# Strategic Product Placement and Pricing in Markets with Inattentive Consumers

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Carsten Dahremöller
aus Bonn

Bonn 2013

| | |
|---|---|
| Dekan: | Prof. Dr. Klaus Sandmann |
| Erstreferent: | Prof. Dr. Matthias Kräkel |
| Zweitreferent: | Prof. Dr. Dezsö Szalay |

Tag der mündlichen Prüfung: 24.04.2013

# Acknowledgements

In preparing this thesis I received support from many people to whom I am grateful. First and foremost, I want to thank Matthias Kräkel who has been a great supervisor. He provided very detailed and constructive comments to all three chapters of my thesis. My sincere thanks are also due to Dezsö Szalay for acting as a referee on my thesis committee.

Moreover, I am particularly grateful to Markus Fels. He has not only been a good friend, our joint research has been a fruitful, usually fun, and inspiring activity. Chapter III of this thesis is an achievement of this cooperation.

I also owe many thanks to Simon Dato, who also contributed to Chapter II. Our countless discussions significantly improved this thesis.

Chapter I benefited a lot from comments and suggestions made by Paul Heidhues. I have learned a lot from him and he has been extremely supportive. In additition, he was the one who aroused my interest for behavioral economic theory.

The Bonn Graduate School of Economics (BGSE) provided not only financial support but also, and more importantly, a stimulating environment with many young researchers. I would like to thank all people who have contributed to this organization, especially Urs Schweizer. Moreover, I am indebted to Paul Heidhues for motivating me to apply for the BGSE. I also would like to thank the Max Planck Institute for Research on Collective Goods for additional financial support.

The topics courses of the Bonn Graduate School of Economics improved my understanding of several fields of Economics and provided useful hints for my own research. Therefore, I am grateful to the lecturers of the courses I attended: Susanne Goldlücke, Daniel Krähmer, Sebastian Kranz, Marc Le Quement, Dezsö Szalay and Alessia Testa.

I have benefited from discussions with and comments made by my office mates: Christoph Roling, Stephan Luck, Jasmin Gider, Anke Becker, Fabian Kosse and Paul Schempp. Moreover, I owe thanks to Fabian Herweg and Daniel Müller.

Finally, I owe much to Annet, my family and friends, without whom I would not have gone so far, and who gave me unconditional support for all my endeavors.

# Contents

# List of Figures

# Introduction

The theory of Industrial Organization has a long tradition in the field of Economics. Relatively novel to this field is the introduction of behavioral biases of consumers that are incorporated into this theory.[1] If one would like to segment this literature, it could be categorized by the different forms of behavioral consumer biases that are considered. Examples of the behavioral biases that have been incorporated into the Industrial Organization literature are, e.g., prospect theory, limited attention, overconfidence, time inconsistent behavior, and limited foresight. This dissertation aims at contributing to the strand of Behavioral Industrial Organization that investigates the impact of limited consumer attention on firm behavior in retail markets. Prominent examples among the numerous works in this strand of literature are Gabaix and Laibson (2006), Armstrong and Chen (2009), Carlin (2009), Wilson (2010), Carlin and Manso (2011), and Ellison and Wolitzky (2011).

This dissertation contains three seperate chapters. Within the first two chapters we assume that the specifications of consumer attention are exogeneously given. In the presented models, this means that some myopic consumers do not take add-ons and their respective prices into account when they make their purchase decisions. In contrast to this exogeneously given attention heuristic that neglects certain product aspects, the process of attention allocation is endogenized in the third chapter of this dissertation. We achieve this by introducing an attention allocation process that is 'built-in' for each consumer. One interpretation of this process of attention allocation is to see it as the solution of some evolutionary optimization process, as we argue in the Appendix. This process of attention allocation determines how, given a certain set of products, consumers allocate their attention among the different attributes of the available products. This endogenization of attention allows for more realistic results on consumer focus and on the ways in which firms can manipulate consumer attention and behavior.

Chapter I is a reprint of Dahremöller (forthcoming). It studies a market in which competing firms sell both a base good and an add-on. Initially, the prices of the base goods are known to all consumers while the prices of the add-ons are shrouded and

[1] See Spiegler (2011) for an excellent overview.

cannot be observed by consumers. The firms compete over a consumer population that is, in part, myopic. This means that some of the consumers do not incorporate the add-on into their purchase decisions. The particular composition of the population, however, can be manipulated by the competing firms. If one firm decides to unshroud the add-on, some of the myopic consumers get educated and all add-on prices become publicly observable. The model is strongly inspired by the seminal work of Gabaix and Laibson (2006), where the authors show that, if firms simultaneously set their prices and decide whether to unshroud the add-on, an equilibrium exists in which all firms shroud the add-on. This was a novel result as traditional models of voluntary disclosure of information could not explain the observation that in some markets firms seem to actively engage in obfuscation.[2]

The model of Chapter I applies some slight modifications to the model of Gabaix and Laibson (2006) and shows that these have substantial implications on the gameplay and on the incentives of firms. The most central modification is the assumption that unshrouding and price-setting do not take place simultaneously. In particular, I assume that firms first decide about unshrouding and before setting their prices. Another modification is the assumption that firms are heterogeneous in their add-on profitability. Applying this model, I find that the shrouding equilibrium vanishes as there always exists one firm that has an incentive to wreck the other firms attempts at shrouding. By doing so, the rival firm can influence the shrouding firms profit in the add-on market and thereby profitably soften competition in the base good market.

These results suggest that the central driver behind the incentives of firms to unshroud the add-on is the longevity of the unshrouding mechanism. If unshrouding has only short-term effects, like in Gabaix and Laibson (2006), firms are expected to leave the add-on shrouded, conditional on the existence of a sufficient number of myopic consumers. If, in contrast, unshrouding has long-term effects, like in Dahremöller (forthcoming), firms are expected to unshroud the add-on for all variations of the consumer population, conditional on the firms having sufficiently strong comparative differences in add-on profitability.

Chapter II is based on joint work with Simon Dato. The paper further extends the models of Gabaix and Laibson (2006) and Dahremöller (forthcoming) by applying a related framework in an infinitely repeated game. In every period of the game a number of firms sell both a base good and an add-on and decide whether or not to unshroud the add-on. As in Dahremöller (forthcoming), unshrouding is strategic in the sense that it has an impact on the future structure of the game and thereby potentially also on future profits of firms. In the particular model, this means that once a firm has unshrouded

[2] Among the traditional works in this field are Grossman (1981), Milgrom (1981), and Shapiro (1995).

the add-on, all myopic consumers become sophisticated for all future periods of the game. We show that, given sufficient parameter conditions, two types of equilibria exists. In the so-called unshrouding equilibrium at least one firm unshrouds the add-on and all firms behave competitively without engaging in collusion. In the so-called shrouding equilibrium all firms shroud the add-on and collude on monopoly prices. It holds that, while there is a short-term individual incentive to unshroud the add-on, there is a collective long-term incentive to leave the add-on shrouded. In particular, there are two incentives for firms to leave the add-on shrouded. First, shrouding may be a requirement for collusion between firms. This is because shrouding the add-on makes it less rewarding for firms to deviate from collusion, implying that collusion is more stable. Second, even if collusion is already stable, shrouding increases joint firm profits.

These results have several implications on governmental policies and on potential regulatory interventions. In particular, the fact that shrouding may be a requirement for collusion has two main implications. First, the degree of obfuscation in a market may be a novel indicator that can be used to detect illegal cartels. Second, if the regulator can intervene to unshroud the add-on, collusion might be destabilized, leading to stronger competition between the firms in the market.

Chapter III is based on joint work with Markus Fels. We consider a monopolist that serves a population of consumers who exhibit a particular form of limited attention. Consumer attention is limited in the sense that consumers have trouble when comparing products with many different attributes and incur a cognitive cost when trying to do so. If consumers face such a choice between complex products, they use a heuristic to allocate their attention to specific attributes of the available choices. In particular, we show that there exists a natural ordering that determines how the attention is allocated. Our functional form of the cognition costs is closely related to two other papers. First, the idea that consumers pay higher attention to attributes that exhibit high dispersion within the choice set is also present in Kőszegi and Szeidl (2012). Second, the idea that consumers use an attention allocation heuristic that is the result of an evolutionary optimization process is similar to Gabaix (2011). One of the novelties of our attention allocation is that the cognitive cost of attention is dependent on the number of attributes that are already taken into account. This implies that consumers can easily distinguish between products that differ in only one attribute (e.g., between standard goods that only differ in their price). However, if consumers have to choose between very complex products, any new attribute that is added to the product receives little attention or is even completely neglected.

After having derived and motivated our attention allocation mechanism, we analyze its implications for attention allocation as well as the means by which consumer attention can be manipulated. We apply this attention allocation mechanism by analyzing how a profit-maximizing firm reacts to such cognitive constraints of the consumer population. First we assume that the firm can only produce one product and derive the optimal design of this product. Then we continue by allowing for several products. We show that, if the firm can produce several products, it has an incentive to extend its product line by offering premium quality products that are not intended for sale, but have the sole function of manipulating consumer attention. We then discuss the optimal design of these so-called bait goods and outline the underlying intuitions.

# I. Unshrouding for Competitive Advantage

*In a market with hidden product details and systematic consumer biases, firms have the possibility to unshroud and thereby to rectify such market obliquities. While the classical view was that firms will have an incentive to unshroud, Gabaix and Laibson (2006) show that there exist constellations in which firms prefer to leave the market shrouded. Building on that model I introduce a more strategic and long-term dimension of unshrouding which turns out to fundamentally alter the underlying incentives to unshroud. In particular, I show that there exists an incentive to unshroud that stems from differences in add-on profitability and that it is dependent on parameter constellations whether a more profitable or a less profitable firm will want to unshroud.*

# 1. Introduction

Consumers in many markets have difficulties making rational decisions or face barriers when trying to collect all information that is relevant for their consumption choice. In these markets firms might either foster or alleviate such market obliquities. For example, in the case of retail financial markets, Liebman and Zeckhauser (2004) identify an array of contract characteristics that can influence the information reception of consumers. Among them are nonlinear pricing, schedule complexity, frequent revisions of schedules, complex vocabulary, delayed costs or bundled consumption. By employing or refraining from these contract characteristics firms are able to manipulate the perceptions of consumers.

In the following I propose a model of such markets in which the degree of consumer sophistication can be manipulated by the competing firms. My model is inspired by the seminal work of Gabaix and Laibson (2006) (henceforth GL), who model an add-on market with hidden add-on prices and limited consumer attention. Firms set prices for their base goods and add-ons and have the option to unshroud the add-on. In their model, both the pricing and the unshrouding decisions are taken simultaneously. If any firm decides to unshroud, all add-on prices get revealed and a part of the consumer population gets educated. Building on this model GL show that, for certain parameter constellations, equilibria exist in which none of the firms unshrouds the add-on.

In the following, I will enrich the model of GL by assuming that firms are heterogeneous in their add-on profitability and, more importantly, by applying a different modeling of consumer education. In particular, I will assume that firms can shroud or unshroud the add-on before they set their prices. This allows unshrouding to have more long-term consequences than in the model of GL.

Using this framework I show that for almost all parameter constellations no shrouding equilibrium exists. At least one firm has an incentive to unshroud the add-on in an attempt to wreck the rival firm's attempts at shrouding. By doing so, the unshrouding firm can influence the shrouding firm's profits within the add-on market and thereby profitably soften competition in the base good market. This incentive to unshroud the add-on is driven by the differences between the firms with regard to their add-on profitability. Thereby it is dependent on parameter constellations whether it is the comparatively efficient firm or the comparatively inefficient firm that unshrouds the add-on.

To motivate my results I first want to give some anecdotal evidence for the depicted firm behavior, in particular for the prediction that there exist cases in which either an efficient firm or an inefficient firm unshrouds the add-on. An example of unshrouding by an efficient firm could be Citibank if it advertises the convenience of its dense coverage of branches. It is efficient because due to its size it has a large number of branches, and due to its market share, it profits from its economies of scale and its per-customer costs of providing local services are low. In contrast, smaller banks have higher cost and might have to rely on cash points from other banks which is costly for its customers. If Citibank would then advertise its dense branch and cash point coverage, this might induce customers to take such features into account when making their purchase decision and increase their likelihood to relocate their account to Citibank.

An example of unshrouding by an inefficient firm was the bank Barclays when it introduced iShares, which are a collection of exchange traded funds that are passively managed and track the value of certain underlying assets, for example the S&P 500. Such investment vehicles are usually offered as supplementary services to bank account holders. On average, such passively managed funds perform better than actively managed funds (Gruber 1996), which were and still are a very popular investment vehicle. Barclays, in contrast to other banks, did not make much profit from actively managed funds due to a lack of scale. Although Barclays continued to offer actively managed funds after the introduction of iShares, they launched an advertisement campaign emphasizing the benefits of their new iShares product, but also pointing out disadvantages and negative attributes of actively managed funds.[1] This can be seen as an attempt to sabotage the market for actively managed funds. In other words, Barclays tried to acquit itself from its comparative disadvantage by trying to dissolve the add-on market in which it is inefficient.

The results of this paper also give an indication what markets are particularly susceptible to shrouding. As said before, I show that at least one firm will unshroud the add-on if unshrouding is costless. Evidently, in reality there might be costs associated with unshrouding. For example, there might be costs of setting up an advertisement campaign, costs of providing information to consumers, or costs for professionals which customers can consult. Such costs make unshrouding less attractive. Now recall that the incentive to unshroud is driven by the differences between the firms with regard to their profitability in the add-on dimension. Therefore, unshrouding should be more likely to be observed in markets in which firms have considerable differences in add-on profitability while shrouding should be observed in markets with rather homogeneous add-on profitability.

An example for a market with relatively homogeneous production costs is the hotel market. It is unlikely that hotels' costs of providing services such as telephone calls or minibar supplies differ substantially. Hence, this market is unlikely to be unshrouded

[1] See Carrell (2008) for further details. The advertisement campaign reported the hidden costs, intransparency, illiquidity and comparatively low performance of actively managed funds.

by any of the participating firms.

A market with substantial cost differences would be the banking and financial industries sector. Due to different sizes and strategic orientations, these industries differ significantly in terms of marginal production costs. For example, if a bank manages a portfolio of investment funds which are highly in demand, the costs of serving an additional customer are small. This is different for small local banks which do not have a well-established portfolio of investment vehicles. Hence, I conjecture that in such markets unshrouding is particularly likely.

The remainder of the paper is organized as follows. The next section introduces the main model and results. Section 3 analyzes various extensions and further implications. Section 4 will give a short overview of the related literature. Section 5 will conclude.

## 2. The Model

There exist two firms that both produce a base good and an add-on. Firms are rational, risk-neutral and aware of the behavioral biases of the consumer population. One of the firms is inefficient in the add-on dimension, which is captured by having higher add-on production costs than its competitor. Add-on production costs are denoted by $\hat{c}_i$ and for simplicity I will assume that $\hat{c}_1 = 0$ and $\hat{c}_2 = \hat{c} > 0$. Hence, firm 1 is more efficient in producing the add-on than firm 2. Firms set prices $p_i$ for the base good and $\hat{p}_i$ for the add-on. The prices of the base goods are common knowledge, but initially the prices of the add-ons are shrouded and consumers cannot observe them. The population consists of $\alpha \in (0, 1)$ myopic consumers and $(1-\alpha)$ sophisticated consumers. The sophisticated consumers know about the existence of the add-on and form Bayesian posteriors about its price. In contrast, myopic consumers do not know about the existence of the add-on or, equivalently, do not anticipate that they will need it.

An example for a market with myopic consumers would be the printer market where a consumer may not anticipate that she will probably need to buy cartridges with lifetime costs typically significantly exceeding the cost of the printer.[2] Another example are retail financial markets where the base good is typically a bank account which offers many add-on services such as overdrafts, money transfers, financial consultancy and investment services. Cruickshank (2000) reports that consumers are often poorly informed when selecting their current account provider and pay only little attention to add-on fees and conditions. Such behavior is potentially myopic in the sense that add-on fees are a substantial part of the account expenses and hence should be considered

[2] According to the Office of Fair Trading (2002) consumers spend 2 to 17 times more on cartridges than on the printer itself.

by consumers.

In line with GL, I assume that instead of buying the add-on, consumers can engage in an upfront substitution. For example, holders of a bank account could apply for a bank loan when running into financial distress instead of using the last minute solution of an expensive overdraft. Since sophisticates anticipate their future need for the add-on, they will engage in this substitution if the firm from which they bought the base good employs high add-on prices. In contrast, myopic consumers do not anticipate their need for the add-on and hence will never substitute. Evidence for such behavior of myopic consumers is given by Massoud et al. (2007) who report that a substantial fraction of credit card consumers pay penalty fees although they had enough cash on their current account to balance their credit card debts, and by Stango and Zinman (2009) who report that the median household in their data set pays $500 of credit card fees per year while over half of these costs could have easily been avoided. Both works reckon that many consumers potentially exhibit myopic behavior as they just forget or do not consider transferring money from their current account to their credit card account to balance it.

Each firm can decide to unshroud the add-on at zero cost. If at least one firm does so, all add-on prices become observable and a fraction of $\lambda \in (0,1)$ myopes gets educated and henceforth behaves like the sophisticated consumers. Define $\alpha' \equiv \alpha(1-\lambda)$ to be the remaining fraction of uneducatable myopes.

Unlike GL, I assume that firms decide about their shrouding behavior before they set their prices. Since the shrouding strategies of firms are determined at the beginning of the game, consumer education potentially is a strategic variable with long-term implications. One example in which consumer education has such long-term implications are markets in which firms interact repeatedly with the same population of consumers. If consumers learn about potential pitfalls related to add-ons, they are likely to keep that knowledge in mind and incorporate it in future purchase decisions. In a repeated game this would mean that once myopic consumers get educated, they retain their sophistication for future periods. For example, in a study about credit card markets Agarwal et al. (2011) report that the quality of decisions a consumer makes increases with her experience in that market. Along these lines, Agarwal et al. (2009) argue that financial literacy is persistent over long periods of time. In contrast to sticky consumer sophistication, prices tend to be much more flexible and can be adjusted quickly. In terms of the model this means that current pricing has no implications for future periods. This repeated game can then be boiled down to a one-shot game in which firms decide about unshrouding before they set their prices.[3]

[3]I will argue later in more detail that the qualitative effects of the repeated game and the sequential

After firms choose their shrouding behavior they set their prices for the base good and the add-on. At this stage, firms have full knowledge of the preceding actions of the competing firms. Hence, they know what the fractions of (remaining) myopic and sophisticated consumers are. This allows firms to adapt their pricing decisions to the (un)shrouding decisions that were taken at the beginning of the game.

Note here that the assumption that unshrouding educates (some) myopic consumers and reveals add-on prices closely follows the design of GL. This will be used to show that the two major modeling changes, namely timing of unshrouding and competitive advantages, will lead to completely novel incentives to unshroud. However, assuming that prices can be made observable before they are fixed may not be reasonable in all market settings. Therefore, I will discuss several alternative modelings with regard to price observability in the discussion section and show that these do not qualitatively change the results.

In the base good market firms compete via Hotelling competition.[4] A unit mass of consumers is uniformly distributed on the interval $[0, 1]$. Each consumer purchases at most one unit of the base good. The firms are located on the boundaries of the market, firm 1 at position 0 and firm 2 at position 1. Let $t$ denote the consumers' cost of travelling one unit of distance. Intuitively, the position of a consumer represents her "ideal variety" and the travelling costs represent her disutility from consuming a product that is different from her ideal. Hence, such a Hotelling setting can be interpreted as a model of consumer brand preferences with $t$ being the degree of brand differentiation. In addition, I assume that consumers' valuation for the good is high enough such that in equilibrium all consumers buy the base good.

As outlined before, consumers can decide whether to buy the add-on from the firm from which they purchased the base good or whether to use an outside substitution at cost $e$ (to be aquired in advance). I assume the tie-breaking rule that consumers decide to buy the add-on and not to substitute if they are indifferent. If a consumer did not substitute, she can only buy the add-on from the firm from which she bought the base good. Similar to GL, I assume that the add-on price is bounded above by $\bar{p}$, which can be interpreted as the maximum willingness to pay, the price of a last minute substitution or a regulatory price ceiling ($e < \bar{p},\ \hat{c} < \bar{p}$). I further assume that $\hat{c} < 3t$, which ensures that the inefficient firm 2 makes positive equilibrium profits and is thus willing to participate in the market.

one-shot game are indeed the same.

[4] See, among many others, Fudenberg and Tirole (2000) or Ellison (2005) for modeling consumer preferences in this way.

The timeline of the game is as follows:

- **Period 0:** Firms decide whether or not to unshroud the add-on.
- **Period 1:** Firms observe all (un)shrouding decisions and set prices $p_i$ for the base good and $\hat{p}_i$ for the add-on.
- **Period 2:** Consumers choose a firm to purchase the base good. Consumers can initiate substitution effort at cost $e$.
- **Period 3:** All consumers observe the add-on prices and purchase the add-on if they did not substitute in period 2.

The central point of the following analysis is to determine whether or not an equilibrium exists in which all firms shroud the add-on. Recall that one unshrouding firm suffices to reveal add-on prices and to educate the educatable consumers.

**Definition I.1.**

- *Shrouding equilibrium: An equilibrium in which all firms shroud the add-on.*
- *Unshrouding equilibrium: An equilibrium in which at least one firm unshrouds the add-on.*

The game is solved by backward induction. Calculating prices and profits in the shrouded and in the unshrouded subgame allows me to determine whether or not firms have an incentive to unshroud the add-on. At the pricing stage, the optimal pricing decision is dependent on the fractions of sophisticates and (remaining) myopes. This composition of the population is determined by the shrouding decisions of firms in the first stage of the game. In particular, if no firm unshrouded the add-on, the composition remains at its initial state and there are $\alpha$ myopic consumers in the market. If in contrast any firm unshrouded the add-on, there are $\alpha' < \alpha$ myopes remaining.

Firms compete via Hotelling competition. Myopic consumers do not consider the add-on and hence anticipate to pay only the price of the base good $p_i$. In contrast, sophisticates consider the add-on and substitute if they anticipate a high add-on price. Hence, they anticipate to pay $p_i + \min\{e, E[\hat{p}_i]\}$. To compute the value of this expression, the following lemma can be applied:

**Lemma I.1.** *Firms will choose add-on prices $\hat{p}_i$ only from the set $\{e, \bar{p}\}$.*

Suppose first that the add-on is unshrouded and hence its price is observable with $E[\hat{p}_i] = \hat{p}_i$. For an add-on price $\hat{p}_i < e$ sophisticates will not substitute but rather buy the add-on. Then, it holds that the firm could increase its profit by lowering $p_i$ by a small amount and raising $\hat{p}_i$ by the same amount. This leaves the demand and profit from sophisticates unchanged, but increases the demand and profit from myopes. For an add-on price $e < \hat{p}_i < \bar{p}$ only myopes buy the add-on. Then, a raise in the add-on price $\hat{p}_i$ has no effect on the demand but a positive effect on add-on revenues and thereby also on the firm's profit. Now suppose that the add-on is shrouded and hence its price is unobservable. In this case the demand of both types of consumers only depends on the base good price. Again it holds that for an add-on price $\hat{p}_i < \bar{p}$ the firm can increase its profit by raising this price and leaving the demand unchanged. Hence, firms will only set add-on prices such that $\hat{p}_i \in \{e, \bar{p}\}$. □

Now let us calculate the demand functions of consumers. Sophisticated consumers form Bayesian posteriors and hence correctly anticipate or observe add-on prices (depending on the shrouding decisions). Hence, they will substitute if the add-on price is $\bar{p}$. This implies that sophisticates will always pay $e$ because they either buy a low-priced add-on at price $e$ or substitute. Now I can apply the usual Hotelling logic. I try to find a consumer who is indifferent between buying at firm 1 or at firm 2. Denote the position of this indifferent consumers as $a \in [0, 1]$. Having found such a consumer, it holds that all consumers "left" to the indifferent one will prefer to buy from the "left" firm 1 while all consumers "right" to the indifferent one will buy from the "right" firm 2. Since the consumers are uniformly distributed on the interval $[0, 1]$, it follows that the demand of firm 1 is $D_1(p_1, p_2) = a$ and the demand of firm 2 is $D_2(p_1, p_2) = 1 - a$. A sophisticated consumer with location $a$ is indifferent if $p_1 + e + ta = p_2 + e + t(1 - a)$. In contrast to sophisticated consumers, myopic consumers only consider the price of the base good. Hence, a myope with location $a$ is indifferent between firms if $p_1 + ta = p_2 + t(1 - a)$. It then holds that the locations of the indifferent consumers are identical for both myopic and sophisticated consumers. The demand functions for the base good then take the following form:

$$D_1(p_1, p_2) = \frac{p_2 - p_1 + t}{2t} \qquad \text{and} \qquad D_2(p_1, p_2) = \frac{p_1 - p_2 + t}{2t}.$$

Now I can calculate the profits that firms earn in equilibrium. First suppose that the add-on is shrouded and hence consumers cannot observe add-on prices. Then, consumers cannot condition their base good choice or their substitution behavior on add-on prices. Hence, in the last stage when the add-on is purchased each firm has monopoly power over all its customers who did not substitute. Therefore, firms will employ high add-on prices $\hat{p}_i = \bar{p}$. This is correctly anticipated by sophisticated consumers who will engage in substitution and each firm will sell the add-on only to its $\alpha$ myopic consumers. Hence, the profit functions take the form:

$$\pi_1(p_1, p_2) = \frac{p_2 - p_1 + t}{2t}(p_1 + \alpha\bar{p}) \qquad \text{and} \qquad \pi_2(p_1, p_2) = \frac{p_1 - p_2 + t}{2t}(p_2 + \alpha(\bar{p} - \hat{c})).$$

Each firm individually maximizes its profit, yielding optimal base good prices $p_1 = t - \alpha\bar{p} + \frac{\alpha\hat{c}}{3}$ and $p_2 = t - \alpha\bar{p} + \frac{2\alpha\hat{c}}{3}$. Inserting these prices into the profit functions yields the following profits:

$$\pi_1^S = \frac{1}{2t}\left(t + \frac{\alpha\hat{c}}{3}\right)^2 \quad \text{and} \quad \pi_2^S = \frac{1}{2t}\left(t - \frac{\alpha\hat{c}}{3}\right)^2.$$

To determine the incentives of firms to unshroud the add-on, these shrouded profits have to be compared with the profits in the unshrouded subgame. If any firm has unshrouded the add-on, sophisticates and educated myopes can observe add-on prices. Hence, firms may have an incentive to set lower add-on prices to prevent their customers from substituting. First consider the efficient firm 1. Given the prices $p_2$ and $\hat{p}_2$, firm 1 can choose between the following profit functions:[5]

$$\pi_1^e = \frac{p_2 - p_1 + t}{2t}(p_1 + e) \quad \text{or} \quad \pi_1^{\bar{p}} = \frac{p_2 - p_1 + t}{2t}(p_1 + \alpha'\bar{p}).$$

Then, it holds that $\max_{p_1} \pi_1^e > \max_{p_1} \pi_1^{\bar{p}}$ if and only if $e > \alpha'\bar{p}$. Hence, the efficient firm 1 will choose low add-on prices $e$ if $\alpha' < \frac{e}{\bar{p}}$ and high add-on prices $\bar{p}$ if $\alpha' > \frac{e}{\bar{p}}$.[6] Applying a similar argument yields that the inefficient firm 2 will choose $\hat{p}_2 = e$ if $\alpha' < \frac{e-\hat{c}}{\bar{p}-\hat{c}}$ and $\hat{p}_2 = \bar{p}$ if $\alpha' > \frac{e-\hat{c}}{\bar{p}-\hat{c}}$. This yields the following profits in the unshrouded subgame:

| | $\hat{p}_1$ | Profit of firm 1 | $\hat{p}_2$ | Profit of firm 2 |
|---|---|---|---|---|
| $\alpha' < \frac{e-\hat{c}}{\bar{p}-\hat{c}}$ | $\hat{p}_1 = e$ | $\pi_1^{U1} = \frac{1}{2t}\left(t + \frac{\hat{c}}{3}\right)^2$ | $\hat{p}_2 = e$ | $\pi_2^{U1} = \frac{1}{2t}\left(t - \frac{\hat{c}}{3}\right)^2$ |
| $\frac{e-\hat{c}}{\bar{p}-\hat{c}} < \alpha' < \frac{e}{\bar{p}}$ | $\hat{p}_1 = e$ | $\pi_1^{U2} = \frac{1}{2t}\left(t + \frac{\alpha'\hat{c}}{3} - \frac{\alpha'\bar{p}-e}{3}\right)^2$ | $\hat{p}_2 = \bar{p}$ | $\pi_2^{U2} = \frac{1}{2t}\left(t - \frac{\alpha'\hat{c}}{3} + \frac{\alpha'\bar{p}-e}{3}\right)^2$ |
| $\frac{e}{\bar{p}} < \alpha'$ | $\hat{p}_1 = \bar{p}$ | $\pi_1^{U3} = \frac{1}{2t}\left(t + \frac{\alpha'\hat{c}}{3}\right)^2$ | $\hat{p}_2 = \bar{p}$ | $\pi_2^{U3} = \frac{1}{2t}\left(t - \frac{\alpha'\hat{c}}{3}\right)^2$ |

Note here that the assumption $\hat{c} < 3t$ ensures that the profits of the inefficient firm 2 are positive in all cases.

Having determined all potential profits now allows me to state the following proposition:

**Proposition I.1.** *With exception of the knife-edge case $\alpha' = \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$, there does not exist a shrouding equilibrium for all parameter constellations of $\alpha$ and $\alpha'$: (i) if $\alpha' < \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$, the efficient firm 1 will unshroud the add-on; (ii) otherwise, the inefficient firm 2 will unshroud the add-on.*

[5] Note that due to unshrouding the fraction of myopic consumers has shrunk to $\alpha'$.

[6] I will show in the Appendix that for special parameter constellations there also exists an equilibrium in which firms mix over add-on prices. However, this mixing has no effect on profits or on the incentive to unshroud.

First suppose it holds that there are only very few uneducatable myopes ($\alpha' < \frac{e-\hat{c}}{\bar{p}-\hat{c}}$).[7] Then, it holds that $\pi_1^{U1} > \pi_1^S$ and hence the efficient firm 1 has an incentive to unshroud the add-on. To get an intuition for this result consider the choice of add-on prices. In the shrouded subgame both firms set a high add-on price $\bar{p}$. In the unshrouded subgame both firms set add-on prices $e$. This change in add-on prices is worse for the inefficient firm since, due to its production costs, the add-on profits of the inefficient firm 2 fall more heavily with the reduction in the add-on price than profits of the efficient firm 1. In addition, the inefficient firm 2 profits less from the increase in add-on sales. This means that, compared to the shrouded subgame, unshrouding increases the profit differences between firms. Now recall that one of the implications of the existence of the add-on is that firms subsidize base good prices because they know that they will generate extra add-on revenues from customers that are "locked in" by having bought the base good.[8] This means that unshrouding lowers the value of attracting another customer for the inefficient firm 2. In other words, firm 2 is less willing to push add-on sales through low base good prices and firm 2 will compete less fiercely in the base good market. Following this logic, the efficient firm 1 anticipates that the reduction of competitive pressure leads to an increase of its own profits, thereby creating an incentive to unshroud the add-on. Note here that the incentives of firms run in opposite directions. The advantage that firm 1 creates for itself by shrouding constitutes a disadvantage for firm 2. Hence, in this case the inefficient firm 2 would prefer the add-on to be shrouded.

Now consider the case with a high fraction of uneducatable myopes ($\frac{e}{\bar{p}} < \alpha'$). In this case it holds that $\pi_2^{U3} > \pi_2^S$. The intuition goes as follows: In such a situation both firms will find it worthwhile, even in the unshrouded subgame, to set high add-on prices $\bar{p}$ in order to exploit the high fraction of uneducatable myopic consumers. As before, due to its inefficiency, firm 2 profits less from add-on sales than the efficient firm 1. Now recall that unshrouding reduces the fraction of myopic consumers from $\alpha$ to $\alpha'$. The resulting reduction in add-on sales is more damaging to the "high-margin" firm 1 than for the "low-margin" firm 2. Therefore, after firm 2 has unshrouded the add-on, the efficient firm 1 is competing less fiercely in the base good market. Note here that firm 2 also loses add-on profits, however the reduced add-on profits are offset by the slackened competition in the base good market. This means that the inefficient firm 2 can use unshrouding as a tool to reduce the competitive advantage of firm 1. In

[7] Note that this case might not exist since $\frac{e-\hat{c}}{\bar{p}-\hat{c}}$ is negative if $e < \hat{c}$.

[8] See Massoud et al. (2011) who argue that penalty fees in the credit card business (price of the add-on) are a direct substitute for interest rates (price of the base good). The authors show that firms try to attract additional customers by employing low interest rates while at the same time conserving their profits high by employing high penalty fees.

other words, unshrouding serves as a kind of sabotage device to diminish the size of the add-on market and to consequently reduce the competitive differences. Note again that while firm 2 has an incentive to unshroud the add-on, the efficient firm 1 would prefer the add-on to be shrouded.

Finally, consider the case where the fraction of uneducatable myopes is intermediate ($\frac{e-\hat{c}}{\bar{p}-\hat{c}} < \alpha' < \frac{e}{\bar{p}}$). In this case the efficient firm 1 will set a low add-on price $e$ while the inefficient firm 2 will set a high price $\bar{p}$. The logic behind the incentive to unshroud is similar to the two cases above, namely that unshrouding serves as a tool to amplify or to mitigate competitive differences. For the lower part of the interval ($\alpha' < \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$) it holds that $\pi_1^{U2} > \pi_1^S$ and hence the efficient firm 1 has an incentive to unshroud the add-on.[9] In contrast, for the upper part of the interval ($\frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}} < \alpha'$) it holds that $\pi_2^{U2} > \pi_2^S$ and consequently the inefficient firm 2 has an incentive to unshroud the add-on. □

Note here that it is not guaranteed that the region in which firm 1 unshrouds does exist since this requires $e > \alpha\hat{c}$. Hence, if $\hat{c}$ is sufficiently low, only firm 2 will have an incentive to unshroud. However, this has no impact on the result that for almost all possible parameter constellations at least one firm has a strict incentive to unshroud the add-on.

I have shown that for almost all fractions of (uneducatable) myopic consumers $\alpha$ and $\alpha'$ one of the firms will unshroud. This stands in contrast to the result of GL who conclude that a shrouding equilibrium exists for a sufficiently large fraction of myopic consumers, i.e. for $\alpha \geq \frac{e}{\bar{p}}$. The difference in results mainly stems from the fact that the underlying reasons for unshrouding are very different in the two setups. To see this suppose that firms would play a shrouding equilibrium. In the model of GL the reason to deviate to unshrouding lies in the intent to inform sophisticates about low add-on prices and thereby to prevent them from substituting. This potentially creates a short-term gain from unshrouding which is due to the incapability of the other firm to react to the deviation. This deviation is worthwhile if there are enough sophisticates in the market, or equivalently, if there are sufficiently few myopes $\alpha < \frac{e}{\bar{p}}$. However, this short term gain is not attainable if firms are able to condition their pricing behavior on shrouding decisions. Therefore, the result of GL is not robust with regard to the proposed strategic dimension of unshrouding.

[9] The Appendix contains a proof showing that if $\alpha' = \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$, it must hold that $\frac{e-\hat{c}}{\bar{p}-\hat{c}} < \alpha' < \frac{e}{\bar{p}}$. Hence, the threshold $\alpha' = \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$ indeed lies in the considered interval.

## 3. Discussion

I will continue by analyzing various alternative modeling assumptions. First I will analyze the impact of the two major modeling assumptions in which this paper differs from the work of GL, namely the timing of unshrouding and the heterogeneity in add-on profitability. Then I will continue by conducting additional robustness checks and analyzing further model extensions.

### 3.1. The Impact of Timing

First consider the timing of unshrouding and assume, like GL do, that shrouding and pricing decisions are taken simultaneously. In this case the strategic dimension of unshrouding vanishes. Suppose firms play an equilibrium in which all firms shroud the add-on. If any firm unilaterally deviates to unshrouding, there will be no impact on the prices of the competing firms due to the simultaneity. Also, since the game is played only once, there is no incentive to influence the profits made by other firms. Therefore, the only incentive to unshroud results from its impact on the demand of the deviating firm. It turns out that a firm will only unshroud if it sets a low add-on price such that, as in GL, the firm uses unshrouding to inform sophisticates about its low add-on price to prevent them from substituting. It then turns out that add-on production costs have no effect on the result of GL.

**Proposition I.2.** *In the game with simultaneous shrouding and pricing decisions and heterogenous add-on profitability, there exists a shrouding equilibrium if and only if the fraction of myopic consumers is sufficiently large:* $\alpha \geq \frac{e}{\bar{p}}$.

The proof is contained in the Appendix.[10]

### 3.2. The Impact of Asymmetric Add-on Production Costs

The second key difference to GL lies in the heterogeneity of add-on production costs, or more general, in the heterogeneity of add-on profitability. Suppose there would be no differences in add-on profitability and all firms earn the same equilibrium profits from add-on sales, say \$1 per sold add-on. When setting their base good prices firms know that they will make \$1 extra add-on profit from each customer. Now suppose that in a market without add-ons firms would set a price of \$10 for the base good.

[10] For completeness note that the efficient firm 1 has a stronger incentive to deviate from shrouding. In particular, firm 1 deviates from shrouding if $\alpha < \frac{e}{\bar{p}}$, while firm 2 deviates from shrouding if $\alpha \geq \frac{e-\hat{c}}{\bar{p}-\hat{c}}$. Since it holds that $\frac{e-\hat{c}}{\bar{p}-\hat{c}} < \frac{e}{\bar{p}}$, the range of parameters for which firm 2 wants to deviate is smaller than for firm 1.

Then, with the existence of the add-on, the equilibrium price will simply be \$9 for the base good since firms know they will get additional \$1 from the add-on. In effect, the existence of the add-on and the profits derived from add-on sales have no consequences for total firm profits. This is the classic Chicago school argument of add-on irrelevance, formally shown by Lal and Matutes (1994).

Indeed, when assuming that both firms are equally profitable in the add-on dimension (e.g. $\hat{c} = 0$), firm profits would be $\pi_i = \frac{1}{2t}(t)^2 = \frac{t}{2}$ in both the shrouded and the unshrouded subgame. It turns out the shrouding has an impact on the size of the add-on market and on add-on profits, but it does not have an impact on total firm profits. Therefore, firms are indifferent between shrouding or unshrouding.

### 3.3. Equivalence of the Model to a Repeated Play of the GL Setting

As indicated before, the results of the game with successive shrouding and pricing decisions are qualitatively similar to a repeated play of the game with concurrent shrouding and pricing. When shopping repeatedly, a consumer is likely to gain experience and to make wiser purchase decisions. For example, if a consumer bought several overpriced printer cartridges in the past and currently needs a new printer, she might be aware of high add-on costs and thus take measures to avoid or reduce them. Hence, in terms of the model, consumers are likely to maintain or even to increase their level of sophistication over their lifetimes. Cruickshank (2000), for example, confronted particularly experienced consumers with their actual bank account fees. Less than 10% of them said that the charges for international transfers, bill paying or cash machine withdrawals were higher than they had expected.

In terms of a repeated game setting this would mean that once myopes have been educated, they retain their level of sophistication for future periods. Therefore, firms would anticipate that their current shrouding decisions will have a future impact. Contrary to sticky consumer sophistication, firms can change prices quite flexibly, or in terms of a model they could change them in every period. Thus, such a repeated game setting and the incentives to unshroud are very similar to the proposed game with successive shrouding and pricing decisions, and indeed the qualitative results are similar.[11][12]

[11] The main difference stems from a lagged reaction to equilibrium deviations. If firms play a shrouding equilibrium and if one firm deviates, the other firm can react and change its price in the next period at the earliest. Hence, there is a possible short-term gain from deviating, which is not present in my model.

[12] This reasoning abstracts from any collusion that may arise in a repeated game.

### 3.4. The Impact of Observable Add-on Prices

Another natural variation of the model would be to assume that sophisticated consumers can always observe add-on prices, even if the add-on is shrouded. In this case the only effect of unshrouding would be to turn the educatable myopes into sophisticates. This seems to be a natural extension to my model since firms decide about unshrouding before prices are set and hence might not be able to inform consumers about particular add-on prices. As before, I now check whether any of the firms has an incentive to unshroud the add-on.

First note that if sophisticates can always observe add-on prices, firms have different add-on pricing incentives than in the main model. Up to now, consumers in a shrouded subgame could not observe add-on prices and hence could not condition their base good choice on add-on prices. Thus, add-on prices had no effect on base good demand or substitution behavior and firms had monopoly power in the add-on market over all consumers who did not substitute. This was correctly anticipated by sophisticated consumers who therefore engaged in substitution. This mechanism is obviously different if sophisticates can observe add-on prices in the shrouded subgame. For example, if a firm in a shrouded subgame sets a low add-on price $e$, sophisticates can observe this and will refrain from substituting. Therefore, firms in a shrouded subgame are able to sell the add-on to sophisticated consumers and will do so for certain parameter constellations.

First suppose that the fraction of initially myopic consumers is low ($\alpha \leq \frac{e-\hat{c}}{\bar{p}-\hat{c}}$). Then, in a shrouded subgame both firms would set low add-on prices $e$ and serve all consumers. Since unshrouding always lowers the fraction of myopic consumers, firms will also set low add-on prices in the unshrouded subgame. In both cases firms sell the add-on to all consumers and hence are indifferent about the distribution of the consumer population. Therefore, in this case no firm has a strict incentive to unshroud the add-on.

In all other cases ($\frac{e-\hat{c}}{\bar{p}-\hat{c}} < \alpha$) the population composition has an impact on profits. Note here that profits are symmetric in the sense that the advantage in profits of the efficient firm 1 is equal to the disadvantage in profits of the inefficient firm 2. Therefore, unshrouding, by affecting the consumer population, either increases or decreases these profit differences. Hence, for $\frac{e-\hat{c}}{\bar{p}-\hat{c}} < \alpha$ one of the firms will have an incentive to unshroud the add-on.[13] The efficient firm 1 will unshroud if this increases the differences in profits, while the inefficient firm 2 will unshroud if this decreases the profit differences.[14] This

[13] As in the main model, there exists one knife-edge case in which firms are indifferent, namely $\alpha > \frac{e}{\bar{p}}$ with $\alpha' = \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$.

[14] Note that, as before, the assumption $\hat{c} < 3t$ suffices to ensure positive profits for the inefficient firm 2.

leads to the following proposition:

**Proposition I.3.** *If sophisticates can always observe add-on prices, there exists a shrouding equilibrium if and only if*

$$\alpha \leq \frac{e-\hat{c}}{\bar{p}-\hat{c}} \text{ with } \alpha' < \alpha \qquad \text{or} \qquad \alpha > \frac{e}{\bar{p}} \text{ with } \alpha' = \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}.$$

The proof is contained in the Appendix. With exception of the knife-edge case $\alpha' = \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$, there only exists a shrouding equilibrium if the fraction of myopic consumers is low enough ($\alpha \leq \frac{e-\hat{c}}{\bar{p}-\hat{c}}$). Again, this result stands in contrast to the findings of GL who conclude that firms do not unshroud the add-on if the fraction of myopic consumers is large enough ($\alpha > \frac{e}{\bar{p}}$).

Note that for the constellations in which a shrouding equilibrium exists ($\alpha \leq \frac{e-\hat{c}}{\bar{p}-\hat{c}}$) both firms set low add-on prices $e$. Thus, firms will only refrain from unshrouding if they are not exploiting myopic consumers anyway. For all other cases in which at least one firm would set a high add-on price in the shrouded subgame, one firm will unshroud the add-on. In other words, there does not exist a shrouding equilibrium in which shrouding has negative consequences for myopic consumers.

### 3.5. The Impact of Unobservable Add-on Prices

I have assumed in the main model that a firm can reveal add-on prices before they are fixed. For instance, a firm could commit itself in advance to reveal its add-on price, for example by creating a website that transparently displays all fees charged by the firm. However, it might not always be reasonable to assume that the firm can also commit itself to reveal the add-on prices of the other firms. In terms of the model this would mean that unshrouding only reveals the add-on price of the firm that unshrouds.

The intuition and the qualitative results for this case are basically the same as in the main model. Firms have an incentive to strategically manipulate the composition of the consumer population in order to soften competition in the base good market. Due to this logic, at least one firm will have an incentive to unshroud the add-on. The only difference to the result of the main model is that firms potentially have an additional incentive to unshroud the add-on. This additional incentive to unshroud stems from the fact that if a firm has set a low add-on price $e$, it wants to inform consumers about it to prevent them from substituting. It will turn out that due to this reasoning there exist cases in which both firms have an incentive to unshroud the add-on.

To prove this result we first have to determine consumer expectations about the add-on prices of shrouding firms. Suppose that a firm did not unshroud and therefore its add-on price is unobservable. In this case, consumer decisions are independent of the unobservable add-on price and the firm has monopoly power in the add-on market.

As argued before, this implies that a shrouding firm will set an add-on price $\bar{p}$, which will be correctly anticipated by sophisticated consumers.

Now let us compare the pricing alternatives of firms and the corresponding optimal shrouding behavior. Suppose both firms set low add-on prices $e$. If both firms do so, they will unshroud since they want to inform sophisticates about their low add-on price to prevent substitution. The resulting profits are $\pi_1^{U1}$ and $\pi_2^{U1}$. Now consider the case in which firms set add-on prices $\hat{p}_1 = e$ and $\hat{p}_2 = \bar{p}$. This can only be an equilibrium if firm 1 unshrouds, thereby disclosing its add-on price and also educating the educatable myopes. The resulting profits are $\pi_1^{U2}$ and $\pi_2^{U2}$.[15] The last possibility is that both firms set high add-on prices $\hat{p}_1 = \hat{p}_2 = \bar{p}$. If no firm unshrouds, profits will be equal to $\pi_1^S$ and $\pi_2^S$. If any firm decides to unshroud, profits will be $\pi_1^{U3}$ and $\pi_2^{U3}$.

It follows that in equilibrium at least one firm has an incentive to unshroud the add-on. The easiest way to see this is by observing that firms can make the same profits as in the main model, namely $\pi_i^{U1}$, $\pi_i^{U2}$, $\pi_i^{U3}$, or $\pi_i^S$. By applying the same calculations as in the main model it holds that at least one firm has an incentive to deviate from a shrouding-strategy. Hence, as before, there does not exist a shrouding equilibrium. Note that, as said before, it might be the case that *both* firms have a *strict* incentive to unshroud the add-on. For example, if there are very few uneducatable myopes, then both firms set low add-on prices $e$ in equilibrium and both firms wish to unshroud in order to inform consumers of their add-on prices and prevent substitution.

### 3.6. The Impact of Specific Demand Frameworks

Until now I have assumed that firms compete via Hotelling competition. This assumption is more specific than in the model of GL who use a more general form of price competition with an arbitrary number of firms. Due to their symmetry assumptions, they are able to calculate equilibrium profits for very general forms of demand functions. However, to account for asymmetries, I had to use a more specific demand function. Using the Hotelling framework allowed me to calculate equilibrium profits in the given setting, which were then used to determine which firm wants to unshroud the add-on for given parameter constellations. I will now argue that the result that at least one firm has an incentive to unshroud the add-on is not an artefact of the Hotelling framework. I will first argue that the result transfers to other frameworks of duopolistic pricing competition. Then I will show that the result also transfers to competition between a larger number of firms.

The central assumption that both GL and I use is that each consumer buys exactly

[15] I am focusing on cases with $\hat{p}_1 \leq \hat{p}_2$. It is straightforward to check that a case with $\hat{p}_1 = \bar{p}$ and $\hat{p}_2 = e$ cannot be an equilibrium.

one unit of the base good. This assumption leads to Lemma I.1, namely that firms will choose add-on prices only from the set $\{e, \bar{p}\}$. Hence, sophisticates anticipate to pay $\min\{\hat{p}_i, e\} = e$ for the add-on or substitution. Myopes do not consider the add-on, or equivalently, anticipate that it is costless. Hence, for both consumer groups the anticipated costs for the add-on or substitution are equal for all firms. It follows that add-on pricing has no effect on the demand for the base good.

Given Lemma I.1, it follows that if firms are symmetric, the Chicago school argument of Lal and Matutes (1994) applies. Hence, all add-on revenues are used to subsidize the base good and the existence of the add-on has no impact on equilibrium profits. However, this is different if firms are asymmetric. A firm that has a comparative advantage in the add-on market will have higher equilibrium profits than a firm with a comparative disadvantage. However, the add-on revenues of the firms will still be influenced by the composition of the consumer population or, respectively, by the shrouding decisions of the firms. Hence, depending on parameter values, unshrouding will either increase or decrease competitive differences. On the basis of the same argument as in the main model this means that if both firms make positive equilibrium profits, one of them has an incentive to unshroud the add-on. In particular, the efficient firm will unshroud if this increases competitive differences while otherwise the inefficient firm will unshroud.

Now I can argue that this reasoning holds for a general number of firms. Suppose several asymmetric firms compete in prices. Firms sell both a base good and an add-on and, as before, every consumer buys exactly one unit of the base good. Due to their asymmetries, some firms will have a comparative advantage in the add-on market while others will have a comparative disadvantage. As before, an inefficient firm would unshroud if this decreases its comparative disadvantage. Similarly, an efficient firm would unshroud if this increases its comparative advantage.[16] Again, at least one firm has an incentive to unshroud the add-on.

### 3.7. Further Remarks

Still, several potential strategies are neglected in the above analysis. For example, in reality firms can voluntarily bind themselves to certain terms, i.e. a bank could advertise that its monthly fees for its current accounts will be kept at zero for the next two years. With such a commitment the bank would discard its ability to react to

[16]This reasoning holds as long as at least two firms make positive equilibrium profits and unshrouding does not change the number of firms that make positive equilibrium profits. If unshrouding increases the number of firms that participate in the market, the incentive to unshroud might be weakened.

actions of its competitors, and hence the competitors would have less scope for strategic maneuvering.[17] Another possible extension of the model is to allow for the invention of new products. Hence, once consumers start to understand certain products, they can be replaced by new products which customers do not yet understand. For example, Carlin and Manso (2011) base a model on this idea. Another extension of the presented model could be to allow for heterogeneous products, which would make it harder for consumers to understand all available products.[18]

## 4. Review of the Literature

This paper contributes to the growing literature of Behavioral Industrial Organization (see Spiegler (2011) for an overview). In particular, this paper studies the impact of limited consumer attention on firm incentives and behavior. Other papers that specifically study limited consumer attention in an Industrial Organization context are GL (2006), Armstrong and Chen (2009), Carlin (2009), Wilson (2010), Carlin and Manso (2011), and Ellison and Wolitzky (2011).

My work builds on the paper by GL, which discusses the impact of consumer myopia in add-on markets. Firms can costlessly unshroud and thereby reveal add-on prices and increase the fraction of informed consumers. GL show that there exist equilibria in which firms will not unshroud. The innovation in my work lies in allowing for a strategic dimension of unshrouding with wide-ranging implications on the results of the model.

Armstrong and Chen (2009) present a model in which some consumers know product prices, but are unaware of product qualitities. The authors show that firms will engage in both price and quality dispersion. However, the model of Armstrong and Chen (2009) does not allow firms to influence the level of consumer sophistication. In contrast, the models of Carlin (2009) and Carlin and Manso (2011) allow for such an influence of consumer sophistication. In their models, one part of the consumer population can observe prices and can therefore choose the cheapest firm. The remaining consumers cannot observe prices and randomly buy from any firm. The exact population distribution is dependent on the obfuscatory efforts of firms. Due to the existence of the uninformed consumers firms can sell overpriced products to them. The profit created by this conduct increases with the fraction of inattentive consumers, which

[17]Note that such a model of price-commitment would be related to a Stackelberg-game in which one firm moves before the other. In this case each firm would have an incentive to be the first mover, regardless of whether or not this includes educating the consumer population.

[18]For example, some banks offer to exchange domestic currencies into foreign currencies while other banks offer to exchange domestic currencies into Travelers Cheques.

consequently creates an incentive to increase the fraction of these consumers. This incentive to obfuscate is different in my model since, although the add-on is potentially overpriced, the exploitation in the add-on dimension is competed away in the base good dimension. Hence, in my model the incentive for shrouding does not stem from the aim to create a large group of consumers which can be exploited, but from the incentive to strategically increase or decrease the competitive advantages between firms.

Both Wilson (2010) and Ellison and Wolitzky (2011) show that obfuscatory equilibria exist in search-theoretic frameworks. Both works assume that it is costly for some consumers to learn the price of a firm. The authors show that this creates incentives for firms to increase the search costs that consumers incur. This can be interpreted as an attempt to increase obfuscation. However, the applied search-theoretic frameworks are very different from the model of pricing competition applied in this paper.

Work on add-on pricing has been conducted by Lal and Matutes (1994), Shapiro (1995), and Ellison (2005). Lal and Matutes (1994) formalize the classic Chicago school argument, arguing that profits earned with add-ons are competed away at the base good level. This leads to their "irrelevance result", which is also present in the model of GL, stating that the existence of the add-on has no impact on the profits of symmetric firms. This effect is not present in my model since the differences in add-on profitability lead to differences in firm profits. Shapiro (1995) argues that firms in markets with shrouded add-ons will always have an incentive to unshroud because this attracts customers. The central difference to my work is the availability of a substitution which offsets this particular customer-winning effect of unshrouding. Ellison (2005) models an add-on market in which consumers have heterogenous valuations for quality. The author then shows that add-ons can be used to price-discriminate between the consumer types. In an extension Ellison also proposes a behavioral interpretation of the assumed consumer characteristics, which was an inspiration for the work of GL.

## 5. Conclusion

I have shown that shrouding is not sustainable in a competitive market environment when firms exhibit heterogeneity in add-on profitability. Firms use unshrouding as a strategic device to alter their own demand structures and the ones of their competitors, thereby creating advantages for themselves. This strategic incentive arises from the intention to enhance or to mitigate competititive (dis)advantages. The incentive to unshroud the add-on depends on the population composition and, in particular, on the percentage of myopic consumers who cannot be educated by unshrouding. Dependent on this population composition either the efficient or the inefficient firm has an incentive to unshroud the add-on. For both cases there exist anecdotal examples which fit into the derived logic.

The model applies to market situations in which pricing is more flexible than consumer education or to situations in which a purchase decision is repeated over a consumer's lifetime. What is left out in the model are costs of unshrouding, which naturally decrease the incentive to educate consumers. Hence, unshrouding is likely to be observed in markets with significant differences in profitability between firms and in markets with relatively cheap unshrouding mechanisms.

# II. Collusive Shrouding and Cartelization

*We investigate the impact of consumer myopia on competition and firm behavior. In our model, firms repeatedly sell a primary good and a respective add-on. We study the impact of consumer myopia in the add-on market on pricing and on the ability of firms to engage in collusion. We show that in a situation in which firms cartelize and charge monopoly prices, limited attention makes deviation from such collusive behavior less rewarding and hence facilitates collusion. In particular, we determine the incentives of firms to educate consumers. We find that a shrouded market in which no firm educates consumers is a sign for cartelization. Hence, if obfuscation is observed in a market, it can serve as a proxy signal for illegal industry agreements.*

## 1. Introduction

In many markets product information is not easily available and consumers have difficulties when trying to collect the information that is relevant for their shopping decisions. In the terminology of the literature, building on the seminal paper of Gabaix and Laibson (2006), these are termed shrouded markets. Firms that participate in these markets can either foster or alleviate the degree of obfuscation. In the economic literature there is ongoing debate about what firms should do in such a situation. Gabaix and Laibson (2006) argue in a model of add-on markets that firms will not reveal add-on prices if there are enough myopic consumers, i.e. if there are enough consumers that do not incorporate all information that is available to them. In their model, if any firm unshrouds the add-on, some of the myopic consumers get educated and add-on prices get revealed. However, it turns out that the education of consumers does not have any strategic effect in their model since it happens so late that it does not have any impact on the game and on the incentives of the competing firms. Dahremöller (forthcoming) picks up this point and shows that, if the education of consumers has strategic implications for the game, firms rather have an incentive to unshroud the add-on.

We further expand this framework by designing an infinitely repeated game in which consumer education is a strategic variable in the sense that it has an effect on the payoffs and strategies of firms. We show that the natural equilibrium[1] of the game is one in which firms set competitive prices and unshroud the add-on. However, for sufficiently high discount factors there also exists an equilibrium in which firms collude on monopoly pricing. In particular, if all consumers are sophisticated and if the discount factor is high enough, firms can collude on monopoly pricing.

One of our central findings is that the existence of myopic consumers makes the collusive equilibrium more stable in the sense of lowering the critical discount factor for which collusion is sustainable. In terms of the model this implies that, dependent on the market constellation, firms have an incentive not to unshroud the add-on in order to keep the fraction of myopic consumers high enough for collusion to be stable. In addition, even if collusion is already stable, shrouding increases monopoly profits.

The fact that shrouding makes collusion more stable and, in many cases, is even a prerequisite for collusion has strong implications for competition analysis and antitrust regulations. First, since, dependent on parameter constellations, shrouding is a requirement for collusion, the regulator might intervene to decrease market obfuscation in order to destabilize collusion. Second, since again shrouding might be a requirement

[1] The term 'natural equilibrium' refers to an equilibrium in which firms play the equilibrium strategy of a respective one-shot game in each period of the game.

for collusion, the degree of obfuscation in a market might be a proxy signal for ongoing collusion in the market. Hence, since it is usually difficult to detect agreements on collusive pricing, it might be helpful for the regulator to consider the degree of obfuscation as an additional indicator for collusion.

One example of firms that coordinated on intransparency with regard to their products was the Lombard Club, which was a cartel of Austrian banks that was detected by the European Commision (see European Commission 2004). For example, at a meeting of the involved banks in 1994 "everyone agreed that, if questioned by the press or by the Association for Consumer Information for rate comparison purposes, they should in future stick to communicating only the (official) rates posted at the counter and not answer any further questions." In another agreement in 1996, the involved banks coordinated on valuing and pricing their portfolio lists only in Austrian Schilling while dropping any reference to the Euro. It is documented that the involved banks agreed that valuing and pricing in both Euro and Austrian Schilling would be more transparent to consumers, but they deemed that competition in this dimension should be avoided. In 1999 the involved banks agreed not to publicize a comparison of their savings products since this would open a way to "fresh competition."

More evidence for the connection between shrouding and collusive profits is presented by Brandenburger and Nalebuff (1998) for the U.S. airlines industry. In 1992, from an initial situation in which all pricing schemes in the industry were rather opaque and intransparent, American Airlines started a new pricing scheme "which emphasizes simplicity and equity and value." Its competitor United Airlines responded within forty-eight hours and in the following all other major competitors also quickly adopted simplified pricing schemes. Just three days later, another competitor, Trans World Airlines, revised its pricing schedule and severely undercut industry prices, which was again followed by quick price cuts of all major competitors. This example indicates that there might be a close relationship between the increase of transparency by the first firm and the following cascades of increased transparency and price cuts by the other firms in the market. This linkage between obfuscation and pricing will also be present in our results.

In our model, we analyze markets for goods whose total price consists of more than one element, for example markets for a base good and an add-on. If both the base good and the add-on are consumed, the effective total price of the product bundle is the sum of both prices. For instance, if a consumer considers buying a printer, she will not only have to pay the immediate price for the printer, but most likely have future expenses for compatible refill cartridges. Another application for our framework are goods that trigger future payments, for example subscriptions for which the total price is the (discounted) sum of all payments.

If a consumer wants to correctly calculate the effective total price of the product bundle, she needs to possess all relevant information and therefore a high degree of attention. If the attention of the consumer is limited though, she possibly does not fully recognize the effective total price. Consumers that exhibit this bias are called myopic consumers. As a result of their myopia, they may not be able to make rational consumption decisions. For example, Cruickshank (2000) reports that users of a current account seldomly fully understand all details of the contract and in most cases pay only little attention to add-on fees or other contract specifications. In line with these findings are the results from a survey considering consumer empowerment in the European Union which was conducted by TNS (2011). Addressing the question whether European citizens are sufficiently empowered as consumers, it is reported that almost six out of ten interviewees did not fully read the terms and conditions of the latest service contract that they signed (including contracts for gas, electricity, mobile phones, bank accounts, or insurances). Building on these results we assume that the consumer population contains a positive fraction of myopic consumers.

Considering such markets, it seems reasonable to assume that firms can exert some influence on the degree of obfuscation. Note here that in many markets there is little scope for obfuscation in the base good market since transparency in this dimension is necessary to attract consumers to the market. This is different for add-on markets which usually offer greater possibilities for firms to shroud information and prices. In our model, if all firms shroud the add-on, a fraction of $\alpha \in (0, 1]$ consumers is myopic and does not consider the add-on price. If any firm unshrouds the add-on, all myopic consumers get educated and behave like the sophisticated consumers for the current and all future rounds of the game.[2] In contrast to Gabaix and Laibson (2006) and Dahremöller (forthcoming) we assume that the total demand for the product bundle is not fixed. Instead, consumers have a personal valuation for the product bundle, which is heterogenous over the consumer population. If the valuation of a consumer is lower than the anticipated price of the product bundle, she will prefer not to buy. Note here that for given prices the total demand is higher if there are myopic consumers than if there are only sophisticated consumers. This is because myopic consumers underestimate the total price of the product bundle and hence are more likely to participate in the market.

Our main finding is that the existence of myopic consumers facilitates collusion. In

[2]Note here that we abstract from the possibility that firms can actively increase the fraction of myopic consumers $\alpha$. In reality it seems reasonable that firms can somehow increase $\alpha$. While we do not explicitly model this possibility, it will turn out that if a shrouding equilibrium exists, firms will have an incentive to increase $\alpha$.

terms of the model this implies that the critical discount factor for which collusion is stable is decreasing in the fraction of myopic consumers. The main driver of this finding is the fact that a deviation from collusion is less rewarding if many consumers are myopic. To get an intuition for this result consider a situation in which firms collude on monopoly pricing. In a model with only sophisticated consumers a firm that deviates and undercuts the collusive price attracts all consumers in the market and earns the entire monopoly profits. This is different if some consumers are myopic. Suppose first that a firm deviates by only lowering its price. Then myopic consumers do not perceive the change in the price and therefore will not switch to the deviating firm. Suppose second that a firm deviates by lowering its price and by unshrouding the add-on. Such a deviation would attract all consumers that still participate in the market, but would lower the total demand since some myopes realize that prices are higher than they anticipated. Both these effects make a deviation from collusion less rewarding and hence collusion is more stable. In other words, the existence of myopic consumers facilitates collusion.

In addition, even if collusion is already stable for a given population composition, firms have an incentive to continue increasing the fraction of myopic consumers. Since myopic consumers underestimate the total price of the product bundle, shrouding may trick these consumers into consumption. Hence, the total demand for the product bundle and the profits of the firms are rising with the fraction of myopic consumers.

Our analysis also yield several insights on welfare. If the consumers valuation for the product bundle is too low, a decision to buy the product bundle is inefficient and will be regretted by the consumer ex post. Therefore we find that shrouding has a negative impact on consumer welfare. This result has implications when applying our results to a regulatory perspective. We find that a regulatory intervention with the aim to unshroud the add-on is increasing total welfare if it can lower the fraction of myopic consumers $\alpha$ sufficiently. In addition, we find that a regulation to unshroud the add-on is always increasing the consumer surplus.

When considered from a regulatory perspective, our results suggest that regulatory tools with the aim to unshroud the add-on can impede collusion. If a regulatory intervention can decrease the fraction of myopic consumers $\alpha$ sufficiently, collusion can potentially be prevented or destabilized. Examples for such intervention would be informational campaigns to increase consumer sophistication or regulations that enhance market transparency. Even if such efforts to unshroud cannot decrease the fraction of myopic consumers $\alpha$ sufficiently in order to prevent collusion, they still increase consumer welfare as they either prevent consumers from making irrational choices or decrease the prices that consumers pay for the product bundle.

In addition to suggesting tools to impede collusion, our results also suggest new tools to detect collusion. We find that in many parameter constellations shrouding is necessary for collusion to be stable. Hence, a shrouded market is a potential indicator for illegal industry agreements. These markets then are candidates for increased scrutiny and inspections by governmental trustbusters. Traditional antitrust provisions like unannounced inspections or leniency policies were used in order to detect, prove, and prevent collusive industry behavior. However, historic evidence suggests that these tools were only partly successful in preventing collusion and cartelization. In particular, unannounced inspections and leniency policies were mostly targeted at disintegrating existing cartels. Our results suggest a new approach to detect active and intact cartels and prevent future cartel formation.

The remainder of this paper will proceed as follows. In Section 2 we give a short overview over the related literature. In Section 3 we present the main analysis and results. Section 4 will conclude.

## 2. Review of the Literature

The economic discourse on information disclosure by competitive firms was started by Grossman (1981) and Milgrom (1981). In their works rational consumers are imperfectly informed about product attributes and firms can credibly reveal the missing information. Within this framework, the authors show that competitive firms indeed have an incentive to reveal the missing information since this has a positive effect on their demands and profits.

One of the first works on obfuscation in add-on markets is provided by Shapiro (1995). He argues that there does not exist an equilibrium in which firms shroud the add-on. Shapiro argues verbally that there is a customer winning effect of unshrouding, which implies that a shrouding equilibrium is not stable.

Gabaix and Laibson (2006) were the first to question that view. Building on the work of Shapiro (1995), they use a model of add-on markets and assume that a given fraction of myopic consumers does not consider the add-on price in their purchase decisions. However, each firm can educate the myopic consumers and thereby help them to make more sophisticated decisions. The authors show that if the fraction of myopic consumers is large enough, there exists an equilibrium in which no firm has an incentive to educate consumers and all firms shroud the add-on.

This point is picked up by Dahremöller (forthcoming) who shows that the results of Gabaix and Laibson (2006) are strongly based on their modeling of unshrouding. In particular, Gabaix and Laibson (2006) use a single-period model in which firms can unshroud only in the last stage of the game. This implies that if any firm unshrouds the add-on, the other firms cannot react to this deviation. Also, unshrouding has no impact on any future payoffs since the game ends at this point. In terms of the model this implies that the education of consumers has no consequence for the play of the game and for the strategies of firms.[3] Examining these effects, Dahremöller (forthcoming) shows that if the education of consumers is modeled to have strategic implications for the game, a shrouding equilibrium no longer exists.

If one would transfer the framework of Gabaix and Laibson (2006) to a repeated game, unshrouding would have only short-term effects, while any long-term effects would be neglected. However, in our model we will follow Dahremöller (forthcoming) by assuming that unshrouding has long-term implications. Examples for long-term effects of consumer education include that firms condition their behavior on the play of the previous rounds and include that unshrouding permanently alters the consumer structure in future periods.

Kosfeld and Schüwer (2010) analyze the framework of Gabaix and Laibson (2006) from a regulatory perspective. They show that a regulatory intervention with the aim to increase consumer education can have positive as well as negative effects on welfare.

# 3. The Model

## 3.1. Model Setup

We model an infinitely repeated game. In each period $n \geq 3$ symmetric firms produce a base good and an add-on at zero costs. Each firm $i$ sets a base-good price $p_i$ and an add-on price $\hat{p}_i$. The common discount factor of firms is $\delta \in [0, 1]$. The consumer population has mass 1 and consists of $\alpha \in (0, 1]$ myopic consumers and $1 - \alpha$ sophisticated consumers. The fraction of myopic consumers only considers the base good prices $p_i$ and neglects the add-on prices $\hat{p}_i$. The remaining fraction of sophisticated consumers is fully informed, rational, and considers both the base good prices $p_i$ and the add-on prices $\hat{p}_i$. We assume that there exists a maximum price $\bar{p}$ for the add-on which can be interpreted as the cost of a last minute substitution or a regulatory usury ceiling.

As outlined before, we assume that in each period each firm can unshroud the add-on. If one firm does so, the myopic consumers get educated, which means that they behave like sophisticated consumers for the current round and all remaining rounds of the game. Consumers derive utility $v$ from consuming one unit of the base good

[3] In essence, Gabaix and Laibson (2006) show that firms potentially do not have an incentive to reveal their add-on prices. This result holds irrespective of whether this revelation of add-on prices is linked to an education of the consumer population.

and the respective add-on in a given period. If a consumer abstains from buying the product bundle, she gets zero utility. Each consumer buys at most one unit of the base good and one unit of the add-on. The "realized" utility of consumers is $U = v - p_i - \hat{p}_i$. However, myopic consumers mistakenly anticipate to get a utility from buying at firm $i$ of $U^m = v - p_i$. Hence, the myopic consumers do not anticipate their future need for the add-on or, equivalently, anticipate that the add-on price is zero. In contrast, sophisticated consumers correctly anticipate that their utility from buying at firm $i$ is $U^s = v - p_i - \hat{p}_i$. The consumption utility $v$ of each consumer is stochastic with $v \sim U[0, \overline{v}]$. The cdf of $v$ then takes the following form:

$$F(z) = \begin{cases} 0 & \text{if } z \leq 0, \\ \frac{z}{\overline{v}} & \text{if } 0 < z < \overline{v}, \\ 1 & \text{if } \overline{v} \leq z. \end{cases}$$

Since firms compete via Bertrand competition, a consumer buys the bundle at the firm that yields her the highest anticipated utility. If there are several firms that yield the highest anticipated utility, the consumer will choose any one of them with equal probability. In addition, if no firm yields positive anticipated utility, the consumer abstains from buying. The well known result of a one-period game with Bertrand competition is that firms earn zero profits. However, the infinite repetition of a game usually allows for a plethora of equilibria and firm strategies[4] and there exists no general mechanism for equilibrium selection. However, in the following, we assume that if firms collude on prices or on shrouding, they will coordinate on the equilibrium that yields the highest profit per firm.[5] In particular, since firms are symmetric, we focus only on these collusive equilibria that yield the highest aggregate firm profit and assume that this will be split equally among the colluding firms.

One implication of the existence of myopic consumers is that firms have an incentive to set high add-on prices along with low base good prices. To see this recall the utility functions $U^m$ and $U^s$. For given prices $p_i$ and $\hat{p}_i$, a firm can always increase the attractiveness of its product bundle for myopics while leaving the attractiveness of the bundle for sophisticates unchanged. The firm can achieve this by lowering the base good price $p_i$ by a small amount and increasing the add-on price $\hat{p}_i$ by the same small amount.

However, this logic of lowering base good prices and raising add-on prices is potentially limited. The reason is that there are several arguments that lead to a lower bound

[4] See, for example, Fudenberg and Maskin (1986).

[5] We will later show that an equilibrium with less-than-optimal profits would not be more stable in terms of collusion. Hence, there is no obvious reason for firms to collude on less-than-optimal profits.

for base good prices. For example, in real-world markets, base good prices cannot be negative due to potential arbitrage opportunities. If consumers receive money for the purchase of a good, they will buy as many units of the good as they can, creating unlimited profits for themselves and a loss for the firm that sells the good. In addition to the condition that prices must be non-negative, there potentially also exist reasons for positive lower price limits. One reason for a lower price bound is the possibility to resell parts of the base good. For instance, if consumers could buy a printer at zero costs and sell the copper wires or other parts of the printer for a profit, they might exploit this as an arbitrage opportunity. Another argument for price boundaries in the base good dimension is brought forward by Miao (2010) who argues that if the base good and the add-on are substitutes, there will be a lower limit for the base good price. Printers, for example, are sold with a starting cartridge. If a cartridge runs low, the consumer has the choice between buying a new cartridge or buying a new printer that is already equipped with a new starting cartridge. If the printer is very cheap compared to the refill cartridge, firms will not be able to sell their high priced refills. Miao (2010) shows that this creates a lower limit for the base good price.

Following the above argumentation, we impose a lower bound for the base good price. For simplicity we set this limit to 0. We will later show that this lower bound for the base good is reached if the following condition holds:

**Assumption II.1.** $\overline{v} \leq \bar{p}$.

We will assume this condition to hold for the remainder of the paper. Note that the assumption of a lower bound for the base good price creates results that stand in contrast to the traditional Chicago school argument on add-on pricing. Formalized, for example, by Lal and Matutes (1994) and Gabaix and Laibson (2006), the Chicago school argument reckons that high profits in the add-on dimension are fully competed away in the base good dimension. So suppose that, in a market with only a base good, firms would charge an equilibrium price of $p^\dagger$. Now suppose that an add-on is introduced, yielding an equilibrium add-on price of $\hat{p}^\dagger$. Then the Chicago school argument predicts that the new base good price will simply be the old base good price minus the new add-on price, i.e. $p^\dagger - \hat{p}^\dagger$. In other words, the Chicago school argument predicts that the base good fully subsidizes the add-on. Obviously, such a cross-product subsidization is not always possible if a lower boundary for the base good price exists.

### 3.2. Analysis

To determine the effect of the existence of myopic consumers we now want to compare a situation in which all consumers are sophisticated to a situation in which a fraction of $\alpha > 0$ consumers is myopic. We will then show that equilibria exist in which firms collude on shrouding. In particular, there exist constellations in which shrouding is necessary to allow firms to cartelize and thereby to jointly earn monopoly profits.[6]

### Only Sophisticated Consumers

First suppose that all consumers are sophisticated. Sophisticates take both the base good prices and the add-on prices into account. We now want to determine for which parameter constellations firms can cartelize. If firms can coordinate on monopoly pricing, they will set monopoly prices $p^M$ and $\hat{p}^M$ and earn aggregate monopoly profits $\pi^M$. The profits $\pi^M$ are split up equally between the firms such that each firm earns a profit of $\frac{1}{n}\pi^M$. If a firm expects to earn $\frac{1}{n}\pi^M$ for all future periods, the present value of these cash flows is $\frac{\pi^M}{n(1-\delta)}$. If any firm deviates from monopoly behavior and undercuts marginally, it will attract all consumers and make a deviation profit of $\pi^{dev}$. Since the deviating firm can undercut the monopoly prices only marginally and thereby attract all consumers, it would earn a deviation profit of the entire monopoly profits $\pi^{dev} = \pi^M$. This deviation will trigger a grim-trigger punishment by the other firms.[7] Hence, after such a deviation collusion breaks down and from that point onwards firms will compete via Bertrand competition and make zero equilibrium profits $\pi^{NC} = 0$ for all following periods ($NC$=non-collusive). The present value of the cash flows after deviation hence is $\pi^{dev} + \frac{\delta\pi^{NC}}{1-\delta}$. Now we want to determine the critical discount factor, which is the discount factor for which firms are indifferent between sticking to collusion and deviating from collusion. Applying our results, the critical discount factor is given by the solution of the following equation:

$$\frac{\frac{\pi^M}{n}}{1-\delta} = \pi^{dev} + \frac{\delta\pi^{NC}}{1-\delta} \tag{II.1}$$

$$\Rightarrow \delta^* = \frac{n - \frac{\pi^M}{\pi^{dev}}}{n} = \frac{n-1}{n}.$$

Thus, for all discount factors $\delta \geq \delta^*$ collusion is sustainable.[8]

[6] Note here that the main focus of our model is to test whether market obfuscation facilitates collusion. It may be the case that there exist other equilibria, which are not considered in our model, in which firms do not coordinate on prices but still shroud the add-on.

[7] See, for example, Mailath and Samuelson (2006) for a detailed discussion of grim-trigger strategies.

[8] Note here that a cooperation would not be more stable if firms would coordinate on profits that are lower than the monopoly profits. A deviating firm can always earn the aggregate collusion profit by undercutting marginally. Examining condition (II.1), there exists no coordination on lower-than-optimal profits that makes a deviation less rewarding. Furthermore, coordinating on higher prices than in the monopoly case does not help either since this would destabilize collusion as a deviating firm could undercut and set monopoly prices, thereby earning higher than collusive profits.

#### The Impact of Myopic Consumers

Now suppose that a fraction of $\alpha > 0$ consumers is myopic. Recall that, in contrast to sophisticated consumers, myopes do not consider the add-on prices $\hat{p}_i$. If all firms cooperate and charge prices $p$ and $\hat{p}$, their aggregate profit is given by:

$$\pi^M(p, \hat{p}) = [\alpha\,(1 - F(p)) + (1 - \alpha)(1 - F(p + \hat{p}))](p + \hat{p}).$$

It follows that if firms collude, they will set the base good price equal to its lower bound:

**Lemma II.1.** *Suppose $\overline{v} \leq \bar{p}$ holds. Then colluding firms will set their base good price at its lower bound, i.e. $p^M = 0$.*

The proof is contained in the Appendix. The intuition for this result is as follows. Suppose that firms set prices such that some sophisticates still participate in the market ($p + \hat{p} < \overline{v}$). Then for every base good price $p > 0$, the firms have an incentive to lower the base good price $p$ and to increase the add-on price $\hat{p}$. This would leave the total bundle price and the demand from sophisticates unchanged, but would increase the demand from myopic consumers. Suppose in contrast that firms set prices such that no sophisticated consumer participates in the market ($p + \hat{p} > \overline{v}$). Then in turns out that the add-on is profitable enough such that firms do not want to decrease demand by increasing the price of the base good.

Applying the result of Lemma II.1 to the aggregate profit function yields the following collusive profit:

$$\pi^M(\hat{p}) = [\alpha + (1 - \alpha)(1 - F(\hat{p}))]\hat{p}.$$

In the following analysis we have to distinguish between an 'inner solution' in which both types of consumers buy the product bundle ($\hat{p} < \overline{v},\ F(\hat{p}) < 1$) and a 'corner solution' in which only myopic consumers buy the product bundle ($\hat{p} \geq \overline{v},\ F(\hat{p}) = 1$). If we have an inner solution, both consumer groups have positive demand for the product bundle and the product bundle yields positive utility to some consumers. In contrast, the corner solution is characterized by an add-on price $\hat{p}$ that, if it would be fully considered, exceeds every consumer's valuation. In this case only myopics possibly consume the product bundle.

Now we want to derive the global maximum of the profit function. First suppose that firms play an inner solution. Then, the aggregate profit of firms takes the following form:

$$\pi^M(\hat{p}) = \left[\alpha + (1 - \alpha)\left(1 - \frac{\hat{p}}{\overline{v}}\right)\right]\hat{p}.$$

The aggregate profit is maximized by charging the monopoly price $\hat{p}^M = \frac{\overline{v}}{2(1-\alpha)}$. Given this price, profits are given by

$$\pi^M_{inner} = \frac{\overline{v}}{4(1-\alpha)}. \tag{II.2}$$

Now consider the corner solution with $\hat{p}^M > \overline{v}$. In this case firms only sell the product bundle to myopic consumers. Then the profit function of firms is given by:

$$\pi^M(\hat{p}) = [\alpha]\,\hat{p}.$$

Obviously, it is optimal to charge the highest possible add-on price $\hat{p}^M = \bar{p}$, yielding a profit of:

$$\pi^M_{corner} = \alpha\bar{p}. \tag{II.3}$$

It will then depend on parameter constellations whether firms will prefer the equilibrium with the inner solution or the equilibrium with the corner solution.

Note here that the optimal price of the inner solution $\hat{p}^M = \frac{\overline{v}}{2(1-\alpha)}$ is larger than the maximum valuation $\overline{v}$ if $\alpha > \frac{1}{2}$. If $\alpha > \frac{1}{2}$ it holds that $1 - F(\hat{p}^M) = 0$ and therefore no sophisticated consumer buys the product bundle. Hence, for $\alpha > \frac{1}{2}$ the inner solution is not feasible in the sense that the derived maximum does not lie in the specified interval. If that is the case, the corner solution will be the global profit maximum.

To get an intuition for the form of the profit function we have depicted two possible constellations in the following graphs:

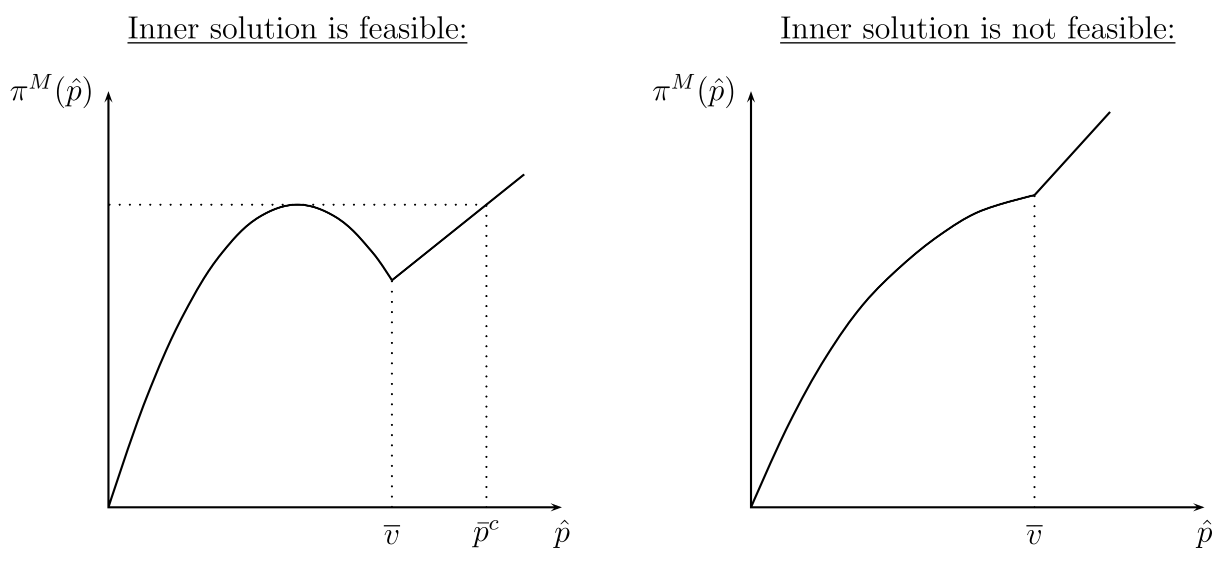


Figure II.1.: Two possible functional forms of the aggregate firm profits.

For a sufficiently low share of myopic consumers $\left(\alpha \leq \frac{1}{2}\right)$, the inner solution is feasible in the sense that the profit-maximizing price of the inner solution does not exceed the maximum valuation $\overline{v}$. The inner solution then corresponds to the global maximum of the profit function if the upper bound for the add-on price is not too large, i.e. if $\bar{p} \leq \bar{p}^c$. If, however, the fraction of myopic consumers is large enough with $\alpha > \frac{1}{2}$, the maximum of the inner solution is not feasible anymore. In this case, only the corner solution can be optimal.

When examining the profit functions, we see that the aggregate profit for both the inner solution and the corner solution are increasing in the fraction of myopic consumers $\alpha$. This leads to the following result:

**Proposition II.1.** *Monopoly profits are increasing in the share of myopic consumers* $\alpha$.

This finding mainly stems from the impact that myopic consumers have on the demand function. Myopic consumers are always more likely to buy the product bundle since they underestimate its total price. Hence, in both the inner and the corner solution, the demand for the product bundle is increasing in the fraction of myopic consumers. This directly implies that the monopoly profits of firms are also increasing in the share of myopic consumers.

Now let us analyze how firms would collude and how a potential deviation strategy would look like. If firms cartelize and coordinate on monopoly prices, they are able to maximize aggregate profits, which will then be split up equally among them. Clearly, all firms prefer these monopoly profits over perfect competition with zero equilibrium profits. Nevertheless, there may be an individual short-term incentive to deviate from monopoly pricing: A firm may deviate by either unshrouding the add-on and/or by setting a lower or a higher price than the one that was set in the collusive phase. Lemma II.1 implies that if a firm wants to deviate from collusion and attract further customers, it can only do so by lowering it's add-on price, but not by lowering its base good price. Note that it is not obvious whether an optimal deviation involves unshrouding the add-on. This is because unshrouding potentially has partly negative effects since, if myopes are turned into sophisticates, they might refrain from buying the product bundle. Therefore, for given prices, unshrouding is decreasing the demand for the product bundle. In the following, we will show that despite its negative effect on demand, an optimal deviation from collusion comprises unshrouding the add-on. To see this recall that, due to low base good prices and high add-on prices, firms generate their profit through add-on sales. Myopic consumers do not incorporate these add-on prices into their purchase decisions. Hence, myopes do not react if the deviating firm changes its add-on price. This creates an incentive to unshroud since, the education induces formerly myopic consumers to take the the change in the add-on price into account. In the following, we will show that this effect dominates the aforementioned reduction in demand and hence a deviating firm will unshroud the add-on.[9]

First note that firms will never deviate by increasing the add-on price. If firms play a corner solution and charge the highest possible add-on price, raising the add-on price is not feasible. If firms play an inner solution and a firm deviates by raising its add-on price, sophisticates will prefer to buy from the other firms and only myopic consumers buy the product bundle from the deviating firm. This is because myopes do not take the add-on price into account and therefore will not change their behavior after a change in the add-on price. Hence, the deviating firm will optimally set the maximum add-on price $\hat{p}^{dev} = \bar{p}$, yielding a deviation profit of $\pi^{dev} = \frac{1}{n}\alpha\bar{p}$. It then holds that $\pi^{dev} \leq \frac{1}{n}\pi^{M}$ and a deviation yields lower profits than the profits earned by sticking to the collusive play.[10] Therefore, increasing the add-on prices is not a profitable deviation.

We can conclude that if a firm decides to deviate, it will undercut the collusive add-on price. Recall that it is not possible to undercut the base good price since it is already at its lower bound. Hence, the firm can only undercut in the add-on price dimension. If a firm undercuts the add-on price, it attracts all sophisticated consumers. In particular, the deviant firm has two possibilities to undercut the collusive add-on price:

The first possibility is that the firm undercuts the add-on price and unshrouds the add-on, thereby educating all myopic consumers. This lures a larger share of consumers to the deviant firm because the fraction of price sensitive consumers has increased. At the same time unshrouding potentially crowds many formerly myopes out of the market because, by taking the add-on into account and learning about higher than anticipated add-on prices, some myopes realize that they would receive negative utility from buying and therefore decide to refrain from the market.

The deviation profit that results if a firm unshrouds the add-on takes the following

[9]Note here that the result that a deviating firm will always unshroud the add-on may be an artefact of our particular population distribution. However, while this result is convenient for the analysis, it is not necessary for our results. To get an intuition for this, recall that if the add-on is unshrouded, firms make zero profits in the competitive equilibrium. Now suppose that firms initially collude and any firm deviates, but does not unshroud the add-on. Then the other firms could unshroud the add-on as part of their grim-trigger strategies. Since there are at least two firms that do so ($N \geq 3$), they have no individual incentive not to unshroud the add-on since there exists at least one other 'punishing' firm that still unshrouds. Hence, if any firm deviates, but does not unshroud, the non-deviating firms will react by unshrouding the add-on and the following non-collusive profits will again be zero.

[10]This result follows from the fact that we looked at a situation in which the inner solution was optimal in the collusive play ($\frac{\bar{v}}{4(1-\alpha)} \geq \alpha\bar{p}$).

form:[11]

$$\pi_i^{dev} = \left[1 - \frac{\hat{p}_i^{dev}}{\overline{v}}\right] \hat{p}_i^{dev}. \tag{II.4}$$

Maximization of (II.4) yields:

$$\hat{p}_i^{dev} = \frac{\overline{v}}{2}.$$

This price is feasible regardless of the strategies that firms played in the collusive phase since $\hat{p}_i^{dev} \leq \overline{v}$ and $\hat{p}_i^{dev} \leq \hat{p}^M$. Inserting the price into the profit function then yields a deviation profit of:

$$\pi_i^{dev} = \frac{\overline{v}}{4}.$$

The second possibility is that the firm undercuts the add-on price but decides against unshrouding the add-on, leaving the fraction of myopic consumers at $\alpha$. In this case, the deviant firm will only attract sophisticated consumers because myopes do not take notice of the change in the add-on price. The deviation profit in this case is equal to:

$$\pi_i^{dev} = \left[\frac{\alpha}{n} + (1-\alpha)(1 - F(\hat{p}_i^{dev}))\right] \hat{p}_i^{dev}. \tag{II.5}$$

If a firm deviates and decides not to educate consumers, there again could be an 'inner' solution and a 'corner' solution. The 'inner' solution is the case for which $1 - F(\hat{p}_i^{dev}) > 0$. Maximization of (II.5) then yields:

$$\hat{p}_i^{dev} = \frac{\overline{v}}{2(1-\alpha)} \left(1 - \alpha\frac{n-1}{n}\right),$$

yielding a deviation profit of:

$$\pi_i^{dev} = \frac{\overline{v}}{4(1-\alpha)} \left(1 - \alpha\frac{n-1}{n}\right)^2. \tag{II.6}$$

The 'corner' solution of the deviation is the case for which $1 - F(\hat{p}_i^{dev}) = 0$. In this case it is optimal to set $\hat{p}_i^{dev} = \bar{p}$. The deviating firm would attract no additional consumers. It would only attract myopic consumers such that the total demand for the firm would be $\frac{1}{n}\alpha$. This deviation might yield a higher profit than (II.6) if the upper bound on the add-on price is extremely high. The add-on price $\hat{p}_i^{dev} = \bar{p}$ would then yield a deviation profit of

$$\pi_i^{dev} = \frac{\alpha\bar{p}}{n}. \tag{II.7}$$

[11] Note that it suffices to consider the case $\hat{p}_i^{dev} \leq \overline{v}$: If a deviating firm decides to unshroud the add-on, all consumers are sophisticated and only potentially purchase the add-on if $\hat{p}_i^{dev} \leq \overline{v}$. Hence, deviating with the corner solution cannot be optimal.

Closer inspection of (II.7) yields that $\hat{p}_i^{dev} = \bar{p}$ can never be a profitable deviation. Note that playing $\hat{p}_i^{dev} = \bar{p}$ and deciding not to educate consumers only corresponds to an actual deviation from collusive behavior if the colluding firms played the inner solution with profits of $\pi^M_{inner} = \frac{\overline{v}}{4(1-\alpha)}$. It follows directly that the deviation profit (II.7) must be lower than the shared collusive profits.[12]

To sum up, if a deviation is profitable, a deviating firm that unshrouds the add-on earns a maximum deviation profit of $\pi_1^{dev} \equiv \frac{\overline{v}}{4}$. The maximum profit that a deviating firm can obtain if it does not unshroud is $\pi_2^{dev} \equiv \frac{\overline{v}}{4(1-\alpha)}\left(1-\alpha\frac{n-1}{n}\right)^2$. As we show in the Appendix it holds that $\pi_1^{dev} > \pi_2^{dev}$. Hence, the following result applies:

**Proposition II.2.** *If a profitable deviation exists, a firm that deviates from collusive play will unshroud the add-on.*

As we have mentioned above, there exist two opposing effects of unshrouding. On the one hand, unshrouding increases the number of sophisticated consumers who notice the deviation. On the other hand, unshrouding crowds out formerly myopic consumers. At first glance it was not clear which of these effects is generally stronger, but now we can argue that the positive effect dominates the negative one.

Now we want to determine the effect that the existence of myopic consumers has on the stability of collusion. To do this we determine the critical discount factors. If firms play an inner solution in the collusive phase, the critical discount factor is given by:

$$\delta^{inner} = \frac{n - \frac{1}{1-\alpha}}{n},$$

which is falling in the share of myopic consumers $\alpha$. If collusion was characterized by a corner solution, the critical discount factor takes the form:

$$\delta^{corner} = \frac{n - \frac{4\alpha\bar{p}}{\overline{v}}}{n},$$

which is also falling in $\alpha$. Now we want to show that the critical discount factor is globally falling in $\alpha$. Since $\delta$ is falling piecewise, it suffices to show that $\delta$ has no 'jump' when the optimal monopoly strategy changes from the inner solution to the corner solution. At the point of at which firms are indifferent between inner solution and corner solution, it holds that $\frac{\overline{v}}{4(1-\alpha)} = \alpha\bar{p}$. Then it immediately follows that for this parameter constellation $\delta^{inner} = \delta^{corner}$. This suffices to ascertain that the critical discount factor $\delta$ is continuous in $\alpha$ and we can conclude:

**Proposition II.3.** *The critical discount factor $\delta$ is globally falling in the fraction of myopic consumers $\alpha$.*

[12]To see this recall that the inner solution is only optimal in collusion if $\alpha\bar{p} < \frac{\overline{v}}{4(1-\alpha)}$.

The intuition behind this finding lies in the fact that a deviation from collusion is less rewarding with the existence of myopic consumers. First, recall that monopoly profits and, with it, individual collusion profits are increasing in the fraction of myopic consumers. This is because myopic consumers underestimate the price of the product bundle and therefore are more likely to buy it. Hence, for given prices, the total demand is increasing in $\alpha$. Second, we have shown that a deviating firm optimally unshrouds the add-on. Since then all consumers are sophisticated, total demand is independent of $\alpha$ and for given prices lower than in the collusive phase. Hence, the higher the initial share of myopic consumers, the less attractive a deviation gets when compared to the collusive play, and therefore the more stable is collusion.

Up to now we have implicitly assumed that collusion cannot be made more stable by coordinating on lower-than-maximal collusive profits. That this holds true was straightforward to prove in the case with only sophisticated consumers. However, we can show that it also holds in the case with myopic consumers.

**Proposition II.4.** *Collusion cannot be made more stable by coordinating on other than monopoly profits.*

The proof for this result is contained in the Appendix. The basic idea is to show that if the firms would coordinate on lower-than-optimal collusive profits, the deviation profit would decrease sufficiently such that a deviation is not more attractive than in the case in which firms coordinate on maximum collusive profits.

We now have established our formal results regarding the critical discount factor $\delta$. Now we want to discuss the resulting implications for cartel stability. As we have shown before, the critical discount factor $\delta$ is falling in $\alpha$. The central implication of this result is that the higher the share of myopic consumers $\alpha$ is, the easier it is to sustain collusion. This implies that firms may have an incentive to raise the fraction of myopic consumers. If, initially, there does not exist a collusive equilibrium, active shrouding can potentially decrease the critical discount factor sufficiently, such that collusion becomes sustainable. In addition, shrouding results in higher collusive profits and hence is beneficial for firms even if collusion is already sustainable without additional obfuscation.

We can now conclude that in a market with a positive fraction of myopic consumers the critical discount factor is always strictly lower than it would be if the whole consumer population was sophisticated. This follows directly from the result that the critical discount factor is globally falling in $\alpha$.

**Corollary II.1.** *The existence of myopic consumers facilitates collusion.*

Another result that directly follows from closer inspection of the critical discount factors is the following:

**Corollary II.2.** *The critical discount factor rises with the number of firms.*

The result that collusion is less stable if the number of firms rises is not particularly new. However, our results indicate that, if the number of firms has an impact on the stability of collusion, the number of firms also has an impact on the whether or not firms shroud the add-on. This is because a breakdown of the collusive shrouding equilibrium leads to an unshrouding of the market. One study that supports this result is Miravete (2007), who presents a study about the U.S. cellular telephone industry. He shows that the entry of new firms to the market tends to 'lift the fog' and leads to more transparent pricing schemes. This finding is in line with our results that an increase in the number of firms may destabilize collusion, which in turn leads to unshrouded and transparent markets. Our results also are along the lines of Brandenburger and Nalebuff (1998) who observed in the airline industry that one initial initiative to increase transparency from one firm triggered a cascade of increased transparency and price cuts by all other firms in the market. In terms of our model this can be interpreted as a deviation from collusive play by one firm, which resulted in a punishment strategy by the other firms, leading to a breakdown of collusion and an increase in market transparency.

### 3.3. Welfare Analysis and Regulatory Intervention

We will now analyze the welfare implications of potential regulatory interventions. To answer this question we have to determine the effects that the firm behavior, in particular shrouding the add-on, has on welfare. We argued before that the regulator may want to prevent collusion and shrouding, but the exact effects on welfare have not been thoroughly derived yet. In line with many previous authors like O'Donoghue and Rabin (1999), we deem the true consumer utility to be the relevant measure of consumer welfare. This true consumer utility stands in contrast to the anticipated consumer utility that we interpret as determining the choice of consumers, but as having only a distorted connection to the real utility of consumers.[13]

The total welfare is simply the sum of the valuations of all consumers that buy the product bundle. The price that consumers pay for the product bundle has no impact on welfare because it simply is a redistribution from consumer surplus to industry profit. If all consumers buy the product bundle, for example if firms play the unshrouded competitive equilibrium, welfare would be $E[v] = \frac{\overline{v}}{2}$. If, however, firms play the shrouded

[13] This implies that myopic consumers do not act according to their own best interest. In other words, their myopia is not a sign of different taste, but a sign of a particular malfunction of their decision behavior.

collusive equilibrium, not all consumers buy the product bundle, which results in a lower total welfare. Hence, the following Lemma applies:

**Lemma II.2.** *The total welfare is higher in the unshrouded competitive equilibrium than in the shrouded collusive equilibrium.*

The result that shrouding is detrimental for welfare is in line with the findings of Gabaix and Laibson (2006). Note, however, that their results stem from the assumption that there exists a substitution that can replace the add-on. In particular, these papers assume that the substitution is lost in terms of welfare. However, in many cases it seems plausible that some part of its price is not completely lost.[14] In this case, total welfare would be independent of the shrouding decisions of firms.

Now we want to determine the effects of firm behavior on the consumer surplus. Obviously, the case in which firms charge competitive prices is better for consumers than the case in which firms charge monopoly prices and shroud the add-on. In addition, the higher the fraction of myopic consumers, the higher tends to be the charged price and the more consumers buy the product although it yields negative utility to them. Hence, the following Lemma applies:

**Lemma II.3.** *The consumer surplus is higher in the unshrouded competitive equilibrium than in the shrouded collusive equilibrium. The consumer surplus is falling in the fraction of myopic consumers* $\alpha$.

The derived results have wideranging implications for regulatory policies. The most obvious regulation would be to force firms to offer their products at marginal cost. Needless to say, this might not be enforcable in real world markets. However, there are other kinds of regulations that can also have positive effects on welfare. For example, our results give new insights into the usefulness of regulations with regard to consumer education and market transparency. The traditional reason for such regulations was that these should enable consumers to make wiser purchase decisions, which in turn was supposed to increase consumer welfare. Our paper presents another reason for such regulations that, in terms of the model, are intended to reduce the fraction of myopic consumers. If the regulator can unshroud the add-on and thereby force firms to play the unshrouded competitive equilibrium,[15] both consumer surplus and total

[14] Consider, for example, the case of hotel telephones. If a consumer brings her cell phone with her to save the costs of the hotel line, the calling costs for the cell phone are not lost in terms of welfare, but are part of third-party firm profits.

[15] For example, if the regulator can decrease the fraction of myopic consumers $\alpha$ to zero, the critical discount factor increases to $\delta^* = \frac{n-1}{n}$. Hence, if the common discount factor of firms is not sufficiently high, unshrouding the add-on is likely to destabilize collusion.

welfare can be increased. Note here that a regulatory intervention is also increasing the consumer surplus if the intervention cannot decrease the fraction of myopic consumers to zero. First, a regulatory intervention that decreases $\alpha$ also increases the critical discount factor $\delta^*$, which might make collusion infeasible. Second, even if collusion still is stable, a regulatory intervention nevertheless increases consumer welfare since the consumer surplus decreases in the fraction of myopic consumers.

Another innovation of our model lies in its predictions on the detection of cartels. Traditional competition policy had to watch out for active arrangements or coordination between firms in order to detect collusive behavior. We have argued that the level of obfuscation may be artificially increased by firms in order to stabilize collusion. Hence, the regulator can use the degree of obfuscation in a market as a proxy for the degree of cartelization. This relation seems a useful extension to traditional antitrust monitoring since obfuscation is usually far easier to detect than active coordination between firms.

Apart from active consumer education, the regulator can also intervene by reducing barriers to entry for new firms or by employing other measures that increase the number of firms that participate in the market. This may inhibit collusion because, as we have shown, the critical discount factor is increasing in the number of firms.

## 4. Conclusion

We have proposed a model of limited attention in which competitive firms can either shroud or unshroud the add-on market. We have shown that two kinds of equilibria exist. In one equilibrium firms collude on monopoly pricing and shroud the add-on. In the other equilibrium firms set prices competitively and unshroud the add-on. The equilibrium in which firms shroud the add-on is only stable if the discount factor of the firms is above a critical discount factor. It turns out that this critical discount factor is decreasing in the fraction of myopic consumers. Hence, firms might try to increase the degree of obfuscation and thereby increase the fraction of myopic consumers, which in turn will tend to stabilize collusion. Another incentive to increase the level of obfuscation is that the profit that firms earn when they collude is increasing in the fraction of myopic consumers.

These results suggest several implications for welfare analysis and regulatory intervention. We find that welfare is maximized if firms do not collude and do not shroud the add-on. Hence, the regulator might employ measurements to increase consumer sophistication. If these measurements are sufficiently efficient, the collusive equilibrium breaks down and only the natural competitive equilibrium remains.

Our results also suggest new insights into competition policy. We have shown that shrouding might be used by firms as a tool to stabilize collusion. Hence, the degree of obfuscation in a market might be a proxy for the degree of collusion and hidden industry agreements. Markets with high obfuscation then are candidates for further investigation by the antitrust agencies.

# III. Product Lines, Product Design, and Limited Attention

*We analyze how firms design their product lines when facing customers with limited attention. We assume that consumers simplify complex decision problems by neglecting several of the relevant aspects. Whether and to what extent a customer pays attention to an attribute of a product depends on the importance of the attribute as well as its dispersion in the set of alternatives. A firm may thus influence its customers' attention through the range of products it makes available. We show that a firm can increase its profit by introducing goods that have the sole function of manipulating consumer attention. We derive several results on how a firm can profitably employ such manipulating goods. In particular, even with a homogenous consumer population our model can explain product differentiation.*

## 1. Introduction

We propose a model of limited attention and analyze its implications for product design. As most consumer products have several characteristics, any purchase decision is a complex problem of trading off advantages in some dimensions against disadvantages in others. We posit that the way in which a decision-maker pays attention to different aspects of the problem reflects a need to simplify such complex decisions. The resulting attention allocation embodies both the decision-maker's valuation of the different aspects of a product and the extent to which the available options (products and outside option) differ in a dimension. We investigate how a firm optimally designs its product line when facing customers whose attention is determined in such a way. We start with the optimal design of a single product that is offered by a monopolist. The optimal design will not only reflect the extent to which customers value a particular characteristic, but will also reflect to which extent customers consider a particular attribute (the attention allocation) and how a particular design changes this attention allocation. We then investigate how a firm can profit from expanding its product line. We show that there is an incentive to offer multiple products to a set of customers with homogeneous preferences. This incentive to offer differentiated products stems from the property that the attention allocation is determined by the choice set. Finally, we discuss the optimal design of a product line.

We show that limited attention tends to decrease the complexity of the offered products. This means that products are simpler in the sense that they are equipped with fewer features and hence are easier to grasp for consumers. One of the central propositions of the paper is that firms will offer at least two kinds of products. There is a primary good that is intended to be sold to consumers. In addition, firms produce what we call bait goods. These are not intended to be sold but have the sole purpose of manipulating consumer attention in a way that is favorable to the firm.[1] We find that the optimal attention manipulation weighs the incentive to redirect attention to more profitable attributes against the incentive to maximize the attention that is paid to each attribute of the product. We find that, in contrast to much of the previous literature, it turns out that the firm tends to suffer from consumers having limited attention. This is due to the fact that limited attention tends to decrease the willingness-to-pay of consumers, which in turn lowers the price that a firm can charge for a given product.

[1] The assumption that a bait good is only used to manipulate consumer attention is useful to simplify the analysis. We later present an extension in which the bait good can also be profitably sold to some consumers. One example would be luxury goods that are sold to a small population of rich consumers and at the same time serve as a bait good to the larger population of less wealthy consumers.

While customers may not fully consider the price either, we show that a firm is not able to exploit this inattention.

In our model each product has a set of attributes. For example, a car can be equipped with features like a sunroof, electric windows, or with certain levels of horsepower or safety. Apart from these rather salient characteristics there are many others that are less noticeable. Examples would be the average durability of the gear box, the seating comfort, or the design of the cigarette lighter. Hence, such complex products like a car have a plethora of different attributes that all add to the overall quality of the product. Although information is freely available for lots of these attributes, consumers usually focus on a few key variables. Such systematic neglect of important information can arguably be classified as limited attention.

We join an emerging strand of the economic literature that incorporates such limited attention into economic models. Gabaix (2011) develops a model of limited attention in which the attention allocation is the solution to an optimal attention allocation problem. Our paper draws heavily on his work, as the attention allocation we assume can be derived from an attention allocation problem with cognitive costs similar to the one proposed by Gabaix (2011). However, our modeling of attention allocation also differs from his approach as we employ some modified assumptions we deem appropriate to reflect the costs of complexity. A more detailed comparison is deferred to later. In addition, the focus of our paper are the implications of limited attention on market outcomes like product design and price. Kőszegi and Szeidl (2012) propose a model of focus-weighted utility which exhibits some parallels to our model. They show that a firm has an incentive to concentrate the advantages of its products in few dimensions while spreading the disadvantages across as many dimensions as possible. While such an incentive is also present in our model (yet for different reasons, as we will explain in the next section), we derive more detailed results regarding the optimal product design. Zhou (2007) studies a monopolist's optimal product design if an advertising technology is available that highlights some of the product's characteristics. He investigates the potential and consequences of screening if customers are differently susceptible to such advertising. In our model, the way how customers allocate their attention is not determined by advertising, but by the attributes of the offered products. As this describes a more specific manipulation "technology", we are able to investigate the scope and limitations of such a manipulation. Also working on limited attention, Spiegler (2006a) and Spiegler (2006b) study the optimal industry behavior if consumers act according to the $S(1)$ sampling routine developed by Osborne and Rubinstein (1998). Beyond the attention heuristic employed, we deem the most important difference to be the welfare implications of limited attention. In Spiegler (2006a) and Spiegler (2006b) customers can be exploited by firms which obfuscate their products. Similarly, Rubinstein (1993) describes a firm's incentive to use complex pricing schemes to extract additional profits from boundedly rational customers. In contrast, we highlight that limited attention may primarily hurt the firm while benefiting customers, despite a firm's ability to manipulate attention. Eliaz and Spiegler (2011) propose a model that also features products whose sole function is to attract attention. Consumers only consider the products of a subset of the firms in the market. Therefore, a firm uses attention grabbers if it wants consumers to consider its products. In contrast we investigate a firm's potential to attract or distract attention from product characteristics (including the price), thereby manipulating the desirability of a purchase. Bordalo, Gennaioli, and Shleifer (2010) and Bordalo (2011) develop a framework of limited attention to account for choice set effects. Their idea of limited attention is inspired by psychological findings concerning the perception of alternatives. While these papers focus on the explanation of several experimental results, we seek to develop predictions concerning the impact of limited attention - which in our framework refers to cognitive restrictions, not errors in perception - on market outcomes.

There is a large literature on choice set effects (Simonson 1989; Huber, Payne, and Puto 1982) and their impact on behavior in various settings (Herne 1997; McFadden 1999; Benartzi and Thaler 2002). Several explanations for compromise effects[2] have been proposed - ranging from extremeness aversion (Simonson and Tversky 1992) to information inference from choice sets (Wernerfelt 1995; Kamenica 2008). Kamenica (2008) shows that information inference creates an incentive to offer premium loss leaders. Though not explicitly relating his model to the compromise effect Vikander (2010) proposes a model of status effects and describes a firm's incentives to advertise premium products to an audience which is not able to afford the purchase. As an example, Vikander (2010) presents some anecdotes of advertisement campaigns for premium goods. For instance, Audi advertised its $118,000 R8 in the half-time of Super Bowl XLII. This advertisement spot cost Audi six million dollars. What is puzzling about this story is that only a minority of Super Bowl viewers are able to afford such a car. So presumably if Audi wanted to advertise this car to people who are able to buy it, there would be other far less costly marketing channels to reach the particular target audience. We will show that in our model there is an incentive to offer premium products which are not intended for sale. Yet in our framework this incentive is based on a firm's ability to manipulate its customers' attention, and not

[2]The compromise effect posits that expanding the choice set by a product which is more extreme in one attribute than any of the previously available options makes products which are mediocre in that attribute look more favorably.

on a firm's attempt to signal product value (Kamenica 2008) or increase its products' prestige value (Vikander 2010).

Johnson and Myatt (2006) describe how a firm may optimally design its product(s) to increase or decrease the dispersion of customer valuation. They find that a firm wants to concentrate its product's value in a single characteristic if a firm is confined to offer products with fixed expected value and there exists one characteristic for which customer tastes vary strongly. Furthermore, they investigate a monopolist's incentives to expand or contract its product line as the taste dispersion changes. In contrast, we discuss an incentive to expand the product line without any taste dispersion.

The remainder of this paper will proceed as follows. In Section 2 we introduce and analyze the underlying attention process. Section 3 is dedicated to the derivation of the optimal product design if only one product can be supplied. This will be extended in Section 4 where firms can introduce an additional product that is designed to manipulate consumer attention. Section 5 analyzes the effects of a firm offering several of these manipulating products. We continue by discussing some possible extensions in Section 6. Section 7 concludes.

## 2. The Role of Attention

This section introduces the model that we use to depict the customers' way of deciding between alternatives in their choice set. A main feature is that there is a difference between *experienced utility* and *decision utility*. Experienced utility measures the satisfaction or welfare that customers derive from a choice. It depends on the attributes an alternative features and the value a customer assigns to these attributes. In contrast, decision utility depicts the way in which customers choose between alternatives. Decision utility thus not only depends on the welfare a customer derives from an alternative, but also on the way choices are made, and thus encompasses choice procedures, perceptions of alternatives (at the moment of choice), salience of attributes and alternatives, and the like. The distinction is supposed to depict the contrast between utility as a measure of welfare and utility as a tool to model choice (behavior).[3] In our model the salience of attributes results in a difference between experienced utility and decision utility. The reason behind this discrepancy is the limited cognitive ability of humans to decide optimally on the plethora of information that is available to them. To be able to focus solely on the limited ability to process information, we do not model any costs of obtaining information or costs of searching. Instead, we will assume

[3] For a distinction of the concepts of experienced utility and decision utility, see Kahneman and Tversky (1984) and Kahneman (2000).

that all product information is readily available, but that consumers have problems of converting product information into an overall assessment of desirability. While we introduce the process of attention allocation and consumer choice and the underlying intuitions in the main text, we will prove in the Appendix that the proposed attention allocation can itself be derived from an optimization problem.

We assume experienced utility to take the following structure, where the experienced utility of an alternative $a \in A$ is denoted by

$$u(a) = \sum_{i=1}^{n} v_i x_i^a.$$

The term $v_i \in \mathbb{R}$ measures the (marginal) value of attribute $i$ to the decision maker (henceforth DM). The variable $x_i^a$ measures the extent to which alternative $a$ features attribute $i$. We will assume that $v_i \neq v_{i'}, \ \forall i \neq i'$.

The consumer does not base her choice on experienced utility, but on decision utility. Formally, the decision utility of an alternative is a function of the experienced utility and the salience of each attribute:

$$\tilde{u}(a) = \sum_{i=1}^{n} m_i v_i x_i^a. \tag{III.1}$$

The term $m_i \in [0,1]$ is the attention parameter associated with alternative $i$. If $m_i = 0$, attribute $i$ is completely neglected. In this case, any differences between alternatives in attribute $i$ will be irrelevant for the decision. We normalize the attention such that $m_i \leq 1, \ \forall i$.[4] Since attention is normalized such that $m_i \leq 1$, an attention allocation of $m_i = 1$ means that the DM pays full attention to attribute $i$. Note here that we would be back in the rational model if $m_i = m_j > 0, \ \forall i, j$.

We now want to motivate and analyze the process that determines the attention parameters $m_i$. First, note that we do not model a problem of strategic attention allocation. While an optimization problem may, to some extent, underly the way in which attention is distributed, we assume it to be given at the moment of choice. Instead, we assume a particular rule how the salience parameters $m_i$ are determined and give empirical and analytical reasons why this rule is sensible. We then seek to investigate how a firm designs its products when faced with customers who allocate attention as described.

Note that we do not intend to model perceptional mistakes. The DM is able to perfectly determine differences between alternatives in each dimension. The limitations in the cognitive process arise when the DM needs to integrate the information

[4] Note that this normalization is without loss of generality since a decision utility with some attention vector $\mathbf{m}$ yields the same choice behavior as a decision utility with an attention vector $\alpha \cdot \mathbf{m}$ for $\alpha > 0$.

about differences between alternatives in multiple dimensions with his own evaluation concerning the importance of each dimension in order to reach an evaluation of each alternative. This task usually includes making judgments as to how an advantage in one dimension trades off against a disadvantage in another dimension. As the number of relevant dimensions increases, so does the cognitive effort of evaluating all resulting trade offs. We seek to model one way in which a decision-maker may deal with such a complex decision problem, namely by simplification through neglect and prioritization of aspects of the decision problem.[5]

We want the attention allocation to reflect a need to simplify a complex problem in order to be able to solve it. We thus want the attention allocation to follow some basic rules:

(a) The level of attention that a DM allocates to an attribute rises with its importance to the DM, i.e. with the level of $v_i$.

(b) The more the alternatives that the DM faces differ in a given attribute, i.e. the larger the contrast within one dimension, the higher is the attention that the DM allocates to that attribute.

(c) The cost of considering another attribute rises with the number of attributes that are already considered.

We will now give a short overview of the assumptions and their impact on the attention process.

Assumption (a) is present in most models of limited attention. The DM allocates more attention to those dimensions that she deems more important. When making a decision she tends to focus on those dimensions from which she derives most value.

Assumption (b) captures the property that the salience of a dimension increases in the extent to which the alternatives differ in the particular attribute dimension. Hence, the consumer's attention is responsive to the set of available alternatives. Therefore, attention is neither fixed nor random but depends on the context. In particular, we will assume that attention is a function of the total dispersion in an attribute. This is in line with models like Kőszegi and Szeidl (2012). Also note that the general intuition of a decision-maker neglecting small differences between options can already be found in Rubinstein (1988). Rubinstein argues that decisions as observed in the well-known

[5] This might be the largest conceptual difference from other models of salience and limited attention. In our model, the reason why the weightings $m_i$ vary across attributes is not that differences between alternatives are perceived as being larger or smaller than they actually are. In contrast, the weightings $m_i$ express the extent to which differences between alternatives are appreciated in the decision process.

Allais paradox can be explained by decision-makers treating similar probabilities or payoffs as equal.[6]

Assumption (c) describes the idea that the DM finds it increasingly difficult to consider an additional attribute when the complexity of the problem increases. This assumption implies that there is an attention hierarchy as it matters which attributes are considered "first".[7] Attributes that are ranked higher in the attention hierarchy are more likely to be taken into account. In addition, we assume that this increased complexity cost is reflected in a lower attention weight given that the attribute is in fact considered. Thus, differences in attributes that are ranked higher are appreciated to a greater extent. Put shortly, each attribute has a rank in the attention hierarchy and a higher rank means that the attribute receives more attention.

In particular we assume that the ranking is strict in the sense that for two attributes that get positive attention, the attribute that ranks higher in the attention hierarchy is allocated strictly higher attention than the lower ranking attribute ($\forall m_i > 0, m_j > 0: \; m_i \neq m_j$). This is important because it implies that behavior is indeed distorted by limited attention. To see this, recall from the functional form of the decision utility (III.1) that if attention would be uniformly dampened (e.g. with $m_i = 0.5, \; \forall i$), decision utility would just be a uniform transformation of experienced utility. In this case, a decision based on experienced utility is always the same as one based on decision utility. For limited attention to have a behavioral effect we need at least two attributes $i$ and $j$ which are allocated different levels of attention, i.e. $m_i \neq m_j$. Our assumptions make sure that any two attributes which are considered receive a different level of attention.

To construct an attention hierarchy which satisfies both (a) and (b), let $\mu_i$ measure the product of the valuation of attribute $i$ and the maximal difference in attribute $i$ between any two alternatives in the choice set:

$$\mu_i = v_i \left( \max_{a \in A} x_i^a - \min_{a \in A} x_i^a \right).$$

$\mu_i$ is the maximal difference in experienced utility a DM faces in dimension $i$ in any binary comparison of two alternatives in the choice set. We will assume that an at-

[6] Empirical results of Chetty, Looney, and Kroft (2009) indicate that consumers indeed often tend to pay higher attention to more dispersed attributes. They report that consumers are more sensitive to changes in the excise tax, which is included in the posted price, than to changes in sales tax, which is added at the register. As the posted price is larger than the additional tax that is added at the register, the two alternatives "buying"/"not buying" differ more in the posted price dimension than they differ in the additional tax-dimension. Thus, the more salient dimension seems to be the more dispersed one.

[7] Note that the attention hierarchy is not meant to literally depict the timely sequence of how the DM takes attributes into account. Instead the attention hierarchy depicts how important an attribute is in the decision.

tribute's rank $r(i)$ in the attention hierarchy will be determined by the measure $\mu_i$. Let $r : i \to \{1, ..., n\}$ be the function that assigns an attention rank to each attribute such that:

$$\mu_i > \mu_{i'} \Rightarrow r(i) < r(i'),$$
$$\mu_i = \mu_{i'} \text{ and } v_i > v_{i'} \Rightarrow r(i) < r(i'),$$

where $r(i)$ denotes the attention rank of the attribute $i$.[8] Having determined the attention order we now turn to a cardinal measure of attention based on the attention hierarchy. Let $m_i$ denote the attention paid to attribute $i$ and define:[9]

$$m_i = \max\left\{0, 1 - \kappa_{r(i)}/\mu_i\right\}.$$

The threshold $\kappa_{r(i)}$ is the minimum level of $\mu$ which an attribute with attention rank $r$ needs to have in order to be taken into account. The threshold $\kappa_{r(i)}$ can be thought of as the cognitive cost of considering the $r$-th dimension of a problem. This cognitive cost should reflect the fact that consumers have difficulties when having to choose between several product that are different in many attributes. We make the following assumptions with regard to the thresholds $\kappa_r$:

**Assumption III.1.**
*(i)* $\kappa_1 = 0$*: There is always one attribute that is fully considered.*
*(ii)* $\kappa_r < \kappa_{r+1}, \ \forall r \in \{1, ..., n-1\}$*: The attention threshold is increasing with each additional attribute that is considered.*

Part (i) of the assumption ensures that there is always at least one attribute to which the DM pays full attention. The assumption captures that we model the DM to have difficulties with the complexity of comparing several attributes. If there is only a single dimension, the DM should face no problems. Part (ii) captures that the difficulty of considering an additional attribute increases with the number of attributes that are already considered.

Assumption III.1 highlights the importance of the order of alternatives which is depicted by $r(i)$. First, the higher an attribute ranks in this order, the more likely it is attended to. If an attribute ranks low in the hierarchy, it is associated with a higher threshold $\kappa_{r(i)}$, and hence it is more likely to be neglected. Attention is a function of both *value* ($v_i$) - more valued attributes are more likely to be considered - and *contrast* - attributes in which alternatives differ greatly are more likely to be considered.

[8] Note that $r(i)$ is well-defined since we assumed $v_i \neq v_j, \ \forall i \neq j$.

[9] To let $m_i$ always be well-defined let $m_i = 0$ if $\mu_i = 0$ and $\kappa_{r(i)} > 0$, and let $m_i = 1$ if $\mu_i = 0$ and $\kappa_{r(i)} = 0$.

### 2.1. Comparison to Other Models of Limited Attention

The proposed model of attention allocation is conceptually very close to the one of Gabaix (2011). It takes over the idea of sparsity, meaning that some attributes are neglected by the DM if they are not "important enough". A major difference to the attention process of Gabaix is that the threshold $\kappa_{r(i)}$ associated with an attribute $i$ depends on an attribute's *relative importance*, i.e. on how large its value and dispersion are compared to other attributes' values and dispersions. While we retain the assumption that the vector of thresholds $(\kappa_1, ..., \kappa_n)$ is exogenous, we impose a structure that we deem appropriate to reflect the notion of complexity costs. First, the decision-making process should not be distorted if the problem is not complex. We ensure this property by assuming $\kappa_1 = 0$. Second, we argue that an increase in complexity should be reflected by an increased difficulty to consider more and more dimensions of a problem. This motivates our assumption of increasing thresholds $\kappa_{r(i)}$.

A second difference is the distinction between the salience, captured by $m_i$, and the value of an attribute, $v_i$. This distinction might seem superfluous at first glance as neither can be observed in isolation. Yet this simple distinction spares us the need to normalize our parameters to make salience independent of scaling. We find this feature desirable as we are concerned that some of the behavioral implications which Gabaix derives are based on this rescaling. In that sense certain behavior is predicted to occur not because people are inattentive but because this inattention is argued to be scale-independent.

The model we use also shares a central property with the model of focus-dependent utility of Kőszegi and Szeidl (2012). They posit that the weighting of an attribute ($m_i$) is a strictly increasing function of the maximal utility difference associated with that attribute ($\mu_i$). Our model shares this feature conditional on an attribute being considered, but not fully considered ($0 < m_i < 1$). However a major difference is that our model also allows attributes to be fully neglected. Kőszegi and Szeidl (2012) apply their model to a monopolist's problem to design a single product. They find an incentive to concentrate value of a product in one attribute and an incentive to split the price (and thus the disadvantage of a purchase) into multiple items. We are concerned that the first result is not an effect of the assumed focus-weighting alone but also depends strongly on the assumed cost function. In contrast, in our model there is an incentive to concentrate value that is entirely due to the weighting of attributes, i.e. in our context due to inattention. The reason is that if customers tend to neglect attributes of little importance, firms have an incentive to concentrate on few attributes in which they excel. The incentive to split the price into several items is also present in our model and is further strengthened by the consumers' tendency to neglect.

Bordalo, Gennaioli, and Shleifer (2010) present a model of salience-based choice. In contrast to our model, the modeling of salience is inspired by observations on perception (and not cognition) and assumes choice to be the result of multiple binary comparisons. More technically, the weighting does not only depend on the available choice set but also on the current binary comparison. Thus, an attribute's salience in an alternative may be quite different depending on which two alternatives are currently examined.

Closer to our understanding of attention is the model of rational inattention of Sims (2003). Like him, we do not seek to model costly information acquisition or information production (if one wants to understand perceptional biases in this way), but the problems associated with processing available information. Yet, while he focuses on the limits of data processing-capability, we seek to model the cognitive costs associated with solving a complex optimization problem.[10]

## 3. Optimal Product Design of a Monopolist

In this section we investigate the problem of a monopolist who wants to design a single product that is intended to be sold to customers with the described attention process. What is central to our analysis is that attention is a function of the choices that are available. This implies that a firm is able to manipulate attention since it can influence the choice set that a consumer faces. We will begin our analysis by restricting the monopolist to a product with a single quality. We proceed by asking whether the monopolist would like to improve this product by introducing additional qualities. Thereby we derive the optimal design of the product. We then discuss under what conditions and in which way a firm may profit from introducing further products, and how the optimal product line is designed.

Suppose there are $n$ qualities (save the price) that a product can have. Together with the price a product thus features $n+1$ attributes.[11] For now suppose that the level of each quality can take any non-negative real value: $x_i \in \mathbb{R}_+,\ \forall i = 1, ..., n$. There is one attribute $x_p$ which denotes the wealth of the decision maker. W.l.o.g. we normalize initial wealth to zero such that a value $x_p^a \leq 0$ means that alternative $a$ is associated

[10] As an illustration of Sims' idea, think of a savings problem to which the optimal rule is to consume half the income: $c_t = 0.5y_t$. Suppose income is a random variable and takes on the value 10.458376 at some $t$. In Sims' model the adaptation of $c_t$ to the optimal value (5.229188) is costly as is requires the processing of the 8-digit input $y_t$. Note however that in Sims (2003), *finding* the optimal rule itself is not subject to cognition cost (though the anticipation of processing cost may alter the optimal solution itself). In contrast, we focus on the impact of limited attention on the *derivation* of the optimal solution.

[11] Note here that we differentiate between the $n$ qualities and the $n+1$ attributes, which are all qualities and the price.

with a price of $P = -x_p^a$. Before we turn to the optimal design problem define a *null good* as an alternative with $x_i^0 = 0,\ \forall i = 1, ..., n$.

A monopolist seeks to design a product that maximizes profits subject to the customer's willingness to buy it. To ensure that the customer is willing to purchase the good, the decision utility of the good (alternative $a$) must be weakly higher than the decision utility of abstaining from the purchase (alternative $b$). Note that alternative $b$ is equivalent to a null good that is free of charge. Therefore, not buying is associated with a decision utility (and experienced utility) of zero. Part (i) of Assumption III.1 then implies that the monopolist cannot extract a positive profit by selling a null good at a positive price. Thus the product the monopolist designs must actually feature some positive qualities.

Let the costs of producing quality level $x_i$ be $c(x_i) = \frac{1}{2}c_i x_i^2$. The monopolist then maximizes his profit subject to the decision utility of alternative $a$ being non-negative:

$$\max_{P,\ x_i} \quad P - \frac{1}{2}\sum_{i=1}^{n} c_i x_i^2,$$

$$\text{subject to} \quad \tilde{u}(a) \geq 0.$$

First, let us look at the case in which a monopolist employs a single quality to equip his product with. In this case, the set of feasible prices is constrained by the fact that the price must rank second in the attention hierarchy: $v_i x_i \geq v_p P \Rightarrow r(i) = 1, r(p) = 2$. Otherwise the decision utility from buying the product would be strictly negative. This yields a maximum price $P^* = v_i x_i / v_p$.[12] Then it follows that the profit maximizing choice of the quality level $x_i$ is:

$$x_i^* = \frac{v_i}{v_p c_i}, \tag{III.2}$$

while the resulting profit is

$$\Pi^* = \frac{v_i^2}{2c_i(v_p)^2}. \tag{III.3}$$

Now let us define the profitability measure $\pi_i$:

$$\pi_i \equiv \frac{v_i^2}{2c_i(v_p)^2},\ \forall i = 1, ..., n.$$

The value $\pi_i$ is the profit that a monopolist can extract from producing a product that features only quality $i$. We will refer to it as the *profitability of quality* $i$ as it also denotes the maximum additional profit a monopolist could make by adding

[12] We assume here that $v_i > v_p$, so the tie due to $\mu_i = \mu_p$ is broken in favor of the quality dimension. If this is not the case the price is set marginally below $v_i/(v_p c_i)$.

this quality to his product under full attention. It is straightforward to see that a monopolist who is confined to produce only a single-quality product will choose the quality with the largest $\pi_i$. It is interesting to compare this result to the one under full attention (full attention is equal to the case $\kappa_{r(i)} = 0,\ \forall r(i)$). First, note that if only one quality is produced, price, product design, and profit are equal to the solution under full attention. The sole difference is that under limited attention the decision utility of purchasing the product is strictly positive. While the price extracts all the value created by the quality, it is not perceived to do so (because it is not fully taken into account). Still it is not possible for the monopolist to exploit this and increase the price further. If the price was increased slightly above the derived price $P^*$, the decision value created by the quality would drop due to a reallocation of attention from quality to price, thereby yielding negative decision utility and preventing the purchase altogether.

Now suppose the monopolist contemplates to introduce a second quality. This is only profitable if it allows to increase the price sufficiently to cover the cost of producing an additional quality. Yet, recall that the optimal price of the optimal single-quality product was already at the threshold of receiving attention rank 1, $v_p P^* = v_i x_i^*$. It follows that any price increase beyond the price of the single-quality product moves the price dimension to the top of the attention hierarchy. In other words, if the optimal product has at least two qualities, its price will rank first in the attention hierarchy. This shift in the attention rank lowers the decision utility derived from the first quality that the product already features and thus also lowers the profit that the monopolist can extract from the first quality.

As before, the maximal price the monopolist can charge is given by the non-negativity constraint on the decision value of the product. Thus the monopolist will charge a price

$$P^* = \frac{1}{v_p}\left(m_i v_i x_i + m_j v_j x_j\right).$$

The optimal levels of the two qualities then are:

$$x_i = \begin{cases} \frac{v_i}{v_p c_i} & \text{if } m_i(x_i = v_i/(v_p c_i)) > 0, \\ 0 & \text{otherwise,} \end{cases}$$

$$x_j = \begin{cases} \frac{v_j}{v_p c_j} & \text{if } m_j(x_j = v_j/(v_p c_i)) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

It is straightforward that the introduction of an additional quality is only profitable if both qualities receive positive attention. But even if this is the case, the resulting profit may fall below the maximal profit of a product with a single quality. This is not because the new quality distracts attention from the quality the product was already equipped with. This may be the case but it does not change profit. Instead, the potential decrease in profits results from the fact that the price increase necessary to make the introduction of a new quality worthwhile distracts attention from the qualities and attracts attention to the price. Formally, the profit of the product with two qualities (if both of them are considered) is given by:

$$\Pi^* = \pi_i + \pi_j - \kappa_2/v_p - \kappa_3/v_p. \tag{III.4}$$

Thus, it is more profitable to sell a two-quality product if and only if:

$$\pi_j \geq \frac{1}{v_p}(\kappa_2 + \kappa_3).$$

The introduction of an additional quality is only worthwhile if the revenue that can be extracted from the second quality is sufficient to cover the physical costs of providing this second quality, the cognitive costs of considering this second quality, and the decrease in profits that is caused by the downgrade in the attention rank of the first quality.

Note that from the profit expression (III.4) it becomes clear that the firm is not benefiting from limited consumer attention. If consumers were fully attentive, the firm could make a profit of $\Pi = \pi_i + \pi_j$. The fact that the firms makes lower profits if consumers have limited attention stems from consumers not fully appreciating the qualities of the product. Therefore, consumers with limited attention have a lower willingness-to-pay than fully attentive consumers have.

In the following we try to exemplify our results with short numerical examples. These examples seek to illustrate the mechanisms that drive our results. Let us first consider an example to illustrate the obstacles of introducing a second quality.

**Example III.1.** *Suppose there are two possible qualities with value parameters* $v_1 = 10$ *and* $v_2 = 9$. *We normalize the value of money to* $v_p = 1$. *Initially, assume there are no cost differences between the two characteristics,* $c_1 = c_2 = 3$. *Finally, assume the cognition costs are* $\kappa_{r(i)} = 2^{r(i)-1} - 1$, *yielding* $\kappa_1 = 0$, $\kappa_2 = 1$, *and* $\kappa_3 = 3$.

*In this setting it is optimal for the monopolist to equip his product with qualities at the level* $x_1^* = 3.33$ *and* $x_2^* = 3$, *such that* $m_1 = 0.97$ *and* $m_2 = 0.89$. *This yields an optimal price of* $p = 56.33$ *and a profit of* $\Pi = 26.17$. *If the monopolist would offer a product with only one attribute, profits would be* $\Pi = 16.67$ *when using quality 1 and* $\Pi = 13.5$ *when using quality 2. Hence, the monopolist will offer a product with both qualities.*

*Now suppose that the second quality is more costly to produce with* $c_2 = 15$, *all else equal. If the monopolist produces both qualities, the profit is* $\Pi = 15.37$ *which is less* *than the profit from equipping the product with only quality 1. Note that this is not the result of the customer not attending to quality 2. If a good with both attributes is produced, the optimal level of quality 2 would be* $x_2^* = 0.6$ *which yields positive attention factor* $m_2 = 0.44 > 0$*. However, the profitability of the second quality* $\pi_2 = 2.7$ *is not sufficient to cover the decrease in willingness-to-pay that is induced by the cognitive costs of attending the second quality* $\kappa_3 = 3$ *and the increase in cognitive costs associated with quality 1. Thus for a low profitability of attribute 2 it is optimal to equip the product with only attribute 1.*

Similar considerations apply if the product can be equipped with more than two qualities. As we have argued before, the price associated with a multi-quality product must be ranking first in the attention hierarchy. Yet, this also means that the price increase associated with adding a third quality to a two-quality product does not result in the price receiving more attention. As soon as the price receives full attention, i.e. as soon as the product features at least two qualities, price increases are no longer associated with an increased salience of the price dimension. Hence, when considering adding a third quality to the product, the necessary price increase will no longer be associated with a rise in attention allocated to the price. Thus, additional qualities will be introduced (naturally in the order of their profitability) if their profitability is sufficient to cover the cognitive costs of their consideration.

Denote by $\pi^{(t)}$ the $t$-th most profitable attribute (ties may be broken according to $v_i$). Then we find the following result.

**Proposition III.1.**
*If a monopolist intends to supply a single good to the market, the optimal design will feature:*
*(i) only the most profitable quality* ($i : \pi_i = \pi^{(1)}$) *if and only if* $\nexists\, m \in \{2, ..., n\} : \sum_{t=2}^{m} \left(\pi^{(t)} - \frac{1}{v_p}\kappa_{t+1}\right) \geq \frac{1}{v_p}\kappa_2$.
*(ii) Otherwise, it will feature the m most profitable qualities for which* $\sum_{t=1}^{m} \left(\pi^{(t)} - \frac{1}{v_p}\kappa_{t+1}\right) \geq 0$*, while* $\pi^{(m+1)} - \frac{1}{v_p}\kappa_{m+2} < 0$.
*All qualities* $i \in \mathcal{I}$ *that the product features are produced at a level* $x_i = v_i/(v_p c_i)$.

*If the optimal good features only one quality, the price extracts the whole valuation of the consumer, with* $P = (v_i x_i)/v_p$*. In this case, the price ranks second in salience and is thus not receiving full attention* ($m_p < 1$)*. The decision utility of purchasing the product is strictly positive, i.e.* $\tilde{u}(a) > 0$*. If the optimal good features several qualities, the price ranks first in salience* ($m_p = 1$) *and extracts the total decision value of the qualities, i.e.* $\tilde{u}(a) = 0$*. Yet it does not extract the entire value of the product in terms* *of experienced utility:*

$$P = \frac{1}{v_p} \sum_{i \in \mathcal{I}} m_i v_i x_i < \frac{1}{v_p} \sum_{i \in \mathcal{I}} v_i x_i,$$

*with $\mathcal{I}$ being the set of qualities that the product features.*

Intuitively, one can understand Proposition III.1 as showing how the limited attention of the customers translates into an augmented cost function of the monopolist. The limited attention reduces the willingness-to-pay for complex products, i.e. for products featuring more than one quality. This is equivalent to a monopolist facing additional costs when offering complex products. These costs include variable costs $\kappa_{t+1}/v_p$ associated with introducing the $t$-th quality ($t > 1$), as well as fixed costs $\kappa_2/v_p$ associated with offering a complex product in the first place.

A couple of things are noteworthy. Given the product features a particular quality, its optimal level is the same as the one under full attention. This is due to two effects resulting from limited attention. First, limited attention reduces a firm's incentives to invest in quality as the created value is not fully taken into account by the customer. Second, the endogeneity of attention increases a firm's incentives to invest in quality as any additional unit of quality increases the attention paid to that attribute and thus the decision value of any unit of quality already invested.[13] In our model, these two effects cancel out each other perfectly as the maximization problem of the monopolist yields the same level of product quality as in the case with unlimited attention for those qualities which are considered. However, the product generally features less qualities than under full attention and the price is lower if the product features more than a single quality. Note here that the exact canceling of the above-mentioned effects hinges on our modeling of limited attention, in particular our use of the range ($\max_{a \in A} x_i^a - \min_{a \in A} x_i^a$) as a measure of attribute dispersion. Yet, we claim that the second effect prevails as long as one retains the assumption of a higher attribute dispersion attracting attention. As this second effect compensates the first, limited attention does not necessarily imply the production of lower quality levels.

To motivate our next section recall the profit (III.4) that the firm makes when it offers a product that features more than one quality. It holds that, as soon as the product features more than one quality, the monopolist cannot extract the whole experienced utility that the product yields to the customer. This is because consumers decide on the basis of decision utility. In the case of multiple qualities the price receives full attention while all qualities receive less than full attention. Yet, the attention that is allocated

[13]Note that the firm's incentive to invest in quality in order to increase attention is conditional on the attribute being considered. If the attribute is not considered, small changes in quality might be insufficient in order to lift attention to a positive level.

to the qualities depends on the available choice set. That is what motivates the next section. We will analyze a way in which the monopolist can increase the customer's attention and thus the decision value of the product that is sold to consumers without changing the product itself.

## 4. Optimal Product Design with Introduction of a Bait Good

Recall that the consumer attention is a function of both the valuation $v_i$ and the dispersion within the choice set ($\max x_i - \min x_i$). A firm might not be able to change the valuation $v_i$ of an attribute, but it can manipulate the dispersion. If the firm can produce several goods, it might have an incentive to produce goods that have the sole purpose of increasing the dispersion of attributes, thereby making these attributes more salient. Note that the quality levels are bounded below by the option of not buying ($\min x_i = 0$). This implies that in order to increase dispersion, products that are designed to manipulate attention must feature high quality levels. Such products will increase dispersion by increasing $\max x_i$. Yet, if these products are not intended for sale, they must be unattractive to consumers. This could, for example, be achieved by a very high price. In the following, we will use the term *bait good* for those goods that have the sole purpose of manipulating consumer attention. Because these bait goods are designed to be unattractive, consumers still buy the main good which we will henceforth call *primary good*.[14] We will now analyze how a bait good is optimally designed.

In the analysis we will assume that the firm does not incur any costs for designing and producing the bait good. This is obviously not realistic in most circumstances. We maintain this assumption to concentrate on the question whether it is possible to increase the willingness-to-pay of consumers for the primary good, and thus the profit made from its sale.

In addition, there exist further arguments why the cost of the bait good might be negligible. First, if the bait good is never actually sold to the customer, it only needs a single item of the bait good that can be (unsuccessfully) offered to each of a large number of customers. If, as we will show, the bait good increases profits, the additional profits reaped from each customer may in sum be sufficient to cover the cost of bait good production. A second avenue to accommodate the cost of producing the bait good is to allow for customer heterogeneity in their intrinsic valuation of money. As we

[14] Obviously, the primary good still has to yield positive decision utility to the consumer in order to be bought.

will argue in the discussion section, a product designed for a richer customer segment is a prime candidate to serve as a bait good for a poorer customer segment. In this way, the bait good is produced both for sale (to a rich customer segment) as well as for the manipulation of attention (of a poorer customer segment).

As outlined before, bait goods will be designed to have premium quality. In reality, maximum quality levels are of course bounded at some level. For example, there are technological and physical limitations on the horsepower that a car can have. Therefore, we will now assume technical frontiers for each quality. Formally, all $x_i \in [0, \bar{x}_i]$ with $\bar{x}_i \in \mathbb{R}^+$.

Now let us look at the introduction of a bait good that is designed to attract attention to the qualities of the primary good, but without changing the attention hierarchy of the qualities embodied in the primary good. Note first that if the monopolist is confined to producing a good with only a single quality, there is nothing to be gained from manipulating attention. This is because the primary good is designed such that the single quality employed ranks first in the attention hierarchy and thus receives full attention. This is different if the primary good features several qualities where, as we have shown before, the price gets full attention while the qualities do not get full attention. A bait good will then derive its value through increasing the attention paid to the qualities, which were not fully considered or neglected altogether. In this case, the following result applies:

**Proposition III.2.** *Suppose a monopolist's profit-maximizing product design features at least two qualities: $|\mathcal{I}| \geq 2$. If at least one of these qualities is not produced at the highest feasible level, the monopolist can strictly increase profits by using a bait good.*

*Proof.* Consider a monopolist whose product features $|\mathcal{I}| \geq 2$ qualities. Further suppose that the optimal level of one of these qualities is below its technological frontier, i.e. $\exists j \in \mathcal{I} : \; x_j^* = \frac{v_j}{v_p c_j} < \bar{x}_j$. Let the monopolist offer a second product, called bait good, which has the same level of qualities as the initial product, i.e. $x_i^b = x_i, \; \forall i$. The price of the bait good is set sufficiently high such that it is unattractive to consumers. This introduction of the bait good has no impact on the attention levels with $m_p = 1$ and $m_i = 1 - \kappa_{r(i)}/(v_i x_i^b), \; \forall i$. However, the profit from the sale of the primary good now has the following form:

$$\Pi^* = P^* - \sum_i c(x_i) = \sum_{i \in \mathcal{I}} \left[ \frac{1}{v_p} m_i v_i x_i - \frac{1}{2} c_i x_i^2 \right],$$

with

$$m_i = \begin{cases} \max\left\{0, 1 - \kappa_{r(i)}/(v_i x_i^b)\right\} & \text{for } x_i \leq x_i^b, \\ \max\left\{0, 1 - \kappa_{r(i)}/(v_i x_i)\right\} & \text{for } x_i > x_i^b. \end{cases}$$

Then it holds that the new optimal level for quality $j$ of the primary good is:

$$x_j^* = \frac{1}{v_p c_j} \left( v_j - \kappa_{r(j)} \frac{1}{x_j^b} \right) < \frac{v_j}{v_p c_j}.$$

Since the original level of the quality was still feasible ($x_j = \frac{v_j}{v_p c_j}$) and the attention allocation under the new design is the same as under the old design, it must hold by revealed preference that the firm makes higher profits with the introduction of the bait good. □

The bait good thus increases the profit that can be made from the sale of the primary good while keeping the attention hierarchy among attributes intact. Note here that we have just shown that a bait good can increase profits under fairly simple conditions. We have not yet looked at the profit maximizing design of the bait good.One obvious venue to further increase profits is by increasing the qualities of the bait good which in turn increases attention for these qualities. This procedure is outlined in the following example.

**Example III.2.** *First, let us consider the case where* $v_1 = 10$*,* $v_2 = 9$*,* $v_p = 1$*,* $\kappa_1 = 0$*,* $\kappa_2 = 1$*,* $\kappa_3 = 3$*, and* $c_1 = c_2 = 3$ *as in Example III.1. In addition, let us assume technological frontiers of* $\bar{x}_1 = \bar{x}_2 = 6$*. If the monopolist designs a bait good with qualities* $x_1^b = \bar{x}_1$*,* $x_2^b = \bar{x}_2$ *and a prohibitively high price, the customers attention for both qualities of the primary good will increase. The high price ensures that the consumer is not being tempted to buy the bait good itself. Given this bait good, the monopolist's optimal design of the primary good features both qualities at levels* $x_1^* = 3.28$ *and* $x_2^* = 2.83$*, such that* $m_1 = 0.98$ *and* $m_2 = 0.94$*. This yields a profit of* $\Pi = 28.16$ *that exceeds the maximal profit of* $\Pi = 26.17$ *that the firm made without the bait good.*

Also note that Proposition III.2 is only a sufficient condition for the profitability of the bait good. For example, suppose that without usage of a bait good, it was optimal to produce a primary product with only one quality. Then the introduction of the bait good may allow the firm to profitably introduce further qualities.

**Example III.3.** *Let us reconsider the second case in Example III.1 in which the higher costs of the second quality* $c_2 = 15$ *made its introduction unprofitable. Using the same bait good as above in Example III.2, the monopolist can make a profit of* $\Pi = 18.52$ *by equipping the primary good with both qualities. This exceeds the profit that the firm can make by offering a primary good with only a single quality. The use of a bait good can thus make the introduction of qualities profitable which were unprofitable without the bait good.*

In addition to merely increasing the attention paid to each attribute of the primary good, a bait good can be employed to shift the attention order, e.g. to push a more profitable characteristic to a more salient position. We will now investigate under what conditions this is profitable. Suppose for simplicity that the primary good features two qualities $i = 1, 2$ and assume that $m_1 > m_2$. Further assume that the bait good in this setup features qualities $x_1^b = \bar{x}_1$ and $x_2^b = v_1\bar{x}_1/v_2 < \bar{x}_2$. Thus the bait good is designed in such a way that the primary good yields maximal profit given that the first quality ranks higher in the attention order than the second quality. Increasing $x_2^b$ above $v_1\bar{x}_1/v_2$ would change the attention order: the first quality would become less, the second quality more salient. Recall that, with the introduction of the bait good, the profit of the firm takes the form:

$$\Pi^* = \sum_{i \in \mathcal{I}} \pi_i m_i^2.$$

Then it is easy to see that the change in the attention order is profitable if the additional profit made from quality 2 exceeds the reduction in profits from quality 1:

$$\pi_2 \left( m_{2,new}^2 - m_{2,old}^2 \right) > \pi_1 \left( m_{1,old}^2 - m_{1,new}^2 \right).$$

This equation is likely to hold if (a) quality 2 is more profitable than quality 1, and/or (b) the technology frontier $\bar{x}_2$ of quality 2 greatly exceeds the restriction $v_1\bar{x}_1/v_2$ that the bait good would have to satisfy in order to retain the old attention order.

Let us consider an example that highlights that it may be optimal to produce a bait good that changes the ordering of the attention hierarchy. The example also illustrates that it may be optimal to produce a bait good with some qualities below their technological boundaries.

**Example III.4** (Profitability)**.** *Suppose customer valuations are* $v_1 = 10$, $v_2 = 9$, *and* $v_p = 1$ *as before, but costs are* $c_1 = 6$ *and* $c_2 = 3$. *The profitability values are* $\pi_1 = 8.33$ *and* $\pi_2 = 13.5$ *and hence quality 1 is less profitable than quality 2. Suppose the bait good was constructed as in Examples III.2 and III.3 by setting* $x_1^b = \bar{x}_1 = 6$ *and* $x_2^b = \bar{x}_2 = 6$. *Then the less profitable quality would rank higher (at rank 2, behind the price) than the more profitable quality (rank 3). This yields a profit of* $\Pi = 20.10$. *From the profit formula for the case with bait goods* $\Pi = \sum_{i \in \mathcal{I}} \pi_i m_i^2$ *it follows that a monopolist may profit from manipulating the attention hierarchy such that the most profitable quality gets the most attention. Suppose* $x_1^b$ *was lowered to* $5$. *This would make quality 2 rank second in the attention order, and quality 1 rank third, yielding a profit of* $\pi = 20.36$.[15]

[15] Actually, it would be sufficient (and thus optimal) to lower $x_1^b$ slightly below $(v_2\bar{x}_2)/v_1 = 5.4$. If, e.g., $x_1^b = 5.3$ was chosen, profit would be $\Pi = 20.42$.

*This example illustrates that the bait good cannot only be used to attract attention, but also to distract attention. This is attractive if a firm wants to distract customers' attention from a more valuable, but less profitable quality to a less valuable, but more profitable quality.*

Example III.4 highlights that firms usually have an incentive to make the most profitable qualities most salient. However, this is not always the case. We will now give an example in which the firm prefers to rank a high-technology attribute higher than a high-profitability attribute. Consider the following example.

**Example III.5** (Technology differences)**.** *Assume* $v_1 = 10$*,* $v_2 = 9$*,* $v_p = 1$*, and* $c_1 = c_2 = 3$ *as in Example III.1. However, assume that* $\bar{x}_1 = 6$ *and* $\bar{x}_2 = 60$*. As in Example III.1, quality 1 is more valuable to the customer and more profitable to the firm than quality 2. It is optimal to set* $x_1^b = \bar{x}_1 = 6$ *irrespective of the attention rank that quality 1 is assigned in the end. Suppose it is assigned rank 2 (the price has rank 1) while quality 2 is assigned rank 3. Then* $x_2^b$ *is restricted to the value* $v_1\bar{x}_1/v_2 = 10 \cdot 6/9 = 6.67$*, because otherwise quality 2 ranks higher than quality 1. The resulting maximal profit for the firm is* $\Pi = 28.30$*. If, however, the firm sets* $x_1^b = \bar{x}_1 = 6$ *and* $x_2^b = \bar{x}_2 = 60$*, the less profitable quality 2 will rank second in the attention hierarchy, while the more profitable quality 1 ranks third. This yields a profit of* $\Pi = 28.49$ *which is an improvement. In this example, for all values of* $\bar{x}_2$ *above approximately* 12.5 *it is optimal for the firm to make the higher-valued and more profitable quality less salient. It therefore illustrates that a firm may sometimes have an incentive to use a bait good to distract from a more profitable quality. This should, however, only occur in situations in which there are small differences in profitability but huge differences in the technological limitations associated with the two qualities.*

As shown in the various examples, the bait good is effective in attracting attention when it features qualities at a high level. However, when contemplating the optimal attention manipulation, the firm faces the following trade-off: On the one hand, the firm has an incentive to let profitable qualities rank high in the attention order (see Example III.4). This allows to attract a lot of attention to these characteristics, which enables the firm to extract more of the surplus created in this dimension. Yet, assigning a particular rank to a quality imposes a restriction. In order for a quality to keep its assigned rank, the extent of attention that can be attracted to all qualities at lower ranks is limited. This may conflict with the incentive to attract as much attention as possible to all, specifically lower-ranking, qualities by exploiting technological frontiers.

Until now we have assumed that the firm can produce only one bait good. In the next section we will investigate under which circumstances it might be optimal to employ several bait goods.

## 5. Usage of Several Bait Goods

As derived above, the purpose of the bait good is to increase attention for the qualities that are featured by the primary good. This increase in attention allows the firm to charge higher prices since it leads to an increase in the consumer's willingness-to-pay. In particular, the bait good increases attention of some attributes if it is equipped with extremely high levels of these quality. However, since the firm does not actually want to sell the bait good, it has to charge a very high price for it in order to make it less attractive to consumers than the primary good. Recall that this implied that the price kept the highest rank in the attention hierarchy. Since the price already ranked first under the optimal design of a single product (if $|\mathcal{I}| \geq 2$), thus being fully considered with $m_p = 1$, the high bait good price had no influence on the salience of the price dimension.

When the firm is allowed to produce more than one bait good, it can use a different approach towards attention manipulation. Suppose the firm initially produces a primary good that features several qualities. One potential bait good that the firm could introduce is one that features a high level of the first quality and no additional qualities. This would increase the attention that consumers pay to the first quality. Since the new bait good only features one quality (and the primary good features many), the bait good's price does not need to be very high in order to make its purchase unattractive. Then the firm could produce several bait goods that each feature one premium quality. As we will show, the usage of several specialized bait goods can make it feasible to make some qualities even more salient than the price. This might be profitable because if the price is less salient, the willingness-to-pay of consumers increases which allows the firm to raise the price of the primary good. Hence, as we will show below, the usage of several bait goods can be even more profitable than the usage of a single high-priced bait good.

Now we want to determine how a specialized bait good can be used to make a quality more salient than the price. In the following, we assume w.l.o.g. $v_i > v_{i+1}, \ \forall i = 1, ..., n-1$. Consider the case in which the firm wants to make the most valuable quality 1 more salient than the price.[16] When employing a specialized bait good, the firm has to meet several restrictions for the design of the bait good. First, it must still hold that the bait good is less attractive than the primary good:

$$\tilde{u}^{primary} \geq \tilde{u}_1^{bait}. \tag{III.5}$$

[16] It is not per se clear which quality should be pushed to the first attention rank. Hence, the following derivations concerning quality 1 simply serve as an example on how a specialized bait good can be used.

In addition, the bait good should not increase the salience of the price, i.e. $p_1^{bait} \leq p^{primary}$. Since the bait good should be unattractive, we can focus on the case $p \equiv p_1^{bait} = p^{primary}$. For a bait good with quality 1 we can then write (III.5) as:

$$\begin{aligned} \sum_{i=1}^{n} m_i v_i x_i^{primary} - m_p v_p p &\geq m_1 v_1 x_1^{bait} - m_p v_p p \\ \Leftrightarrow \sum_{i=1}^{n} m_i v_i x_i^{primary} &\geq m_1 v_1 x_1^{bait}. \end{aligned} \tag{III.6}$$

Since the objective is to push the price to a lower attention rank, the bait good must be designed such that quality 1 ranks higher than the price. So if a bait good features only quality 1, the bait good must suffice

$$v_1 x_1^{bait} \geq v_p p. \tag{III.7}$$

Combining (III.6) and (III.7), it must hold that:

$$\sum_{i=1}^{n} m_i v_i x_i^{primary} \geq m_1 v_1 x_1^{bait} \geq m_1 v_p p. \tag{III.8}$$

Note here that the price was initially at the highest attention rank. Hence, if quality 1 gets pushed in front of the price, it is now ranking first with $m_1 = 1$. Then, (III.8) can be written as:

$$\sum_{i=1}^{n} m_i v_i x_i^{primary} \geq v_1 x_1^{bait} \geq v_p p. \tag{III.9}$$

By setting $x_1^{bait}$ and $p$ such that all three terms in (III.9) have equal value, the firm can charge a maximum price of:

$$p = \frac{1}{v_p} \sum_{i=1}^{n} m_i v_i x_i^{primary}.$$

This expression looks like the pricing formula in the case with only a single bait good. Note, however, that $m_1$ has increased relative to the case with only one bait good (because quality 1 has gained one rank in the attention hierarchy). Hence, by employing a specialized bait good the firm raised the price it can charge for the primary good.

Until now we have described how attribute 1 can be pushed in front of the price in the ranking of salience. We can proceed by doing the same with qualities 2, 3, and so on. If quality 2 shall gain attention rank 2, the bait good for quality 2 must be designed such that it satisfies:

$$v_1 x_1^{bait} \geq v_2 x_2^{bait} \geq v_p p. \tag{III.10}$$

If this condition can be fulfilled, the second bait good increases $m_2$ because quality 2 would gain one rank in the attention hierarchy.

In principle, this process can be continued for all other qualities. For all qualities that are more salient than the price it must hold that $v_i x_i^{bait} = v_{i+1} x_{i+1}^{bait}$. Note here that since $v_i > v_{i+1}$, each additional bait good must be produced to have a higher level of the quality on which it is specialized, i.e. $x_i^{bait} < x_{i+1}^{bait}$. This may not be feasible for all qualities due to the technological bounds $\bar{x}_i$. Hence, it may not be possible to make all qualities more salient than the price. However, even for qualities that rank below the price attention can be increased by using bait goods. Just as before, qualities that rank below the price can still be made more salient relative to the no-bait-good-case by producing a bait good that has a quality level for the given attributes that is above the quality level of the primary good. This effect is basically the same as in the case with only one bait good.

**Example III.6** (Several bait goods)**.** *Suppose that we have $v_1 = 10$, $v_2 = 9$, $v_p = 1$, $c_1 = 3$, and $c_2 = 3$. In addition, we assume that $\bar{x}_1 = \bar{x}_2 = 6$. We have seen before that, without the usage of a bait good, the firm would make a profit of $\Pi = 26.17$. By using a single bait good that featured both qualities, the firm was able to increase its profit to $\Pi = 28.16$. The firm achieved this increase in profit by producing a bait good with $x_1^{bait} = \bar{x}_1 = 6$ and $x_2^{bait} = \bar{x}_2 = 6$. This leads to an increase in attention factors $m_1$ and $m_2$, relative to the case without a bait good, and thereby increases the consumers' willingness-to-pay. However, since the bait good was not intended for sale, the firm had to set a high bait good price which implied that the price remained in attention rank 1.*

*If the firm can produce two bait goods, it turns out that it is easier to make them unattractive. The reason is that the primary good features two qualities and each bait good features only one quality. Let us consider the optimal bait good design. The firm wants to achieve that (some) qualities are more salient than the price dimension. This can be achieved by producing a bait good that features a high level of its particular quality. However, since the price is dependent on the resulting attention structure, we have to solve this problem recursively. This is necessary because the bait good design is dependent on the price, but the price is also dependent on the bait good design.*

*Solving this problem, we find that the optimal qualities of the primary good are $x_1^{primary} = 3.33$ and $x_2^{primary} = 2.83$ while the price is $p = 57.42$. The bait good that features quality 1 is produced with $x_1^{bait} = 5.74$ and $p^{bait} = p = 57.42$. This ensures that $v_1 x_1^{bait} \geq v_p p$ and hence quality 1 is more salient than the price. In principle, the firm would also like to make quality 2 more salient than the price. However, this would require $v_2 x_2^{bait} \geq v_p p$, yielding $x_2^{bait} \geq 6.38$, which is not feasible due to $\bar{x}_2 = 6$. Thus the firm is confined to offer a second specialized bait good at the highest technological* *level with* $x_2^{bait} = 6$. *Hence, quality 1 will rank first in the attention hierarchy, followed by the price and quality 2. This yields attention parameters* $m_1 = 1$, $m_p = 0.98$, *and* $m_2 = 0.94$. *The resulting profit of the firm is* $\Pi = 28.71$, *which is higher than in the case with only a single bait good.*

*This example illustrates several effects that occur when using several bait goods. Note that the quality of the first bait good is restricted to* $x_1^{bait} = 5.47$ *in order to not be more attractive than the primary good. A similar restriction would be in place for all qualities that are intended to be made more salient than the price. For example, quality 2 would have to be no higher than* $x_2^{bait} = 6.66$.[17]

*Note that for all qualities that do not rank first in the attention hierarchy these restrictions limit the attention that can be drawn to these qualities. Suppose that the technological limits are very high (say* $\bar{x}_i = 50$*). Then the firm could design two bait goods such that qualities 1 and 2 are more salient than the price. This however would mean that the attention drawn to quality 2 is limited by equation (III.10). It will turn out that even if it is feasible to make a quality more salient than the price, it might not always be optimal to do so. In the case with* $\bar{x}_i = 50$ *it is more profitable to employ a single bait good and exploit the technological boundaries. Although this implies that qualities 1 and 2 receive a lower rank (2 instead of 1 and 3 instead of 2), the increase in attention for quality 2 will offset the decrease in attention for quality 1.*

*To see this consider the outlined example with* $\bar{x}_i = 50$. *In this case, the bait good levels that would be chosen in order to make attributes 1 and 2 more salient than the price would be* $x_1^{bait} = 6$ *and* $x_2^{bait} = 6.66$. *This would yield attention parameters* $m_1 = 1$, $m_2 = 0.98$, $m_p = 0.95$, *and a profit of* $\Pi = 29.72$. *If, however, the firm would use a single bait good with maximum quality levels* $x_1^{bait} = x_2^{bait} = 50$, *the attention parameters would be* $m_p = 1$, $m_1 = 0.998$, *and* $m_2 = 0.99$. *In this case, the profit of the firm would be* $\Pi = 29.92$, *which is an improvement.*

We conclude that the usage of several specialized bait goods is only profitable if the technological limits on quality levels are neither too low (making qualities more salient than the price allows only very small prices) nor too high (usage of a single bait good that exploits the technological limits is more profitable than trying to distract attention from the price through specialized bait goods).

[17]The beforementioned value $x_2^{bait} = 6.38$ is sufficient to increase the attention rank of quality 2. However, due to the increased attention, it is then optimal to increase the quality 2 of the primary good a bit. This increases the attractiveness of the primary good such that $x_2^{bait}$ can be further increased. Solving this recursively, the maximum quality that the second bait good can have in order to still be less attractive than the primary good is $x_2^{bait} = 6.66$.

# 6. Discussion

This paper seeks to highlight how a firm may employ certain products to manipulate the attention allocation of its customers. There are a couple of issues and potential extensions of the model that are worth further discussion.

## 6.1. Heterogeneous Consumers

Until now we have assumed that consumers are homogeneous in both their preferences and their cognitive abilities. This assumption could be weakened in several ways. One interesting possibility is to allow customers to have different levels of wealth. For simplicity assume there are two groups of customers, very loosely labeled some are *poor* and some are *rich*. Let the first group assign higher importance to the money dimension (and thus the price) of a product: $v_p^{poor} > v_p^{rich}$. This will result in a firm designing different products to cater to both groups of customers. Recalling Proposition III.1, the rich group's ideal product may feature more qualities than the poor group's product. More interestingly, it features all the qualities the poor group's product has on a higher level. This makes the rich group's product an ideal candidate to be a bait good for the poor group. It increases the attention paid to all the qualities that the product for the poor features and thereby increases the willingness-to-pay of the poor. In this way, a firm may employ products designed for richer customer segments as a bait good for poorer customer segments. This might explain one of the examples we discussed in the introduction: the advertisement of expensive cars to an audience of which a majority is not able to afford it.

If customers differ in the cognitive constraints they face, i.e. if they differ in their cognition costs $\kappa_i$, a firm may cater to these different groups with products differing in their degree of complexity. One reason for different cognitive constraints could be that one group has to act under stronger time pressure. If this is the case, the firm could, for example, offer products that differ in the number of qualities they feature. The hurried customer segment is then offered a product with only a few essential features, while customers with more time prefer more elaborate products.

Finally, customers may differ in their valuation for different qualities. As these valuations influence both a quality's profitability and the relative ease of attention attraction, a firm may have strong incentives to segment the customer population and design an appropriate product line for each segment separately. Still, analogous to the discussion of differing wealth levels, a firm may have an incentive to offer a product designed for one customer segment to a different segment despite that segment's unwillingness to purchase the product. Offering a sports car to a family father may increase his willingness-to-pay for horse power despite his general focus on car safety.

### 6.2. Negative "Qualities"

So far we have assumed that all qualities (save the price) are valued by the customer ($v_i \geq 0, \forall i$). There are certainly some characteristics that a product may feature which customers dislike, e.g. the level of exhaust fumes of cars, or the level of sugar and trans fats in food. While the customer dislikes these characteristics and their presence reduces consumers' willingness-to-pay, a reduction and/or replacement of these attributes may be costly to the firm. Employing bait goods can thus have further beneficial effects since, in addition to attracting attention to positive characteristics, the firm could distract from negative characteristics. Also, a firm might have incentives to focus on reducing few negative characteristics a lot instead of reducing all negative characteristics a little bit. A label "sugar free" or "low carb" may effectively distract from other negative characteristics, for example high levels of trans fats.

### 6.3. Attribute Dependence

We assume attributes to be independent. One might consider relaxing this condition to allow for complementarity and/or substitutability between attributes in production and/or consumption. In the case of dependencies in consumption, the importance of an attribute $v_i$ would be a function of the level of complementary/substitute attributes. An attribute's level may thus decrease the attention paid to a substitute attribute even if both attributes are several ranks apart in the attention hierarchy.

There is a complicating issue connected with allowing for dependencies in consumption between attributes. The question is whether the DM is aware of all dependencies between attributes or only of the dependencies between the attributes that receive positive attention. What if there is a dependency between an attribute that is considered and an attribute that is neglected? In addition, one might wonder why a DM seeks to decrease the complexity of a decision problem by neglecting some attributes but has no problems taking into account complex interactions between attributes.

## 7. Conclusion

We have introduced a new approach to model limited attention and applied it to the problem of optimal product design. The proposed attention heuristic fulfills several desirable properties that we think are realistic in real world markets.

Using this framework, we have shown that limited attention has far-reaching implications for product design and in general also for product lines. We have started with the case in which the monopolist can offer only a single product. In general, firms tend to produce products that have fewer attributes than they would have if consumers were fully attentive. However, attributes that have positive quality may actually be produced at the same level as in a fully rational model.

We find that a monopolist would actually prefer consumers to be fully attentive. This is because the consumers' willingness-to-pay is lower under limited attention since they do not fully appreciate all the qualities inherent in a product. Since the monopolist profits from an increase in attention, there is an incentive to introduce goods of premium quality that are not intended for sale, but which increase the attention of the consumers. These bait goods can be used to increase attention, but they can also distract consumer attention from attributes that are less profitable to the monopolist. Furthermore, they can be used in such a way that the price appears less salient to consumers, which further increases both the consumers' willingness-to-pay and the monopolist's profits.

# A. Appendices

## 1. Appendix to Chapter I

### 1.1. Proof of Proposition I.1:

The proof is completed by the following observations:

- Claim: For most parameter constellations there does not exist an equilibrium in which firms mix over add-on prices. If such a mixing equilibrium exists, it has no implications on profits (and also no implications on the incentive to unshroud):

  It was proven in Lemma I.1 that firms will set add-on prices such that $\hat{p} \in \{e, \bar{p}\}$. Now suppose that with probability $\gamma$ the efficient firm 1 sets an add-on price $e$ and a corresponding base good price $p^e$. With the remaining probability $1-\gamma$, the efficient firm 1 sets an add-on price $\bar{p}$ and a corresponding base good price $p^{\bar{p}}$. Now suppose the inefficient firm 2 sets an add-on price $e$. Then, firm 2 faces the following profit function:

  $$\pi_2^e(p_2) = \gamma \left[ \frac{p_1^e - p_2 + t}{2t}(p_2 + e - \hat{c}) \right] + (1-\gamma) \left[ \frac{p_1^{\bar{p}} - p_2 + t}{2t}(p_2 + e - \hat{c}) \right].$$

  If, in contrast, firm 2 sets an add-on price $\bar{p}$, the following profit function applies:

  $$\pi_2^{\bar{p}}(p_2) = \gamma \left[ \frac{p_1^e - p_2 + t}{2t}(p_2 + \alpha'(\bar{p} - \hat{c})) \right] + (1-\gamma) \left[ \frac{p_1^{\bar{p}} - p_2 + t}{2t}(p_2 + \alpha'(\bar{p} - \hat{c})) \right].$$

  Comparing these profits yields

  $$e - \hat{c} > \alpha'(\bar{p} - \hat{c}) \quad \Leftrightarrow \quad \max_{p_2} \ \pi_2^e(p_2) > \max_{p_2} \ \pi_2^{\bar{p}}(p_2).$$

  Hence, for $e - \hat{c} > \alpha'(\bar{p} - \hat{c})$ firm 2 will set an add-on price $\hat{p}_2 = e$, while for $e - \hat{c} < \alpha'(\bar{p} - \hat{c})$ firm 2 will set $\hat{p}_2 = \bar{p}$. Note that with exception of the knife-edge case $e - \hat{c} = \alpha'(\bar{p} - \hat{c})$ firm 2 is not indifferent between add-on prices $e$ or $\bar{p}$. Hence, for $e - \hat{c} \neq \alpha'(\bar{p} - \hat{c})$ firm 2 will not mix over add-on prices. In addition, for $e - \hat{c} = \alpha'(\bar{p} - \hat{c})$ firm 2 will choose the same base good price regardless of the chosen add-on prices ($p_2^e = p_2^{\bar{p}}$). Hence, the potential mixing over add-on prices in the knife-edge case has no effect for firm 1 since its profits are only influenced by the base good price of its competitor.

The same logic as above is applicable for firm 1. Firm 1 will set an add-on price $\hat{p}_1 = e$ if $e > \alpha'\bar{p}$ and an add-on price $\hat{p}_1 = \bar{p}$ if $e < \alpha'\bar{p}$. For $e = \alpha'\bar{p}$, firm 1 is indifferent and may randomize between the add-on prices with any positive probability. However, for $e = \alpha'\bar{p}$ it holds that $p_1^e = p_1^{\bar{p}}$, and hence mixing has no effect on the profit of firm 2.

- Claim: Suppose $\alpha' = \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$. Then, it holds that $\frac{e-\hat{c}}{\bar{p}-\hat{c}} < \alpha' < \frac{e}{\bar{p}}$:
  It obviously holds that $\frac{e-\hat{c}}{\bar{p}-\hat{c}} < \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$. To prove that in this case it also holds that $\frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}} < \frac{e}{\bar{p}}$, consider the following transformation of equations:

$$\begin{aligned} \alpha' &= \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}} \\ \Leftrightarrow\ e-\alpha'\bar{p} &= \underbrace{\hat{c}(\alpha-\alpha')}_{>0} \\ \Rightarrow\ \alpha' &< \frac{e}{\bar{p}}. \end{aligned}$$

  Hence, it follows that if $\alpha' = \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$, $\alpha'$ must lie in the considered interval $[\frac{e-\hat{c}}{\bar{p}-\hat{c}}, \frac{e}{\bar{p}}]$.

### 1.2. Proof of Proposition I.2:

The proof is a straightforward generalization of the main proposition of GL. As before, sophisticated consumers anticipate high add-on prices if firms shroud the add-on. For reasons similar to Lemma I.1, firms will only set add-on prices of either $e$ or $\hat{p}$. This implies that consumers are indifferent between add-on prices and shop only on the basis of the base good price.

Now suppose that all firms shroud the add-on. Then, sophisticates will substitute and firms will set high add-on prices $\bar{p}$. The profit functions of firms then take the form:

$$\pi_i = D_i(p_i, p_j) \cdot (p_i + \alpha(\bar{p} - \hat{c}_i)).$$

The motive for a deviation would be to increase the fraction of consumers that buy the add-on. Sophisticates will only buy the add-on if the firm unshrouds (making add-on prices observable) and sets a low add-on price $e$. Then, the profit function is

$$\pi_i = D_i(p_i, p_j) \cdot (p_i + e - \hat{c}_i).$$

When comparing these profits, the only difference lies in the add-on markup. Since shrouding and pricing take place simultaneously, the non-deviating firm cannot adjust its base good price $p_j$. Furthermore, the base good demand $D_i(\cdot)$ is independent of add-on prices and also independent of the composition of the consumer population.

Hence, for $\alpha = \frac{e-\hat{c}_i}{\bar{p}-\hat{c}_i}$ both profits are equal and firm $i$ is indifferent between shrouding and unshrouding. For $\alpha < \frac{e-\hat{c}_i}{\bar{p}-\hat{c}_i}$ the firm has an incentive to deviate from shrouding, while for $\alpha > \frac{e-\hat{c}_i}{\bar{p}-\hat{c}_i}$ the firm will prefer to leave the market shrouded. Note here that since firms have different add-on production costs ($0 = \hat{c}_1 < \hat{c}_2 = \hat{c}$), they have different incentives about shrouding. In particular, the firm with the lowest add-on production cost has the strongest incentive to unshroud the add-on (the highest critical value of $\alpha$). It holds that no firm has an incentive to deviate to unshrouding if

$$\alpha \geq \max\left\{\frac{e-\hat{c}_1}{\bar{p}-\hat{c}_1}, \frac{e-\hat{c}_2}{\bar{p}-\hat{c}_2}\right\} = \max\left\{\frac{e}{\bar{p}}, \frac{e-\hat{c}}{\bar{p}-\hat{c}}\right\} = \frac{e}{\bar{p}}.$$

### 1.3. Proof of Proposition I.3:

If sophisticates can observe add-on prices, firms in a shrouded subgame adjust their add-on prices to the consumer population just like in an unshrouded subgame. The only difference is that they either face $\alpha$ myopes in a shrouded subgame or $\alpha'$ myopes in an unshrouded subgame. Applying previous results, this implies that in a shrouded subgame, firms will set add-on prices $\hat{p}_1 = \hat{p}_2 = e$ for $\alpha < \frac{e-\hat{c}}{\bar{p}-\hat{c}}$, add-on prices $\hat{p}_1 = e, \hat{p}_2 = \bar{p}$ for $\frac{e-\hat{c}}{\bar{p}-\hat{c}} < \alpha < \frac{e}{\bar{p}}$ and add-on prices $\hat{p}_1 = \hat{p}_2 = \bar{p}$ for $\frac{e}{\bar{p}} < \alpha$. Now we can compare profits to determine the incentives of firms to unshroud the add-on.

First consider the case $\alpha < \frac{e-\hat{c}}{\bar{p}-\hat{c}}$, in which firms will set low add-on prices $e$ in the shrouded subgame. This yields profits of

$$\pi_1^{S1} = \frac{1}{2t}\left(t + \frac{\hat{c}}{3}\right)^2 \quad \text{and} \quad \pi_2^{S1} = \frac{1}{2t}\left(t - \frac{\hat{c}}{3}\right)^2,$$

which are independent of $\alpha$. Since $\alpha' < \alpha$, firms will make the same profits $\pi^{U1}$ $(= \pi^{S1})$ in the unshrouded subgame. Hence, no firm has a strict incentive to unshroud the add-on and therefore a shrouding equilibrium exists.

Now suppose it holds that $\frac{e-\hat{c}}{\bar{p}-\hat{c}} < \alpha < \frac{e}{\bar{p}}$. Then, if the add-on is shrouded, firms will set add-on prices $\hat{p}_1 = e$ and $\hat{p}_2 = \bar{p}$. This yields profits of

$$\pi_1^{S2} = \frac{1}{2t}\left(t + \frac{\alpha\hat{c}}{3} - \frac{\alpha\bar{p}-e}{3}\right)^2 \quad \text{and} \quad \pi_2^{S2} = \frac{1}{2t}\left(t - \frac{\alpha\hat{c}}{3} + \frac{\alpha\bar{p}-e}{3}\right)^2.$$

Now suppose that any firm unshrouded the add-on, thereby reducing the fraction of myopes to $\alpha'$. This implies that firms would earn profits of either $\pi^{U1}$ or $\pi^{U2}$ in an unshrouded subgame.[1] Examining these profits, it holds that $\pi_1^{S2} > \pi_1^{U1}$ if and only if $\pi_2^{S2} < \pi_2^{U1}$. Equivalently, $\pi_1^{S2} > \pi_1^{U2}$ holds if and only if $\pi_2^{S2} < \pi_2^{U2}$. Hence, one of the firms makes higher profits in the unshrouded subgame and therefore has an incentive to unshroud the add-on.[2]

[1] It always holds that $\alpha' < \alpha$. Hence, $\alpha < \frac{e}{\bar{p}}$ implies $\alpha' < \frac{e}{\bar{p}}$ and the case $U3$ can be neglected.

[2] Firms are indifferent about shrouding at the lower bound of the interval $\alpha = \frac{e-\hat{c}}{\bar{p}-\hat{c}}$.

For the last case $\frac{e}{\bar{p}} < \alpha$, firms in a shrouded subgame will set high add-on prices $\bar{p}$. Hence, this case is equal to the analysis of the main model and a shrouding equilibrium only exists at the knife-edge case $\alpha' = \frac{e-\alpha\hat{c}}{\bar{p}-\hat{c}}$.

# 2. Appendix to Chapter II

## 2.1. Proof of Lemma II.1:

Following the derivations in the main text, the aggregate profit of the colluding firms is equal to:

$$\pi^M(p,\hat{p}) = [\alpha\,(1-F(p)) + (1-\alpha)(1-F(p+\hat{p}))](p+\hat{p}).$$

To derive the desired result we need to make two case distinctions:

- Suppose that $p+\hat{p} \leq \overline{v}$:
  Then it holds that some sophisticates still participate in the market and hence $1-F(p+\hat{p}) > 0$. Now we want to derive the optimal base good price:

  $$\frac{\partial \pi^M}{\partial p} = [-\alpha f(p) - (1-\alpha) f(p+\hat{p})](p+\hat{p}) + \alpha(1-F(p)) + (1-\alpha)(1-F(p+\hat{p})) \overset{!}{=} 0$$
  $$\Leftrightarrow \; p = \frac{1}{2}(\overline{v} - (2-\alpha)\hat{p}). \tag{A.1}$$

  Recall that we have assumed that some sophisticates participate in the market. Then, if the add-on price is below its maximum value, the firms can increase their profit by lowering the base good price marginally while increasing the add-on price by the same amount. This leaves the total price of the product bundle and the demand from sophisticates unchanged but increases the demand from myopic consumers. Hence, it is profitable to increase the add-on price $\hat{p}$ until it reaches its upper bound $\bar{p}$ or until the base good price $p$ reaches its lower bound. We now want to argue that the base good price reaches its lower bound before the add-on price reaches its lower bound. Suppose the add-on price reaches its upper bound first. Then the value of the base good price (A.1) would be:

  $$p = \frac{1}{2}(\overline{v} - (2-\alpha)\bar{p}).$$

  Since we have assumed that $\overline{v} \leq \bar{p}$, the above expression would yield a negative price of the base good, i.e. $p < 0$. This is not allowed in terms of the model due to the lower bound for the base good. Hence, the optimal base good price is $p^M = 0$.

- Suppose that $p+\hat{p} > \overline{v}$:
  This implies that no sophisticated consumer participates in the market, i.e. $1-F(p+\hat{p}) = 0$. Note here that in this case the profit is strictly increasing in the add-on price $\hat{p}$ and firms will set the add-on price equal to the maximum add-on

price, i.e. $\hat{p} = \bar{p}$. Maximizing the profit of firms over the base good price then yields:

$$\begin{aligned} \frac{\partial \pi^M}{\partial p} &= -\alpha f(p)(p+\bar{p}) + \alpha(1-F(p)) \stackrel{!}{=} 0 \\ \Leftrightarrow \quad p &= \frac{\overline{v}}{2}\left(1-\frac{\bar{p}}{\overline{v}}\right). \end{aligned}$$

This base good price is negative if $\overline{v} < \bar{p}$, as was assumed in Assumption II.1. Hence, due to the lower bound for the base good price, it is again optimal to set $p^M = 0$.

## 2.2. Proof of Proposition II.2:

To prove that an optimal deviation from collusion includes an unshrouding of the add-on, we have to compare the different possibilities how firms can collude and how a firm could deviate from the collusion.

- First consider the case in which the deviating firm decides to unshroud the add-on. As we have already shown in the main text, the deviating firm will set an add-on price of $\hat{p}_i^{dev} = \frac{\overline{v}}{2}$ and earn a deviation profit of $\pi_i^{dev} = \frac{\overline{v}}{4}$. The add-on price $\hat{p}_i^{dev} = \frac{\overline{v}}{2}$ is feasible since $\overline{v} \leq \bar{p}$, which ensures that $\hat{p}_i^{dev} \leq \bar{p}$.

- Now suppose that the deviating firm does not unshroud the add-on. In this case, we have to distinguish whether the deviating firm charges an add-on price below $\overline{v}$ (inner solution) or an add-on price above $\overline{v}$ (corner solution).

  - If the deviating firm charges an add-on price below $\overline{v}$, the optimal add-on price was $\pi_i^{dev} = \left[\frac{\alpha}{n} + (1-\alpha)(1-\frac{\hat{p}}{\overline{v}})\right]\hat{p}$, yielding a deviation profit of $\pi_i^{dev} = \frac{\overline{v}}{4(1-\alpha)}\left(1-\alpha\frac{n-1}{n}\right)^2$.

    Note that, for some parameter constellations, it might be the case that $\hat{p}_i^{dev} = \frac{\overline{v}}{2(1-\alpha)}\left(1-\alpha\frac{n-1}{n}\right) > \overline{v}$. In this case, the derived add-on price that a deviating firm sets is not an inner solution since no sophisticated consumer will buy the product bundle. It holds that the add-on price of the optimal deviation is feasible if:

    $$\begin{aligned} \hat{p}_i^{dev} =& \frac{\overline{v}}{2(1-\alpha)}\left(1-\alpha\frac{n-1}{n}\right) \leq \overline{v} \\ \Leftrightarrow \alpha & \qquad\qquad \leq \frac{n}{n+1}. \end{aligned}$$

    Hence, the deviation add-on price of the inner solution $\pi_i^{dev} = \left[\frac{\alpha}{n} + (1-\alpha)(1-\frac{\hat{p}}{\overline{v}})\right]\hat{p}$ is only feasible if $\alpha \leq \frac{n}{n+1}$.

– Now suppose that the deviating firm charges an add-on price above $\overline{v}$. In this case, the profit function of the deviating firm is:

$$\pi_i^{dev} = \left[\frac{\alpha}{n}\right]\hat{p}.$$

It is easy to verify that the add-on price that maximizes the above profit function is $\hat{p} = \bar{p}$, which would yield a profit of $\pi_i^{dev} = \frac{1}{n}\alpha\bar{p}$. As already argued in the main text, this cannot be a profitable deviation since this deviation profit is lower than the profit under collusion. Hence, it will never be a profitable deviation to charge an add-on price above $\overline{v}$.

In summary, we can conclude that a profitably deviating firm that decides not to unshroud the add-on cannot do better than obtaining a profit of $\pi_i^{dev} = \frac{\overline{v}}{4(1-\alpha)}\left(1-\alpha\frac{n-1}{n}\right)^2$. However, recall that the corresponding add-on price is only feasible if $\alpha \leq \frac{n}{n+1}$.

Now we can check what the optimal deviation strategy looks like. Comparing the deviation profit that a firm can obtain by unshrouding to the profit without unshrouding yields:

$$\begin{aligned}
&\pi_{i,shrouding}^{dev} && \leq \pi_{i,unshrouding}^{dev} \\
\Leftrightarrow\ &\frac{\overline{v}}{4(1-\alpha)}\left(1-\alpha\frac{n-1}{n}\right)^2 && \leq \frac{\overline{v}}{4} \\
\Leftrightarrow\ &\left(1-\alpha\frac{n-1}{n}\right)^2 && \leq 1-\alpha \\
\Leftrightarrow\ &\frac{2-n}{n}+\alpha\left(\frac{n-1}{n}\right)^2 && \leq 0 \\
\Leftrightarrow\ &\alpha && \leq \frac{n(n-2)}{(n-1)^2}. \qquad \text{(A.2)}
\end{aligned}$$

Recall that a deviation without unshrouding the add-on can only be profitable if $\alpha \leq \frac{n}{n+1}$. Hence, when we want to check whether a deviation without unshrouding can be optimal, we can focus on cases with $\alpha \leq \frac{n}{n+1}$. Then the inequality (A.2) holds if the following relation is fulfilled:

$$\begin{aligned}
&\frac{n(n-2)}{(n-1)^2} && \geq \frac{n}{n+1} \\
\Leftrightarrow\ &n && \geq 3,
\end{aligned}$$

which is fulfilled by assumption. Therefore, we can conclude that if a profitable deviation exists, unshrouding will be part of the optimal deviation strategy.

### 2.3. Proof of Proposition II.4:

Suppose firms coordinate on prices $\hat{p}^{coll}$, yielding some collusive profit $\pi^{coll}$: In analogy to the above analysis, the critical discount factor is then given by:

$$\delta^* = \frac{n - \frac{\pi^{coll}}{\pi^{dev}}}{n},$$

which is decreasing in the collusion-to-deviation profit ratio $\frac{\pi^{coll}}{\pi^{dev}}$. Hence, the less attractive a deviation is relative to the collusive play, the more stable collusion is.

Recall that if a profitable deviation exists, it is optimal to set $\hat{p}^{dev} = \frac{\overline{v}}{2}$ and unshroud the add-on, thereby earning $\pi_i^{dev} = \frac{\overline{v}}{4}$. This deviation profit is feasible only if firms coordinated on prices $\hat{p}^{coll} \geq \frac{\overline{v}}{2}$ in the collusive play. Then the critical discount factor is minimized by coordinating on maximum profits, which are $\pi^M = \frac{\overline{v}}{4(1-\alpha)}$ or $\pi^M = \alpha\bar{p}$. Hence, coordinating on lower profits cannot stabilize collusion.

Now suppose firms coordinate on prices $\hat{p}^{coll} < \frac{\overline{v}}{2}$, which yields an aggregate profit of $\pi(\hat{p}) = [\alpha + (1-\alpha)(1-\frac{\hat{p}}{\overline{v}})]\hat{p}$. Playing $\hat{p}^{dev} = \frac{\overline{v}}{2}$ does not correspond to an undercutting anymore and a firm would then optimally deviate by undercutting the collusive price marginally, irrespective of the (un)shrouding decision. This actually follows from the fact that deviation profits are increasing in $\hat{p}^{dev}$ for all add-on prices $\hat{p}^{dev} < \frac{\overline{v}}{2}$.

We will now show that a deviating firm still optimally decides to unshroud the add-on, thereby making use of the fact that the optimal deviation price does not depend on the (un)shrouding decision:

$$\begin{aligned}
& \pi_{i,unshrouding}^{dev} && > \pi_{i,shrouding}^{dev} \\
\Leftrightarrow \quad & \left[1 - \frac{\hat{p}_i^{dev}}{\overline{v}}\right] \hat{p}_i^{dev} && > \left[\frac{\alpha}{n} + (1-\alpha)(1 - \frac{\hat{p}_i^{dev}}{\overline{v}})\right] \hat{p}_i^{dev} \\
\Leftrightarrow \quad & 1 - \frac{\hat{p}^{dev}}{\overline{v}} && > \frac{1}{n}
\end{aligned}$$

This holds since $n \geq 3$ and $\hat{p}^{dev} \leq \frac{\overline{v}}{2}$. Now we that know that a deviating firm will optimally unshroud the add-on and undercut the collusive price marginally, it remains to check whether the critical discount factor can be lowered by coordinating on add-on prices $\hat{p}^{coll} < \frac{\overline{v}}{2}$. Since colluding with the monopoly price leads to a collusion-to-deviation profit ratio of at least $1/(1-\alpha)$, it must hold that:

$$\begin{aligned}
\frac{\pi^{coll}}{\pi^{dev}} = \frac{[\alpha + (1-\alpha)(1-\frac{\hat{p}}{\overline{v}})]\hat{p}}{\left[1-\frac{\hat{p}}{\overline{v}}\right]\hat{p}} &> \frac{1}{1-\alpha} \\
\Leftrightarrow\ (1-\alpha)\left[1-(1-\alpha)\frac{\hat{p}}{\overline{v}}\right] &> 1-\frac{\hat{p}}{\overline{v}} \\
\Leftrightarrow\ \left[1-(1-\alpha)^2\right]\frac{\hat{p}}{\overline{v}} &> \alpha \\
\Leftrightarrow\ \alpha\,(2-\alpha)\,\frac{\hat{p}}{\overline{v}} &> \alpha \\
\Leftrightarrow\ \hat{p} &> \frac{\overline{v}}{2-\alpha},
\end{aligned}$$

which contradicts the assumption that firms colluded with prices $\hat{p}^{coll} < \frac{\overline{v}}{2}$. Note that it might be more profitable for a deviating firm to play the corner solution and earn $\pi^{dev} = \alpha\bar{p}$ than to unshroud and undercut the collusive price marginally. But since this would only decrease the collusion-to-deviation profit ratio and result in a higher critical discount factor, collusion would be further destabilized. We can therefore conclude that coordinating on other than monopoly profits cannot stabilize collusion.

# 3. Appendix to Chapter III

## 3.1. Derivation of an Optimal Attention Allocation under Cognitive Constraints

Suppose a decision maker faces the problem of choosing between a finite number of alternatives from the set $A$. Each alternative is described by a finite vector of attributes $i \in I$. Let $x_i^a$ denote the extent to which alternative $a$ features attribute $i$. The experienced utility of each alternative $a \in A$ is expressed by

$$u(a) = \sum_i v_i x_i^a, \tag{A.3}$$

where $v_i$ denotes the value the DM ascribes to an additional unit $x_i$ of attribute $i \in I$. Thus the choice problem can be expressed by $(A, v)$ where $A = (x_i^a)_{a \in A, i \in I}$ and $v = (v_i)_{i \in I}$. Suppose a decision-maker (DM) faces cognitive constraints such that she incurs cognitive costs whenever she faces a choice between multi-attribute alternatives. i.e. $|I| \geq 2$. The DM faces no information problem, she perfectly knows the values $A$ and $v$. She, however, faces problems whenever she needs to integrate this information in order to make a choice. She thus imperfectly considers or takes into account the information, and thus evaluates each alternative by its decision utility given by

$$\tilde{u}(a) = \sum_i \left[ m_i v_i x_i^a + (1 - m_i) v_i x_i^d \right], \tag{A.4}$$

where $m_i \in [0, 1]$ can be thought of as an attention parameter. $m_i = 1$ denotes full attention, while $m_i = 0$ denotes complete neglect. When neglecting the information in one dimension/attribute, the DM ascribes some value $x_i^d$ to each alternative. Depending on the assumptions one wants to make, this default value may differ. If the DM is "on average right", one may choose $x_i^d = \bar{x}_i$ where $\bar{x}_i$ is the average value of $x_i$ across the available alternatives. If the DM is pessimistic, $x_i^d = \min_{a \in A} x_i^a$ might be a good assumption. If the DM has some default alternative, $x_i^d = x_i^{default}$ could be reasonable. Regardless of these assumptions, if the DM neglects a dimension, i.e. if $m_i = 0, \ \forall i \in I$, she is not able to discriminate between alternatives along this dimension.[3]

[3] The exact assumptions about $x_i^d$ are irrelevant for the behavior of the DM. For any value of $m_i \in [0, 1]$ and for any modeler's choice of $(x_i^d)_{i=1,\ldots,n}$, a constant $\sum_i (1 - m_i) v_i x_i^d$ is added to the decision utility of each alternative. For a given vector $(m_i)_{i=1,\ldots,n}$, this constant is identical across alternatives. It may differ *across* sets of alternatives because $m_i$ is a function of this set, as we are about to derive. Yet, it does not differ across alternatives for a given set of alternatives. It has thus no impact on the desirability of one alternative over another. With no behavioral impact, we drop it in the main section without loss of generality.

Now, let's look at the error from imperfectly considering the dimensions:

$$\sum_{a\in A}\left[\sum_i \left(v_i x_i^a - m_i v_i x_i^a - (1-m_i)v_i x_i^d\right)\right] = \sum_i \left[(1-m_i)v_i \sum_{a\in A}(x_i^a - x_i^d)\right].$$

An optimal attention allocation will weigh losses from an erroneous representation against losses from incurred cognition costs.[4] Remember that we seek to model cognitive costs associated with complexity. It should thus be straightforward to assume that considering a single dimension is costless, as there is no complexity involved. Considering a second dimension involves the need to make a first trade-off consideration and should thus be associated with some positive cognition costs. Taking into account additional dimensions should become increasingly costly as the number of trade-off considerations that need to be made rises exponentially. It thus matters which dimensions are considered "first". Denote by $r: I \to \{1, ..., n\}$ the order in which the attributes are considered. We will henceforth refer to it as the attention hierarchy. Given some place in the hierarchy, each dimension is associated with some cognitive effort cost $\kappa_{r(i)}$.

The exact loss function is, again, a modeling choice. It should include losses from an imperfect problem representation and a loss from exerting cognitive effort. Let us consider the following maximization problem:

$$\max_{m_i}(-L) = -\frac{1}{2}\sum_i (1-m_i)^2\mu_i - \sum_i \kappa_{r(i)}|m_i|, \tag{A.5}$$

where $\mu_i$ could have several forms depending on the modeling choice. Consider, for example, $x_i^d = \bar{x}_i$ and losses from errors that are quadratic in each dimension. Then one might use $\mu_i = v_i^2\sigma_i^2$, where $\sigma_i^2$ is the variance of attribute $i$ in the set of alternatives. Consider the case $x_i^d = \min_{a\in A} x_i^a$. Then $\mu_i$ is the product of $v_i$ and the sum (or the average) increase in value by considering the true value of an alternative. Whatever the modeling choice, as $x_i^d$ is independent of $m_i$, $\mu_i$ is a term capturing a product of $v_i$ and some measure of dispersion of attribute $x_i$ in the set of alternatives.

The term capturing cognitive effort cost in a dimension $i$ is the $\ell_1$ norm of the attention parameter $m_i$ times the cognitive cost parameter associated with the rank of $i$.[5]

[4] One may argue that the DM's objective is to make the right decision, not to form a correct representation of the world. So, one might want to insert the loss from taking the wrong action into the objective function. Yet, to determine that loss, one needs to know the right action. The problem would amount to choosing the optimal attention with hindsight. Without the information about the correct action, the best thing one can do is to optimize the representation of the world one bases decisions on. An accurate representation of an alternative's desirability is vital to make the correct choice.

[5] For a more elaborate discussion of modeling cognitive costs with such a cost function, see Gabaix (2011).

The optimal solution of problem (A.5) is then given by

$$m_i = \max\{0, 1 - \kappa_i/\mu_i\}. \tag{A.6}$$

Hence, whenever $\mu_i < \kappa_{r(i)}$, dimension $i$ is neglected. One could thus interpret $\mu_i$ as a measure of importance of dimension $i$ to the DM. If the dimension is important enough compared to the cognitive costs associated with its consideration, it will be taken into account. And, given that it is taken into account ($m_i > 0$), the extent to which a dimension is taken into account rises in the importance of the dimension.

It is important to note that the attention $m_i$ which a dimension $i$ receives is crucially determined by the cognitive costs $\kappa_{r(i)}$ associated with its consideration. As was discussed before these costs shall reflect the rising difficulty of solving increasingly complex problems. We will thus assume

$$\begin{aligned} \kappa_r &= 0, \qquad \text{for } r = 1, \\ \kappa_{r+1} &> \kappa_r, \quad \forall r \in \{1, ..., n-1\}. \end{aligned} \tag{A.7}$$

The attribute which is considered first, i.e. the attribute which receives rank 1 in the attention hierarchy, is considered without cognitive effort. Considering additional attributes becomes increasingly costly. The attention hierarchy $r$ is thus crucial for the eventual attention allocation. Consider the following 2-step procedure. First, the DM needs to select an attention hierarchy $r : I \to \{1, ..., n\}$ which associates each dimension with some consideration costs $\kappa_{r(i)}$. One can think of this as the problem to determine which dimension to consider first, which second, and so on. After assigning a rank to each dimension, the DM solves the above described problem of optimal attention allocation given some assignment of consideration costs. The problem can then be solved backward. Given any assignment $r(i)$, the optimal attention allocation is given by (A.6). Plugging this back into the objective function (A.5) yields

$$\begin{aligned} (-L) &= \sum_{i:m_i>0} \left[ -\frac{1}{2}\left(\frac{\kappa_{r(i)}^2}{\mu_i^2}\right)\mu_i - \kappa_{r(i)} + \frac{\kappa_{r(i)}^2}{\mu_i} \right] - \frac{1}{2}\sum_{j:m_j=0} \mu_j \\ &= \sum_{i:m_i>0} \left[ \frac{\kappa_{r(i)}^2}{\mu_i} - \kappa_{r(i)} \right] - \frac{1}{2}\sum_{j:m_j=0} \mu_j. \end{aligned}$$

The objective at the first stage, anticipating the result of the second stage, is thus

$$\max_{r(i)} \sum_{i:m_i>0} \left[ \frac{1}{2}\frac{\kappa_{r(i)}^2}{\mu_i} - \kappa_{r(i)} \right] - \frac{1}{2}\sum_{j:m_j=0} \mu_j. \tag{A.8}$$

Now we can state the following result:

**Proposition A.1.** *The optimal assignment* $r^*(i)$ *will satisfy:*

$$\mu_i > \mu_j \Rightarrow r(i) < r(j). \tag{A.9}$$

*The optimal assignment will thus assign higher attention ranks to more important dimensions.*

*Proof.* Under an optimal assignment $r^*(i)$ interchanging the ranks of any two attributes $i, j \in I$ may not lead to an increase in $(-L)$. Note that the objective function $(-L)$ is additively-separable across attributes. We may thus confine attention to the parts of the objective function that depend on the two attributes $i$ and $j$.

Suppose $\mu_i = \mu_j$. It is easy to see that interchanging their ranks has no effect on the objective function. We will thus only look at cases in which $\mu_i > \mu_j$. Take any ranking $r$. Under this ranking attribute $i$ and $j$ are associated with some cognitive costs $\kappa_{r(i)}, \kappa_{r(j)}$. Denote by $\kappa_h = \max\left\{\kappa_{r(i)}, \kappa_{r(j)}\right\}$ and $\kappa_l = \min\left\{\kappa_{r(i)}, \kappa_{r(j)}\right\}$. Whichever attribute is assigned $\kappa_l$ has a higher rank under $r$. We will now show that if $r$ does not assign $\kappa_l$ to attribute $i$ (the one with strictly higher importance), $r$ cannot be a maximizer for $(-L)$ for some set of cognitive costs $(\kappa_1, \kappa_2, ...)$ satisfying our assumption (A.7).

Let us distinguish four cases:

(i) $\kappa_h > \kappa_l > \mu_i > \mu_j$. Both attributes are neglected before and after interchanging the rank. The objective function is thus invariant to such a change in ranking.

(ii) $\mu_i > \mu_j > \kappa_h > \kappa_l$. Both attributes are taken into account at the lower rank. However, it is better to assign attribute $i$ the higher rank (and thus $\kappa_l$) if

$$\begin{aligned}
&\frac{1}{2}\frac{\kappa_l^2}{\mu_i} - \kappa_l + \frac{1}{2}\frac{\kappa_h^2}{\mu_j} - \kappa_h && > \frac{1}{2}\frac{\kappa_l^2}{\mu_j} - \kappa_l + \frac{1}{2}\frac{\kappa_h^2}{\mu_i} - \kappa_h \\
\Leftrightarrow\ &\frac{\kappa_l^2}{\mu_i} + \frac{\kappa_h^2}{\mu_j} && > \frac{\kappa_l^2}{\mu_j} + \frac{\kappa_h^2}{\mu_i} \\
\Leftrightarrow\ &(\kappa_h^2 - \kappa_l^2)\mu_i && > (\kappa_h^2 - \kappa_l^2)\mu_j \\
\Leftrightarrow\ &\mu_i && > \mu_j.
\end{aligned}$$

(iii) $\kappa_h > \mu_i > \mu_j > \kappa_l$. Both attributes are considered at the higher rank but neglected at the lower rank. It is better to assign attribute $i$ the higher rank (and thus $\kappa_l$) if

$$\begin{aligned}
&\frac{1}{2}\frac{\kappa_l^2}{\mu_i} - \kappa_l - \frac{1}{2}\mu_j && > \frac{1}{2}\frac{\kappa_l^2}{\mu_j} - \kappa_l - \frac{1}{2}\mu_i \\
\Leftrightarrow\ &\kappa_l^2\mu_j - \mu_j^2\mu_i && > \kappa_l^2\mu_i - \mu_i^2\mu_j \\
\Leftrightarrow\ &\mu_i\mu_j(\mu_i - \mu_j) && > \kappa_l^2(\mu_i - \mu_j) \\
\Leftrightarrow\ &\mu_i\mu_j && > \kappa_l^2, \quad \text{which is true since } \mu_i > \mu_j > \kappa_l.
\end{aligned}$$

(iv) $\mu_i > \kappa_h > \mu_j > \kappa_l$. Attribute $i$ is considered both at the higher and lower rank. Attribute $j$ is only considered at the higher rank but neglected at the lower rank. Still, it is better to assign attribute $i$ to the higher rank if

$$\frac{1}{2}\frac{\kappa_l^2}{\mu_i} - \kappa_l - \frac{1}{2}\mu_j > \frac{1}{2}\frac{\kappa_l^2}{\mu_j} - \kappa_l + \frac{1}{2}\frac{\kappa_h^2}{\mu_i} - \kappa_h.$$

This indeed holds true, since

$$\begin{aligned}
\frac{1}{2}\kappa_l^2\left(\frac{1}{\mu_i} - \frac{1}{\mu_j}\right) - \frac{1}{2}\mu_j - \frac{1}{2}\kappa_h^2\frac{1}{\mu_i} + \kappa_h &= \\
\frac{1}{\mu_i\mu_j}\left[\frac{1}{2}\kappa_l^2(\mu_j - \mu_i) - \frac{1}{2}\mu_i\mu_j^2 - \frac{1}{2}\kappa_h^2\mu_j + \kappa_h\mu_i\mu_j\right] &= \\
\frac{1}{2\mu_i\mu_j}\left[\kappa_l^2(\mu_j - \mu_i) + \left(\kappa_h\mu_i\mu_j - \mu_i\mu_j^2\right) + \left(\kappa_h\mu_i\mu_j - \kappa_h^2\mu_j\right)\right] &= \\
\frac{1}{2\mu_i\mu_j}\left[\kappa_l^2(\mu_j - \mu_i) + \mu_i\mu_j(\kappa_h - \mu_j) + \kappa_h\mu_j(\mu_i - \kappa_h)\right] &> \\
\frac{1}{2\mu_i\mu_j}\left[\kappa_l^2(\mu_j - \mu_i) + \kappa_h\mu_j(\kappa_h - \mu_j) + \kappa_h\mu_j(\mu_i - \kappa_h)\right] &= \\
\frac{1}{2\mu_i\mu_j}\left[(\mu_j\kappa_h - \kappa_l^2)(\mu_i - \mu_j)\right] &> 0,
\end{aligned}$$

in which the first inequality (line 4) holds since $\mu_i > \kappa_h > \kappa_j$, and hence the middle term is replaced by strictly lower term. The final inequality holds since $\mu_i > \kappa_h > \mu_j > \kappa_l$.

We have now shown that an optimal solution is to set $r(i)$ according to (A.9). One may argue that this rule may not be obeyed for dimensions for which case (i) holds. While this is true, one can counter that (A.9) is optimal for a choice problem $(A, v)$, and thus for a given vector of $(\mu_i)_{i=1..n}$, for *all* cognitive cost vectors $(\kappa_1, ..., \kappa_n)$ for which $\kappa_{r+1} > \kappa_r$, $\forall r = 1, ..., n$. The optimal assignment (A.9) is thus invariant to changes in the cost vector (e.g. due to changes in cognitive resources for some given choice task). In addition, even if case (i) may hold for some attributes for a given cost vector, it cannot hold for all attributes as long as $\kappa_1 = 0$. □

### Characteristics of the Attention Function

This section seeks to discuss some characteristics of the attention function, that we derived.

First, more important attributes receive (weakly) more attention than less important attributes, $\partial m_i/\partial \mu_i \geq 0$. The attention each attribute receives thus depends positively on its value to the decision-maker. In addition to this internal factor, the attributes dispersion within the choice set $A$ increases attention. The attention an attribute receives thus depends on the choice environment. It is thus possible to **attract attention** to an attribute by varying the choice set appropriately.

Second, as one attribute gains importance it may eventually gain rank in the attention hierarchy. As another attribute receives a lower rank this other attribute loses attention, $\partial m_i/\partial \mu_j \leq 0$, with strict inequality only if $\partial m_j/\partial \mu_j > 0$. It is hence possible to **distract attention** from an attribute. It is noteworthy that this distraction effect only works through the attention hierarchy and is thus discontinuous. While this feature might be mathematically undesirable, it has some desirable effects which we will discuss shortly.

Next, the attention process features **neglect**, or, in Gabaix' terminology, the attention vector is sparse. Technically, for any decision problem $(A, v)$ there exist vectors of cognitive costs $\kappa$ satisfying our assumptions such that there exist attributes $i \in I : m_i = 0$ whenever $|I| \geq 2$. So, for any complex choice problem, that is one which involves at least two dimensions, cognitive costs may lead to the neglect of at least one dimension. Similarly, for any vector of cognitive costs $\kappa$ satisfying our assumptions there exist choice problems $(A, v) : |I| \geq 2$, such that at least one of the dimensions is ignored.

In addition, due to our assumptions on $\kappa$, for any choice problem $(A, v)$ there exist attributes $i \in I : m_i > 0$. So, there is **no complete neglect**. As we seek to model the need to simplify a complex choice problem, the DM always considers at least one dimension as this amounts to solving a simple choice problem. This directly implies that complexity costs, as modeled here, will never lead to strictly dominated choices.

The attention hierarchy is not just implicit. Any two attributes that are considered receive a **different weight**: $m_i \neq m_j$, $\forall i, j \in I : m_i, m_j > 0$. More specifically, for any two attributes that are considered, the higher ranking attribute receives strictly more attention.

This, together with the impossibility of complete neglect, implies that the attention process *always* features **over- and underweighting**. Let $\bar{m} = \frac{1}{n} \sum_i m_i$. Then for any complex choice problem and cognitive cost vector $\kappa$ satisfying our assumptions there exist attributes which are overweighted and attributes which are underweighted. Formally, $\forall (A, v, \kappa) : \exists i \in I : m_i > \bar{m}$ and $\exists j \in I : m_j < \bar{m}$. This is important as it implies that under the derived attention function the decision utility of an alternative is not just an affine transformation of experienced utility.

### Final Remarks

The optimal attention allocation rule

$$m_i = \max\left\{0, 1 - \frac{\kappa_{r(i)}}{\mu_i}\right\}$$
$$\mu_i > \mu_j \Rightarrow \kappa_{r(i)} < \kappa_{r(j)}, \ \forall i, j \in I$$

we derived is, of course, more general than the attention allocation rule we employ in the main part of this paper. We have decided to opt for $\mu_i = \max_{a \in A} - \min a \in A$ as a simple measure of dispersion as it yields unique results for the optimal product design. The results will remain qualitatively unchanged when employing a different measure of dispersion. In addition, we have assumed a tie-breaking rule $\mu_i = \mu_j$ and $v_i > v_j \Rightarrow \kappa_{r(i)} < \kappa_{r(j)}, \ \forall i, j \in I$ for the determination of the attention hierarchy. Assuming a different tie-breaking rule would not change our results qualitatively.

# Bibliography


Agarwal, S., J. Driscoll, X. Gabaix, and D. Laibson (2009). The age of reason: Financial Decisions over the Lifecycle. *Brookings Papers on Economic Activity 2*, 51–117.

Agarwal, S., J. Driscoll, X. Gabaix, D. Laibson, L. Center, and M. Room (2011). Learning in the Credit Card Market. Working Paper.

Armstrong, M. and Y. Chen (2009). Inattentive Consumers and Product Quality. *Journal of the European Economic Association 7*(2-3), 411–422.

Benartzi, S. and R. Thaler (2002). How Much is Investor Autonomy Worth? *The Journal of Finance 57*(4), 1593–1616.

Bordalo, P. (2011). Choice-Set Effects and Salience Theory. Working Paper.

Bordalo, P., N. Gennaioli, and A. Shleifer (2010). Salience Theory of Choice Under Risk. Working Paper.

Brandenburger, A. and B. Nalebuff (1998). *Co-opetition*. Crown Business.

Carlin, B. I. (2009). Strategic Price Complexity in Retail Financial Markets. *Journal of Financial Economics 91*(3), 278–287.

Carlin, B. I. and G. Manso (2011). Obfuscation, Learning and the Evolution of Investor Sophistication. *Review of Financial Studies 24*, 754–785.

Carrell, L. (2008). *ETFs for the Long Run*. Wiley.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and evidence. *The American Economic Review 99*(4), 1145–1177.

Cruickshank, D. (2000). Competition in UK Banking: A Report to the Chancellor of the Exchequer. London: HM Treasury.

Dahremöller, C. (forthcoming). Unshrouding for Competitive Advantage. *Journal of Economics & Management Strategy*.

Eliaz, K. and R. Spiegler (2011). On the strategic use of attention grabbers. *Theoretical Economics 6*(1), 127–155.

Ellison, G. (2005). A Model of Add-On Pricing. *Quarterly Journal of Economics 120*(2), 585–637.

Ellison, G. and A. Wolitzky (2011). A Search Cost Model of Obfuscation. Working Paper.

European Commission (2004). 2004/138/EC: Commission Decision: Austrian Banks - Lombard Club. `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:%2032004D0138:EN:NOT`.

Fudenberg, D. and E. Maskin (1986). The Folk Theorem for Repeated Games with Discounting or with Incomplete Information. *Econometrica 54*, 533–554.

Fudenberg, D. and J. Tirole (2000). Customer poaching and brand switching. *RAND Journal of Economics 31*(4), 634–657.

Gabaix, X. (2011). A Sparsity-Based Model of Bounded Rationality. Working Paper.

Gabaix, X. and D. Laibson (2006). Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets. *Quarterly Journal of Economics 121*(2), 505–540.

Grossman, S. (1981). The Informational Role of Warranties and Private Disclosure about Product Quality. *Journal of Law and Economics 24*(3), 461.

Gruber, M. (1996). Another Puzzle: The Growth in Actively Managed Mutual Funds. *Journal of Finance 51*(3), 783–810.

Herne, K. (1997). Decoy alternatives in policy choices: Asymmetric domination and compromise effects. *European Journal of Political Economy 13*(3), 575–589.

Huber, J., J. Payne, and C. Puto (1982). Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. *The Journal of Consumer Research 9*(1), 90–98.

Johnson, J. and D. Myatt (2006). On the Simple Economics of Advertising, Marketing, and Product Design. *The American Economic Review 96*(3), 756–784.

Kahneman, D. (2000). Experienced Utility and Objective Happiness: A Moment-Based Approach. In D. Kahneman and A. Tversky (Eds.), *Choices, Values, and Frames*, pp. 673–692. Cambridge University Press.

Kahneman, D. and A. Tversky (1984). Choices, Values, and Frames. *American Psychologist 39*(4), 341–350.

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

Kosfeld, M. and U. Schüwer (2010). Add-on Pricing, Consumer Myopia and Regulatory Intervention. Working Paper.

Kőszegi, B. and A. Szeidl (2012). A Model of Focusing in Economic Choice. Working Paper.

Lal, R. and C. Matutes (1994). Retail Pricing and Advertising Strategies. *Journal of Business 67*(3), 345.

Liebman, J. and R. Zeckhauser (2004). Schmeduling. Working Paper.

Mailath, G. and L. Samuelson (2006). *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, USA.

Massoud, N., A. Saunders, and B. Scholnick (2007). Who Makes Credit Card Mistakes? Working Paper.

Massoud, N., A. Saunders, and B. Scholnick (2011). The cost of being late? The case of credit card penalty fees. *Journal of Financial Stability 7*(2), 49–59.

McFadden, D. (1999). Rationality for Economists? *Journal of Risk and Uncertainty 19*(1), 73–105.

Miao, C. (2010). Consumer myopia, standardization and aftermarket monopolization. *European Economic Review 54*, 931–946.

Milgrom, P. (1981). Good News and Bad News: Representation Theorems and Applications. *Bell Journal of Economics 12*(2), 380–391.

Miravete, E. (2007). The Doubtful Profitability of Foggy Pricing. Working Paper.

O'Donoghue, T. and M. Rabin (1999). Doing It Now or Later. *American Economic Review 89*(1), 103–124.

Office of Fair Trading (2002). Consumer IT goods and services.

Osborne, M. and A. Rubinstein (1998). Games with Procedurally Rational Players. *The American Economic Review 88*(4), 834–847.

Rubinstein, A. (1988). Similarity and Decision-making under Risk (Is There a Utility Theory Resolution to the Allais Paradox?). *Journal of Economic Theory 46*(1), 145–153.

Rubinstein, A. (1993). On Price Recognition and Computational Complexity in a Monopolistic Model. *Journal of Political Economy 101*(3), 473–484.

Shapiro, C. (1995). Aftermarkets and Consumer Welfare: Making Sense of Kodak. *Antitrust Law Journal 63*, 483–511.

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

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

Sims, C. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*, 665–690.

Spiegler, R. (2006a). Competition over agents with boundedly rational expectations. *Theoretical Economics 1*(2), 207–231.

Spiegler, R. (2006b). The Market for Quacks. *Review of Economic Studies 73*(4), 1113–1131.

Spiegler, R. (2011). *Bounded Rationality and Industrial Organization*. Oxford University Press.

Stango, V. and J. Zinman (2009). What Do Consumers Really Pay on Their Checking and Credit Card Accounts? Explicit, Implicit, and Avoidable Costs. *American Economic Review 99*(2), 424–429.

TNS (2011). Special Eurobarometer 342: Consumer empowerment. TNS Opinion & Social.

Vikander, N. (2010). Targeted Advertising and Social Status. Working Paper.

Wernerfelt, B. (1995). A Rational Reconstruction of the Compromise Effect: Using Market Data to Infer Utilities. *Journal of Consumer Research 21*(4), 627–633.

Wilson, C. (2010). Ordered Search and Equilibrium Obfuscation. *International Journal of Industrial Organization 28*(5), 496–506.

Zhou, J. (2007). Advertising, Misperceived Preferences, and Product Design. Working Paper.