute in the choice set, Δ_z, $z = q, p$; we call the weight function $w(\Delta_z)$. By N1, $w(\Delta_t)$ is strictly decreasing in Δ_z. By N2, $w(\Delta_z)\Delta_z$ is strictly increasing in Δ_z.

Suppose that firm k offers only two products, a and t. Fooling requires that $u_a > u_t$ (see Lemma 2) while inside the store $\hat{u}_t \geq \hat{u}_a$ by incentive compatibility (IC). We show that the norming assumptions in BRS rule out such a preference change if a and t are the only products in the product-line.

Assume $u_a > u_t \Leftrightarrow q_a - p_a > q_t - p_t$. Then either (1) $q_a > q_t$ and $p_a > p_t$, or (2) $q_a < q_t$ and $p_a < p_t$, or (3) $q_a > q_t$ and $p_a < p_t$. If (1) is true, then $u_a > u_t \Leftrightarrow \Delta_q > \Delta_p$. N2 then implies $w(\Delta_q)\Delta_q > w(\Delta_p)\Delta_p$, which is equivalent to $w(\Delta_q)s_a - w(\Delta_p)p_a > w(\Delta_q)q_t - w(\Delta_p)p_t$. Thus, the same product is preferred outside and inside the store and fooling is not possible. If (2) is true, $u_a > u_t \Leftrightarrow \Delta_q < \Delta_p$ leads to a similar contradiction. Then N2 implies

[21] See the previous footnote.

$w(\Delta_q)\Delta_q < w(\Delta_p)\Delta_p$ and product a (now being the less qualitative option) will be preferred both inside and outside the store. Finally, if (3) is true, product a dominates product t in both attributes implying that a is strictly preferred over t for any (positive) attribute weights. Again, product a is preferred both inside and outside the store and fooling is not possible.

These results readily extend to Assumption RT. Case (3) is immediate. For case (1), note that t is the less-qualitative product. A preference change towards t would thus require $\theta^k = P$, that is, $\frac{\Delta_q}{\Delta_p} > \kappa_{RT} \geq \beta$.[22] But with only two products spanning the attribute range Δ_q, $\Delta_q > \beta\Delta_p$ implies that the product with higher quality is preferred at the store. That is, $\hat{u}_a^k > \hat{u}_t^k$, which contradicts incentive compatibility (IC). With case (2) we get a similar contradiction. This concludes the proof of step 1.

Step 2: Fooling is always possible with a third product (a single decoy is sufficient). Assume that Assumption RT holds and that firm k offers three products, a, t, and d. The firm may either fool with $(\theta_a, \theta_t) = (Q, Q)$ or $(\theta_a, \theta_t) = (P, P)$.

- Assume that the firm wants to fool with $(\theta_a, \theta_t) = (Q, Q)$. Fix any characteristics (q_a, p_a) and (q_t, p_t) that imply that the consumer is fooled if $(\theta_a, \theta_t) = (Q, Q)$. Then $u_t < u_a$, $q_t > q_a$, and $p_t > p_a$. (For example, choose the characteristics defining the best response in the proof of Proposition 2.) Choose $q_d = q_t$ and $p_d > p_a + \kappa_{RT}(q_t - q_a) > p_t$.[23] Then by Assumption RT, $\frac{\Delta_p^k}{\Delta_q^k} > \kappa_{RT} \Leftrightarrow \theta^k = Q \Rightarrow (\theta_a, \theta_t) = (Q, Q)$. Note that product d is strictly dominated by target t and thus, does not violate incentive compatibility of the fooling regime: The firm can sell product t according to the best response defined in the proof of Proposition 2.

- Assume that the firm wants to fool with $(\theta_a, \theta_t) = (P, P)$. Fix any characteristics (q_a, p_a) and (q_t, p_t) that imply that the consumer is fooled if $(\theta_a, \theta_t) = (P, P)$. Then $u_t < u_a$, $q_t < q_a$, and $p_t < p_a$. (For example, choose the characteristics defining the best response in the proof of Proposition 2.) Choose $p_d = p_t$ and $q_d < q_a - \kappa_{RT}(p_a - p_t)$, which implies $q_d < q_t$. Then by Assumption RT, $\frac{\Delta_q^k}{\Delta_p^k} > \kappa_{RT} \Leftrightarrow \theta^k = P \Rightarrow (\theta_a, \theta_t) = (P, P)$. Note that product d is strictly dominated by target t and thus, does not violate incentive compatibility of the fooling regime: The firm can sell product t according to the best response defined in the proof of Proposition 2.

This concludes the proof of step 2. Similar to the Focusing framework, this result is generic to the model by BRS and does not depend on our rank-based implementation. The

[22]Recall from Assumption RT that $\kappa_{RT} \geq \beta$ is some (exogenously defined) threshold that measures the level of stimulus necessary for a preference distortion.

[23]See the previous footnote.

framework of Relative Thinking assumes utility weights to be a function of the attribute spread Δ_a^k. Most naturally, such spreads are open to manipulation by a single option, i.e., a single decoy.

It follows that under Assumption RT, the characterization of products a and t corresponds to the equilibrium defined in Proposition 2, part a). Holding more than three products is unnecessary yet costly which implies that the fooling equilibrium of Proposition 2 will be realized with exactly three products of which one is a decoy.

Assumption S (Salience)

We have shown in the proof of Proposition 2 that if firms can choose θ_j^k for each product individually, the unique weakly undominated best response is to fool with $a \neq t$ and choose $(\theta_a, \theta_t) = (P, Q)$. We show that under Assumption S, this choice is possible iff the firm adds a third product (i.e., a single decoy) to the product-line.

Step 1: A decoy is necessary. The specifications of products a and t that a best-responding firm will choose are given in the proof of Proposition 2. We show that a distortion $(\theta_a, \theta_t) = (P, Q)$ with these product specifications cannot be constructed without the help of additional (decoy) products. Note first that the specification Proposition 2 implies $q_a > q_t$ and $p_a > p_t$. Thus, none of the two products is dominated. Suppose that the firm only holds these two products. Then the reference quality is given by $z_R^k = \frac{(q_a+q_t)}{2}$ and the reference price is given by $p_R^k = \frac{(p_a+p_t)}{2}$. Because $(q_j - q_R^k)(p_j - p_R^k) > 0$ for $j \in \{a, f\}$, we can exploit Proposition 1 in BGS: The "advantageous" attribute of product j—higher quality *or* lower price relative to the reference—is overweighted if and only if $\frac{q_j}{p_j} > \frac{q_R^k}{p_R^k}$. Also, iff $\frac{q_j}{p_j} < \frac{q_R^k}{p_R^k}$, then the "disadvantageous" attribute of product j is overweighted, while iff $\frac{q_j}{p_j} = \frac{q_R^k}{p_R^k}$, consumers weigh both attributes equally.

Assume towards a contradiction that the firm can construct $(\theta_a, \theta_t) = (P, Q)$. For t being quality-salient, by $q_t < q_R^k$ and Proposition 1 in BGS,

$$\frac{q_t}{p_t} < \frac{q_R^k}{p_R^k} \Leftrightarrow \frac{q_t}{p_t} < \frac{q_a}{p_a}.$$

But for a being price-salient, by $q_a > q_R^k$ and Proposition 1 in BGS,

$$\frac{q_a}{p_a} < \frac{q_R^k}{p_R^k} \Leftrightarrow \frac{q_t}{p_t} > \frac{q_a}{p_a},$$

a contradiction. This concludes the proof of step 1.

Step 2: A single decoy is sufficient. Assume that firm k wants to fool using distortion $(\theta_a, \theta_t) = (P, Q)$ and chooses the specifications of product a and t according to the best-response defined in the proof of Proposition 2 part b). Note that by this specification $q_a > q_t \geq \underline{q} > 0$ and $p_a > p_t > 0$.

- Assume that $\frac{q_t}{p_t} > \frac{q_a}{p_a}$. We construct a reference point using one additional product d that satisfies the following properties: (1) $p_R^k = p_t$, (2) $q_R^k < q_t$ and (3) $\frac{q_a}{p_a} < \frac{q_R^k}{p_R^k} < \frac{q_t}{p_t}$. The construction is illustrated in Figure II. With such a reference point,

  1. Product t is quality-salient: By $p_R^k = p_t$, the salience of p_t is $\sigma(p_t, p_t)$. By homogeneity of degree zero, $\sigma(\alpha p_t, \alpha p_t) = \sigma(p_t, p_t)$ for any $\alpha > 0$. Let $\alpha = \frac{q_t}{p_t} > 0$, then $\sigma(p_t, p_t) = \sigma(q_t, q_t)$. By ordering, $\sigma(q_t, q_t) < \sigma(q_t, q_R^k)$ because $q_R^k < q_t$. Thus, $\sigma(q_t, q_R^k) > \sigma(p_t, p_R^k)$: product t is quality-salient.
  2. Product a is price-salient: By $q_R^k < q_t < q_a$ and $p_R^k = p_t < p_a$, $(q_a - q_R^k)(p_a - p_R^k) > 0$, and product a neither dominates nor is dominated by the reference good. Thus, Proposition 1 in BGS applies. Because $q_a > q_R^k$, by $\frac{q_R^k}{p_R^k} > \frac{q_a}{p_a}$, product a is price-salient.

  To satisfy property (1), choose $p_d = 2p_t - p_a$, which implies $p_d < p_t$. To satisfy property (2) and (3), choose $q_d < 2q_t - q_a$, which implies $q_d < q_t$. It remains to be shown that the decoy d does not violate fooling conditions. Note that $q_d - p_d < 2q_t - q - a - (2p_t - p_a) \Leftrightarrow u_d < 2u_t - u_a$. Because $u_t < u_a$ by the specifications of a and t, this implies that $u_d < u_t < u_a$. We first show that IC is not violated: Because t is quality-salient, $\hat{u}_t^k = \beta q_t - p_t > u_t$. But then, if (i) $\theta_d^k = N$, $\hat{u}_t^k > \hat{u}_d^k$ follows from $\hat{u}_t^k > u_t > u_d = \hat{u}_d^k$, if (ii) $\theta_d^k = Q$, $\hat{u}_t^k > \hat{u}_d^k$ follows from $q_d < q_t$, $p_d < p_t$ and $u_t > u_d$, if (iii) $\theta_d^k = P$, then $\hat{u}_t^k > \hat{u}_d^k$ iff $\hat{u}_a^k > \hat{u}_d^k \Leftrightarrow q_a - q_d > \beta(p_a - p_d)$ by $\hat{u}_t^k = \hat{u}_a^k$. To prove that $q_a - q_d > \beta(p_a - p_d)$, note that $q_a - q_d > q_a - (2q_t - q_a) = 2(q_t - q_a)$ by $q_d < 2q_t - q_a$ and $p_a - p_d = p_a - (2p_t - p_a)$ by $p_d = 2p_t - p_a$. Thus $q_a - q_d > \beta(p_a - p_d)$ if $2(q_a - q_t) > 2\beta(p_a - p_t) \Leftrightarrow (q_a - q_t) > \beta(p_a - p_t)$. But the latter inequality is true by $\hat{u}_t^k = \hat{u}_a^k \Leftrightarrow q_a - \beta q_t = \beta p_a - p_t$. Thus, $\hat{u}_t^k > \hat{u}_d^k$. Finally, we have to show that PCC is not violated, i.e., that $E_{\tilde{\beta}}\left[\hat{u}_a^k\right] > E_{\tilde{\beta}}\left[\hat{u}_d^k\right]$. To see that this is true note that we have shown that $u_a > u_t > u_d$ and $\hat{u}_a^k = \hat{u}_t^k > \hat{u}_d^k$. Because $E_{\tilde{\beta}}\left[\hat{u}_a^k\right]$ is between $\hat{u}_a^k$ and u_a and $E_{\tilde{\beta}}\left[\hat{u}_d^k\right]$ is between $\hat{u}_d^k$ and u_d (both by $\tilde{\beta} < \beta$) it follows that $E_{\tilde{\beta}}\left[\hat{u}_a^k\right] > E_{\tilde{\beta}}\left[\hat{u}_d^k\right]$.

- Assume that $\frac{q_t}{p_t} < \frac{q_a}{p_a}$. We construct a reference point using one additional product d that satisfies the following properties: (1) $q_R^k = q_a$, (2) $p_R^k > p_a$ and (3) $\frac{q_a}{p_a} > \frac{q_R^k}{p_R^k} > \frac{q_t}{p_t}$. The construction is illustrated in Figure II. With such a reference point,

1. Product t is quality-salient: By $q_R^k > q_t$ and $p_R^k > q_t$, $(q_t - q_R^k)(p_t - p_R^k) > 0$, and product t neither dominates nor is dominated by the reference good. Thus, Proposition 1 in BGS applies. Because $q_t < q_R^k$, by $\frac{q_R^k}{p_R^k} > \frac{q_t}{p_t}$, product t is quality-salient.
2. Product a is price-salient: By $q_R^k = q_a$, the salience of q_a is $\sigma(q_a, q_a)$. By homogeneity of degree zero, $\sigma(\alpha q_a, \alpha q_a) = \sigma(q_a, q_a)$ for any $\alpha > 0$. Let $\alpha = \frac{p_a}{q_a} > 0$, then $\sigma(q_a, q_a) = \sigma(p_a, p_a)$. By ordering, $\sigma(p_a, p_a) < \sigma(p_a, p_R^k)$ because $p_R^k > q_t$. Thus, $\sigma(q_a, q_R^k) < \sigma(p_a, p_R^k)$: product a is price-salient.

To satisfy property (1) choose $q_d = 2q_a - q_t > q_a$. To satisfy property (2) and (3), choose $p_d > 2p_a - p_t$. It remains to be shown that the decoy d does not violate fooling conditions. But note that $p_d > p_a = b$: The decoy has a price above the maximum willingness to pay and thus, will never be chosen (and can therefore not violate fooling conditions).

This concludes the proof of step 2. We conclude: Under Assumption S, the characterization of products a and t corresponds to the equilibrium defined in Proposition 2, part b). Holding more than three products is unnecessary yet costly which implies that the fooling equilibrium of Proposition 2 will be realized with exactly three products of which one is a decoy.

■

Proof of Proposition 4 (Equilibrium with Fooling and Non-Fooling Firms). The result is trivial if either *all* consumers are optimistic or *all* consumers are sophisticated/pessimistic. These cases were covered by our earlier Propositions. So assume that there exists a positive mass of consumers with beliefs $\tilde{\beta} < \beta$ and a positive mass of consumers with beliefs $\tilde{\beta} \geq \beta$.

Let $\beta > 1$. Fix a market offering according to the Proposition. There exists two types of stores with strictly positive demand, k^L and k^H. Type k^L is a fooling firm that supplies products according to the equilibrium defined in Proposition 2 and k^H is a non-fooling firm that supplies products according to the equilibrium defined in Proposition 1. There exists at least 2 firms of each type. All other firms choose $(M^k, \Theta^k) = \emptyset$. All firms make zero profits. Note that conditional on purchasing at type k^H, all consumers expect to purchase q^* at price $p^* = c^*$. At the same time, conditional on purchasing at type k^L, all sophisticated and pessimistic consumers (correctly) expect to purchase the target at the fooling firms (see Lemma 2 for a proof), while all optimists (falsely) expects to purchase the attraction product. We prove that a competitive equilibrium with this market supply exists and that it defines the unique competitive market supply.

Existence. Assume that we have an equilibrium. Firms of type k^L fool and sell quality $q_t \neq q^*$ at $p_t = c(q_t)$ to the optimists, while firms of type k^H are truthful and sell q^* at $p^* = c^*$ to the sophisticates/pessimists. We have to check whether consumers or firms want to deviate. Consider first the optimistic population that are assumed to purchase at k^L. They have the alternative to purchase (q^*, c^*) at k^H instead of (q_a, p_a) at k^L (of course, they only *expect* to buy product a, while they really buy the target t). However, because $u_a > u^*$ in the candidate equilibrium defined above, purchasing at k^L always promises a higher payoff and optimists will not switch to k^H:

1. Consider an equilibrium according to Proposition 2 part b). Then $u_a = q_a - p_a = \beta(q^Q + b) - c^Q - b$. Note that $(\beta - 1)b > 0$ by $\beta > 1$. Strict convexity of the cost function then implies $\beta q^Q - q^Q > \beta q^* - c^* > q^* - c^*$ and thus, $u_a > u^*$.

2. Consider an equilibrium according to Proposition 2 part a) where $(\theta_a, \theta_t) = (Q, Q)$. Then $u_a = q_a - p_a = \underline{q} - c^Q + \beta(q^Q - \underline{q})$. Note that by assumption, $\underline{q} < q^*$ and thus, $u_a > q^* - c^Q + \beta(q^Q - q^*)$ by $\beta > 1$. It follows that $u_a > u^*$ because $q^* - c^Q + \beta(q^Q - q^*) > q^* - c^* \Leftrightarrow \beta q^Q - c^Q > \beta q^* - c^*$ by strict convexity of the cost-function.

3. Consider an equilibrium according to Proposition 2 part a) where $(\theta_a, \theta_t) = (P, P)$. Then $u_a = q_a - p_a = q^P + (\beta b - c^P) - b$ and $u_a > u^* \Leftrightarrow q^P - \beta c^P + \beta b - b > q^* - c^*$ $\Leftrightarrow q^P/\beta - c^P + (1 - 1/\beta)b > q^*/\beta - c^* + (1 - 1/\beta)c^*$. Note first that $b > c^*$, so $(1 - 1/\beta)b > (1 - 1/\beta)c^*$. Strict convexity of the cost function further implies that $q^P/\beta - c^P > s^*/\beta - c^*$. So $u_a > u^*$.

Pessimists also do not want to switch to k^L. They correctly expect to buy product t at k^L (for a proof see Lemma 2) which yields surplus $u_t = q_t - c(q_t)$. Because $q^* = \arg\max(q - c(s))$ and $q_t \neq q^*$ by strict convexity of $c(q)$, $u^* = q^* - c^* > u_t$ and shopping at type k^H yields higher surplus. Finally, firms of either type have no incentive to deviate. By Proposition 2, no firm can find a more profitable strategy when serving optimistic agents if there are at least 2 firms of type k^L. By Proposition 1, no firm can find a more profitable strategy when serving pessimistic agents if there are at least 2 firms of type k^H. The only strategy that yields non-negative profits when not generating demand is $(M^k, \Theta^k) = \emptyset$. This strategy yields zero profits as well and thus, does not constitute a deviation incentive. Hence, this is an equilibrium. Note also that any firm that is not of type k^L or k^H must choose $(M^k, \Theta^k) = \emptyset$ by above reasoning, confirming part c) of the proposition. (q.e.d.)

Uniqueness. The proofs of Propositions 1 and 2, respectively, show that unless there exist at least 2 firms supplying products according to Proposition 1 as well as at least 2 firms supplying products according to Proposition 2, there exists a deviation incentive to a strategy with strictly positive profits. In particular, by the uniqueness and continuity of the best response conditional on attracting only sophisticated/pessimistic consumers (Proposition 1), there must exist at least 2 firms supplying a product with expected surplus $\bar{u}^H \geq u^* = q^* - c(q^*)$ to consumers of type $\tilde{\beta} \geq \beta$. Otherwise, at least one firm could attract the entire population of types $\tilde{\beta} \geq \beta$ at strictly positive profit. Similarly, there must exist at least 2 firms supplying a product with expected surplus $\bar{u}^L \geq u_a = q_a - p_a$ to consumers of type $\tilde{\beta} < \beta$, where q_a and p_a are defined by the equilibrium characterized in Proposition 2. Otherwise, at least one firm could attract the entire population of types $\tilde{\beta} < \beta$ at strictly positive profit. By the strict difference of u_a and u^* (in particular, $u_a > u^*$, see the existence proof above), 1 firm cannot satisfy both of these conditions at the same time (attracting both groups of consumers with positive probability), even if it would play a mixed strategy: Such a firm would either have to make negative profits in expectation (to attract both groups without generating a deviation incentive for other firms) or generate an offer that (for at least one of the two groups of consumers) could be profitably undercut by other firms. It follows that at least 2 firms satisfying the respective condition must exist for each group *separately*. Because each firm only serves one group of consumers, the only possibility to satisfy the respective condition without making negative profit is for each firm to choose market supply according to Propositions 1 and 2, respectively. It follows that any competitive equilibrium must have the characteristics listed in the Proposition. (q.e.d.) ■