

NBER WORKING PAPER SERIES

SALIENCE AND CONSUMER CHOICE

Pedro Bordalo  
Nicola Gennaioli  
Andrei Shleifer

Working Paper 17947  
<http://www.nber.org/papers/w17947>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
March 2012

We are grateful to David Bell, Tom Cunningham, Matt Gentzkow, Bengt Holmstrom, Daniel Kahneman, Drazen Prelec, Jan Rivkin, Josh Schwartzstein, Jesse Shapiro, Itamar Simonson, Dmitry Taubinski and Richard Thaler for extremely helpful comments. Gennaioli thanks the Barcelona GSE Research Network and the Generalitat de Catalunya for financial support. Shleifer thanks the Kauffman Foundation for research support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Saliency and Consumer Choice  
Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer  
NBER Working Paper No. 17947  
March 2012  
JEL No. D03,D11

### **ABSTRACT**

We present a theory of context-dependent choice in which a consumer's attention is drawn to salient attributes of goods, such as quality or price. An attribute is salient for a good when it stands out among the good's characteristics, in the precise sense of being furthest away in that good from its average level in the choice set (or more generally, an evoked set). A local thinker chooses among goods by attaching disproportionately high weights to their salient attributes. When goods are characterized by only one quality attribute and price, saliency tilts choices toward goods with higher ratios of quality to price. We use the model to account for a variety of disparate bits of evidence, including decoy effects in consumer choice, context-dependent willingness to pay, balance of qualities in desirable goods, and shifts in demand toward low quality goods when all prices in a category rise. We then apply the model to study discounts and sales, and to explain demand for low deductible insurance.

Pedro Bordalo  
Department of Economics  
Royal Holloway  
University of London  
Egham Hill, Egham, TW20 0EX  
United Kingdom  
pedro.bordalo@rhul.ac.uk

Andrei Shleifer  
Department of Economics  
Harvard University  
Littauer Center M-9  
Cambridge, MA 02138  
and NBER  
ashleifer@harvard.edu

Nicola Gennaioli  
CREI  
Universitat Pompeu Fabra  
Ramon Trias Fargas 25-27  
08005 Barcelona (Spain)  
ngennaioli@crei.cat

# 1 Introduction

Imagine yourself in a wine store, choosing a red wine. You are considering a French syrah from the Rhone Valley, selling for \$20 a bottle, and an Australian Shiraz, made from the same grape, selling for \$10. You know and like French syrah better, you think it is perhaps 50% better. Yet it sells for twice as much. After some thought, you decide the Australian shiraz is a better bargain and buy a bottle.

A few weeks later, you are at a restaurant, and you see the same two wines on the wine list. Yet both of them are marked up by \$40, with the French syrah selling for \$60 a bottle, and the Australian shiraz for \$50. Now, you think the French wine is 50% better, but only 20% percent more expensive. At the restaurant, it is a better deal. You splurge and order the French wine.

This example illustrates what perhaps has happened to many of us, namely thinking in context and figuring out which of several choices represents a better deal in light of the options we face. In this paper, we try to formalize the intuition behind such thinking. The intuition generalizes what we believe goes through a consumer's mind in the wine example: at the store, the price difference between the cheaper and the more expensive wine is more salient than the quality difference, encouraging the consumer to opt for the cheaper option, whereas at the restaurant, after the markups, the quality difference is more salient, encouraging the consumer to splurge. We argue that this kind of thinking can help account for and unify a broad range of disparate thought experiments, field experiments, and even field data that have been difficult to account for in standard models, and certainly in one model.

Consider a few examples. A car buyer would prefer to pay \$17,500 for a car equipped with a radio to paying \$17,000 for a car without a radio, but at the same time would not buy a radio separately for \$500 after agreeing to buy a car for \$17,000 (Savage 1954). In a related vein, experimental subjects thinking of buying a calculator for \$15 and a jacket for \$125 are more likely to agree to travel for 10 minutes to save \$5 on the calculator than to travel the same 10 minutes to save \$5 on the jacket (Kahneman and Tversky 1984).

When faced with a choice between a good toaster for \$20, and a somewhat better one for \$30, most experimental subjects choose the cheaper toaster. But when a marginally superior

toaster is added to the choice set for \$50, many consumers switch to the middle toaster, violating the axiom of Independence of Irrelevant Alternatives (Tversky and Simonson 1993).

Imagine sunbathing with a friend on a beach in Mexico. It is hot, and your friend offers to get you an ice-cold Corona from the nearest place, which is a hundred yards away. He asks for your reservation price. In the first treatment, the nearest place to buy the beer is a beach resort. In the second treatment, the nearest place is a corner store. Many people would pay more for a beer from a resort than for one from the store, contradicting the fundamental assumption that willingness to pay for a good is independent of context (Thaler 1985, 1999).

When gasoline prices rise, many people switch from higher to lower grade gasoline (Hastings and Shapiro 2011).

Stores often post extremely high regular prices for goods, but then immediately put them on sale at substantial discounts. The original prices and percentage discounts are displayed prominently for consumers. In some department stores, more than half the revenues come from sales (Ortmeyer, Quelch and Salmon 1991).

Consumers opt for insurance policies with small deductibles even though the implied claim probabilities (by comparison with high deductible policies) are implausibly high (Sydnor 2010, Barseghyan, Molinari, O'Donoghue and Teitelbaum, 2011).

In this paper, we suggest that these and several other phenomena can be explained in a unified way using a model of salience in decision making. As described by psychologists Taylor and Thompson (1982), “salience refers to the phenomenon that when one’s attention is differentially directed to one portion on the environment rather than to others, the information contained in that portion will receive disproportionate weighing in subsequent judgments”. Bordalo, Gennaioli, and Shleifer (2012, hereafter BGS 2012) apply this idea to understanding decisions under risk, and present a model in which decision makers overweigh salient lottery states. They find that many anomalies in choice under risk, such as frequent risk-seeking behavior, Allais paradoxes, and preference reversals obtain naturally when salience influences decision weights. We follow BGS (2012) in stressing the interplay of attention and choice, and extend the concept of salience to riskless choice among goods with different attributes, which may include various aspects of quality, but also prices. We then describe decision making by a consumer who overweighs in his choices the most salient

attributes of each good he considers, and show that many of the phenomena just described, as well as several others, obtain naturally in such a model.<sup>1</sup>

In our model, a good’s salient attributes are those that stand out in the sense of being furthest from their average value in the choice context. Following Kahneman and Miller’s (1986) Norm Theory, we capture the choice context by the “evoked set,” which is the set of goods that come to the agent’s mind when making his choice. We call “reference good” the good with average attributes in the evoked set. The evoked set thus determines the attribute levels the decision maker views as normal, or reference, in a situation. The salient attributes are those attributes whose levels are unusual or surprising relative to the reference. The consumer focuses on those when making his choice.

Most of our results are obtained by assuming that the evoked set coincides with the choice set itself. In these cases, the application of the model is straightforward and the reference good captures Bodner and Prelec’s (1994) idea of “centroid reference.” We also explore situations in which the choice context prompts the agent to think about goods that are not in the choice set, and in particular to think about the same goods sold at historical, or normal, prices. In Thaler’s beer example, people seem to be thinking about normal beer prices at the resort or at the store. Likewise, in the Hastings-Shapiro gasoline evidence, buyers seem to be recalling previous gasoline prices. In these cases, product features such as prices that are surprising relative to prior expectations become very salient, and the consumers focuses on them when making his choice.<sup>2</sup>

The dependence of choice on external reference points is a central feature of many behavioral models. Most prominently, in Kahneman and Tversky’s Prospect Theory decision makers evaluate risky prospects by comparing them to reference points. Koszegi and Rabin (2006, 2007) suggest that reference points correspond to the decision maker’s expectations. Here we adopt the general perspective of this work, but propose perhaps a simpler notion of the reference good as having average attributes of goods in the evoked set. More importantly,

---

<sup>1</sup>We are continuing to model the phenomenon of local thinking (Gennaioli and Shleifer 2010, BGS 2012), which refers to individuals focusing on and incorporating into their decisions some aspects of their environment to a much greater extent than others. Other research that pursued a related strategy includes Mullainathan (2002), Schwartzstein (2012), Gabaix (2011) and Woodford (2012).

<sup>2</sup>Our approach is related to situations in which decision makers evaluate their options using mental accounts (Thaler 1980). The marketing literature also stresses the effect of evoked sets on choice (see Roberts and Lattin 1997 for a review).

apart from the role of the reference good in determining attribute salience, the consumer's preferences are completely standard.

We show that salience provides powerful intuitions to account for the disparate phenomena described above, and delivers several new predictions. In a broad range of situations, salience creates a tendency for consumers to focus on the relative advantage of goods having a high quality to price ratio. The model thus delivers the fundamental intuition that buyers look for bargains, whether expressed in high quality (relative to price) or low prices (relative to quality). This principle also implies that the same price difference looms smaller to a local thinker when it occurs at higher prices, explaining the choice of wines in the store vs restaurant, as well as the radio and the jacket/calculator problems: going to another shop to save \$5 looks like a good deal for the \$15 calculator, but not for the \$125 jacket.

This logic helps provide a unified explanation for:

- Decoy effects: when a bad deal such as a very expensive but marginally superior toaster is added to the choice set, the second best toaster looks like a bargain and its quality becomes salient. This leads the consumer to revise his original choice. Compromise effects, namely the preference for goods having balanced qualities in the choice set, arise in a similar way.
- Context-dependent willingness to pay: recalling that beer is expensive at resorts makes a sunbather more willing to pay a higher price (while still viewing the quality of that beer as salient) than he would if he was thinking about store prices for beer.
- Hastings-Shapiro evidence: when gas prices rise, current gas grades look like a bad deal relative to normal gas prices. The price of more expensive grades becomes salient to the consumer, who therefore switches to cheaper grades.
- Sales as a manifestation of decoy effects: the original price of a good acts as a decoy in increasing the salience of the quality of the good on sale. This perspective explains why retailers might use frequent sales, why they would put expensive rather than cheap goods on sale, and why sales do not work in the case of standard goods.

- Evidence on demand for insurance: since the percentage variation in deductibles across insurance policies is larger than the percentage variation in their premia, differences in deductibles are salient. This tilts the consumer towards buying a low deductible policy, even though doing so is unjustified by the underlying risk.

Economists have tried several more standard approaches in accounting for some of the experimental evidence we discuss here. Wernerfelt (1995) and Kamenica (2008) explain the decoy effects by suggesting that decoys indirectly provide consumers with information about the quality of the products. The standard analysis of sales is also information-theoretic; it focuses on intertemporal price discrimination and seller selection of customers depending on their willingness to wait (Varian 1980, Lazear 1986, Sobel 1984). The present model offers two advantages. First, it can account for a broad range of context-dependent choices in a unified framework based on attribute salience. Second, it can account for some evidence that we see as dumbfounding from the standard perspective, such as Thaler’s beer example.

This paper is not the first to propose a psychologically based account of context phenomena. Huber, Payne and Puto (1982) and Simonson (1989) introduced some of the most striking experimental evidence and Simonson and Tversky (1992) and Bodner and Pr-elec (1994) proposed theoretical explanations based on loss aversion. Other papers relate to context dependence more broadly: Spiegler (2011) reviews several models where bound-edly rational consumers exhibit context dependent preferences (such as default bias), and embeds them in standard market settings. Koszegi and Rabin (2006) explore a model of reference-dependent preferences, and in particular how expectations influence willingness to pay. Heidhues and Koszegi (2008) propose a psychological model of sales based on loss aversion. Two papers most closely related to ours are Cunningham (2011) and Koszegi and Szeidl (2011); we discuss both of them after presenting the model.

## 2 The Model

### 2.1 Setup

A consumer evaluates all goods in an evoked set  $\mathbf{C} \equiv \{\mathbf{q}_k\}_{k=1,\dots,N}$  with  $N > 1$  goods. Each good  $k$  is a vector  $\mathbf{q}_k = (q_{1k}, \dots, q_{mk}) \in \mathbb{R}^m$  of  $m > 1$  quality attributes, where  $q_{ik}$  ( $i = 1, \dots, m$ ) measures the utility that attribute  $i$  generates for the consumer. The last attribute  $i = m$  stands for the price of good  $k$ , which gives the consumer a disutility  $q_{mk} = -p_k$ . The consumer has full information about the attributes of each good.<sup>3</sup> Most of the results in this paper are derived using the simplest setting where a good is identified by a single quality attribute and a price, namely  $\mathbf{q}_k = (q_k, -p_k)$ .

Absent salience distortions, a consumer values  $\mathbf{q}_k$  with a separable utility function:<sup>4</sup>

$$u(\mathbf{q}_k) = \sum_{i=1}^m \theta_i q_{ik}, \quad (1)$$

where  $\theta_i$  is the weight attached to attribute  $i$  in the valuation of the good ( $\theta_m$  is the weight attached to the numeraire and hence to the good's price).<sup>5</sup> We normalize  $\theta_1 + \dots + \theta_m = 1$ , which allows us to handle the relative utility weights of different attributes:  $\theta_i$  captures the importance of attribute  $i$  for the overall utility of the good (i.e., the strength/frequency with which a certain attribute is experienced during consumption), and  $\theta_i/\theta_j$  is the rational rate of substitution among attributes  $j$  and  $i$ .

A local thinker departs from (1) by inflating the relative weights attached to the attributes that he perceives to be more salient. As in BGS (2012), we say that attribute  $i$  is salient for good  $\mathbf{q}_t$  if the value of  $q_{it}$  "stands out" - relative to  $\mathbf{q}_t$ 's other attributes - with respect to the average level  $\bar{q}_i = \sum_k \frac{q_{ik}}{N}$  of the same attribute in  $\mathbf{C}$ . We think of  $\bar{q}_i$  as the *reference*

---

<sup>3</sup>If attribute  $i$  is a car's speed,  $q_{ik}$  measures the consumer's pleasure from driving fast in car  $k$ . By considering multiple quality attributes we develop a formalism that can be applied to models of product characteristics (e.g. Rosen 1974, Keeney and Raiffa 1976).

<sup>4</sup>Adopting additive representations of preferences is appropriate when attributes are independent in a specific sense (see Keeney and Raiffa (1976)). Additivity enables us to apply the formalism developed in BGS (2012), allowing for a stark characterization of the effects of salience. One could extend the formalism of salience to the case of non-additive preferences, but a fuller analysis is best left for future research.

<sup>5</sup>We have not included the income  $w$  of the consumer in the numeraire good (from which the consumer obtains total utility  $w - p_k$ ). This is because  $w$  is not an attribute of the good and thus its evaluation is not distorted by salience. The term  $\theta_m w$  is then just an additive constant in the evaluation of any good in  $\mathbf{C}$ .



level of attribute  $i$  in  $\mathbf{C}$ , and of  $\bar{\mathbf{q}} = \{\bar{q}_1, \dots, \bar{q}_N\}$  as the consumer's *reference good* (which may not be a member of  $\mathbf{C}$ ).

When the set  $\mathbf{C}$  is equal to the choice set, the reference good is simply the average good available to the consumer. We also consider the case in which  $\mathbf{C}$  includes goods that come to the consumer's mind in specific choice contexts. In the spirit of Kahneman and Miller (1986), this allows the reference good to be influenced also by the choice options that the consumer views as normal in a given context. To capture this idea, we study the case where the consumer thinks about the historical prices at which these goods he is currently choosing from were available in the past. Thus, the evoked set  $\mathbf{C}$  may include the choice set  $\mathbf{C}_{choice}$  as a subset. Many of our results are obtained when the choice and the evoked set coincide,  $\mathbf{C}_{choice} = \mathbf{C}$ , but historical prices play an important role in generating context-dependent willingness to pay and other anchoring-like effects. We return to the distinction between evoked and choice set in Section 2.2.

Given a reference good  $\bar{\mathbf{q}}$ , we formalize the salience of a good's attributes as follows.

**Definition 1** *The salience of attribute  $q_{it}$  for good  $\mathbf{q}_t$  is measured by a symmetric, continuous function  $\sigma(q_{it}, \bar{q}_i)$ , satisfying:*

1) *Ordering. For any  $(q, \bar{q})$  and  $(q', \bar{q}')$  such that  $[\min(q, \bar{q}), \max(q, \bar{q})] \subset [\min(q', \bar{q}'), \max(q', \bar{q}')]$ , we have:*

$$\sigma(q, \bar{q}) < \sigma(q', \bar{q}'). \quad (2)$$

2) *Diminishing sensitivity. For any  $q, \bar{q} > 0$  and all  $\epsilon > 0$ , we have:*

$$\sigma(q + \epsilon, \bar{q} + \epsilon) < \sigma(q, \bar{q}). \quad (3)$$

3) *Reflection. For any  $q, \bar{q}, q', \bar{q}' > 0$  we have:*

$$\sigma(q, \bar{q}) < \sigma(q', \bar{q}') \Leftrightarrow \sigma(-q, -\bar{q}) < \sigma(-q', -\bar{q}'). \quad (4)$$

To illustrate these three properties, consider the salience function employed in BGS

(2012), which sets:<sup>6</sup>

$$\sigma(q_{it}, \bar{q}_i) = \frac{|q_{it} - \bar{q}_i|}{|q_{it}| + |\bar{q}_i|}, \quad (5)$$

for  $|q_{it}|, |\bar{q}_i| \neq 0$ , and  $\sigma(0, 0) = 0$ .

According to ordering, salience increases in contrast: attribute  $i$  is more salient for good  $\mathbf{q}_t$  if  $q_{it}$  is farther from its reference level  $\bar{q}_i$  in the evoked set. An attribute is salient when it is very different from, or surprising relative to, its reference value. In (5), this is captured by the numerator  $|q_{it} - \bar{q}_i|$ . Diminishing sensitivity says that salience decreases as the value of an attribute uniformly increases in absolute value across all goods. In (5), this is captured by the denominator  $|q_{it}| + |\bar{q}_i|$ . Finally, reflection says that salience is shaped by the magnitude of attributes, so that negative attributes such as prices are treated similarly to positive attributes. In (5), reflection takes the strong form  $\sigma(q, \bar{q}) = \sigma(-q, -\bar{q})$ .

To see the intuition behind Definition 1, consider the salience of a good's price. Ordering implies that if good  $\mathbf{q}_t$  is more expensive than the reference good (i.e.  $p_t > \bar{p}$ ), an increase in its price  $p_t$  raises the extent to which the good's price is salient in the evoked set. Conversely, if good  $\mathbf{q}_t$  is cheaper than the reference good (i.e.  $p_t < \bar{p}$ ), an increase in  $p_t$  reduces the salience of the price for that good. On the other hand, diminishing sensitivity implies that if the prices of all goods rise, price becomes less salient for all goods. Intuitively, when the price level is high, price differences across goods are less noticeable.

Given a salience function  $\sigma$ , a local thinker ranks a good's attributes and distorts their utility weights as follows:

**Definition 2** *Attribute  $i$  is more salient than attribute  $j$  for good  $\mathbf{q}_t$  if and only if  $\sigma(q_{it}, \bar{q}_i) > \sigma(q_{jt}, \bar{q}_j)$ . Let  $r_{it}$  be the salience ranking of attribute  $i$  for good  $\mathbf{q}_t$ , where lower  $r_{it}$  corresponds to higher salience. Attributes with equal salience receive the same (lowest possible) ranking. The local thinker then evaluates good  $\mathbf{q}_t$  by transforming the weight  $\theta_i$  attached to attribute*

---

<sup>6</sup>In BGS (2012) the reference value of attribute  $i$  for good  $t$  was assumed to be the average level  $\bar{q}_{i,-t}$  of such attribute across all goods other than  $t$ . The current specification is slightly more tractable but yields the same results. Another specification is adopted in Bordalo (2011), whereby salience is the average pairwise contrast between good  $t$  and its alternatives  $r \neq t$ ,  $\sigma(q_{it}) = \left| \sum_{r \neq t} c(q_{it}, q_{ir}) \right|$ , where contrast  $c(\cdot, \cdot)$  satisfies properties 1 and 2 of Definition 1.

$i \in \{1, \dots, m\}$  into:

$$\hat{\theta}_i^t = \theta_i \cdot \frac{\delta^{r_{it}}}{\sum_j \theta_j \delta^{r_{jt}}} \equiv \theta_i \omega_i^t, \quad (6)$$

where  $\delta \in (0, 1]$ . Thus, the local thinker over-weights attribute  $i$  if and only if  $\omega_i^t > 1$ . The local thinker's (LT) evaluation of good  $\mathbf{q}_t$  is then given by:

$$u^{LT}(\mathbf{q}_t) = \sum_{i=1}^m \hat{\theta}_i^t \cdot q_{it}. \quad (7)$$

Relative to the rational case, the local thinker evaluates  $\mathbf{q}_t$  by over-weighting the utility impact of attribute  $i$  if the latter is more salient than average (i.e.  $\delta^{r_{it}} > \sum_j \theta_j \delta^{r_{jt}}$ ), and under-weighting it otherwise. Parameter  $\delta$  captures the degree of local thinking. As  $\delta \rightarrow 1$ , the local thinker converges to the rational case (i.e.  $\omega_i^t \rightarrow 1$ ). As  $\delta \rightarrow 0$ , the local thinker focuses only on the most salient attribute and neglects all others.

To see how the model works, return to the wine example from the Introduction. A consumer is evaluating two bottles of wine characterized by their known quality and price. Suppose that the consumer is considering a high end wine  $\mathbf{q}_h = (q_h, -p_h)$  and a low end wine  $\mathbf{q}_l = (q_l, -p_l)$ , where  $q_h > q_l$  and  $p_h > p_l$ . Since  $\mathbf{C} \equiv \{\mathbf{q}_h, \mathbf{q}_l\}$ , the reference wine has quality  $\bar{q} = (q_h + q_l)/2$  and price  $\bar{p} = (p_h + p_l)/2$ . Using the salience function (5), quality is salient for the high end wine  $\mathbf{q}_h$  if and only if  $\frac{q_h - (q_l + q_h)/2}{q_h + (q_l + q_h)/2} > \frac{p_h - (p_l + p_h)/2}{p_h + (p_l + p_h)/2}$ , namely when the deviation of wine  $\mathbf{q}_h$  from the average wine is larger, in percentage terms, along the quality than the price dimension. The quality  $q_h$  of the high end wine is thus salient when:

$$\frac{q_h}{p_h} > \frac{q_l}{p_l}, \quad (8)$$

namely, when the high end wine has a higher quality/price ratio than the low end wine. It is easy to see that when (8) holds, quality is salient for the low end wine as well. If instead the high end wine has a lower quality/price ratio than the low end wine (i.e.  $q_h/p_h < q_l/p_l$ ), then price is the salient attribute for both wines.

In this example: i) the same attribute (quality or price) is salient for both wines, and ii) the salient attribute is the relative advantage of the good with the highest  $q/p$ . As we show in Section 3, when the evoked set includes more than two options, different attributes

can be salient for different goods. This good-specific salience helps account for violations of Independence of Irrelevant Alternatives.

For a given salience ranking and utility weights  $\theta_1$  and  $\theta_2$  attached to quality and price, respectively, Definition 2 implies that the consumer's valuation of wine  $k = h, l$  is given by:

$$u^{LT}(\mathbf{q}_k) = \begin{cases} \theta_1 \cdot \left(\frac{1}{\theta_1 + \delta\theta_2}\right) \cdot q_k - \theta_2 \cdot \left(\frac{\delta}{\theta_1 + \delta\theta_2}\right) \cdot p_k & \text{if } q_h/p_h > q_l/p_l \\ \theta_1 \cdot \left(\frac{\delta}{\delta\theta_1 + \theta_2}\right) \cdot q_k - \theta_2 \cdot \left(\frac{1}{\delta\theta_1 + \theta_2}\right) \cdot p_k & \text{if } q_h/p_h < q_l/p_l \\ \theta_1 \cdot q_k - \theta_2 \cdot p_k & \text{if } q_h/p_h = q_l/p_l \end{cases} \quad (9)$$

If quality is salient, the relative weight of quality increases,  $\hat{\theta}_1^k = \theta_1 \cdot \left(\frac{1}{\theta_1 + \delta\theta_2}\right) > \theta_1$ , and the relative weight of price decreases,  $\hat{\theta}_2^k = \theta_2 \cdot \left(\frac{\delta}{\theta_1 + \delta\theta_2}\right) < \theta_2$ , as compared to the rational consumer's evaluation. If in contrast price is salient, its relative weight increases at the expense of that of quality. Thus, the consumer's evaluation of any wine  $k$  increases relative to the rational benchmark,  $u^{LT}(\mathbf{q}_k) > u(\mathbf{q}_k)$ , when its quality is salient, and decreases when its price is salient, in which case  $u^{LT}(\mathbf{q}_k) < u(\mathbf{q}_k)$ .

Through its impact on evaluation, salience affects the choice among wines. When prices are salient, namely when  $q_h/p_h < q_l/p_l$ , Expression (9) implies that the low end wine  $\mathbf{q}_l$  is chosen over the high end wine  $\mathbf{q}_h$  provided:

$$\delta\theta_1 \cdot (q_l - q_h) - \theta_2 \cdot (p_l - p_h) > 0, \quad (10)$$

which is easier to meet than its rational counterpart, with  $\delta = 1$ . Intuitively, when price is salient, the local thinker undervalues both wines, but he undervalues the high end wine more; this is because the local thinker focuses on the dimension, price, along which the low end wine does better.

Analogously, when quality is salient, namely when  $q_h/p_h > q_l/p_l$ , Expression (9) implies that the low end wine  $\mathbf{q}_l$  is chosen over the high end wine  $\mathbf{q}_h$  provided:

$$\theta_1 \cdot (q_l - q_h) - \delta\theta_2 \cdot (p_l - p_h) > 0, \quad (11)$$

which is harder to meet than its rational counterpart, with  $\delta = 1$ . Intuitively, when quality

is salient, the local thinker overvalues both wines, but overvalues the high quality wine more. Thus, he is less likely to choose the low end wine than in the rational case.

Saliency tilts the local thinker’s preferences toward the wine offering the highest quality/price ratio. When the high end wine has the highest quality/price ratio, the consumer focuses on quality and is more likely to choose  $\mathbf{q}_h$ . When the low end wine has the highest quality/price ratio, the consumer focuses on price and is more likely to pick  $\mathbf{q}_l$ . In marketing and psychology, it has long been recognized that consumers are drawn to goods with a high quality/price ratio (or value per dollar). This notion has been explained by assuming that the consumer experiences a distinct “transaction utility” (Thaler 1999), in that he derives direct pleasure from making a good deal (Jahedi 2011). In our example, the consumer does not derive any special utility from making good deals. Instead, the quality/price ratio affects choice by determining whether a good’s relative advantage is salient.

The quality/price ratio in (9) creates two forms of context dependence in our model. The first one concerns the consumer’s sensitivity to changes in a good’s attributes. For instance, an increase in  $q_h$  always increases the valuation of the high end wine, but the effect is particularly strong when  $q_h$  becomes so high relative to  $q_l$  that quality becomes salient for wine  $\mathbf{q}_h$ . The second form of context dependence is that the evaluation of a good depends on the alternatives of comparison. For instance, a reduction in the quality  $q_l$  of the low end wine can boost the valuation of the high end wine  $\mathbf{q}_h$  by rendering the latter’s quality salient.

Given the intuitive appeal of the quality/price ratio, we now consider the class of saliency functions under which  $q/p$  is the critical driver of saliency. Take an evoked set  $\mathbf{C}$  consisting of  $N > 1$  goods characterized by their quality and price [i.e.,  $\mathbf{q}_k = (q_k, -p_k)$ ] and by a reference good  $\bar{\mathbf{q}} = (\bar{q}, -\bar{p})$ . We find:

**Proposition 1** *Let  $\mathbf{q}_k$  be a good that neither dominates nor is dominated by the reference good  $\bar{\mathbf{q}}$ . The following two statements are then equivalent:*

- 1) *The advantage of  $\mathbf{q}_k$  relative to  $\bar{\mathbf{q}}$  is salient if and only if  $q_k/p_k > \bar{q}/\bar{p}$ .*
- 2) *Saliency is homogeneous of degree zero, i.e.  $\sigma(\alpha x, \alpha y) = \sigma(x, y)$  for all  $\alpha > 0$ .*

When the saliency function is homogenous of degree zero, a good’s advantage relative to the reference is salient provided the good has a favourable quality/price ratio. To see

this, suppose that  $\mathbf{q}_k$  has higher quality and price than average, namely  $q_k > \bar{q}$ ,  $p_k > \bar{p}$ . Then, its advantage relative to the reference good is quality  $q_k$ . This quality is salient provided  $\sigma(q_k, \bar{q}) > \sigma(p_k, \bar{p})$ . Under homogeneity of degree zero this condition is equivalent to  $\sigma(q_k/\bar{q}, 1) > \sigma(p_k/\bar{p}, 1)$ . By ordering, this is met precisely when  $\mathbf{q}_k$  has a higher quality/price ratio than average,  $q_k/p_k > \bar{q}/\bar{p}$ . Conversely, if  $\mathbf{q}_k$  has lower quality and price than average -  $q_k < \bar{q}$ ,  $p_k < \bar{p}$  - its advantage relative to the reference good is price  $p_k$ . This price is then salient provided  $\sigma(p_k, \bar{p}) > \sigma(q_k, \bar{q})$ , which occurs precisely when  $\mathbf{q}_k$  has above average quality/price ratio.

Homogeneity of degree zero is a reasonable property, as it ensures that the salience ranking is scale-invariant, in the sense that it is invariant under linear transformations of the units (utils) in which the attributes are measured. Interestingly, homogeneity of degree zero is related to diminishing sensitivity. The Appendix in fact proves:

**Lemma 1** *If  $\sigma(\cdot, \cdot)$  satisfies the ordering property for positive attribute values, and is homogenous of degree zero, then it also satisfies diminishing sensitivity.*

Although our basic results hold under Definition 1, summarizing salience by a good's quality to price ratio aids both tractability and psychological intuition. In the remainder, we therefore restrict our attention to the case where the following assumption holds:

A.0: The salience function satisfies ordering, reflection and homogeneity of degree zero.

In section 2.2 we provide a psychological justification for this assumption.<sup>7</sup> In light of A.0, we can fully characterize the salience ranking of any good  $\mathbf{q}_k = (q_k, -p_k)$  in the quality price space, including in regions where it either dominates or is dominated by the reference good  $\bar{\mathbf{q}} = (\bar{q}, -\bar{p})$ . The resulting salience rankings are graphically represented in Figure 1 below. Note that there is a trade-off between good  $\mathbf{q}_k$  and the reference good  $\bar{\mathbf{q}}$  in quadrants I ( $q_k < \bar{q}$ ,  $p_k < \bar{p}$ ) and II ( $q_k > \bar{q}$ ,  $p_k > \bar{p}$ ), whereas  $\mathbf{q}_k$  dominates  $\bar{\mathbf{q}}$  in quadrant IV and is dominated by  $\bar{\mathbf{q}}$  in quadrant III.

---

<sup>7</sup>To extend the homogeneity of degree zero property to attribute levels of zero, we interpret  $\sigma(q_{ik}, 0)$  as  $\lim_{\underline{q}_i \rightarrow 0} \sigma(q_{ik}, \underline{q}_i)$ . Moreover, when comparing  $\sigma(q_{ik}, 0)$  and  $\sigma(q_{jk}, 0)$ , we assume the limit then keeps the ratio of hedonic utilities  $\underline{a}_i/\underline{a}_j$  constant at 1. Homogeneity of degree zero is stronger than diminishing sensitivity, as is exemplified by the salience function  $\sigma(x, y) = \frac{|x-y|}{x+y+\theta}$ , with  $\theta > 0$ . In this case  $\sigma(\alpha x, \alpha y) > \sigma(x, y)$  for  $\alpha > 1$ . Thus homogeneity excludes certain weak forms of diminishing sensitivity.

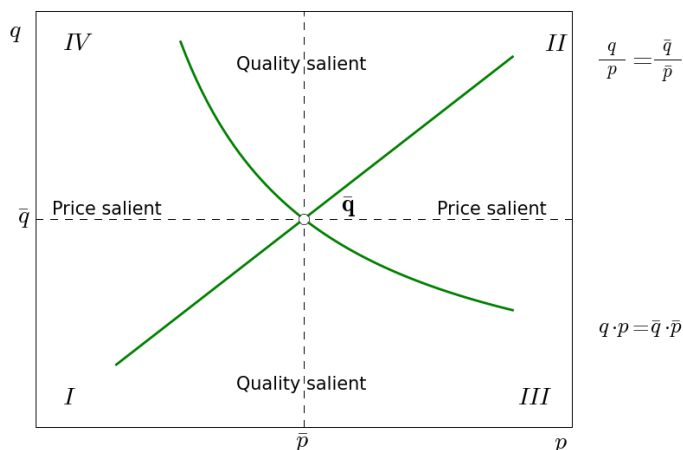


Figure 1: Saliency of attributes of  $\mathbf{q}_k = (q, -p)$  depends on its location relative to  $\bar{\mathbf{q}} = (\bar{q}, -\bar{p})$ .

From the previous discussion, in quadrants I and II the saliency ranking of a good is determined by its location relative to the upward sloping curve  $q/p = \bar{q}/\bar{p}$ , along which the good’s quality/price ratio is equal to that of the reference good. This determines, together with the downward sloping curve  $q \cdot p = \bar{q} \cdot \bar{p}$  in the quadrants III and IV, four regions where either price or quality is salient.<sup>8</sup> To jointly characterize the saliency ranking of all goods in an evoked set  $\mathbf{C}$  we simply need to compute the reference quality and price, and then place the goods in the “windmill” diagram of Figure 1 above. In this diagram, a good’s price  $p_k$  is salient in regions where it is far from the reference price  $\bar{p}$ . Accordingly, the good’s quality  $q_k$  is salient in the regions where it is far from the reference quality  $\bar{q}$ . Figure 1 allows us to develop visual intuitions for the role of saliency in explaining choices.

## 2.2 Discussion of Setup and Assumptions

Our model of context-dependent evaluation hinges on two basic facts about perception: i) our perceptive apparatus is structured to detect changes in stimuli (captured by the ordering property), and ii) changes are better detected when they occur close to a baseline reference level (captured by the diminishing sensitivity property). BGS (2012) provide a fuller de-

<sup>8</sup>To identify the downward sloping curve, note that when  $\mathbf{q}_k$  dominates the reference (i.e.  $q_k > \bar{q}$  and  $p_k < \bar{p}$ ), then  $q_k$  is salient if and only if  $\sigma(q_k/\bar{q}, 1) > \sigma(1, \bar{p}/p_k)$ , namely if and only if  $q_k p_k > \bar{q} \bar{p}$ . Instead, when  $\mathbf{q}_k$  is dominated by the reference, its quality is salient if and only if  $q_k p_k < \bar{q} \bar{p}$ .

scription of these psychological phenomena. That paper introduces a general framework of choice that combines ordering, diminishing sensitivity and reflection and applies it to choices under risk. Here we show how the same assumptions shed light on a wide variety of choice patterns and puzzles in a riskless setting.

Three assumptions of the model merit further comment. The first is homogeneity of degree zero (A.0) of salience, which plays a larger role in the present paper than in BGS (2012). The crucial role of this assumption is to pin down the trade-off between diminishing sensitivity and ordering, in a way closely related to Weber’s law: the salience of an attribute for a good remains constant when the level of that attribute increases in all goods, provided the difference between the good’s level and the reference level increases proportionally. While we do not claim that this assumption is universally applicable, it is supported by an emerging paradigm in psychology stressing that people possess an innate “core number system” which compares magnitudes in terms of ratios.<sup>9</sup>

The key predictions of our model are shaped by diminishing sensitivity and ordering. These properties determine the effect of changing the level of an attribute on its salience. For instance, ordering implies that increasing the price of a good increases the salience of its price, provided that price is above average, while diminishing sensitivity implies that price differences become less salient as the price level increases. These predictions hold for any increasing utility function, and can be tested experimentally. When ordering and diminishing sensitivity are in conflict, as when both price levels and price dispersion increase, homogeneity of degree zero pins down the *relative* importance of each force. It thus allows to make precise predictions on the effect of context on the consumer’s choices, such as the role of the quality to price ratio. These predictions do depend on the consumer’s utility function.

Second, we have assumed that evaluation depends on the attributes’ salience *ranking*. This rank-based discounting aids tractability, but has some shortcomings: i) evaluation is discontinuous at those attribute values where salience ranking changes, and ii) evaluation may be non-monotonic. Non-monotonicity may even lead, in finely tuned examples, to a

---

<sup>9</sup>Feigenson, Dehaene and Spelke (2004): “To sum up, the findings indicate that infants, children and adults share a common system for quantification.” This system exhibits a logarithmic (i.e. ratio based) representation of numerical magnitude: “numerical representations therefore show two hallmarks: they are ratio-dependent and are robust across multiple modalities of input.” Interestingly, the “system becomes integrated with the symbolic number system used by children and adults for enumeration and computation.”



dominated good being preferred over a dominating good. In the Appendix we show that with a continuous salience weighting these shortcomings disappear under general conditions. In the main text, we however stick to the more tractable rank-based discounting.

Third and most important, following Kahneman and Miller (1986), we have defined the reference good as having each attribute’s average level in the evoked set. Since our goal is to study how salience shapes the effect of context, we limit our analysis by taking the evoked set as given. Several authors have recently proposed models that endogenize the set of options that come to the decision maker’s mind, as distinct from the choice set (Rubinstein and Salant 2007, Eliaz and Spiegler 2010, Masatlioglu et al. 2010, Manzini and Mariotti 2010). These models focus on the “consideration set” as it is understood in the marketing literature, namely a typically small subset of all available options that the agent actually considers when making a choice.<sup>10</sup> In contrast, in the examples and applications in this paper, the choice set is small and the evoked set can include other options that are not in effect available (such as historical prices for goods in the choice set).

Several models of consumer choice incorporate loss aversion relative to a reference good, including Tversky and Kahneman (1991), Tversky and Simonson (1992) and Bodner and Pr-elec (1994). A main implication of these models is a bias towards middle-of-the-road options, which avoid large perceived losses in every attribute. This prediction is hard to reconcile with evidence that in many situations consumers do choose extreme options. Moreover, these models do not speak to the other puzzles reviewed in the Introduction, such as the Savage car radio problem, context dependent WTP or the Hastings-Shapiro data.

Recently, several papers have introduced other related models of context dependent evaluation. The literature on relative thinking assumes that valuation of a good depends on the “referent” levels of its characteristics (Azar 2007, Cunningham 2011). The fundamental assumption is that the marginal utility of a characteristic decreases with the level of its referent. This is reminiscent of the diminishing sensitivity property of salience, and in fact Cunningham (2011) reproduces some related patterns of choice, such as the Savage car radio

---

<sup>10</sup>The determination of the choice set is also an important input in (rational) discrete choice models: the predictions of these models depend quantitatively on how the set of alternatives is specified. Moreover, allowing for incomplete consumer information (Goeree 2008) suggests an important role for (un)awareness of available choices.

puzzle. By assuming that valuation changes are driven solely by diminishing sensitivity, Cunningham’s approach implies that all goods’ valuations are distorted in the same way. Thus, it does not account for patterns of choice in which ordering plays a role, such as the taste for balance (section 3.4) or the Hastings-Shapiro evidence on gasoline (section 4.1).

Koszegi and Szeidl (2011) build a model that centrally features the idea of ordering: their consumers are essentially local thinkers who focus on and overweigh those attributes in which options differ the most in terms of utility. Koszegi and Szeidl then use their model to shed light on biases in intertemporal choice. By neglecting diminishing sensitivity, the Koszegi Szeidl model predicts a strong bias towards concentration, namely consumers tend to overvalue options whose advantages are concentrated in a single dimension. This bias seems difficult to reconcile with the evidence on diminishing sensitivity (such as the Savage car radio puzzle), and also with the evident desire of luxury manufacturers to avoid shortcomings in any aspect of their merchandise.

By combining diminishing sensitivity with ordering within the context of an evoked set, our model provides a unified account of several well-known choice patterns and puzzles. It reconciles patterns explored separately by Cunningham (2011) and Koszegi and Szeidl (2011), sheds light on phenomena currently gathered under the banner of mental accounting (such as context dependent willingness to pay), and generates new predictions of interest in economic applications.

### **3 Salience and Choice**

We now examine various implications of our model, motivated by the evidence summarized in the introduction. Section 3.1 considers context effects that occur due to a uniform increase in the level of one attribute (price) across all goods. Section 3.2 investigates context effects that occur when new goods are added to the evoked set  $\mathbf{C}$ . Section 3.3 studies a taste for balance in goods having two positive quality attributes. Finally, Section 3.4 applies these results to examine how historical prices affect the local thinker’s willingness to pay for quality. This is the only case in Section 3 where the choice and the evoked sets differ.

### 3.1 Buying Wine in a Store vs. at a Restaurant

Suppose that a consumer buying wine attaches the same weight to quality and price ( $\theta_1 = \theta_2 = 1/2$ ). In the wine store, the available wines are:

$$\mathbf{C}_{store} = \begin{cases} \mathbf{q}_h = (30, -\$20) \\ \mathbf{q}_l = (20, -\$10) \end{cases} . \quad (12)$$

The rational consumer is indifferent between  $\mathbf{q}_h$  and  $\mathbf{q}_l$  because  $u(\mathbf{q}_h) = 30 - 20 = u(\mathbf{q}_l) = 20 - 10$ . This is not true for the local thinker. Since the quality/price ratio of the low end wine is higher than that of the high end wine (i.e.  $20/\$10 > 30/\$20$ ), price is salient for both wines. It follows from (10) that the high end wine is undervalued relative to the low end wine, so the local thinker strictly prefers  $\mathbf{q}_l$  to  $\mathbf{q}_h$ . In the wine store, price is more salient than quality, so the local thinker is overly sensitive to price differences. He perceives  $\mathbf{q}_l$  to be slightly less good, but a lot cheaper than  $\mathbf{q}_h$ .

Suppose now that the same two wines are offered at a restaurant, with uniformly higher prices:

$$\mathbf{C}_{restaurant} = \begin{cases} \mathbf{q}_h = (30, -\$60) \\ \mathbf{q}_l = (20, -\$50) \end{cases} . \quad (13)$$

The rational consumer is again indifferent between  $q_h$  and  $q_l$ , because  $u(q_h) = 30 - 60 = u(q_l) = 20 - 50$ . Unlike in the store, however,  $\mathbf{q}_h$  now provides a better quality to price ratio than  $\mathbf{q}_l$ , since  $30/\$60 > 20/\$50$ . As a consequence, in the restaurant the consumer focuses on quality and, from (11), the high end wine is chosen over the alternative. At the restaurant the local thinker is less sensitive to price differences and perceives  $q_h$  to be slightly more expensive but significantly better than  $\mathbf{q}_l$ . This occurs even though the quality gradient  $q_h - q_l$  and the price gradient  $p_h - p_l$  are the same in the store and at the restaurant, so that the rational consumer does not systematically change his choice between the two contexts.

Context influences decisions here because the ranking of the quality to price ratio changes from the store to the restaurant. The store displays a higher percentage variation along the price dimension than along the quality dimension, which implies that the cheaper good is the better deal. The reverse is true at the restaurant. As a consequence, the consumer focuses

on price in the store and on quality in the restaurant.

These effects, arising from the diminishing sensitivity of the salience function, naturally deliver a well known feature of consumer behavior: lower price sensitivity for choice among more expensive goods. An example of this phenomenon is Savage’s (1954) car radio problem<sup>11</sup>, in which a consumer is more likely to buy a car radio when the price of the radio is added to the price of the car than when the radio is sold in isolation, after the car purchase. To see this, denote by  $q$  the car’s quality and by  $q + q_r$  its quality when the radio is installed. Denote by  $p$  the car’s price and by  $p_r$  the price of the radio. When choosing whether to buy the car alone or with the radio, the consumer faces  $\mathbf{C}_{bundle} \equiv \{(q, p), (q + q_r, p + p_r)\}$ . The salience of quality for the car with the radio is  $\sigma(q + q_r, q + q_r/2)$ , the salience of its price is  $\sigma(p + p_r, p + p_r/2)$ . When instead the consumer chooses whether to keep his car without the radio or to install a radio in it, he faces  $\mathbf{C}_{isol} \equiv \{(q, 0), (q + q_r, p_r)\}$ . The salience of quality for the car with the radio is still  $\sigma(q + q_r, q + q_r/2)$  while the salience of its price is  $\sigma(p_r, p_r/2)$ . By diminishing sensitivity  $\sigma(p + p_r, p + p_r/2) < \sigma(p_r, p_r/2)$ , so the price of the radio is more salient when the radio is bought in isolation. It is easy to check that this analysis is confirmed by the  $q/p$  logic under assumption A0.

Similarly, our model sheds light on the jacket and calculator problem (Kahneman and Tversky 1984), in which subjects who have decided to buy a bundle ((jacket, \$125), (calculator, \$15)) are willing to travel 10 minutes to save \$5 when the discount applies to the calculator, but not to the more expensive jacket. Intuitively, walking for 10 minutes (vs. not walking at all) has salience  $\sigma(10, 5)$ . Saving 5 dollars on the jacket has salience  $\sigma(120, 122.5)$ , saving them on the calculator has salience  $\sigma(10, 12.5)$ . Since  $\sigma(10, 12.5) > \sigma(120, 122.5)$ , the discount is more likely to be salient if it is applied to the calculator.

These results generalize to choice among an arbitrary number of goods. To see this, suppose that the local thinker is choosing between  $N > 1$  goods located along a rational indifference curve. The indifference condition allows us to identify the effect of salience, abstracting from rational utility differences. Given the quasilinear utility in (1), the  $N$  goods display a constant quality/price gradient, formally  $\theta_1(q_k - q_{k'}) = \theta_2(p_k - p_{k'})$  for all

---

<sup>11</sup>This problem was proposed as a riskless choice counterpart to Allais’ paradox in decision making under risk, illustrating the breakdown of the sure thing principle. Salience accounts for both versions of the problem, see BGS (2012).

$k, k' = 1, \dots, N$ . Assume, without loss of generality, that quality and price increase in the index  $k$  (i.e.  $q_1 < \dots < q_N$  and  $p_1 < \dots < p_N$ ). In the Appendix we prove:

**Proposition 2** *Along a rational linear indifference curve, the local thinker chooses the good with the highest quality/price ratio. In particular:*

- 1) *if  $q_1/p_1 > \theta_2/\theta_1$ , the cheapest good  $(q_1, p_1)$  has the highest  $q/p$  ratio and is chosen;*
- 2) *if  $q_1/p_1 < \theta_2/\theta_1$ , the most expensive good  $(q_N, p_N)$  has the highest  $q/p$  ratio and is chosen;*
- 3) *if  $q_1/p_1 = \theta_2/\theta_1$ , the  $q/p$  ratio is constant and the consumer is indifferent between the goods.*

Saliency tilts the rational linear indifference curves, favoring either the cheapest or the highest quality good. Diminishing sensitivity determines which good is chosen. When, as in case 1), the price level is low relative to the quality level, variation along the price dimension is more salient than that along the quality dimension. As a consequence, the consumer focuses on prices, breaking indifference in favour of the cheapest good. When, as in case 2), the price level is high relative to the quality level, the consumer attends more to quality differences. As a result, he breaks indifference in favour of the highest quality good. In both cases the consumer prefers the good with the highest quality to price ratio, which is either the cheapest or the highest quality good in the choice set.<sup>12</sup>

To visualize Proposition 2, note that with linear utility a rational indifference curve is a positively sloped line in the  $(q, p)$  diagram. If the evoked set consists of a collection of points on an indifference line, then the reference good  $(\bar{q}, \bar{p})$  also lies on that line. Exploiting these features, Figure 2 graphically represents cases 1) and 2) of Proposition 2.

As in the case of the wine store, in the left panel goods vary more along the price than along the quality dimension: price is salient and consumers choose the cheapest good. The reverse holds in the right panel.<sup>13</sup> The shift is saliency ranking from the left to the right

---

<sup>12</sup>The linearity of rational indifference curves (which is due to the quasi linearity of preferences) is useful to obtain such a sharp characterization. For a concave indifference curve, the reference good will lie below the rational indifference curve itself, and so saliency rankings will differ across goods. As we show below, concave evoked sets generate decoy effects.

<sup>13</sup>The local thinker's tendency to choose extreme goods in the choice set generalizes to any evoked set  $\mathbf{C}$  lying on a positively sloped line, even if this line is not a rational indifference curve. Also in this case all goods will have the same saliency ranking, and the good taking the most favourable value of the salient attribute will thus be maximally overvalued (even if it is not necessarily chosen).

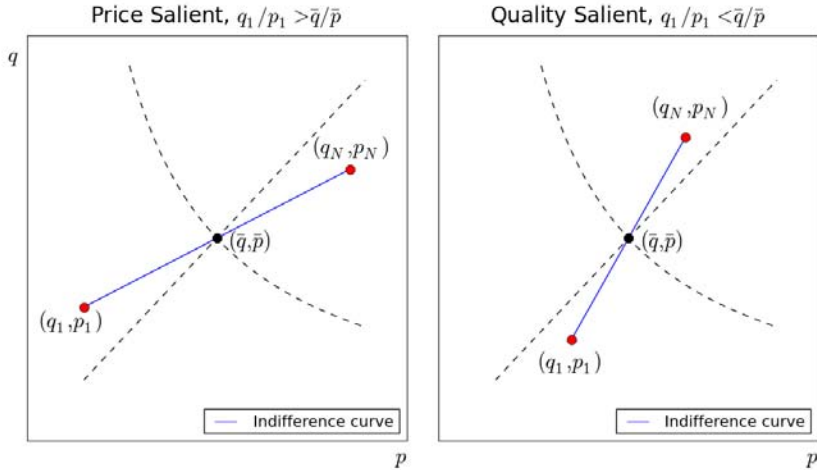


Figure 2: All goods on an indifference curve have the same salience ranking.

panel thus captures changes in context generated by a uniform change in one of the goods' attributes (e.g. price). In Section 4.1 we consider the more general effects of uniform price movements within a sub-category of goods.

### 3.2 Decoy Effects and Violations of IIA

There is ample experimental evidence that manipulation of the choice set alters the preference among existing goods, in violation of independence of irrelevant alternatives (IIA). A well documented anomaly in both marketing and psychology is the so called decoy effect (Huber, Payne and Puto 1983, Tversky and Simonson 1993), in which adding an option dominated by one of two goods boosts the demand for the dominating good. Another well known anomaly is the compromise effect (Simonson, 1989), whereby adding an extreme option to a pairwise choice induces subjects to change their preferences toward the middle of the road, or compromise, option. We now show how our model can account for these phenomena as a result of the impact of the added option on salience.

Consider again the wine example in (12), with a variation in which a third, more expensive

and higher quality wine  $\mathbf{q}_d$  is added to the wine list (and  $\theta_1 = \theta_2$ )

$$\mathbf{C}_0 = \begin{cases} \mathbf{q}_h = (30, -\$20) \\ \mathbf{q}_l = (20, -\$10) \end{cases} \quad \mathbf{C}_{decoy} = \begin{cases} \mathbf{q}_d = (30, -\$30) \\ \mathbf{q}_h = (30, -\$20) \\ \mathbf{q}_l = (20, -\$10) \end{cases} \quad (14)$$

Wine  $\mathbf{q}_d$  is dominated by  $\mathbf{q}_h$ , yielding lower utility than the original options,  $u(\mathbf{q}_d) = 0 < u(\mathbf{q}_h) = u(\mathbf{q}_l) = 10$ . A rational decision maker is indifferent between  $\mathbf{q}_h$  and  $\mathbf{q}_l$  but prefers both to  $\mathbf{q}_d$ . The inclusion of  $\mathbf{q}_d$  in the evoked set does not affect his choice.

As shown in Section 3.1, in  $\mathbf{C}_0$  the local thinker picks the low end wine  $\mathbf{q}_l$  because it has the highest quality/price ratio, so prices are salient. What happens when  $\mathbf{q}_d$  is added to the list? The new wine delivers the highest quality in the choice set, but is much more expensive than the other wines. In particular, the quality/price ratio of  $\mathbf{q}_d$ ,  $30/30$ , is lower than the quality/price ratio of the high end wine  $\mathbf{q}_h$ ,  $30/20$ . Now, by comparison with  $\mathbf{q}_d$ , the high end wine seems a better deal than in the original choice set  $\mathbf{C}_0$ .

To see the implications for choice, note that in the set  $C_{decoy}$ , the reference wine is  $\bar{\mathbf{q}} = (26.7, -\$20)$ . The high end wine  $\mathbf{q}_h$  delivers above reference quality  $30 > 26.7$  at the reference price  $\$20$ . Intuitively, the quality of  $\mathbf{q}_h$  becomes salient. The low end wine still dominates the reference wine along the price dimension, since  $\$10 < \$20$ , and this dimension remains salient because  $\mathbf{q}_l$  is a better deal than  $\bar{\mathbf{q}}$ , formally  $20/10 > 26.7/20$ . As a consequence, after the decoy is added, the low end wine remains price salient but the high end wine becomes quality salient. Under this new salience configuration, the local thinker prefers  $\mathbf{q}_h$  to  $\mathbf{q}_l$ . Our model therefore yields a decoy effect: in pairwise choice the local thinker prefers  $\mathbf{q}_l$  to  $\mathbf{q}_h$  but he switches to  $\mathbf{q}_h$  when an expensive inferior good  $\mathbf{q}_d$  is added, thus violating IIA.<sup>14</sup> The intuition is that when the bad deal  $\mathbf{q}_d$  is added, the quality of  $\mathbf{q}_h$  becomes salient, and  $\mathbf{q}_h$  becomes a good deal.

This argument does not rely on introducing a decoy  $\mathbf{q}_d$  which is dominated by the originally neglected option  $\mathbf{q}_h$ . It relies on the introduction in the choice set of an option which highlights the quality dimension of  $\mathbf{q}_h$  while not being so attractive that it is itself chosen. Take two goods  $\mathbf{q}_l = (q_l, p_l)$ ,  $\mathbf{q}_h = (q_h, p_h)$ , such that  $\mathbf{q}_h$  is chosen if and only if it is quality

<sup>14</sup>As  $\mathbf{q}_d$  lies on a lower indifference curve, and  $\mathbf{q}_h$  is quality salient,  $\mathbf{q}_d$  is never chosen.

is salient. Denoting by  $\Delta u = [q_h - q_l] - [p_h - p_l]$  the rational utility difference between them (with  $\theta_1 = \theta_2$ ), this means

$$-(1 - \delta)[p_h - p_l] \leq \Delta u \leq (1 - \delta)[q_h - q_l] \quad (15)$$

This condition says that preference reversals occur provided the rational utility difference between the goods is sufficiently small: only in this case a change in salience can affect choice among the two goods. We restrict our attention to decoy options  $\mathbf{q}_d$  such that  $\bar{q} \leq q_h$  and  $\bar{p} \leq p_h$ , where  $(\bar{q}, \bar{p})$  is the reference good in  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ . These constraints allow for goods  $\mathbf{q}_d$  which make  $\mathbf{q}_h$  an intermediate good in the enlarged choice set. The appendix then proves that, when Equation (15) holds, we have:

**Proposition 3**

*i) If  $\frac{q_l}{p_l} > \frac{q_h}{p_h}$ , so that price is salient and  $\mathbf{q}_l$  is chosen from  $\{\mathbf{q}_l, \mathbf{q}_h\}$ , then for any  $\mathbf{q}_d$  satisfying  $\frac{q_d}{p_d} < \frac{q_h}{p_h} + \frac{p_l}{p_d} \left[ \frac{q_h}{p_h} - \frac{q_l}{p_l} \right]$ , good  $\mathbf{q}_h$  is quality salient in  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ . Moreover, there exist options  $\mathbf{q}_d$  with  $q_d > q_h$  and  $p_d > p_h$  such that  $\mathbf{q}_h$  is chosen from  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ .*

*ii) If  $\frac{q_l}{p_l} < \frac{q_h}{p_h}$ , so quality is salient and  $\mathbf{q}_h$  is chosen from  $\{\mathbf{q}_l, \mathbf{q}_h\}$ , then there exist no decoy options  $\mathbf{q}_d$  such that  $\frac{q_d}{p_d} \leq \frac{q_h}{p_h}$  and  $\mathbf{q}_h$  is price salient in  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ . In particular, for no  $\mathbf{q}_d$  satisfying these properties is  $\mathbf{q}_l$  chosen from  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ .*

Consider first case *i*), where  $\mathbf{q}_l$  is a good deal when compared to  $\mathbf{q}_h$ , namely  $q_l/p_l > q_h/p_h$  (so that the price dimension is salient) and the consumer prefers  $\mathbf{q}_l$  over  $\mathbf{q}_h$  in a pairwise choice. Then the Proposition identifies a decoy  $\mathbf{q}_d$  sufficient to reverse this preference, namely, when  $\mathbf{q}_d$  has a low enough quality-price ratio. In particular, it must be that  $\frac{q_d}{p_d} < \frac{\bar{q}}{\bar{p}} < \frac{q_h}{p_h}$ , that is, the decoy must be a “bad deal”: the constraints on  $q_d$  imply that the decoy lowers the overall quality-price ratio in the choice set to the point that  $q_h/p_h > \bar{q}/\bar{p}$ . Since the decoy’s quality is low relative to its price, this makes the *quality* of  $\mathbf{q}_h$  salient. Then,  $\mathbf{q}_h$  is chosen as long as the decoy is not too attractive.

The decoy effect is strongest when the new option  $\mathbf{q}_d$  is dominated by  $\mathbf{q}_h$ , with the same or lower quality but a much higher price. This is the case in the example (14). However, preference reversals can also occur when the added option  $\mathbf{q}_d$  is not dominated by  $\mathbf{q}_h$ , including when  $q_d > q_h$  and  $p_d > p_h$ . In this case,  $\mathbf{q}_h$  is perceived as providing intermediate



levels of quality and price. As long as  $\mathbf{q}_d$  provides a relatively larger increase in price than in quality compared to  $\mathbf{q}_h$ , the consumer focuses on the quality of  $\mathbf{q}_h$  and is more likely to choose it. This case provides a rationale for the compromise effect, which in our model arises due to a similar mechanism as the decoy effect.

Figure 3 provides a graphical intuition for the decoy/compromise effect of case *i*). When the new good  $\mathbf{q}_d$  has a sufficiently lower  $q/p$  ratio than existing options, the evoked set becomes concave with respect to prices. As a result, the intermediate good has both higher quality and higher quality/price ratios than the reference good, becoming quality salient.<sup>15</sup>

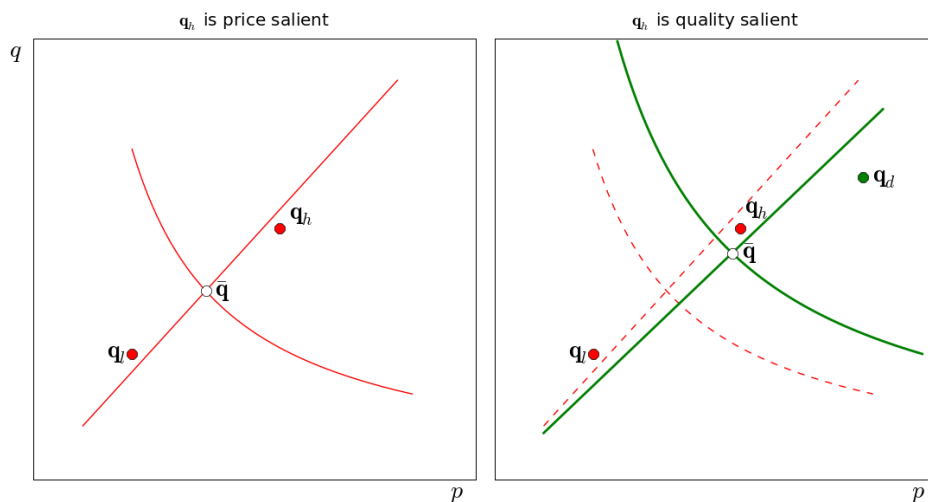


Figure 3: Adding a decoy changes the quality/price ratio of the reference good.

Consider case *ii*) of Proposition 3. Now  $\mathbf{q}_h$ 's quality is already salient in the pairwise comparison with  $\mathbf{q}_l$ . Adding a decoy to the lower quality good  $\mathbf{q}_l$ , namely a bad deal  $\mathbf{q}_d$  with relatively low quality to price ratio (as implied by the condition  $q_d/p_d < q_h/p_h$ ), has no effect on  $\mathbf{q}_h$ 's salience ranking: in fact,  $\mathbf{q}_h$  remains a high quality, high quality-price ratio good, so its quality remains salient. A striking implication is that in this case there is no decoy option that boosts the relative evaluation of the lower quality good  $\mathbf{q}_l$ , even for decoys such that  $\mathbf{q}_l$  is a dominating option ( $q_d < q_l, p_d > p_l$ ) or a compromise option ( $q_d < q_l, p_d < p_l$ ).

<sup>15</sup>In typical illustrations of the compromise effect, the three goods lie on a straight line in attribute space, with the intermediate good equidistant from the other two (Tversky and Simonson, 1993). If utility is concave, this arrangement translates into a concave choice set as in Figure 3.

There are instances, not contemplated in Proposition 3, in which a decoy might increase the relative evaluation of a lower quality good.<sup>16</sup> However, Proposition 3 captures an important asymmetry generated by our model, whereby goods with high quality and high price are more likely to benefit from decoy effects than their low quality, low price competitors. This effect is different from loss aversion (Tversky and Simonson 1993, Bodner and Prelec 1994) in that consumers do not mechanically prefer middle-of-the-road options. It is, however, consistent with Heath and Chatterjee (1995)’s survey of experimental results on decoy effects. The authors find that adding appropriate decoys typically boosts experimental subjects’ demand for high quality goods, but rarely for low quality goods. We explore this asymmetry further in the context of sales, Section 4.2.

### 3.3 Goods with Multiple Positive Quality Attributes

Having examined the tradeoff between quality and price, we now consider the trade-off between two quality dimensions. Several experiments document subjects’ tendency to select options that offer a more balanced combination of positive qualities in the choice set, in accordance with the compromise effect. We now show that this taste for balance arises naturally in our model due to diminishing sensitivity: for unbalanced goods, the salient attributes are their shortcomings rather than their strengths. We also show that this mechanism is richer than standard loss aversion accounts and yields novel predictions.

To do so, we consider goods  $\mathbf{q}_k \equiv (q_{1k}, q_{2k}, p)$  that differ in their qualities but not in their prices, so that price is the least salient dimension. We omit the price for notational convenience. In this setup, Definition 1 implies that  $q_{1k}$  is more salient than  $q_{2k}$  for good  $\mathbf{q}_k$  if and only if  $\sigma(q_{1k}, \bar{q}_1) > \sigma(q_{2k}, \bar{q}_2)$ . Once more, the salience ranking of a good in quality-quality space is determined by its location relative to the reference  $(\bar{q}_1, \bar{q}_2)$ . Good  $\mathbf{q}_k$  presents a trade-off relative to the reference good whenever it has a higher level of one quality but a lower level of the other, namely it lies in quadrants III and IV of the left panel of Figure 1.

Suppose that  $q_{1k} > \bar{q}_1$  and  $q_{2k} < \bar{q}_2$ . Then, homogeneity of degree zero implies that the

---

<sup>16</sup>These include decoys with extremely high quality to price ratios, but very low levels of quality.

upside  $q_{1k}$  of good  $k$  is salient whenever  $\sigma(q_{1k}/\bar{q}_1, 1) > \sigma(1, \bar{q}_2/q_{2k})$ , which is equivalent to:

$$q_{1k} \cdot q_{2k} > \bar{q}_1 \cdot \bar{q}_2.$$

The salience ranking is determined by the quality-quality product  $q_{1k} \cdot q_{2k}$ .<sup>17</sup> In this respect, a version of Proposition 1 carries through: if a good is neither dominated by nor dominates the reference good, its relative advantage is salient if and only if it has a higher quality-quality product than the reference good.

Consider now how salience affects choice along a rational indifference curve. In a quality-quality trade-off, rational indifference curves are downward sloping. Due to diminishing sensitivity, goods that are unbalanced in the sense of having a low quality-quality product  $q_{1k} \cdot q_{2k}$  will have their weak dimension salient. As a consequence, the consumer chooses an intermediate good, namely a good with high enough  $q_{1k} \cdot q_{2k}$ :

**Proposition 4** *Let all goods in a choice set be located on a rational indifference curve with utility level  $u$ . Then:*

- 1) if  $\bar{q}_1 < \frac{u}{2\theta_1}$ , the local thinker chooses the good  $\mathbf{q}_k$  with highest  $q_{1k}$  subject to  $q_{1k} \cdot q_{2k} > \bar{q}_1 \cdot \bar{q}_2$
- 2) if  $\bar{q}_1 > \frac{u}{2\theta_1}$ , the local thinker chooses the good  $\mathbf{q}_k$  with highest  $q_{2k}$  subject to  $q_{1k} \cdot q_{2k} > \bar{q}_1 \cdot \bar{q}_2$
- 3) if  $\bar{q}_1 = \frac{u}{2\theta_1}$ , the local thinker chooses a good “sufficiently close” to  $(\bar{q}_1, \bar{q}_2)$ .

The local thinker picks the good that is most specialized (has the most extreme strength) relative to the reference good, provided that good’s weakness is not so bad that it is noticed. If the reference level of  $\bar{q}_1$  in the choice set is low, case 1, the consumer chooses the good which has the highest value of this attribute but also a sufficiently high value of  $q_{2k}$ . When instead the reference level of  $\bar{q}_1$  is high, case 2, the consumer chooses the good having the highest value of this attribute but also a sufficiently high value of  $q_{1k}$ . When the reference is itself balanced, case 3, the consumer chooses a good close to it. The Appendix fully characterizes the consumer’s choice in the latter case.

The local thinker’s choice balances two forces. On the one hand, keeping the salience ranking fixed, the local thinker tries to maximize the salient quality along the rational

---

<sup>17</sup>This condition can be directly mapped into our previous analysis of the quality-price tradeoff by noting that one can write the product  $q_{1k} \cdot q_{2k}$  as a quality-cost ratio  $q_{1k}/q_{2k}^{-1}$ , which measures the added value of  $q_1$  per unit lost of  $q_2$  needed to keep good  $\mathbf{q}_k$ ’s relative salience constant.

indifference curve. On the other hand, as the good’s strength becomes more pronounced at the expense of its weakness, the latter becomes increasingly salient due to diminishing sensitivity.<sup>18</sup>

At the same time, diminishing sensitivity plays a central role in generating the decoy/compromise effect in the quality-price space described in Proposition 3. In that case, very unbalanced goods are those with high quality and high price. If the choice set is concave with respect to prices, then diminishing sensitivity is very strong for extreme goods, ensuring that their prices are salient. This renders intermediate goods relatively more attractive.

This effect is again different from loss aversion (Tversky and Simonson 1993, Bodner and Prelec 1994) in that consumers do not mechanically prefer middle-of-the-road options. They instead prefer goods that are somewhat specialized in favor of their salient upsides. Unlike in Koszegi and Szeidl’s “bias towards concentration”, specialization here cannot be excessive, because a severe lack of quality in any dimension is highly salient. An uncommonly spacious back seat may enhance consumers’ valuation of a car, but not if this comes at the cost of an extremely small trunk. Producers often specialize a little, rarely a lot.

### 3.4 Salience and Willingness to Pay

Willingness to pay (WTP) is a central concept in economic theory. The WTP for quality  $q$  is defined as the maximum price at which the consumer is willing to buy  $q$  instead of sticking to the outside option of no consumption  $\mathbf{q}_0 = (q_0, p_0)$ , where typically  $q_0 = p_0 = 0$ . In standard theory, knowledge of  $q$  and of  $\mathbf{q}_0$  are sufficient to determine WTP for  $q$  (assuming quasi-linear utility).

In contrast to this prediction, evidence suggests that the willingness to pay for a good can be influenced by contextual factors. In a famous experiment (Thaler 1985), subjects were first asked to imagine sunbathing on a beach on a very hot summer day and then to state their willingness to pay for a beer to be bought nearby and brought to them by a friend.

---

<sup>18</sup>Thus, in quality-quality tradeoffs the local thinker does not go all the way to the extreme good, as he does in quality-price trade-offs. In fact, along a quality-price indifference curve, an increase in quality is matched by an increase in price, so that diminishing sensitivity causes both attributes to become less salient. In contrast, along a quality-quality indifference curve one quality increases at the expense of the other. Due to diminishing sensitivity, the reduction in one quality dimension exerts a stronger effect on salience than the increase in the other quality dimension.

Subjects stated a higher willingness to pay when the place from which a beer is bought was specified to be a nearby resort hotel than when it was a nearby grocery store. Thus, the source of beer influences the subject’s willingness to pay even though the consumption experience is identical in the two scenarios (back at the beach).

Thaler’s explanation for this effect is based on “mental accounting.” First, information about the nearby location prompts the subject to imagine a price for the beer, such as a price experienced in the past at a similar location. This evoked price constitutes a mental account, which the subject uses to assess his WTP. Second, and crucially, the consumer is assumed to derive utility from buying a good below its evoked price (transaction utility). Because the evoked price at the resort is higher, and the transaction utility associated with buying at the resort is *ceteris paribus* also higher, the consumer states a higher WTP for beer from the resort.

In our model, information about the nearby location also prompts the decision maker to imagine a price for beer. However, our explanation does not rely on transaction utility. Instead, this evoked price affects salience. When thinking of the high price at the resort, the local thinker is willing to pay a high price for the beer and still perceive quality as salient. When thinking of the low price in the store, however, the local thinker is not willing to pay a high price for the beer, as that price would be very salient. In other words, in our model the evoked price acts, through salience, as an anchor for the consumer.

To see this formally, suppose that the consumer must state his WTP for quality  $q$  in the evoked set  $\mathbf{C} \equiv \{\mathbf{q}_k\}_{k=0,\dots,N}$ ,  $N \geq 1$ . Good  $\mathbf{q}_0 = (0, 0)$  is the outside option of not consuming  $q$ . Goods  $\mathbf{q}_k = (q, -p_k)$ ,  $k = 1, \dots, N$  identify the context in which WTP is stated. In the beer example, context could be summarized by a single option  $\mathbf{q}_1$  (i.e.  $N = 1$ ), capturing the price recalled by the consumer for the resort or the store, respectively. Alternatively, context might induce the consumer to recall an entire distribution  $\mathbf{q}_1, \dots, \mathbf{q}_N$  of possible resort or store prices for beer of quality  $q$ . Either way, the consumer includes these recalled prices for  $q$  in his evoked set. These prices do not constitute choice options in a strict sense (the only choice the subject is making is to select a WTP for  $q$ ), so this is the first instance in our model where the evoked and the choice sets differ.

Since the consumer is evaluating the good  $\mathbf{q} = (q, -p)$  for a price  $p$ , his full evoked set is

$\mathbf{C} \cup \{(q, -p)\}$ . We define the consumer's willingness to pay for  $q$  in a choice context  $\mathbf{C}$  as:

$$\begin{aligned} \text{WTP}(q|\mathbf{C}) &= \sup p \\ \text{s.t. } & u^{LT}(\mathbf{q}|\mathbf{C} \cup \{(q, -p)\}) \geq u^{LT}(\mathbf{q}_0|\mathbf{C} \cup \{(q, -p)\}). \end{aligned} \quad (16)$$

WTP is still defined as the maximum price  $p$  that the consumer is willing to pay for  $q$  against the prospect of obtaining the outside option  $\mathbf{q}_0 = (0, 0)$ , but the superscript  $LT$  indicates that now the consumer's preferences are distorted by salience. This change has one crucial implication: different values of  $p$  can alter the salience of  $q$ , changing the consumer's valuation of the good. As a consequence, the maximization in (16) tends to select a price  $p$  such that  $q$  is salient.

In the evoked set  $\mathbf{C} \cup \{(q, -p)\}$ , the reference good has quality  $\bar{q} = q \cdot \frac{N+1}{N+2}$  and price  $\bar{p} = \frac{p}{N+2} + \hat{p} \frac{N}{N+2}$ , where  $\hat{p} = \sum_{k=1}^N p_k / N$  is the average price of the alternative goods. We also assume for simplicity that the consumer weighs quality and price equally (i.e.  $\theta_1 = \theta_2$ ). We can then show:

**Proposition 5** *The consumer's willingness to pay for  $q$  depends on the price  $\hat{p}$  as follows:*

$$\text{WTP}(q|\mathbf{C}) = \begin{cases} \delta q & \text{if } \hat{p} \leq \delta q \\ \hat{p} & \text{if } \delta q < \hat{p} \leq \frac{1}{\delta} \cdot q \\ q/\delta & \text{if } \frac{1}{\delta} \cdot q < \hat{p} \leq \frac{1}{\delta} \cdot q \cdot \frac{1}{k(N)} \\ \delta q & \text{if } \hat{p} > \frac{1}{\delta} \cdot q \cdot \frac{1}{k(N)} \end{cases} \quad (17)$$

where  $k(N) = \frac{N(N+1)}{(N+2)^2 - (N+1)} < 1$ . As  $\delta \rightarrow 1$ , the willingness to pay tends to  $q$  and becomes independent of context  $\hat{p}$ .

The price context only affects WTP if the consumer is a local thinker, namely if  $\delta < 1$ . If  $\delta = 1$ , Equation (16) recovers the standard case where WTP equals  $q$  and does not depend on  $\hat{p}$ .

For  $\hat{p} \leq \frac{1}{\delta} \cdot q \cdot \frac{1}{k(N)}$  the consumer's WTP weakly increases in the average price of alternative goods  $\hat{p}$ . In contexts where quality is more expensive, namely  $\hat{p}$  is higher, the consumer is

willing to pay a higher price  $p$  and still view quality as salient.<sup>19</sup> The highest possible WTP is  $q/\delta$ , which is the consumer's valuation when quality is salient. Through salience, a higher price  $\hat{p}$  acts like an anchor, increasing WTP. Drawing the consumer's attention to higher prices at which  $q$  is available induces him to increase the price at which he perceives the same quality as a good deal, increasing his WTP.

Interestingly, Proposition 5 suggests that when the reference price is implausibly high, this effect vanishes. If  $\hat{p}$  is too high, even for a local thinker focusing on quality (i.e.  $\hat{p} > q/\delta$ ), price becomes salient and the consumer's WTP drops. The WTP in (16) is graphically represented in Figure 4.

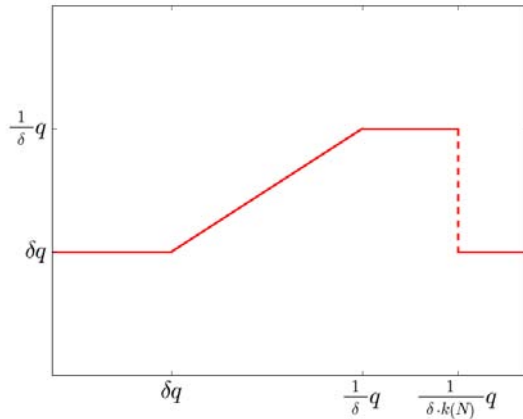


Figure 4: Willingness to Pay for  $q$  as a function of reference price  $\hat{p}$ .

To see how Thaler's example works in our model, imagine that - upon learning that the nearby location is a resort - subjects populate their evoked set by recalling beer prices that they experienced (or expect) in resorts. This generates a reference price  $\hat{p}_{resort} = \sum_{k=1}^N p_{k,resort}/N$ . The reference price for the store is  $\hat{p}_{store} = \sum_{k=1}^N p_{k,store}/N$ . Naturally,  $\hat{p}_{resort} > \hat{p}_{store}$ . The model says that, provided the reference prices do not preclude all trade (i.e.  $\hat{p}_{resort}, \hat{p}_{store} < q/\delta$ ), the consumer's WTP is weakly higher at the resort than in the store, consistent with Thaler's example.

This analysis shows that in our model context shapes evaluation not only through the characteristics of the alternatives of choice, as in Sections 3.1 and 3.2, but also through the

<sup>19</sup>Put differently, as  $\hat{p}$  increases the consumer perceives  $(q, p)$  as a good deal even at higher prices  $p$ .

reference options that enter the consumer’s evoked set. Take for example the choice of wine in a store versus at a restaurant. Although as we showed in Section 3.1 the higher prices at the restaurant induce the consumer to select high quality wines, this is unlikely to happen if wine prices are outrageous even by restaurant standards. Unexpectedly high wine prices at a restaurant will be very salient to the consumer, even if price differences among the actual options of choice are fairly small. In other words, salience is not only shaped by the actual options in the choice set, but also by the extent to which the options of choice differ from the consumer’s past experiences/expectation. We address this mechanism in Section 4.1.

## 4 Applications

We now discuss field evidence on context effects and illustrate how our model can help us think about them in a coherent way.

### 4.1 Context Effects due to Price Changes

Hastings and Shapiro (2011) show that consumers react to parallel increases in gas prices by switching to cheaper (and lower quality) gasoline. One possible explanation for this behavior is mental accounting (Thaler 1999): when thinking about purchasing gas, the consumer thinks about the “gas consumption” account, to which he allocates a fixed monetary budget. The budget is targeted to past gas prices, so that - as gas prices increase - the consumer (who mostly cares about the quantity of gas) shifts his consumption from the expensive to the cheaper gas.

In our model, as in mental accounting, the consumer who purchases gas has in mind a reference gas expenditure, and in particular historical gas prices. In our model, however, the consumer does not allocate a fixed monetary budget to gas consumption. Instead, as prices of all gas grades increase beyond their reference value, prices become more salient than qualities. As a consequence, the consumer becomes overly sensitive to price differences, causing him to switch to lower octane, cheaper gas.

To see this, consider as in Section 3.1 a consumer choosing between a high end wine  $\mathbf{q}_h = (30, -\$60)$  and a low end wine  $\mathbf{q}_l = (20, -\$50)$  at a restaurant. When choosing



among these wines, however, the consumer also recalls the past prices at which he bought the two wines at this (or other) restaurants. The situation (restaurant) evokes consuming the goods at past prices,  $\mathbf{q}_h^{hist} = (30, -p_h^{hist})$  and  $\mathbf{q}_l^{hist} = (20, -p_l^{hist})$ . Rather than reminding the consumer of his “restaurant wine” budget, the situation reminds the consumer of past restaurant prices for wine. The evoked set of the consumer is then  $\mathbf{C} = \{\mathbf{q}_h, \mathbf{q}_l, \mathbf{q}_h^{hist}, \mathbf{q}_l^{hist}\}$ , but the consumer only chooses among  $\mathbf{q}_h$  and  $\mathbf{q}_l$ . Imagine two alternative situations. In the first, the prices found by the consumer at the restaurant are identical to the wines’ historical prices. The evoked set is thus:

$$\mathbf{C}_{hist=actual} = \begin{cases} \mathbf{q}_h = & (30, -\$60) \\ \mathbf{q}_l = & (20, -\$50) \\ \mathbf{q}_h^{hist} = & (30, -\$60) \\ \mathbf{q}_l^{hist} = & (20, -\$50) \end{cases} \quad (18)$$

With this evoked set, salience and choice are identical to the case studied in Section 3.1. When reference prices are equal to actual prices, the consumer chooses solely based on the choice set. As in Section 3.1, the consumer opts for the expensive wine.

Suppose alternatively that the consumer finds the wine prices at this restaurant to be much higher than historical prices for restaurant wine. In particular, suppose that the prices of the available wines are unexpectedly high by \$30, namely  $p_h^{hist} = \$30$ ,  $p_l^{hist} = \$20$ . The evoked set is then:

$$\mathbf{C}_{hist<actual} = \begin{cases} \mathbf{q}_h = & (30, -\$60) \\ \mathbf{q}_l = & (20, -\$50) \\ \mathbf{q}_h^{hist} = & (30, -\$30) \\ \mathbf{q}_l^{hist} = & (20, -\$20) \end{cases} \quad (19)$$

In this case, the reference wine is  $\bar{\mathbf{q}} = (25, -\$40)$ . The high end wine  $\mathbf{q}_h$  still yields above average quality, but given its very high price it has a lower than average quality/price ratio, as  $30/\$60 < 25/\$40$ . As a consequence, the high end wine becomes price salient. Since its price is high, this greatly reduces its value as perceived by the local thinker, and he chooses the low end wine, regardless of its salience ranking (in this numerical example the low end wine is valued correctly because quality and price are equally salient). When the consumer

finds wines at the restaurant to be unexpectedly pricey, he switches to lower quality wines.

This intuition rationalizes the Hastings Shapiro example where increase in gasoline prices induce consumers to switch towards cheaper low octane gas. To the extent that historical prices are fixed in the evoked set, any current increase in prices – particularly in the prices of expensive goods – will be salient. This is due to the ordering property of the salience function: as the price of an expensive good rises, price becomes more salient for that good. Including historical prices in the evoked set provides a natural way to capture the consumer’s adaptation to a reference price: if he has observed a given price sufficiently many times, that price effectively becomes the reference price to which all other prices are compared.

This effect is due to an increase in the prices of a sub-category of goods in the evoked set, regardless of the role of historical prices. Imagine for instance a consumer choosing among different qualities of Bordeaux wines. The more expensive Bordeaux wines are relative to other wines in the wine list, the more salient the price of Bordeaux wines will be. This induces the consumer to substitute towards cheaper Bordeaux, or potentially to leave the category altogether. This effect is thus fundamentally different from the one at play in the restaurant vs store example of Section 3.1. In that case, the price of all wines uniformly increased at the restaurant, making quality salient and inducing the consumer to substitute towards higher quality wines.

In general, it is ambiguous whether consumers react to price hikes in a given category of goods by substituting away from the most expensive items in the category or towards them. To gain traction on this issue, take an evoked set  $\mathbf{C}$  having  $N > 1$  elements and partition it into two subsets  $\mathbf{C}_F$  and  $\mathbf{C}_C$ . Subset  $\mathbf{C}_F$  is the set of goods for which price is held fixed, while  $\mathbf{C}_C$  is the set for which price increases. Denote by  $\omega$  the fraction of goods that belong to  $\mathbf{C}_C$ , by  $\bar{p}$  the average price of goods in  $\mathbf{C}$ , and by  $\bar{p}_X$  the average price in  $\mathbf{C}_X$ ,  $X = F, C$ . We then show:

**Proposition 6** *If  $\bar{p}_C > \bar{p}$ , a marginal increase in the prices of all goods in  $\mathbf{C}_C$  (holding constant the prices in  $\mathbf{C}_F$ ) boosts the salience of price for the most expensive goods in  $\mathbf{C}_C$  only if:*

$$\frac{p_C^{\max} - \bar{p}_C}{\bar{p}_F} < \frac{1 - \omega}{\omega}, \quad (20)$$

where  $p_{\mathbf{C}}^{\max}$  is the highest price in  $\mathbf{C}_{\mathbf{C}}$ . If  $\omega = 1$ , the salience of price decreases for all goods in  $\mathbf{C}_{\mathbf{C}}$ .

When the prices of items in the expensive category  $\mathbf{C}_{\mathbf{C}}$  increase, the most expensive category members become more price salient when two conditions are met. First, the price range  $p_{\mathbf{C}}^{\max} - \bar{p}_{\mathbf{C}}$  in the category must be small. Indeed, as price differences in the category become small, the price hike tends to uniformly draw the consumer's attention to all prices in the category, starting with the most expensive items. As price becomes more salient in  $\mathbf{C}_{\mathbf{C}}$ , the consumer substitutes away from its most expensive goods.

Second, the size  $\omega$  of the category must be small. When  $\omega$  is large, too many goods increase in price, and diminishing sensitivity prevails, rendering quality more salient as in the store vs restaurant example. When instead few prices change ( $\omega$  is small) the forces of ordering prevail, increasing price salience for the items whose price have increased. This is the mechanism at work in the wine example of Equations (18) and (19). Thus, our model yields a testable prediction as to which effect of price hikes should prevail depending on price dispersion.

## 4.2 Salience and “Misleading Sales”

Retailers frequently resort to sales events as a means to sell their products. In 1988, for example, sales accounted for over 60% of department store volume (Ortmeyer, Quelch and Salmon 1991). The standard explanation for sales is price discrimination: sporadic sales allow retailers to lure low willingness to pay customers, whereas high willingness to pay customers who cannot wait for a sale buy at the higher regular prices. It is probably true that low willingness to pay customers tend to sort into sales events, but the high frequency and predictability of sales casts some doubt on the universal validity of the price discrimination hypothesis. In particular, there is growing concern that retailers may deliberately inflate regular prices in order to lure consumers into artificial sales events. The Pennsylvania Bureau of Consumer Protection has successfully pursued retailers for advertising misleading sales prices. In Massachusetts, regulatory changes have tightened rules for price compari-

son claims, for example requiring that retail catalogues state that the “original” price is a reference price and not necessarily the previous selling price.

In this section we show that salience - and in particular the logic of decoy effects - can shed light on these “misleading sales” events, yielding two new testable predictions:

- In a store selling different qualities, misleading sales boost demand only for high quality goods,
- Misleading sales boost demand only for non-standard goods.

To see how the model works, suppose that a consumer is considering whether or not to buy a good of quality  $q$  and price  $p$  in a store. The good is non-standard in the sense that it is only available in this store, so the effective choice set faced by the consumer is  $\mathbf{C}_0 \equiv \{(0, 0), (q, p)\}$ , where  $(0, 0)$  is the outside option of not buying the good. We later consider the case of standard goods, which can be easily found at different stores. As before, the consumer weights quality and price equally,  $\theta_1 = \theta_2$ .

With respect to this purchasing decision, the salience of the good’s quality for the consumer is equal to  $\sigma(q, q/2)$  while the salience of its price is equal to  $\sigma(p, p/2)$ . Given homogeneity of degree zero,  $\sigma(q, q/2) = \sigma(p, p/2)$ , namely quality and price are equally salient for any  $q$  and any  $p$ . Thus, in  $\mathbf{C}_0$  the consumer’s valuation of the good is rational and the maximum price he can be charged for the good is his true valuation, namely  $p = q$ .

Suppose now that there is a sale event in the store. By a sale event we mean that the consumer is offered the same quality  $q$  at the sale price  $p_s$  rather than at the full regular price  $p_f > p_s$ . Crucially, then, when deciding whether or not to buy the good, the regular price becomes part of the consumer’s evoked set, which becomes equal to  $\mathbf{C}_{sale} \equiv \{(0, 0), (q, p_s), (q, p_f)\}$ .<sup>20</sup>

Consider the standing of the option  $(q, p_s)$  in the new evoked set  $\mathbf{C}_{sale}$ . The salience of quality is  $\sigma(q, 2q/3)$ , while the salience of price is  $\sigma(p_s, \frac{p_s+p_f}{3})$ . The crucial issue here is that the retailer can manipulate the salience of price by manipulating the price discount  $p_s/p_f$ . In particular, it is easy to show:

---

<sup>20</sup>This is either because the store displays the regular price at the moment of the sale and/or because the consumer recalls the regular price. Another, less realistic, possibility is that the consumer now considers three options: buy nothing, buy quality  $q$  today, or buy quality  $q$  at the regular price in the future.

**Proposition 7** *The retailer can charge a sale price  $p_s = q/\delta$  and still have the customer buy the product by setting any full price in the interval  $p_f \in (q/\delta, 7q/2\delta)$ .*

By artificially inflating the regular price of the good and by offering at the same time a generous discount, the retailer can extract up to the local thinker’s valuation  $q/\delta$  from the consumer. This is because the consumer views the discount as a good deal, inflating his valuation of quality. The model limits the maximal regular price and thus the maximal discount to  $p_s/p_f \geq 2/7$ . The reason is that an excessively high regular price renders prices salient, reducing the consumer’s valuation.

We now illustrate our first prediction, namely that a “misleading sales” policy should be only effective for a high quality good. Suppose that the store has an expensive high quality good  $\mathbf{q}_h = (q_h, p_h)$  and a cheaper, lower quality good  $\mathbf{q}_l = (q_l, p_l)$ , where  $q_h > q_l$ , and  $p_h > p_l$ . For the sake of illustration, we assume that the prices at which these goods are sold are fixed (e.g. by the producer).<sup>21</sup> The store, however, can try to influence which good is sold by adopting a misleading sales policy. In the case of the high quality good, this amounts to making the good available also at a full price  $p_{fh} > p_h$ . Similarly, for the low quality good, the store can set a full price  $p_h > p_{fl} > p_l$ . Suppose that the goods are such that  $\mathbf{q}_h$  is sold if and only if it is quality salient, implying that condition (15) holds and  $q_h - \delta p_h > 0$ . We then find:

**Proposition 8** *The store can always make the high quality good quality salient, and have the consumer choose it over the low quality good, by holding a sale on  $\mathbf{q}_h$  where the full price  $p_{hf}$  is suitably chosen. In contrast, a sale is ineffectual for the low quality good: if the consumer chooses  $\mathbf{q}_h$  in the absence of a sale, there exists no full price  $p_{fl} \in (p_l, p_h)$  for  $\mathbf{q}_l$  that makes  $\mathbf{q}_h$  price salient, and  $\mathbf{q}_l$  be chosen, in the context of the sale.*

It is always possible to engineer sales inducing the local thinker to overvalue the high quality good  $\mathbf{q}_h$  relative to  $\mathbf{q}_l$ , but not the reverse. The reason is that holding a sale on the good with lowest quality/price ratio unambiguously decreases the quality/price ratio of the reference good. This effect reinforces the salience of quality for the high quality good

---

<sup>21</sup>A general analysis of sales policies, including the case where a store is able to choose the goods’ prices, is left for future work.

and renders the low quality good price salient (for lower price is the advantage of the latter good). As a result, the sale boosts the overvaluation of the high quality good and may cause an undervaluation of the low quality one. Both of these effects imply that sales on the low quality good are unlikely to work.<sup>22</sup>

By contrast, sales work if the high quality good is initially undervalued relative to the low quality good. In this case, holding a sale on the high quality good  $\mathbf{q}_h$  boosts the salience of its quality, increasing this good's valuation relative to  $\mathbf{q}_l$  (regardless of the latter's salient attribute). Thus, sales should be effective specifically for high quality goods that, in the absence of sales, would be price salient.

Consider now the second prediction of our model, namely that sales are unlikely to work when standard goods are involved, for which market prices are well known. To see this, suppose that the consumer wishes to purchase a standard good of quality  $q$ , for instance a metro ticket. There are  $N > 1$  potential sellers of the good. Suppose for the sake of the argument that each of these sellers implements a misleading sales policy consisting of a regular price  $p_f$  and a sales price  $p_s$  for the good, where  $p_f/p_s = k \in (1, 7/2)$  (see Proposition 7 above).

In this case, the consumer's evoked set consists of  $2N$  goods (two goods for each of the  $N$  sales), and the outside option of not buying  $(0, 0)$ . Formally,  $\mathbf{C}_{sale} \equiv \{(0, 0), (q, p_s), \dots, (q, p_f)\}$  where  $(q, p_s)$  and  $(q, p_f)$  are repeated  $N$  times. For the items on sale, then, the salience of quality is  $\sigma(q, q \frac{2N}{2N+1})$ , and that of price is  $\sigma(p_s, p_s \frac{N+Nk}{2N+1})$ . Due to homogeneity of degree zero, these expressions imply that when the number of sellers is sufficiently large, namely when

$$N > \frac{4 + 2\sqrt{2(1+k)}}{4(k-1)},$$

the items on sale have salient price [i.e.,  $\sigma(q, q \frac{2N}{2N+1}) < \sigma(p_s, p_s \frac{N+Nk}{2N+1})$ ], rather than salient quality as in the non-standard good case of Proposition 7. This is true for any given magnitude  $k$  of the sale.

This result is intuitive. As the number of sellers  $N$  increases, the average quality  $\bar{q} = q \frac{2N}{2N+1}$  in the choice set gets arbitrarily close to the quality  $q$  of the standard good. As a result,

---

<sup>22</sup>Blattberg and Wisniewski (1989) present suggestive evidence for this effect, in the context of sales at a grocery chain.

quality becomes non-salient. By contrast, the price variability generated by sales renders prices salient, increasing the consumer's price sensitivity above its rational counterpart. As a result, when deciding where to buy a standardized good the local thinker focuses on price because price is the attribute that varies most across sellers (almost by definition of standardized goods)! This implies that a generalized policy of misleading sales does not work in the case of standardized goods, because it induces consumers to focus on prices, reducing their willingness to pay.

This argument has the additional implication that an individual seller may find it profitable to abandon a misleading sale policy and set a stable price equal to the average price  $\bar{p} = p_s \frac{N+Nk}{2N+1}$  set by other stores across sales and non-sales events. By charging exactly the average price, this store becomes quality salient and consumers switch to it. This holds even though the average price charged by the store is above the sale price at which the standardized good might be available in the market. This can help explain why stores with consistently low prices, called EDLP for everyday low prices, have steadily been gaining market share of standardized goods from mainstream stores that engage in frequent sales in standardized markets such as grocery and general merchandise (see Ortmeier, Quelch and Salmon, 1991).

Our model has further implications for the pricing of standard vs non standard goods. Because the quality of standard goods does not vary across stores, our model predicts that consumers should be more price sensitive for standard than for nonstandard goods (relative to the rational case). This can help explain an empirical regularity uncovered by Lynch and Ariely (2000), who studied online wine markets. The authors found that consumers are very price sensitive for standard wines, which are offered by many sellers, but not for unique wines, sold by one or few sellers. Relatedly, Jaeger and Storckmann (2011) find that price dispersion in wine retail prices increases with price levels (which we explain with diminishing sensitivity), and particularly so for vintage (i.e., non-standard) wines. One possible equilibrium implication of this reasoning can be that standard goods should not only display lower price dispersion than non-standard goods, but they should also have a higher quality to price ratio on average because consumers tend to undervalue price-salient goods relative to their true preferences. Diminishing sensitivity would also induce price

dispersion to increase with the price level.

### 4.3 An application to Insurance Demand

Barseghyan, Molinari, O'Donoghue and Teitelbaum (2011) analyze consumer choice of insurance plans which differ in two dimensions, deductibles and premia. They find evidence that consumers put too much weight on plans' deductibles, relative to their premia. As an illustration, many consumers prefer a home all perils plan with a \$500 deductible and a \$679 premium to a plan with a \$1000 deductible and a \$605 premium, implying that the risk of a claim for a home accident is at least 14.8%, when the mean risk estimated from the data is around 8.4%. Sydnor (2010) finds similar evidence in the choice of home insurance. Both papers stress that the data are at odds with standard risk aversion and suggest an interpretation of the evidence in which that consumers overweight the (small) claim probabilities. We now show how this behavior can be understood in our model, formalizing the intuition that deductible to cost ratio plays a role in insurance choice.

At time  $t_0$ , a consumer decides whether to buy insurance against a loss  $L$  that materializes at time  $t_1$  with probability  $p$ . His consumption utility is linear (we abstract from risk aversion and time discounting), so we can normalize his endowment to zero. A rational consumer with linear utility sees no benefit of buying insurance, so both the demand for insurance and the choice of which plan to buy are driven by salience.

An insurance plan  $I_i = (R_i, P_i)$  has a cost  $P_i$  and covers the amount  $R_i$  in case loss  $L$  materializes. The implied deductible is  $D_i = L - R_i$ . The consumer's utility under plan  $I_i$  is equal to:

$$-pL + pR_i - P_i.$$

The choice of not buying insurance is represented by plan  $I_0 = (0, 0)$ , and the premium at which a rational consumer is indifferent between  $I_i$  and  $I_0$  is equal to the expected coverage  $p \cdot R_i$ .

Following Barseghyan et al (2011), and in line with industry practice, we consider linear



pricing schemes  $P_i = c + \phi \cdot R_i$ . In particular, we assume that:

$$P_i = c + p \cdot R_i. \quad (21)$$

Equation (21) implies that any extra unit of insurance is fairly priced at the margin but the insurance company makes a profit  $c \geq 0$  on the plan.

Consider a local thinker's choice between two plans  $I_H, I_L$ . Plan  $I_H$  provides a higher coverage  $R_H > R_L$  but entails a higher premium  $P_H > P_L$ . Given the pricing Equation (21), the rational consumer is indifferent between  $I_H$  and  $I_L$ . However, this is not so for the local thinker. Consistent with Definition 1, salience is defined on the utility value of the attributes  $(R, P)$  of the insurance policies. The coverage of policies  $I_H$  and  $I_L$  is then more salient than their premiums whenever

$$\frac{R_H}{R_L} > \frac{P_H}{P_L}, \quad (22)$$

namely when the higher coverage granted by plan  $I_H$  in case of accident is higher, in percentage terms, than the extra premium the consumer must pay for it. By exploiting Equation (21), this condition can be written in terms of deductibles as follows:

$$\frac{L - D_H}{L - D_L} > \frac{c + p \cdot (L - D_H)}{c + p \cdot (L - D_L)} \quad (23)$$

It is easy to see that under the pricing policy of Equation (21), condition (22) is always met. This is because the accident happens with probability  $p$  less than one so that – given the profit  $c$  – the premium  $P_i$  increases less than proportionally with the coverage  $R_i$ .

To further illustrate this result, let  $s$  denote the percentage savings guaranteed in case of accident by the generous policy  $I_H$ . By writing  $D_H = (1 - s)D_L$  and by taking the linear approximation of Equation (23) around  $s = 0$ , it is easy to see that deductibles are salient when the percentage decrease  $s$  in the deductible granted by  $I_H$  is larger than its incremental premium  $p \cdot s$ . This condition always holds because  $p < 1$ . Intuitively, the reduction in deductible granted by the generous plan  $I_H$  is much higher, in percentage terms, than the extra price the consumer has to pay for it. As a result, the difference in deductibles across policies “stands out” and draws the consumer's attention when making his decision.

Given the salience of a plan's coverage (and thus of deductibles), the consumer's perceived utility from the insurance is given by:

$$-pL + pR_i - \delta \cdot P_i.$$

Since the no-insurance option  $(0, 0)$  is evaluated correctly (as both dimensions are equal to 0 and thus equally salient), the consumer's WTP for the high coverage policy is:

$$P_H = \frac{p}{\delta} \cdot R_H,$$

which is above the actuarially fair price, justifying a profit margin  $c > 0$ . As the consumer focuses on deductibles, his preferences tilt in favor of low-deductible policies even when their prices are unfavorable from an actuarial perspective.

## 5 Conclusion

We combine two ideas to explain a wide range of experimental and field evidence regarding individual choice, as well as to make new predictions.

The first idea is that choices are made in context and that in particular goods are evaluated by comparison with other goods the decision maker is thinking about. This idea is intimately related to Kahneman and Tversky's (1979) concept of reference points, and is also central to related studies of choice by Tversky and Kahneman (1991), Tversky and Simonson (1993), Bodner and Prelec (1994) and Koszegi and Rabin (2006). In our model, context is often determined by the choice set itself, and the reference good relative to which the options are evaluated has the average characteristics of all the goods in the choice set. In some examples, however, decision makers also recall their previous experiences with goods in the choice set, such as seeing them at historical or normal prices, in which case these experiences also influence context. We use Kahneman and Miller's (1986) concept of the evoked set to describe situations in which prior experiences shape context, and then define the reference good as one having the average characteristics in the evoked set.

The second idea, which extends our earlier work on choice under risk (BGS 2012), holds

that the salience of each good’s attributes relative to the reference good, such as its quality and price, determines the attention the decision maker pays to these attributes as well as their weight in his decision. We argue that ordering and diminishing sensitivity are the two critical properties of salience that together help account for a broad range of evidence.

We show that our model provides insight into several puzzles of consumer choice. The model makes stark predictions for choice in experimental settings, in which the reference good is well defined. First, by showing how irrelevant alternatives change the reference good, the model accounts for two well-known violations of independence of irrelevant alternatives, namely decoy and compromise effects. Moreover, it predicts that these effects differentially benefit more extreme goods (e.g. expensive, high-quality goods). In the design of desirable goods, the model predicts a preference for some specialization as long as a minimum balance across attributes is provided. Moreover, by allowing expected or historical prices to shape the reference good, the model also helps think about context-dependent willingness to pay, exemplified by Thaler’s celebrated beer example. In a companion paper (BGS 2012b), we show that our model also helps account for the endowment effect. Taken together, these predictions suggest that the salience mechanism can be seen as a simpler alternative to loss aversion in generating context effects.

Turning to the field evidence, we show that our model provides a unified way of thinking about several phenomena described as mental accounting, and makes predictions for how consumers would react to changes in the prices of individual goods or whole categories of goods. In particular, we provide a natural explanation of Hastings and Shapiro’s empirical finding that consumer substitute toward lower quality gasoline when all gas prices rise, while at the same time accounting for instances in which consumer substitute toward higher quality goods when prices rise (e.g. the wine example). We present a new theory of sales, based on the idea that the original prices of goods put on sale serve as decoys that attract consumers to these goods. Our approach, unlike the standard model of sales, explains why firms often try to put goods on sale immediately after offering them first, so that “original” prices are in effect reference prices and not the previous selling price (leading to conflict with regulators). It also generates new predictions, such as that a store selling different qualities would only put high quality goods on sale, and that sales are most effective in boosting demand for

non-standard goods. Finally, our model also helps explain some puzzling evidence regarding consumer demand for over-priced insurance with very low deductibles.

In generating our predictions, we took the evoked set as given, without discussing the important problem of its determination. In the applications we considered, the evoked set is either identical to the choice set, in which case the problem is trivial, or else combines the choice set with fairly natural historical experiences that come to the decision makers mind. In other examples, however, the problem of the determination of the evoked set, just like that of the determination of expectations or of reference points, becomes a much more complex matter of what the decision maker recalls and thinks about. We leave this much harder problem to future work.

# Appendix

## A.0 Introduction

As highlighted in Section 2.1, the quality/price ratio in (9) creates two forms of context dependence in our model. The first one is that a consumer is overly sensitive to changes in a good’s salient attributes. The second is that the evaluation of a good depends on the alternatives of comparison. Formally:

**Observation (valuation and choice).** *The local thinker over-values good  $\mathbf{q}_t$ , formally  $u^{LT}(\mathbf{q}_t) > u(\mathbf{q}_t)$ , if and only if:*

$$\mathbf{cov}(\omega_i^t, q_{i,t}) > 0, \quad (24)$$

*while he over-values good  $\mathbf{q}_t$  relative to good  $\mathbf{q}_k$ , formally  $u^{LT}(\mathbf{q}_t) - u^{LT}(\mathbf{q}_k) > u(\mathbf{q}_t) - u(\mathbf{q}_k)$ , if and only if  $\mathbf{cov}(\omega_i^t, q_{i,t}) > \mathbf{cov}(\omega_i^k, q_{i,k})$ , which can be rewritten as:*

$$\mathbf{cov}(\omega_i^t, q_{i,t} - q_{i,k}) + \mathbf{cov}(\omega_i^t - \omega_i^k, q_{i,k}) > 0. \quad (25)$$

To derive (24), note that Definition 2 implies  $u^{LT}(\mathbf{q}_t) = \mathbb{E}_i[\omega_i^t \cdot q_{it}]$ , where the expectation is measured relative to the probability distribution defined by the weights  $(\theta_1, \dots, \theta_{m+1})$ . Expanding the right hand side and using  $\mathbb{E}_i[\omega_i^t] = 1$ , we get  $u^{LT}(\mathbf{q}_t) = u(\mathbf{q}_t) + \mathbf{cov}(\omega_i^t, q_{i,t})$ .

According to (24), salience boosts the valuation of a good when its most salient attributes, namely those having the higher weights  $\omega_i^t$ , are precisely those along which the consumer obtains the highest utility  $q_{i,t}$ . In addition, salience boosts the valuation of good  $\mathbf{q}_t$  relative to that of good  $\mathbf{q}_k$  if the association between salience and utility is more positive for good  $\mathbf{q}_t$ . Equation (25) decomposes this condition into two effects. First,  $\mathbf{q}_t$  is overvalued relative to  $\mathbf{q}_k$  when – for common weights  $\omega_i^t$  across the two goods –  $\mathbf{q}_t$  fares better than  $\mathbf{q}_k$  along the salient attributes [i.e.  $\mathbf{cov}(\omega_i^t, q_{i,t} - q_{i,k}) > 0$ ]. This effect generalizes the wine example above. But with more than two goods, differences in the salience rankings of the goods’ attributes create a second effect:  $\mathbf{q}_t$  tends to be overvalued relative to  $\mathbf{q}_k$  if the salience ranking of  $\mathbf{q}_t$  overweights, relative to  $\mathbf{q}_k$ , those attributes yielding high utility [i.e.  $\mathbf{cov}(\omega_i^t - \omega_i^k, q_{i,k}) > 0$ ].

## A.1 Proofs

**Proposition 1** *Let  $\mathbf{q}_k$  be a good that neither dominates nor is dominated by the average good  $\bar{\mathbf{q}}$ . The following two statements are then equivalent:*

- 1) *The advantage of  $\mathbf{q}_k$  relative to the average good  $\bar{\mathbf{q}}$  is salient if and only if  $q_k/p_k > \bar{q}/\bar{p}$ .*
- 2) *The salience function is homogeneous of degree zero, i.e.  $\sigma(\alpha x, \alpha y) = \sigma(x, y)$  for all  $\alpha > 0$ .*

**Proof.** The salience of  $\mathbf{q}_k$ 's quality is  $\sigma(q_k, \bar{q})$ , while the salience of price is  $\sigma(p_k, \bar{p})$ . Suppose that 1) holds, so that  $\sigma(q_k, \bar{q}) > \sigma(p_k, \bar{p})$  if and only if  $q_k/p_k > \bar{q}/\bar{p}$ , namely  $q_k/\bar{q} > p_k/\bar{p}$ . Consider the implications for  $\sigma(q_k, \bar{q})$ . For any given values of  $p_k, \bar{p}$ , the condition  $\sigma(q_k, \bar{q}) = \sigma(p_k, \bar{p})$  is invariant under scaling of  $q_k$  and  $\bar{q}$ , as it depends only of the ratio  $q_k/\bar{q}$ . As a result,  $\sigma(q_k, \bar{q})$  must only depend on this ratio, and must be proportional to  $\sigma\left(\frac{q_k}{\bar{q}}, 1\right)$ . Setting  $q_k = \bar{q}$  shows the proportionality constant is 1.

Suppose now that 2) holds. Then  $\sigma(q_k, \bar{q}) = \sigma(q_k/\bar{q}, 1)$  and  $\sigma(p_k, \bar{p}) = \sigma(p_k/\bar{p}, 1)$ , where both  $q_k/\bar{q}$  and  $p_k/\bar{p}$  are larger than 1. By the ordering property of salience, then, quality is salient if and only if  $q_k/\bar{q} > p_k/\bar{p}$ . ■

**Lemma 1** *If  $\sigma(\cdot, \cdot)$  satisfies the ordering property for positive attribute values, and is homogenous of degree zero, then it also satisfies diminishing sensitivity.*

**Proof.** Let  $x, y > 0$  and  $\epsilon > 0$ . Under the conditions of the Lemma, we have  $\sigma(x + \epsilon, y + \epsilon) = \sigma(x, \alpha(y + \epsilon))$ , where  $\alpha = \frac{x}{x + \epsilon}$ . For either ordering of  $x, y$ , we have  $\alpha(y + \epsilon) \in (\min\{x, y\}, \max\{x, y\})$ . As a consequence, it follows from ordering that  $\sigma(x + \epsilon, y + \epsilon) < \sigma(x, y)$ . ■

**Proposition 2** *Along a rational linear indifference curve, the local thinker chooses the good with the highest quality/price ratio. In particular:*

- 1) *if  $q_1/p_1 > \theta_2/\theta_1$ , the cheapest good  $(q_1, p_1)$  has the highest quality/price ratio and is chosen;*
- 2) *if  $q_1/p_1 < \theta_2/\theta_1$ , the most expensive good  $(q_N, p_N)$  has the highest quality/price ratio and is chosen;*
- 3) *if  $q_1/p_1 = \theta_2/\theta_1$ , all goods have the same quality/price ratio and the consumer is indifferent between them.*

**Proof.** Consider an indifference curve characterized by  $u(q, p) = \theta_1 q - \theta_2 p = u$ . As in the text, order the elements of the choice set by increasing quality and price, so that  $\mathbf{q}_1 = (q_1, p_1)$  is the cheapest good. The goods' quality-price ratios satisfy  $\frac{q_i}{p_i} = \frac{\theta_2}{\theta_1} + \frac{u}{\theta_1 p_i}$ , and in particular the average good  $(\bar{q}, \bar{p})$  satisfies  $\frac{\bar{q}}{\bar{p}} = \frac{\theta_2}{\theta_1} + \frac{u}{\theta_1 \bar{p}}$ .

1)  $\frac{q_1}{p_1} > \frac{\theta_2}{\theta_1}$  when  $u < 0$ , in which case the price quality/ratio is decreasing as price increases, and price is salient for all goods. This is because price is the relative advantage of cheap goods (whose prices are under  $\bar{p}$  and have high quality/price ratios), while it is the relative disadvantage of expensive goods (whose prices are under  $\bar{p}$  and have low quality/price ratios). Since the cheapest good is the best option along the salient price dimension, it is chosen. Formally, all goods are undervalued,  $u^{LT}(q_i, p_i) = \frac{\delta \theta_1 q_i - \theta_2 p_i}{\delta \theta_1 + \theta_2}$ , but the cheapest good is the least undervalued.

2)  $\frac{q_1}{p_1} < \frac{\theta_2}{\theta_1}$  when  $u > 0$ , in which case the price quality/ratio is increasing as price increases, and quality is salient for all goods. Since the most expensive good is the best option along the salient quality dimension, it is chosen. Formally, all goods are overvalued,  $u^{LT}(q_i, p_i) = \frac{\theta_1 q_i - \delta \theta_2 p_i}{\theta_1 + \delta \theta_2}$ , but the highest quality good is the most overvalued.

3)  $\frac{q_1}{p_1} = \frac{\theta_2}{\theta_1}$  when  $u = 0$ , in which case the price quality/ratio is constant along the indifference curve. As a result, quality and price are equally salient for all goods. The local thinker evaluates each good correctly (as the rational agent) and is thus indifferent between them. ■

Take two goods  $\mathbf{q}_l = (q_l, p_l)$ ,  $\mathbf{q}_h = (q_h, p_h)$ , such that  $\mathbf{q}_h$  is chosen if and only if it is quality is salient. Denoting by  $\Delta u = [q_h - q_l] - [p_h - p_l]$  the rational utility difference between them (with  $\theta_1 = \theta_2$ ), this means

$$-(1 - \delta)[p_h - p_l] \leq \Delta u \leq (1 - \delta)[q_h - q_l] \quad (26)$$

We restrict our attention to decoy options  $\mathbf{q}_d$  such that  $\bar{q} \leq q_h$  and  $\bar{p} \leq p_h$ , where  $(\bar{q}, \bar{p})$  is the reference good in  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ . These constraints allow for goods  $\mathbf{q}_d$  which make  $\mathbf{q}_h$  an intermediate good in the enlarged choice set. When Equation (26) holds, we have:

**Proposition 3**

i) If  $\frac{q_l}{p_l} > \frac{q_h}{p_h}$ , so that price is salient and  $\mathbf{q}_l$  is chosen from  $\{\mathbf{q}_l, \mathbf{q}_h\}$ , then for any  $\mathbf{q}_d$  satisfying  $\frac{q_d}{p_d} < \frac{q_h}{p_h} + \frac{p_l}{p_d} \left[ \frac{q_h}{p_h} - \frac{q_l}{p_l} \right]$ , good  $\mathbf{q}_h$  is quality salient in  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ . Moreover, there exist options  $\mathbf{q}_d$  with  $q_d > q_h$  and  $p_d > p_h$  such that  $\mathbf{q}_h$  is chosen from  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ .

ii) If  $\frac{q_l}{p_l} < \frac{q_h}{p_h}$ , so quality is salient and  $\mathbf{q}_h$  is chosen from  $\{\mathbf{q}_l, \mathbf{q}_h\}$ , then there exist no decoy options  $\mathbf{q}_d$  such that  $\frac{q_d}{p_d} \leq \frac{q_h}{p_h}$  and  $\mathbf{q}_h$  is price salient in  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ . In particular, for no  $\mathbf{q}_d$  satisfying these properties is  $\mathbf{q}_l$  chosen from  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ .

**Proof.** A sufficient condition for reversal between  $\mathbf{q}_l$  and  $\mathbf{q}_h$  is that good  $\mathbf{q}_h$  is chosen if and only if its relative advantage, namely quality, is salient. This means that  $q_h - \delta p_h > q_l - \delta p_l$  and also  $\delta q_l - p_l > \delta q_h - p_h$ . The first expression yields  $\Delta u > -(1 - \delta)(p_h - p_l)$  and the second yields  $\Delta u < (1 - \delta)(q_h + q_l)$ , where  $\Delta u = [q_h - q_l] - [p_h - p_l]$ .

Next, consider case i). Since  $q_l/p_l > q_h/p_h$ , so that good  $\mathbf{q}_h$  has a relatively low quality price ratio, price is salient in  $\{\mathbf{q}_l, \mathbf{q}_h\}$  and  $\mathbf{q}_l$  is chosen. If adding the decoy  $\mathbf{q}_d$  to the choice set makes  $\mathbf{q}_h$  quality salient, then the latter is preferred to  $\mathbf{q}_l$  in  $\{\mathbf{q}_l, \mathbf{q}_h, \mathbf{q}_d\}$ . Good  $\mathbf{q}_h$  becomes quality salient in several different regimes: a) if  $\mathbf{q}_h$  has high quality and high quality/price ratio relative to the reference good,  $\frac{q_h}{p_h} > \frac{\bar{q}}{\bar{p}}$  and  $q_h > \bar{q}$ ,  $p_h > \bar{p}$ . b) if  $\mathbf{q}_h$  dominates the reference good, with higher quality and lower price,  $q_h \cdot p_h > \bar{q} \cdot \bar{p}$  and  $q_h > \bar{q}$ ,  $p_h < \bar{p}$ . c) if  $\mathbf{q}_h$  has low quality and low quality/price ratio relative to the reference good,  $\frac{q_h}{p_h} < \frac{\bar{q}}{\bar{p}}$  and  $q_h < \bar{q}$ ,  $p_h < \bar{p}$ . And d) if  $\mathbf{q}_h$  is dominated by the reference good, with lower quality and higher price,  $q_h \cdot p_h < \bar{q} \cdot \bar{p}$  and  $q_h < \bar{q}$ ,  $p_h > \bar{p}$ .

We are mainly interested in regime a), in which the decoy is located close to the other goods, i.e.  $\bar{q} < q_h$  and  $\bar{p} < p_h$ , and it is a “bad deal”, i.e. it has a low quality-price ratio. In fact, in this regime the condition that  $\mathbf{q}_h$  has quality/price ratio above the reference good reads:

$$\frac{q_d}{p_d} < \frac{q_h}{p_h} + \frac{p_l}{p_d} \left( \frac{q_h}{p_h} - \frac{q_l}{p_l} \right)$$

We can write this as  $q_d < p_d \frac{q_h}{p_h} + p_l \left( \frac{q_h}{p_h} - \frac{q_l}{p_l} \right)$ . So the upper boundary for  $\mathbf{q}_d$  has slope  $q_h/p_h$ , but it is shifted downwards by a factor proportional to  $q_h/p_h - q_l/p_l$ . In particular,  $\frac{q_d}{p_d} < \frac{q_h}{p_h} < \frac{q_l}{p_l}$ . (Both regimes a) and b) impose upper bounds on  $q_d$ . In regime b),  $\bar{q}_d < q_h$ ,  $\bar{p} > p_h$  and the condition on  $q_h \cdot p_h$  yields  $q_d < q_h [3p_h/\bar{p} - 1] - q_l$ . Regimes c) and d) instead



impose lower bounds on  $q_d$ .)

In regime a),  $\mathbf{q}_h$  is quality salient so (26) guarantees it is preferred to  $\mathbf{q}_l$ . To see that the alternative  $\mathbf{q}_d$  is never chosen, two cases are distinguished: either  $\mathbf{q}_d$  has higher quality and lower quality-price ratio than  $\mathbf{q}_h$ , in which case it is price salient; or it has lower quality and lower quality-price ratio than  $\mathbf{q}_h$ , in which case it can either be dominated ( $q_d < q_h$  and  $p_d > p_h$ ) or not. In either case, by being quality salient  $\mathbf{q}_h$  is overvalued relative to  $\mathbf{q}_d$ . Thus, a small enough  $\delta$  can be found such that  $\mathbf{q}_h$  is chosen. A sufficient condition for  $\mathbf{q}_h$  to be chosen, for any  $\delta$ , is that the decoy lies on a lower rational indifference curve than  $\mathbf{q}_h$ . This is guaranteed for dominated  $\mathbf{q}_d$ , and by continuity for some  $\mathbf{q}_d$  with  $q_d > q_h$  as well. In fact, given the assumptions that  $\theta_1 = \theta_2$  and that  $\mathbf{q}_h$  provides positive utility, this holds for all decoys in regime a).

Consider now case *ii*). Since  $q_l/p_l < q_h/p_h$ , so that good  $\mathbf{q}_h$  has a relatively high quality price ratio, quality is salient in  $\{\mathbf{q}_l, \mathbf{q}_h\}$  and  $\mathbf{q}_h$  is chosen. Given the constraints  $\bar{q} < q_h$  and  $\bar{p} < p_h$ , adding a decoy  $\mathbf{q}_d$  to the choice set makes  $\mathbf{q}_h$  price salient when it increases the quality price ratio of the average good to the level where  $q_h/p_h < \bar{q}/\bar{p}$ . However, this is excluded by the condition that the decoy is a “bad deal”, namely  $q_d/p_d < \max\{q_l/p_l, q_h/p_h\}$ .

■

**Proposition 4** *Let all goods in a choice set be located on a rational indifference curve with utility level  $u$ . Then:*

- 1) if  $\bar{q}_1 < \frac{u}{2\theta_1}$ , the local thinker chooses the good  $\mathbf{q}_k$  with highest  $q_{1k}$  subject to  $q_{1k} \cdot q_{2k} > \bar{q}_1 \cdot \bar{q}_2$
- 2) if  $\bar{q}_1 > \frac{u}{2\theta_1}$ , the local thinker chooses the good  $\mathbf{q}_k$  with highest  $q_{2k}$  subject to  $q_{1k} \cdot q_{2k} > \bar{q}_1 \cdot \bar{q}_2$
- 3) if  $\bar{q}_1 = \frac{u}{2\theta_1}$ , the local thinker chooses a good “sufficiently close” to  $(\bar{q}_1, \bar{q}_2)$ .

**Proof.** Consider an indifference curve characterized by  $u(q_1, q_2) = \theta_1 q_1 + \theta_2 q_2 = u$ . The average good  $\bar{\mathbf{q}}$  also lies on the indifference curve, and good  $\mathbf{q}_k$ 's advantage relative to  $\bar{\mathbf{q}}$  is salient whenever  $q_{1k} \cdot q_{2k} > \bar{q}_1 \cdot \bar{q}_2$ . Since qualities are non-negative,  $q_i$  ranges from 0 to  $\frac{u}{\theta_i}$ , and the average qualities along the indifference curve are  $(\frac{u}{2\theta_1}, \frac{u}{2\theta_2})$ . This identifies the central point of the indifference curve, which maximizes the product of qualities, namely  $q_{1k} \cdot q_{2k} \leq \frac{u}{2\theta_1} \cdot \frac{u}{2\theta_2}$  for all  $k$ .

If  $\bar{q}_1$  is low,  $\bar{q}_1 < \frac{u}{2\theta_1}$ , then goods which are closer to the central point satisfy  $q_{1k} \cdot q_{2k} >$

$\bar{q}_1 \cdot \bar{q}_2$ . Thus their relative advantage, namely quality along dimension 1, is salient. If  $q_{1k} > \frac{u}{2\theta_1}$ , then these goods are overvalued, while if  $q_{1k} < \frac{u}{2\theta_1}$  they are undervalued.

For goods farther from the central point,  $q_{1k} \cdot q_{2k} < \bar{q}_1 \cdot \bar{q}_2$ , their relative disadvantage is salient. Goods with low  $q_1$  ( $q_{1k} < \bar{q}_1$ ) have  $q_1$  salient. Since  $\bar{q}_1 < \frac{u}{2\theta_1}$ , these goods are undervalued. Goods with high  $q_1$  and low  $q_2$  ( $q_{1k} > \bar{q}_1 \cdot \bar{q}_2 / q_{2k}$ ) have  $q_2$  salient. Since  $\bar{q}_2 < \frac{u}{2\theta_2}$ , these goods are also undervalued. Crucially, by construction goods far from the central point are more undervalued than those close to the central point. Case 1) follows, and case 2) follows from an identical reasoning.

Consider now case 3), and suppose for simplicity that the average good  $\bar{\mathbf{q}} = \left( \frac{u}{2\theta_1}, \frac{u}{2\theta_2} \right)$  is not part of the choice set. In this case, all available goods satisfy  $q_{1k} \cdot q_{2k} < \bar{q}_1 \cdot \bar{q}_2$ . Thus all goods' relative disadvantages relative to the average are salient. In this case, relative disadvantages coincide with absolute disadvantages, so all goods are undervalued. The least undervalued goods are the ones closest to the average (central) good itself. ■

This tendency for the local thinker to “go to the middle” in quality-quality space can generate violations of IIA, leading in particular to the so called compromise effect. Consider a pairwise choice between goods  $\mathbf{q}_1$  and  $\mathbf{q}_2$ , which have equal rational utility  $u$  and specialize in attribute  $q_2$ , that is  $q_{11}, q_{12} < \frac{u}{2\theta_1}$ . Suppose now that  $\mathbf{q}_2$  is less balanced than  $\mathbf{q}_1$  in the sense that  $q_{12} < q_{11}$ . Then  $\mathbf{q}_1$  is chosen because it has higher levels of the salient attribute  $q_1$ . However, by introducing a good  $\mathbf{q}_3$  which is even less balanced than  $\mathbf{q}_2$  but yields a similar rational utility, it is often possible to transform in the consumer's eyes the previously unbalanced  $\mathbf{q}_2$  into a middle of the road compromise, rendering  $\mathbf{q}_2$ 's advantage  $q_{22}$  salient. In particular,

**Corollary 1** *Let goods  $\mathbf{q}_1, \mathbf{q}_2$  have rational utility  $u$  and satisfy  $\frac{1}{2}q_{11} < q_{12} < q_{11} \leq \frac{u}{2\theta_1}$ . Then: i) the balanced good  $\mathbf{q}_1$  is chosen from the choice set  $\{\mathbf{q}_1, \mathbf{q}_2\}$ , and ii) there exists an extreme good  $\mathbf{q}_3$ , satisfying  $q_{13} \leq \frac{1}{2}q_{11}$  and with rational utility arbitrarily close to  $u$ , such that the intermediate good  $\mathbf{q}_2$  is chosen from  $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ .*

**Proof.** i) In the pairwise choice between  $\mathbf{q}_1$  and  $\mathbf{q}_2$ , the former is the more balanced good. Namely, the average good satisfies  $\bar{q}_1 < \frac{u}{2\theta_1}$  and  $\mathbf{q}_1$  satisfies  $q_{1k} \cdot q_{2k} > \bar{q}_1 \cdot \bar{q}_2$ . Proposition 4 implies that  $\mathbf{q}_1$  is chosen.

ii) Since  $q_{12} < q_{11} < \frac{u}{2\theta_1}$ , both goods have upsides along dimension 2, which is also the advantage of good  $\mathbf{q}_2$  relative to  $\mathbf{q}_1$ . As a consequence, if quality  $q_2$  becomes salient for good  $\mathbf{q}_2$  its valuation becomes larger than that of  $\mathbf{q}_1$ . This can be achieved by adding a third good  $\mathbf{q}_3$  with a sufficiently low quality  $q_3$  such that the resulting average along quality  $q_1$  is close to  $q_{12}$ . The condition  $\frac{1}{2}q_{11} < q_{12}$  ensures that this is possible when qualities are non negative. Note that  $\mathbf{q}_2$  is chosen over  $\mathbf{q}_3$  if the latter has an upside lower than  $q_{22}$ , or if that upside is not salient. ■

The compromise effect for goods with two quality attributes is very similar to the compromise and decoy effects detailed previously in the context of quality/price space, in that adding an irrelevant alternative can render the strength of the intermediate good salient. Unlike in the case of decoys, this effect does not necessarily rely on a dominated (or unattractive) irrelevant alternative. It relies on an irrelevant alternative that is sufficiently unbalanced to make the previously rejected good be perceived as a good compromise. It is by generating a taste for balance that salience creates a compromise effect.

**Proposition 5** *The consumer's willingness to pay for  $q$  depends on the price  $\hat{p}$  as follows:*

$$WTP(q|\mathbf{C}) = \begin{cases} \delta q & \text{if } \hat{p} \leq \delta q \\ \hat{p} & \text{if } \delta q < \hat{p} \leq \frac{1}{\delta} \cdot q \\ q/\delta & \text{if } \frac{1}{\delta} \cdot q < \hat{p} \leq \frac{1}{\delta} \cdot q \cdot \frac{1}{k(N)} \\ \delta q & \text{if } \hat{p} > \frac{1}{\delta} \cdot q \cdot \frac{1}{k(N)} \end{cases} \quad (27)$$

where  $k(N) = \frac{N(N+1)}{(N+2)^2 - (N+1)} < 1$ . As  $\delta \rightarrow 1$ , the willingness to pay tends to  $q$  and becomes independent of context  $\hat{p}$ .

**Proof.** The average quality in  $\mathbf{C} \cup \{(q, -p)\}$  is  $\bar{q} = q \frac{N+1}{N+2}$ . The average price is

$$\bar{p} = \frac{1}{N+2} \left[ p + \sum_{i=1}^N p_i \right] = \frac{1}{N+2} [p + \hat{p}N]$$

where  $\hat{p} = \frac{1}{N} \sum_{i=1}^N p_i$ . Thus, the salience of quality and price of good  $(q, -p)$  are, respectively

$$\sigma \left( 1, \frac{N+1}{N+2} \right), \quad \sigma \left( 1, \frac{1}{N+2} \left[ 1 + \frac{\hat{p}}{p} N \right] \right)$$

Set  $k(N) = \frac{N(N+1)}{(N+2)^2 - (N+1)} < 1$ . It follows that quality is salient when

$$p \in (\hat{p} \cdot k(N), \hat{p}), \quad \text{or} \quad \hat{p} \in \left( p, \frac{p}{k(N)} \right)$$

and accordingly, price is salient when

$$\hat{p} < p \quad \text{and} \quad \hat{p} > \frac{p}{k(N)}$$

Recall the definition of willingness to pay:

$$\text{WTP}(q|\mathbf{C}) = \sup p \text{ s.t. } u^{LT}(\mathbf{q}|\mathbf{C} \cup \{(q, -p)\}) \geq u^{LT}(\mathbf{q}_0|\mathbf{C} \cup \{(q, -p)\}).$$

Consider first the case where the good is expensive relative to the reference price,  $\hat{p} < p$ . Then price is salient, so the consumer buys the good if and only if its discounted quality is sufficiently high,  $\delta q \geq p$ . Thus,  $\text{WTP} = \delta q$  whenever  $\hat{p} < \delta q$ .

Consider now the case where quality is salient, so the good is cheaper than the reference price,  $\hat{p} \geq p$ , but the price is not too low. If quality is salient, the consumer buys the good as long as its inflated quality is above its price,  $\frac{q}{\delta} \geq p$ . Thus, price can be jacked up all the way to  $q/\delta$ , as long as it does not change the salience ranking:  $\text{WTP} = \max \left\{ \frac{q}{\delta}, \hat{p} \right\}$ . As a consequence, for  $\hat{p} \leq \frac{q}{\delta}$ ,  $\text{WTP} = \hat{p}$ . For  $\frac{q}{\delta} \frac{1}{k(N)} > \hat{p} > \frac{q}{\delta}$ , we find  $\text{WTP} = \frac{q}{\delta}$ .

Finally, consider the case  $\hat{p} > \frac{q}{\delta} \frac{1}{k(N)}$ . Now the reference price is so high that even at the highest possible price for the good, namely  $q/\delta$ , its price is salient. As a result,  $\text{WTP}$  goes back down to  $\delta q$ . ■

Take an evoked set  $\mathbf{C}$  having  $N > 1$  elements and divide it into two subsets  $\mathbf{C}_F$  and  $\mathbf{C}_C$  such that  $\mathbf{C}_F \cap \mathbf{C}_C = \emptyset$  and  $\mathbf{C}_F \cup \mathbf{C}_C = \mathbf{C}$ . Denote by  $\omega$  the fraction of goods that belong to  $\mathbf{C}_F$ , by  $\bar{p}$  the average price of goods in  $\mathbf{C}$  and by  $\bar{p}_X$  the average price in subset  $\mathbf{C}_X$ ,  $X = F, C$ . We then prove:

**Proposition 6** Denote by  $p_{\mathbf{C}}^{\max}$  the highest price in  $\mathbf{C}_{\mathbf{C}}$ . Then, a marginal increase in the prices of all goods in  $\mathbf{C}_{\mathbf{C}}$  (holding constant the prices in  $\mathbf{C}_{\mathbf{F}}$ ) boosts the salience of price for the most expensive goods in  $\mathbf{C}_{\mathbf{C}}$  only if  $\omega > 0$  and  $p_{\mathbf{C}}^{\max} > \bar{p}$  and provided:

$$\frac{p_{\mathbf{C}}^{\max} - \bar{p}_{\mathbf{C}}}{\bar{p}_{\mathbf{F}}} < \frac{\omega}{1 - \omega}. \quad (28)$$

If  $\omega = 0$ , the salience of price decreases for all goods in  $\mathbf{C}_{\mathbf{C}}$ .

**Proof.** Suppose the prices of all goods in  $\mathbf{C}_{\mathbf{C}}$  are shifted by a small  $\gamma > 0$ . Then the average price in  $\mathbf{C}$  shifts by  $(1 - \omega)\gamma$ , where  $(1 - \omega)$  is the share of goods in  $\mathbf{C}_{\mathbf{C}}$ . Consider the salience of price for goods in  $\mathbf{C}_{\mathbf{C}}$  which have price  $\hat{p}$ , i.e.  $\sigma(\hat{p} + \gamma, \bar{p} + (1 - \omega)\gamma)$ . Diminishing sensitivity implies that salience decreases in  $\gamma$  whenever  $\omega = 0$ , or when  $\omega > 0$  but  $\hat{p} < \bar{p}$ . This is because in either situation the average payoff level increases but the difference between payoffs weakly decreases.

For salience to increase in  $\gamma$ , it is necessary that the difference in payoffs increases as well, so that the ordering property of salience may dominate over diminishing sensitivity. A necessary condition for salience to increase is thus that  $\omega > 0$  and  $\hat{p} > \bar{p}$ . The precise trade-off between payoff level and payoff difference (i.e. between diminishing sensitivity and ordering) is not pinned down by the properties of salience considered in Definition 1. However, assuming homogeneity of degree zero, we get that

$$\partial_{\gamma} \sigma(\hat{p} + \gamma, \bar{p} + (1 - \omega)\gamma) > 0 \Leftrightarrow \partial_{\gamma} \frac{\hat{p} + \gamma}{\bar{p} + (1 - \omega)\gamma} > 0$$

Replacing  $\hat{p}$  for  $p_{\mathbf{C}}^{\max}$ , we get the condition in the proposition. ■

**Proposition 7** The retailer can charge a sale price  $p_s = q/\delta$  and still have the customer buy the product by setting any full price in the interval  $p_f \in (q/\delta, 7q/2\delta)$ .

**Proof.** As in the text, consider the evoked set  $\mathbf{C}_{\text{sale}} = \{(0, 0), (q, p_s), (q, p_f)\}$ . Consider the evaluation of the good on sale,  $(q, p_s)$ . The salience of its quality is (using homogeneity of degree zero)  $\sigma(q, \frac{2q}{3}) = \sigma(1, \frac{2}{3})$ . The salience of its price is  $\sigma(p_s, \frac{p_s + p_f}{3}) = \sigma\left(1, \frac{1 + \frac{p_f}{p_s}}{3}\right)$ .

Therefore, quality is more salient than price as long as  $\frac{p_f}{p_s} \in (1, \frac{7}{2})$ . In fact, if  $p_f$  is much higher than  $p_s$ , then the price difference among them becomes salient again. For ratios  $p_f/p_s$  at which quality is salient, the willingness to pay is  $p_s = q/\delta$ , from which the result follows.

■

**Proposition 8** *The store can always make the high quality good quality salient, and have the consumer choose it over the low quality good, by holding a sale on  $\mathbf{q}_h$  where the full price  $p_{hf}$  is suitably chosen. In contrast, a sale is ineffectual for the low quality good: if the consumer chooses  $\mathbf{q}_h$  in the absence of a sale, there exists no full price  $p_{fl} \in (p_l, p_h)$  for  $\mathbf{q}_l$  that makes  $\mathbf{q}_h$  price salient, and  $\mathbf{q}_l$  be chosen, in the context of the sale.*

**Proof.** The store can always make the high quality good quality salient by holding a sale with a full price  $p_{fh} = 3p_h - p_l$  (in which case  $p_h$  coincides with the average quality in the choice set).

Instead, by holding a sale on the low quality good, the store lowers the quality-price ratio of the reference good. Thus, as long as  $p_{fl} < p_h$ , this makes it easier for  $\mathbf{q}_h$  to be quality salient, as it has both higher quality and price and also higher quality to price ratio compared to the reference good. In particular, if in the absence of a sale  $\mathbf{q}_h$  is quality salient and chosen by the consumer, holding the sale for  $\mathbf{q}_l$  has no effect on the consumer's choice.

■

## A.2 Continuous Salience Distorsions

The discontinuous nature of salience distortions can give rise to a non-monotonic relation between quality and utility, and similarly between price and utility. When  $q < \bar{q}$ , increasing  $q$  can make price salient and decrease utility. When  $p < \bar{p}$ , increasing  $p$  can make quality salient and increase utility. These effects can be corrected if salience distortions are made continuous. Formally, let:

$$u^{LT}(q, p) = \delta[\sigma(q, \bar{q})] \cdot q - \delta[\sigma(p, \bar{p})] \cdot p$$

where  $\delta[\sigma(\cdot, \cdot)]$  is some increasing function of salience. Then impose  $d_q u^{LT}(q, p) > 0$  and  $d_p u^{LT}(q, p) < 0$ . This yields

$$\partial_q u^{LT}(q, p) > 0 \Leftrightarrow \delta'[\sigma(q, \bar{q})] \cdot \sigma'(q, \bar{q}) \cdot q + \delta[\sigma(q, \bar{q})] > 0 \quad (29)$$

By assumption  $\delta[\sigma(q, \bar{q})] > 0$  and the logic of overweighting salient states implies that  $\delta'[\sigma(q, \bar{q})] > 0$ . Thus, a sufficient condition for evaluation to increase in quality is that

$$\sigma'(q, \bar{q}) = \partial_q \sigma(q, \bar{q}) + \frac{1}{N} \partial_{\bar{q}} \sigma(q, \bar{q}) > 0$$

From the ordering property of salience, the first term is negative only when  $q < \bar{q}$ , while the second term is negative only when  $q > \bar{q}$ . The diminishing sensitivity property of salience ensures that when  $q > \bar{q}$  the above inequality always holds. This is because increasing quality when quality is above average increases its salience and therefore perceived utility. On the other hand, that inequality never holds when  $q < \bar{q}$ . Thus the constraint from (??) is that salience does not fall too fast as  $q$  approaches  $\bar{q}$  from below.

To see this, assume that  $\delta[\sigma(q, \bar{q})] = \exp[(1 - \delta)\sigma(q, \bar{q})]$ , where now  $\delta \in (0, 1]$ , and that the salience function satisfies homogeneity of degree zero. Then from (29) we get:

$$(1 - \delta) \sigma' \left( \frac{q}{\bar{q}}, 1 \right) \cdot \left( 1 - \frac{q}{N\bar{q}} \right) \frac{q}{\bar{q}} + 1 > 0$$

where  $N$  is the number of options in the evoked set. As advertised, this condition restricts how fast salience falls (how negative  $\sigma'$  is) as  $q$  approaches  $\bar{q}$  from below. It is easy to see that, for all  $\delta < 1$ , this condition is satisfied with the standard salience function  $\sigma(q, \bar{q}) = |q - \bar{q}| / (|q| + |\bar{q}|)$ . Naturally, this also ensures that  $d_p u(q, p) < 0$ .

Monotonicity of evaluation also ensures that dominant goods have lower evaluation than the corresponding dominating goods, and are never chosen. In fact, keeping the reference good constant, monotonicity implies that moving a good from a dominated position to a dominating position strictly increases its evaluation.

## References:

- Alchian, Armen and William Allen. 1964. *Exchange and production: competition, coordination and control*. Belmont, CA: Wadsworth Pub. Co.
- Azar, Ofer. 2007. “Relative thinking theory.” *Journal of Socio-Economics*, 36 (1): 1 – 14.
- Barseghyan, Levon, Francesca Molinari, Ted O’Donoghue, and Josh Teitelbaum. 2011. “The Nature of Risk Preferences: Evidence from Insurance Choices”, SSRN Working Paper 1646520.
- Blattberg, Robert and Kenneth Wisniewski. 1989. “Price-Induced Patterns of Competition, *Marketing Science* 8(4), 291 – 309.
- Bodner, Ronit and Drazen Prelec. 1994. “The Centroid Model of Context Dependent Choice”, Unpublished Manuscript, MIT.
- Bordalo, Pedro. 2011. “Salience Theory and Context Effects”, Unpublished Manuscript Royal Holloway, University of London.
- Bordalo, Pedro, Nicola Gennaioli and Andrei Shleifer. 2012. “Salience Theory of Choice under Risk”, *Quarterly Journal of Economics*, Forthcoming.
- Bordalo, Pedro, Nicola Gennaioli and Andrei Shleifer. 2012b. “Salience and Experimental Tests of the Endowment Effect”, *American Economic Review, Papers & Proceedings*, Forthcoming.
- Cunningham, Tom. 2011. “Comparisons and Choice”, Unpublished Manuscript, Harvard University.
- Eliaz, Kfir and Ran Spiegler. 2010. “Consideration sets and competitive marketing”, *Review of Economic Studies*, 78 (1): 235 – 262.
- Feigenson, Lisa, Stanislas Dehaene and Elizabeth Spelke. 2004. “Core systems of number”, *Trends in Cognitive Sciences*, 8 (7): 307 – 314
- Gabaix, Xavier. 2011. “A Sparsity-Based Model of Bounded Rationality”, NBER working paper No. 16911.



- Gennaioli, Nicola and Andrei Shleifer. 2010. “What Comes to Mind”, *Quarterly Journal of Economics*, 125 (4): 1399–1433.
- Goeree, Michelle. 2008. “Limited Information and Advertising in the U.S. Personal Computer Industry”, *Econometrica*, 76 (5): 1017 – 1074.
- Hastings, Justine and Jesse Shapiro. 2011. “Mental Accounting and Consumer Choice: Evidence from Commodity Price Shocks”. Unpublished Manuscript, University of Chicago Booth School of Business.
- Heidhues, Paul and Botond Koszegi. 2008. “Competition and price variation when consumers are loss averse”. *The American Economic Review*, 98 (4): 1245 – 1268
- Huber, Joel, John Payne and Christopher Puto. 1982. “Adding asymmetrically dominated alternatives: violations of the regularity and similarity hypothesis”, *Journal of Consumer Research*, 9 (1): 90 – 98.
- Jaeger, David A. and Karl Storchmann. 2011. “Wine Retail Price Dispersion in the United States: Searching for Expensive Wines?”, *American Economic Review: Papers & Proceedings*, 101 (3): 136 – 141.
- Jahedi, Salar. 2011. “A taste for bargains”, Unpublished Manuscript, University of Arkansas
- Kahneman, Daniel and Dale T. Miller. 1986. “Norm theory: Comparing reality to its alternatives”, *Psychological Review*, 93: 136–153.
- Kahneman, Daniel and Amos Tversky. 1979. “Prospect Theory: an Analysis of Decision under Risk”, *Econometrica*, 47 (2): 263– 292.
- Kahneman, Daniel and Amos Tversky. 1984. “Choices, Values and Frames”, *American Psychologist*, 39: 341 – 350.
- Kamenica, Emir. 2008. “Contextual Inference in markets: on the informational content of product lines.” *The American Economic Review*, 98 (5): 2127 – 2149
- Keeney, Ralph, and Howard Raiffa. 1976. *Decisions with multiple objectives: Preferences and Value Trade-Offs*. Cambridge University Press.
- Koszegi, Botond and Matthew Rabin. 2006. “A Model of Reference-Dependent Preferences”, *Quarterly Journal of Economics*, 121 (4): 1133 – 1165.

- Koszegi, Botond and Matthew Rabin. 2007. “Reference-Dependent Risk Attitudes”, *American Economic Review*, 7 (4): 1047 – 1073.
- Koszegi, Botond and Adam Szeidl. 2011. “A Model of Focusing in Economic Choice”, unpublished manuscript, UC-Berkeley.
- Lazear, E. 1986. “Retail pricing and clearance sales”. *American Economic Review*, 76 (1): 14 – 32.
- Lynch, John and Dan Ariely. 2000. “Wine online: search costs affect competition on price, quality and distribution.”, *Marketing Science*, 19 (1): 83 – 103
- Manzini, Paola and Marco Mariotti. 2007. “Sequentially Rationalizable Choice”, *American Economic Review*, 97 (5): 1824 – 1839.
- Masatlioglu, Yusufcan, Daisuke Nakajima and Erkut Ozbay. 2010. “Revealed Attention”, forthcoming in *American Economic Review*.
- Mullainathan, Sendhil. 2002. “Thinking through categories”, unpublished manuscript, Harvard University.
- Ortmeyer, Gwen, John Quelch and Walter Salmon. 1991. “Restoring credibility to retail pricing”, *Sloan Management Review*, 33 (1): 55 – 66.
- Roberts, John and James Lattin. 1997. “Consideration: Review of Research and Prospects for Future insights”, *Journal of Marketing Research*, 34 (3): 406 – 410.
- Rosen, S . 1974. “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”, *The Journal of Political Economy*, 82 (1): 34 – 55.
- Rubinstein, Ariel and Yuval Salant. 2008. “(A,f): choice with frames”, *Review of Economic Studies*, 75 (4): 1287 – 1296.
- Savage, Daniel. 1954. *The Foundation of Statistics*. New York: Dover Publications.
- Schwartzstein, Joshua. 2012. “Selective attention and learning”, Unpublished Manuscript, Harvard University.
- Simonson, Itamar. 1989. “Choice based on reasons: the case of the attraction and compromise effects”, *Journal of Consumer Research*, 16 (2): 158–174.

- Simonson, Itamar, and Amos Tversky. 1992. "Choice in context: Tradeoff contrast and extremeness aversion", *Journal of Marketing Research*, 29 (3): 281 – 295.
- Spiegler, Ran. 2011. *Bounded rationality and Industrial Organization*. Oxford University Press.
- Sobel, Joel. 1984. "The timing of sales". *Review of Economic Studies*, 51 (3): 353 – 368.
- Sydnor, Justin. 2010. "(Over)insuring Modest Risks". *American Economic Journal: Applied Economics*, 2: 177 – 199.
- Taylor, Shelley and Suzanne Thompson. 1982. "Stalking the elusive vividness effect." *Psychological Review*, 89 (2): 155 – 181.
- Thaler, Richard. 1980. "Toward a Positive Theory of Consumer Choice", *Journal of Economic Behavior and Organization*, 1 (1): 39 – 60.
- Thaler, Richard. 1985. "Mental Accounting and Consumer Choice", *Marketing Science*, 4 (3): 199 – 214.
- Thaler, Richard. 1999. "Mental Accounting Matters", *Journal of Behavioral Decision Making*, 12: 183 – 206.
- Tversky, Amos and Daniel Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference-Dependent Model", *Quarterly Journal of Economics*, 106 (4): 1039 – 1061
- Tversky, Amos and Itamar Simonson. 1993. "Context Dependent Choices", *Marketing Science*, 39 (10): 1179 – 1189
- Varian, Hal. 1980. "A model of sales". *American Economic Review*, 70: 651 – 659.
- Wernerfelt, Birger. 1995. "A rational reconstruction of the compromise effect: using market data to infer utilities", *Journal of Consumer Research*, 21 (4): 627 – 633.
- Woodford, Michael. 2012. "Inattentive valuation and Reference-Dependent Choice", Unpublished Manuscript, Columbia University.